Nordic Journal of Language Teaching and Learning

Exploring teaching practices to stimulate meaningful language use in the Frisian trilingual primary school

Eabele Tjepkema*
NHL Stenden University of Applied Sciences
eabele.tjepkema@nhlstenden.com

Joana Duarte
University of Amsterdam, NHL Stenden University of Applied Sciences & University of Groningen
joana.da.silveira.duarte@nhlstenden.com

Abstract

In Frisian trilingual primary schools, Dutch, Frisian and English are taught through content and language integrated learning (CLIL). CLIL is often characterized as an umbrella term, which can be carried out in various ways, and including regional and/or minority languages which are used to different extents in the out-of-school context (Cenoz et al., 2014). Through the planning of meaning-focused communicative tasks (Ellis, 2009a), using different formats (Llinares & Dalton-Puffer, 2015), the students' language use can be stimulated by co-construction of knowledge. However, little is known about how CLIL is implemented in the Frisian trilingual primary schools and how these practices are related to development in the three languages.

This paper addresses the following research question: Which meaning-focused teaching practices to stimulate target language use can be observed in the three target language lessons (Dutch, Frisian and English) at the end of the Frisian trilingual primary school (students aged 10 - 12 years), and how are these related to student language development in these languages? Data were collected by event sampling observations of 51 content-lessons in the three target languages in seven trilingual primary schools (n = 165 students), and for which an observation instrument was developed (Tjepkema, 2021) focusing on the CLIL teaching practices. Student language development in the three target languages was assessed using vocabulary and reading

Nordic Journal of Language Teaching and Learning, Vol. 11, No. 3, 2023. ISSN: 2703-8629.
https://doi.org/10.46364/njltl.v11i2.1175
*Corresponding Author

comprehension test. Taking a dynamic systems perspective (De Bot et al. 2007), quantitative analysis included the relation between the teaching practices to stimulate meaningful language use in the three languages and student language development in these languages. Results show that higher occurrences of the stimulation of meaningful language use are significantly related only to Frisian language development in terms of vocabulary and reading comprehension. Furthermore, there is evidence for a relation between the stimulation of meaningful language use in the Dutch lessons and Dutch reading comprehension development. However, no relation was found between meaningful use of English in English CLIL lessons and the students' development of English.

Implications of our results for CLIL and bilingual education are related to the importance of developing teacher competences to stimulate meaningful language in content-based minority language education. Other implications for future research and policy making in multilingual education systems are discussed.

***Keywords**:*

trilingual education, CLIL, minority languages, content-based language teaching, task-based language teaching, Frisian

1. Introduction

Fryslân, a bilingual region in the North of the Netherlands, has a population of nearly 650,000 inhabitants. Frisian was the dominant language spoken in this province until the 1960's (Klinkenberg et al., 2018), but, as a result of both migration and a reduction of its relative isolation, the position of Frisian is changing and its use is declining in everyday-life and education (Gorter, 2005; Klinkenberg et al., 2018). Currently, Frisian is mainly spoken as a first language in the region's rural areas. Furthermore, there is an increase in the use of Dutch and migrant languages in the more urban areas of the region, meaning that Frisian is also increasingly spoken as an additional language. Further, English is growingly used in everyday life in the media and tourism (Gorter, 2005). From 1997 onwards, a growing number of Frisian primary schools adopted a model of trilingual education, in which the use of Dutch, Frisian and English is stimulated through 'interactive language use' in meaningful activities related to Content and Language Integrated Learning (CLIL) (Van Ruijven & Ytsma, 2008). Nowadays, about 18% of all primary schools in the region are trilingual (Duarte & Günther-van der Meij, 2018). Regardless of the ratios attributed to the different languages of instruction, little is known about the CLIL-pedagogies implemented to stimulate communicative language use in the three target languages, and how such practices are related to pupils' language development. In this paper, we explore the CLIL-strategies focused on meaning, i.e. concerning 'processing the semantic and pragmatic meaning of utterances' (Ellis, 2009a, p. 223), that can be observed in the lessons related to the three target languages, and how the variation in the focus on meaning is related to student language development in these languages. In addition, we identify CLIL-teaching methods that stimulate target language use and analyse how these are related to student language development in the three target languages.

Language development is often stimulated in communicative and meaning focused tasks (Bruton, 2005; Bygate, 2016b; Ellis, 2014; Ortega, 2015). According to Ellis (2000; 2014), while exercises are form focused, tasks are meaning-focused and aim to achieve communicative abilities by engaging students in natural communicative activities. Following Candlin (2009), tasks promote an optimal environment for communicative language learning, as in tasks students are encouraged to actively respond, explore, and participate in meaningful communication with fellow students. Task-Based Language Teaching (TBLT) is a relatively recent content-based approach, in which language use and language development are stimulated in communicative and meaning focused tasks (Bruton, 2005; Bygate, 2016b; Ellis, 2014;

Ortega, 2015). Similar to other types of content-based education, in TBLT the stimulation of interaction is related to the provision of input, the eliciting of modified output, and the provision of corrective feedback (Pica, 2005; Robinson, 2011). As in TBLT much research has been related to the stimulation of language use by task-design (Bygate, 2016a), this paper is concerned with TBLT as a teaching approach to stimulate the focus on meaning, within the wider context of Content and Language Integrated Learning (CLIL) in Frisian trilingual schools.

There is an array of studies on CLIL-implementation and its effects, for example on oral proficiency and didactic task attainment (Pérez-Cañado, 2018; Ruiz de Zarobe, 2013), on understanding classroom content through scaffolding (Mahan, 2022; Dalton-Puffer, 2007), and on emotional and cognitive benefits. However, many CLIL studies lack a focus on process and the micro-dimensions of the classroom as they concentrate on language outcomes and using test scores, questionnaires, and standardised measures (Ruiz de Zarobe, 2011).

In this paper, we therefore use observational data that allows us to link the CLIL setting in which Dutch, Frisian and English classes take place within Frisian trilingual schools to the micro-level of interaction within task-based language learning and how this is related to language development in the target languages. We address the following research questions:

RQ1

a Which CLIL-strategies aimed at focus on meaning can be observed in the lessons for the three target languages?

b How is the variation in the focus on meaning related to student language development in these languages?

RQ2

a Which CLIL-teaching methods that stimulate target language use can be observed?

b How are these observations related to student language development in the three target languages?

2. Theoretical background

2.1 Task-based language teaching in CLIL contexts to promote language development

2.1.1 CLIL to stimulate additional language use in school

Since its development in the 1990s, CLIL has evolved to become part of the educational literature as a dual-focused and multi-dimensional educational approach for initially foreign language learning (Coyle, 2002), in which incidental and implicit target language learning takes place in meaning focused contexts (Dalton-Puffer et al., 2010). However, CLIL can include different languages and can be carried out in various ways. Nowadays it is often characterised as an umbrella term (Cenoz et al., 2014; Dalton-Puffer et al., 2014; Marsh, 2008). Moreover, Dalton-Puffer et al. (2022) make a distinction between CLIL as a context for foreign language learning and CLIL as an approach focussing on both language development and academic achievement. Exposure to the target language in - but also outside - the school may also be an important factor (Sylvén, 2013). Based on her findings, Günther-van der Meij (2018) argues that in the Frisian context, English language development is primarily related to exposure to English at home and mainly by media (TV, audio devices, and computers). Tjepkema (2021) also demonstrates that students' Frisian vocabulary development is related to peer language use outside the classroom. In this study, pupils who spoke mainly Dutch in the schoolyard performed less well on Frisian vocabulary development than students who spoke predominantly Frisian or equally Dutch and Frisian in the schoolyard.

As such, Sylvén (2013) also pleas for a better considered planning of CLIL, tuned to the students' needs. Furthermore, it is generally accepted that in CLIL lessons additional languages are used as language of communication and instruction for academic content learning (Cenoz, 2015; Llinares & Pastrana, 2013; Ruiz de Zarobe, 2013). Therefore, CLIL can be characterised as a communicative approach stimulating both content learning and implicit language development (Heras & Lasagabaster, 2015). Reported benefits of CLIL on language development reflect this characterization of CLIL as a communicative approach. For instance, Dalton-Puffer (2011) reports that CLIL students demonstrate larger receptive and productive lexicon, with a wider stylistic range and more appropriate use, a wider range of both lexical and morphosyntactic resources, deployed in more elaborate and with more complex structures, and a higher degree of accuracy. Verspoor and Edelenbos (2009) found that CLIL students not only scored better on the average language skills in English, but also demonstrated more authentic language use

than non-CLIL students. Also, Pérez-Cañado (2012) reports significant higher levels for global communicative competence, receptive skills, speaking fluency, morphology use, fluency of writing, lexical and syntactical complexity of writing in the target language. Furthermore, positive effects for reading comprehension skills are found (Admiraal et al., 2006; Lasagabaster, 2008).

2.1.2 Tasks to stimulate language use

Like CLIL and other types of content-based language teaching, in task-based language teaching language use and language development are stimulated in communicative and meaning-focused tasks (Bruton, 2005; Bygate, 2016b; Ellis, 2014; Ortega, 2015). The task-design and teaching involve the provision of multimodal input, the stimulation of higher order thinking skills, and authentic communication, the stimulation of multimodal output through the use of different genres (Llinares, 2015) or modalities, and the provision of both input- and output-scaffolding for successful elaboration. From a socio-cultural view it is theorised that meaningful and communicative tasks can promote students' engagement (Van Oers, 2012). Within this framework, meaningfulness refers to 'what makes personal sense for the students' or a cultural significance rather than content as rational phenomenon. Therefore, meaningful education should address learning outcomes that have a cultural significance, and thus include the use of authentic (problem-solving) tasks related to 'the real-world'. Concerning authentic tasks, a distinction can be made between pedagogic tasks and real-world tasks, both meant to promote collaborative dialogue (Ellis, 2009a; Bygate, 2016a). Pedagogic tasks are planned to stimulate 'interactional authenticity' or 'normal target language use' in the classroom (Bygate, 2016a). In contrary, real-world tasks are aimed at 'situational authenticity' related to language use expected outside the classroom setting (Bygate, 2016a; Long et al., 2016). For the design of such pedagogic and/or real-world tasks, the following general task criteria listed by Ellis (2009b, p. 223) are widely recognized:

- The primary focus should be on 'meaning' (by which is meant that learners should be mainly concerned with processing the semantic and pragmatic meaning of utterances).
- There should be some kind of 'gap' (i.e. a need to convey information, to express an opinion or to infer meaning).
- Learners should largely have to rely on their own resources (linguistic and non-linguistic) in order to complete the activity.

- There is a clearly defined outcome other than the use of language (i.e. the language serves as the means for achieving the outcome, not as an end in its own right).

Related to the first research question, in pedagogic and/or real-world tasks focus on meaning (FoM) is understood as attention to the learning of non-linguistic knowledge for which linguistic and non-linguistic resources are needed (Ellis, 2000). In the literature, FoM is often framed as the opposite of Focus on Form that can be described as the drawing of the student's attention to linguistic aspects (vocabulary, words, grammatical structures, pragmatic patterns, etc.) of the target language in meaning making or communicative activities (Long, 1991). This Focus on Form takes place incidentally and should not be mixed up with focus on forms, which refers to explicit language teaching with the target language as object (Long, 1991).

Both Bygate (2016a; 2016b) and Skehan (1996; 2009) focus on task planning and/or task structuring to stimulate academic language learning integrating complexity, accuracy, lexis, and fluency. Firstly, Bygate (2016a) suggests that task repetition, defined as 'repetition of the same or similar task' (p. 393), can help students to combine a focus on meaning and form, and to produce more complex language accurately and fluently. Secondly, Skehan's proposed a three-phase lesson structure that includes a pre-task, a during-task and a post-task (Skehan, 1996; 2009; Bygate, 2016a; Van den Branden, 2016). From this structure, the pre-task concerns an introduction task in which the students are informed about the nature and purpose of the during-task, and in which input materials are presented and aims are shared. Regarding the planning of activities in these lesson-phases, Llinares and Dalton-Puffer (2015) lists four different formats that are more or less commonly used in CLIL practices in three European countries (Austria, Finland and Spain): whole-class discussion, group-work discussion, student-presentation, and role-play. The format of interview (an individual oral assessment) was additionally planned as an atypical format for research purposes to stimulate student reflection on content learning. As it shows that all of these five formats have different strengths and engage students in the use of different types of language use, Llinares and Dalton-Puffer (2015) recommend using this range of tasks in the CLIL lessons, and especially to include role-play.

The during-task concerns the second lesson phase. One condition in this phase is that students are working independently on interactive and meaning focused tasks, supported by the circulating teacher. The post-task is the final phase of the task structure in which presentation of the outcomes of the during-task are presented, orally and/or written and in the target language. The planning of a post-task in which students are stimulated to subsequently transcribe their own performance

can affect online-planning a second during-task condition (Skehan, 2016), and promote focus on form (Foster & Skehan, 2013). Online-planning concerns the preparation of the presentation, including preparation of formulation of utterances more carefully with positive effects on complexity and accuracy over fluency (Ellis, 2009a). The use of this three-phase lesson structure can also be related to Hulstijn & Laufer's (2001) Involvement Load Hypothesis (ILH), from which the planning of presentation of learning outcomes in the target language in the third phase can provide the pupils a *need* to use the target language. In case the use of new language is required, the search for new language and the evaluation of (new) language is hypothesised to be stimulated.

2.2 Research on language development of pupils in Frisian trilingual primary schools

Some previous research has been conducted on CLIL in trilingual settings such as Fryslân. Studies suggest that CLIL within trilingual programmes can enhance students' proficiency in the three involved languages, especially in terms of improved (oral) communication skills and cognitive development (Dalton-Puffer, 2017; Lasagabaster & Doiz, 2017; Lorenzo & Casal, 2019; Smit, 2013). However, trilingual CLIL poses unique challenges, such as balancing instructional time among three languages, ensuring linguistic equity, and addressing potential language hierarchy issues. In addition, research often delves into how trilingual CLIL influences students' language identities and attitudes towards each language. Factors such as the language of instruction's prestige and societal perceptions play a role (Dalton-Puffer, 2017; Lasagabaster & Doiz, 2017; Lorenzo & Casal, 2019; Smit, 2013). Applied linguistic research on CLIL practices have predominantly addressed the observation of translanguaging and trans-semiotic practices (including among others L1 use and also other semiotics such as visualisations and gestures for meaning making, and also Systemic Functional Linguistics including for instance the use of different genres for content knowledge construction (Dalton-Puffer et al., 2022).

Against this backdrop, trilingual education was developed for the Frisian setting. With the main aim of improving the quality of education, particularly regarding Frisian and English, the Frisian trilingual project – 'Trijetalige Skoalle' – started in 1997/98 with seven primary schools (Van Ruijven & Ytsma, 2008). One of the main aims for the project was to allow pupils to reach the official attainment targets (a general defined level of ability that a pupil is expected to achieve in every subject at each key stage in the national curriculum) for all languages at the end of primary school (Riemersma & de Vries, 2011).

This research on language learning in the trilingual primary schools shows that more attention to additional language learning in the curriculum comes at no cost for Dutch language development (Van Ruijven & Ytsma, 2008). Students in the trilingual primary school showed higher scores for technical language skills in Frisian (spelling and technical reading skills). In contrast, for both Dutch and English language development, no effects were found. Regarding English language use, students in the trilingual primary school reported feeling more able and comfortable to have a short conversation in English than students from the non-trilingual primary schools. As part of a follow-up research in trilingual primary schools, students in grades 6 and 7 demonstrated better oral English communicative competence than students from a regular primary school. However, no effect for oral English communicative competence was found in grade 8. Moreover, monolingual Dutch speaking students showed better results on oral English communicative competence tests than bilingual (Dutch-Frisian) students (Klinkenberg et al., 2017).

In sum, there has been limited research on CLIL and language development in Frisian trilingual primary schools. So far, only a few longitudinal studies have registered positive results on language development for the Frisian minority language, no negative effects from the reduced instruction on Dutch language development, and better communicative competences for English. However, research so far has not focussed on the types of teaching approaches implemented and how these influence pupils' language development in the three instruction languages. With this paper, and specifically focussing on CLIL settings in 9 trilingual classes, we aim at exploring the CLIL strategies aimed at a) focus on meaning and b) stimulation of target language use to determine their influence on student language development in the three target languages.

3. Methodology

The process of language development can be seen as a complex life-long non-linear process over time in which the target is moving (De Bot, 2016; Ellis, 2008; Larson-Freeman, 2015). In this process, language development is related to both socio-psychological and pedagogical factors, which also develop over time depending on the change of student's needs and opportunities. To monitor both language development and pedagogical practices data were collected during two measurement points in 9 classes (grades 7-8, n=165 students) of seven trilingual primary schools over a period of two school years (2012, T1 and 2014, T2). For the observation of the

teaching practices, data were collected by event sampling observations in 51 CLIL lessons (17 lessons for each of the three target languages). In the following sections, the different instruments will be presented in more detail.

For the coding of these data, an observation instrument was developed. Simultaneously to the analysis of teaching practices, vocabulary and reading comprehension tests in each of the three languages were administered in the same classes. Moreover, questionnaires were used to collect self-reported data information regarding the students' home language and the language spoken with peers.

3.1 Sample

Data were collected from a total of 165 students (aged 10-12 years) in 9 different classes. More information about the sample can be found in Table 1. In terms of teaching time of the three languages, all seven schools surveyed in this study implemented a system in which the three languages were used from grade 1-8, and applied ratios of 20% Frisian, 10% English and the rest Dutch as scheduled lessons in grades 7 and 8[1] (Tjepkema, 2021). Independently of the chosen model, the principle remained that the languages should be consciously and consistently separated (Duarte & Günther-van der Meij, 2018). Also, it was expected from both the teacher and the student to consistently use the language that was scheduled at the time, although in practice many instances of language-mixing were observed (Tjepkema, 2021).

To stimulate development of pedagogic approaches in content-based language lessons, the participating teachers are offered an in-service training programme over the period of two years in between the data-collection. This voluntary in-service training programme included the following topics:

- Provision of rich multimodal input;
- Stimulation of output by promoting both teacher-student and student-student interaction;
- Teacher techniques to provide interactional feedback;
- Pedagogic approaches of translanguaging;
- The three-phase lesson model including a pre-phase, a during-phase and a post-phase
- Stimulation of student reflection on language and language use.

[1] Since 1985 Dutch primary schools include 8 grades (grade 1-8), integrating preschool (or early childhood education ages 4-6), elementary education (ages 6-10) and pre-secondary education (ages 10-12).

The training programmes consisted of eight session of two hours each: four sessions in the first running year and four sessions in the second running year. The training included the demonstration of the topics by showing examples and the planning of interactive formats. Additionally, interactive formats are planned to stimulate the sharing of knowledge, the planning of activities in their own classroom, and sharing practices in the additional sessions.

School	Class	Students	Gender (boy/girl)	Teachers grade 7	Teachers grade 8	Completion in-service training	School size	Region	Appr. population village (2012)
1	1	11[1]	5/6	1	1		144	North-east	1,100
	2	6[1]	6/0	1		All 3 teachers			
2	3	24	11/13	2	2	Teacher DL and FL T1	353	North-east	9,500
	4	23	13/10	2	2				
3	5	28	12/16	3	2		196	South-west	13,000
4	6	7[1]	7/0	1	1	Both teachers	63	South-west	350
5	7	19[1]	13/6	2	1[2]	Teacher DL and FL T1 and T2	106	South-west	1,200
6	8	22	13/8	1	1	Teacher T1 and T2	208	South-west	3,300
7	9	26	15/11	2	2[3]	Teacher DL and FL T1 and T2	149	South-west	2,100

[1] Part of multi-graded classes

[2] Same classroom teacher Dutch and Frisian lesson first and second school year

[3] Same classroom teacher Dutch and Frisian lesson and same teacher English lesson first and second school year

Table 1: Number of students, student gender, number of classroom teachers and school contextual factors (T1 denotes first measurement and T2 denotes second measure)

From the observed lessons, all Dutch target language lessons (Dutch lessons) were timetabled as content lessons, and all but one (measurement 2, class 6) of the English target language lessons (English lessons) were scheduled as English language lessons. Of the Frisian target language lessons in both measurements, all but three lessons were time tabled as content lessons in classes 3, 4 and 5.

Regarding student home language and schoolyard language, all respondents completed the questionnaires. For the observations of teaching practices, in classes 3-9 all lessons could be recorded and used for observation of pedagogic performance and student language use in the language lessons. However, because of too much noise a number of video recordings of lessons in class 2 could not be used for the coding of both teacher behaviour and student language use. Because little information about the differences between the first and T2 of these language lessons could be derived, the video-recordings of classes 1 and 2 are not used for data-analyses.

3.2 Instruments: video-observations

To collect information about the CLIL teaching practices, the target language lessons (the lessons aimed at the stimulation of the use of Dutch, Frisian and English respectively) were video-recorded in each class. In line with Ytsma (2001), those target language lessons can include both content lessons and language lessons with content used for language practice.

For the coding of the teaching practices, an observation instrument was developed. With the use of this instrument, pedagogical performance was processed quantitatively by event sampling (i.e. whether an instance could be observed instead of the frequency of that observed instance).

The footage was coded individually by trained teacher trainees (all taking part in NHL Stenden University's degree programme Multilingualism in an International Perspective with 30 ECTS). The coding training took three sessions in which for a group of four raters, the inter-rater reliability could be increased from $\kappa = .59$ to $\kappa = .98$. A random post measurement confirmed the positive values for the inter-rater reliability.

In the first research question, focus on meaning is understood as attention to the learning of non-linguistic knowledge for which linguistic and non-linguistic resources are needed (Ellis, 2000). For the purposes of this study, focus on meaning is a composite variable including the following items implying actions carried out by the teachers:

- attention to content,
- rich multimodal input,

- introduction of new content,
- attention to pre-knowledge,
- the use of real-world tasks,
- reflection on content,
- demonstration.

Introduction of new knowledge and attention to pre-knowledge were observed in the pre-phase of the video-recorded lessons. Real-world tasks concerning situational authenticity and stimulation of language use expected outside the class setting were observed in the during-phase. Lastly, reflection on content and demonstration of content through oral and/or written presentation were observed in the post-phase (Skehan, 1996). Each aspect of the composite variable was scored as observed (1) or not observed (0), and so the summation of scores for focus on meaning vary in a range from 0 to 7.

The second research question addressed the use of teaching methods to stimulate target language use in student-student interaction. The variable teaching methods is a composite variable and includes pedagogic tasks and teaching formats. From these, pedagogic tasks were operationalised as real-world tasks which provide "situational authenticity" (Bygate, 2016a). Within the sub-variable teaching formats, whole classroom discussion, group work, individual seatwork, role-play and demonstration were distinguished. From these formats, group work and role-play were considered to stimulate student language use more effectively (Llinares & Dalton-Puffer, 2015). The format demonstration is in line with Skehan's (1996) format of public performance and can be related to presentation (Llinares & Dalton-Puffer, 2015). However, demonstration can also include written performance of tasks or formats. Thus, the variable teaching methods includes the following items:

- pedagogic tasks,
- group work,
- role-play,
- demonstration.

As a reminder, 'pedagogic tasks' include activities to stimulate authentic language use, and 'demonstration' include activities to promote oral and/or written presentation of knowledge construction.

Regarding the scoring of this variable, each item was scored as observed (1) or not observed (0), and so the summation of scores for teaching methods to stimulate student-student interaction vary in a range from 0 to 4.

3.3 Instruments: language development

To assess the students' language development for the three target languages, tests were used in T1 and T2 that were especially developed for evaluation in the Dutch primary schools. To collect data regarding language development, the assessment for language development of the three languages was limited to vocabulary and reading comprehension, as proxies for the development of lexis and complexity.

Only the Frisian vocabulary test (α = .89) is part of the evaluation system for the Frisian language (Jonkman, 2017). For Frisian reading comprehension, the test *Cito Begripend lêzen Frysk* (groep 8) Start [Reading comprehension Frisian (grade 8) Start] was used (Hemker & Jongen, 2011). For these tests raw data from T1 and T1 were used.

As to not disturb the continuous evaluation of the participating schools, for the assessment of Dutch language development the latest versions of the regular school tests could not be used. Therefore, for the testing of Dutch Vocabulary Development, *Cito Leeswoordenschat E7* [Cito Reading vocabulary E7] and *Cito Leeswoordenschat M8* [Cito Reading vocabulary M8] were used (Verhoeven & Vermeer, 1995). Regarding the assessment of Reading comprehension in Dutch, *Eduforce DLE-test Begrijpend lezen Groep 7 versie A* and *Eduforce DLE-test Begrijpend lezen Groep 8 versie A* [Reading comprehension tests for grades 7 and 8] were used (De Jong, et al., 2002). As different tests are used for T1 and T2, referential scores were used. As no sufficient referential scores for Dutch reading comprehension could be determined, z-scores are computed for both T1 and T2.

For the testing of both English vocabulary and reading comprehension, the *Cito Me2!Engels (groep 7)* [Me2English, grade 7] the parts Vocabulary and Reading Comprehension were used (Alberts et al., 2007). For T2 (end grade 8) the *Cito Me2!Engels (groep 8)* [Me2English, grade 8] the parts Vocabulary and Reading Comprehension are used (Alberts et al, 2006/2007). As different tests are used for T1 and T2, referential scores were administered.

3.4 Data analyses

Regarding data analyses, for each research question, data were first quantitatively analysed for frequency in T1 and T2, then the analyses of conditional means were executed to test the relation

between the teaching practices and student language development. Data analyses were performed using SPSS (version 27). To determine language development, the results for the T2 were controlled for the results in the T1. In order not to lose variance, in the analyses of co-variance the results of T1 were thus considered as covariates. Similarly, for the assessment of both student- and teacher language use, results of T1 for these variables were used as covariates.

For further co-variables, sociolinguistic questionnaires were used to collect self-reported data regarding the students' home language and the languages spoken with peers. As language use and therefore language development can be related to both gender and social-background (Ross & Kasper, 2013; Tedick et al., 2011), in the testing for the relation between both reading comprehension development and vocabulary development both gender and socio-economic status (SES) were taken into account as co-variables. To control data analyses for SES, schools are asked to share information about the educational level of the student's mother as an approximation for SES. Related to the Dutch educational system, the following categories are included:

- No/primary education only,
- Lower vocational education,
- Intermediate vocational education,
- Higher education.

4. Results

4.1 Findings Focus on Meaning strategies in the three target language lessons (RQ1)

For T1 and T2, the teacher FoM pattern scores in the three target language lessons are presented in figures 1a and 1b respectively.

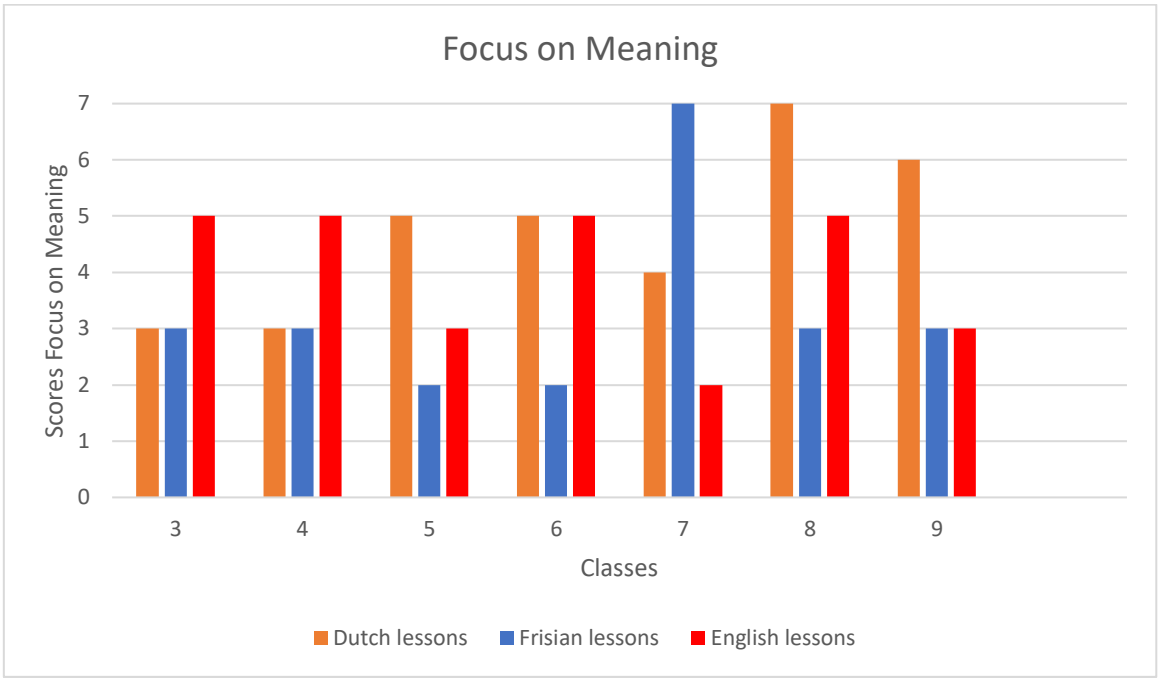

Figure 1a: FoM scores in the three language lessons per class T1

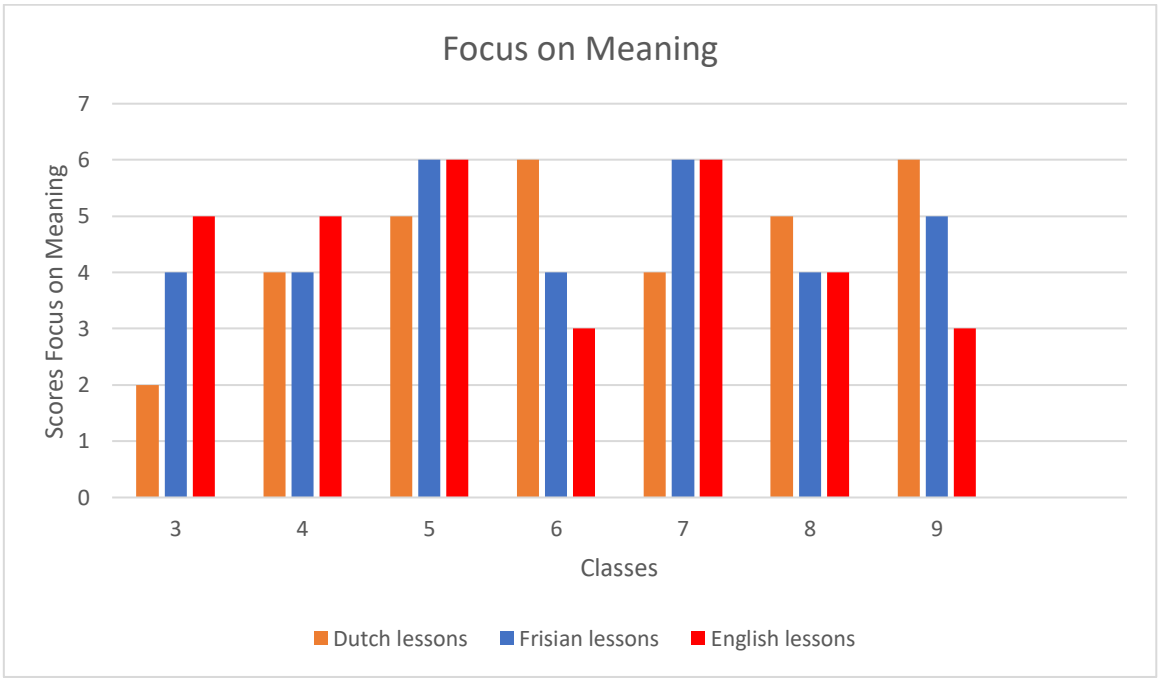

Figure 1b: FoM scores in the three language lessons per class T2

Figures 1a and 1b show differences of FoM scores between the three target language lessons, between classes and between T1 and T2. Table 2 shows that in all three target language lessons differences in FoM scores could be registered between T1 and T2. In addition, the largest increase of FoM mean scores was observed in Frisian lessons and that the smallest increase of FoM mean scores was observed in Dutch lessons.

	Dutch lessons		**Frisian lessons**		**English lessons**	
	T1	T2	T1	T2	T1	T2
Mean	4.4	4.6	3.3	4.7	4.0	4.6
Median	5	5	3	4	5	5
Mode	2[a]	4[a]	3	4	5	3[a]
SD	1.90	1.40	1.70	0.95	1.29	1.27

[a] Multiple modes exist. The smallest is shown.

Table 2: Measures of central tendency scores and standard deviations scores FoM in the three language lessons T1 and T2

Table 2 shows an increase of FoM mean scores in all three target language lessons, that in T1 the highest mean score is observed in Dutch lessons but that in T2 the highest mean score is observed in the Frisian lessons. However, in T1 the most variance for FoM is also observed for Dutch lessons (*SD* 1.90). To anticipate problems with outliers in this small sample, also the median and the mode are evaluated as measures of central tendency. Regarding the languages, in T1 the highest scores FoM are observed in the English lessons. However, in T2 there is not much difference between the three target languages. Regarding the standard deviance, the least variance for FoM between teachers was observed in Frisian lessons (*SD* = 0.95). To examine the FoM scores, Tables 3 and 4 show the scores for the different components.

Blue = similar teacher per class in target language lessons

	Attention to content			Rich Multi-modal input			New Content			Pre-knowledge			Real-world Tasks			Reflection Content			Demonstration		
class	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E
3	x	x	x			x	x		x	x	x	x					x	x			
4	x	x	x			x	x		x	x	x	x					x	x			
5	x	x	x	x			x	x	x	x	x	x							x		
6	x	x	x	x		x		x		x		x	x		x	x		x			
7	x	x	x	x	x		x	x		x	x			x			x			x	x
8	x	x	x	x		x	x	x	x	x	x		x		x	x			x		x
9	x	x	x	x	x		x	x	x	x			x			x					x
Sum	7	7	7	5	2	4	6	5	5	7	5	4	3	1	2	3	3	3	2	1	3

Table 3: Scores FoM components in the three language lessons per class T1

Blue = similar teacher per class in target language lessons

	Attention to content			Rich Multi-modal input			New Content			Pre-knowledge			Real-world Tasks			Reflection Content			Demonstration		
class	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E
3	x	x	x		x	x				x	x	x			x		x				x
4	x	x	x		x	x	x			x	x	x	x		x		x				x
5	x	x	x				x	x	x	x	x	x	x	x	x	x	x	x		x	
6	x	x	x	x	x	x	x	x	x	x	x					x		x	x		
7	x	x	x	x	x	x			x	x	x	x		x	x		x		x	x	x
8	x	x	x	x	x	x		x	x	x		x	x			x				x	x
9	x	x	x	x	x	x	x	x	x	x	x					x			x	x	x
Sum	7	7	7	4	6	6	4	4	5	7	6	5	3	2	4	4	4	2	3	4	5

Table 4: Scores FoM components in the three language lessons per class T2

In Table 2 above it was shown that there is no common pattern for the way FoM scores were related to the different target language lessons. Tables 3 and 4 show that there is also no common pattern for the way FoM scores related to the different target language lessons offered by the same class teacher. Furthermore, results show that in both measurements attention to content is observed in all target language lessons. In both measurements attention to pre-knowledge is most stimulated in Dutch lessons and least in English lessons, and that real-world tasks are least used in Frisian lessons. Regarding the variation in the provided input, Table 3 shows that rich multimodal input was the least observed in T1 of Frisian lessons.

Regarding evaluation of content learning by reflection on content and demonstration, there was little difference between the target language lessons (and no difference between target language lessons in the second measurement). However, demonstration tasks were more used in English lessons and (in T2). Reflection on content was more stimulated in both Dutch and Frisian lessons.

To summarise, the descriptive statistics show that an increase of FoM was observed in all three language lessons. However, the largest increase was observed for the Frisian lessons. With respect to the observation of the different FoM components, it was observed that there was more attention to pre-knowledge in Dutch lessons, and that real-world tasks are least observed in Frisian lessons. Furthermore, demonstration by which students presented the outcomes of the during-phase task in the target language was most observed in English lessons. In contrast, reflection on content learning was the least observed in English lessons.

4.2 The relation between Focus on Meaning and language development (RQ1)

Table 5 presents the mean scores and standard deviation for language development in the three target languages.

	Dutch				**Frisian**				**English**			
	VOC		RC		VOC		RC		VOC		RC	
	Mean (*n*)	*SD*	*Mean* (*n*)	*SD*	*Mean* (*n*)	*SD*	*Mean* (*n*)	*SD*	*Mean* (*n*)	*SD*	*Mean* (*n*)	*SD*
score 1												
score 2	108.4 (23)	34.83	.176 (22)	0.70								
score 3									89.3 (31)	17.75	71.0 (31)	19.49
score 4	115.1 (21)	22.12	-.683 (24)	1.44	22.3 (48)	5.92	46.9 (17)	7.06	83.5 (20)	17.74	72.3 (20)	19.32
score 5	112.9 (48)	15.79	.262 (46)	0.58	29.0 (26)	4.29	45.1 (26)	9.35	86.0 (45)	13.87	74.9 (45)	18.58
score 6	118.6 (32)	26.50	.294 (33)	0.82	22.8 (32)	6.24	48.3 (22)	5.73	87.0 (32)	16.99	73.5 (32)	16.13
total	113.9 (124)	24.10	.074 (125)	0.95	24.1 (106)	6.29	46.7 (65)	7.70	86.6 (128)	16.19	73.2 (128)	18.20

Table 5: Mean scores and standard deviations language development in the three languages related to FoM

Results show that in both Dutch and Frisian language development, the lowest mean scores for vocabulary development (108.4 and 22.3 respectively) were observed in classes in which FoM is least observed (scores of 2 and 4 respectively). Only in Dutch lessons, the highest score for FoM seems to match with the highest mean score for vocabulary development (score 6).

In both Dutch and Frisian language development, a pattern was observed in which the highest score for FoM seemed to align with the highest mean scores for reading comprehension development, but in which the lowest reading comprehension score did correspond to the lowest FoM score. Conversely, in English language development a pattern was found in which the highest mean score for reading comprehension development (71.0) was observed in classes in which a FoM was least observed (score of 3), and in which the highest mean score for reading comprehension does not match the lowest FoM score.

The differences between mean scores were tested for significance. The results of these analyses are depicted in Table 6.

	F	***p***	***df2***	***df2***	**observed power**
Dutch Vocabulary Development	1.47	.23	3	120	.38
Frisian Vocabulary Development	3.62	.03	2	103	.66
English Vocabulary Development	0.04	1.0	3	124	.06
Dutch Reading Comprehension Development	12.43	.00	3	121	1.0
Frisian Reading Comprehension Development	0.42	.52	2	62	.10
English Reading Comprehension Development	0.29	.84	3	124	.10

Table 6: Relation between FoM in the three target languages and vocabulary development in the three languages

For vocabulary development, only in Frisian target language lessons significant differences were found between FoM scores in Frisian lessons ($p = .03$). To examine the difference between the FoM-scores in Frisian vocabulary development, post hoc analyses were carried out. Table 7 shows that there are significant differences between the FoM scores of 4 and 5 ($p = .00$) and the scores of 6 and 4 ($p = .05$). However, no differences were found between the FoM scores of 6 and 5 ($p = .61$). Thus, higher scores for Frisian vocabulary development were found in Frisian lessons in which there was more FoM.

(I) FoM score	(J) FoM scores	Mean difference	*p*	*SE*
5	4	2.57	.03	1.13
	6	0.65	.61	1.28
6	4	1.92	.05	0.97
	5	-0.65	.61	1.28

Table 7: Results post hoc analyses for the relation between FoM and Frisian vocabulary development

For reading comprehension development, only in Dutch lessons differences were found between reading comprehension development scores in FoM (p = .00). Table 8 shows the outcomes of the post hoc analyses to examine the differences between FoM scores in Dutch reading comprehension development. As expected, students from classes in which a FoM score of 4 is observed perform significantly worse on Dutch reading comprehension than students from other classes.

(I) FoM score	(J) FoM scores	Mean difference	*p*	*SE*
4	2	-1.17	.00	0.24
	5	-1.65	.00	0.35
	6	-1.37	.00	0.33

Table 8: Results post hoc analyses for the relation between FoM and Dutch reading comprehension development

We found no evidence that the scores for reading comprehension development were related to FoM in any of the three languages lessons.

4.3 Use of teaching methods in the three target language lessons (RQ2)

Figures 2a and 2b show the total scores for teaching methods (pedagogic tasks, group work, role-play, and demonstration) that are used to stimulate student language use in T1 and T2 respectively.

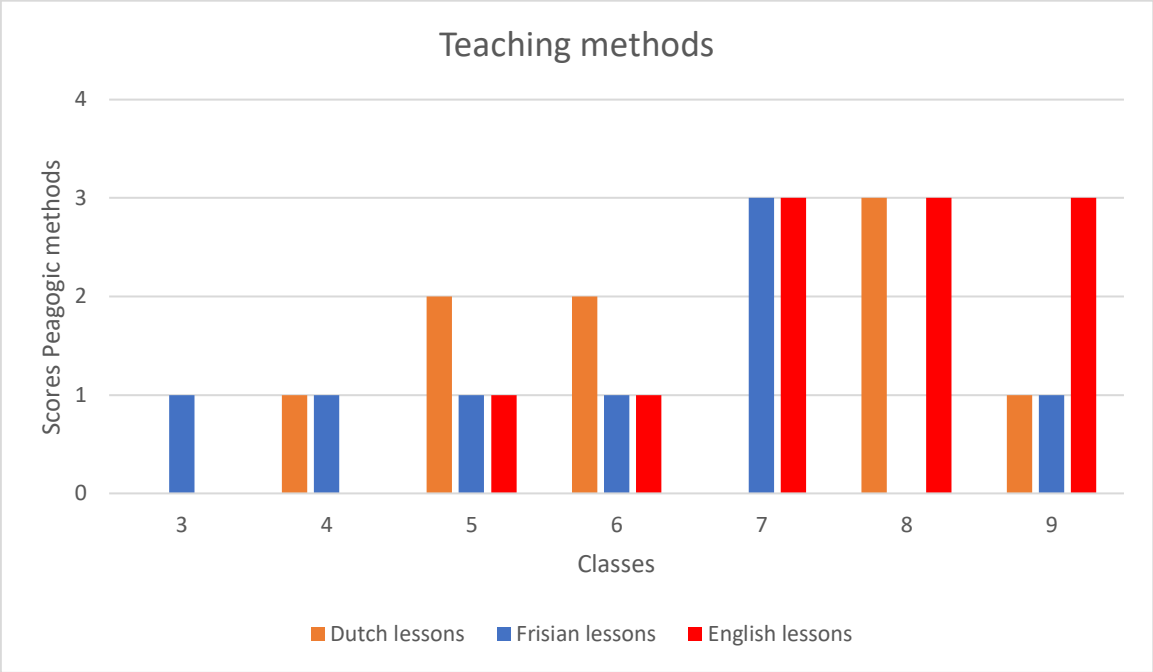

Figure 2a: Teaching methods to stimulate student language use T1 per class

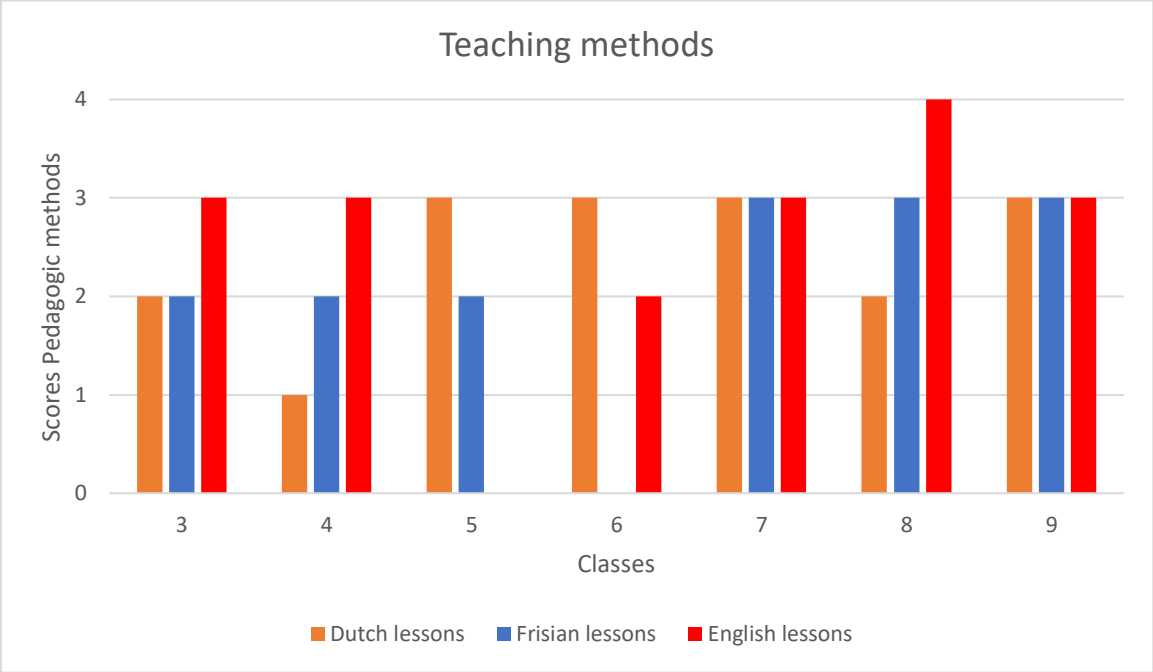

Figure 2b: Teaching methods to stimulate student language use T2 per class

Results show differences between teachers, between target language lessons, and between measurements in the use of teaching methods. Table 9 shows in both T1 and T2 the highest mean scores (1.6 and 2.5 respectively) but also the most variance in the use of teaching methods (*SD* = 1.28 and *SD* = 1.30 respectively) was observed in English lessons. Furthermore, in both measurements lowest mean scores for teaching methods are observed in Frisian lessons.

	Dutch lessons		Frisian lessons		English lessons	
	T1	T2	T1	T2	T1	T2
Mean	1.2	2.4	1.1	2.1	1.6	2.5
Median	1	3	1	2	1	3
Mode	1	3	1	2	3	3
SD	1.02	0.74	0.80	1.00	1.28	1.30

[a] Multiple modes exist. The smallest is shown.

Table 9: Measures of central tendency scores and standard deviations scores teaching methods to stimulate student language use in the three language lessons T1 and T2

To examine the scores for the composite variable teaching methods, data for the different components are presented in tables 10 and 11. For comparison, in addition to the data for pedagogic tasks, group work, role-play and demonstration, data for whole classroom discussion, individual seatwork, and the category other formats are presented as well.

	Pedagogic Tasks			Groupwork			Role-play			Demon-stration			Classroom discussion			Individual seatwork			Other		
class	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E
3					x											x		x			
4	x				x											x		x	x		
5	x	x	x							x			x			x	x				x
6	x	x	x	x									x		x		x	x			
7		x	x		x	x					x	x					x		x		
8	x		x	x		x				x		x	x		x		x				x
9	x	x	x			x						x	x	x	x	x				x	
Sum	5	4	5	2	3	3	0	0	0	2	1	3	4	1	3	4	4	3	2	1	2

T1 Blue = similar teacher per class in the three target language lessons

Table 10: Scores teaching method components to stimulate language use in the three language lessons T1

	Pedagogic Tasks			Groupwork			Role-play			Demon-stration			Classroom discussion			Individual seatwork			Other		
class	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E	D	F	E
3	x	x	x		x	x	x					x		x	x		x			x	
4	x	x	x		x	x						x		x	x	x	x		x	x	
5	x	x		x			x				x						x	x			
6	x		x	x		x				x				x	x						
7	x	x	x	x	x	x				x	x	x						x			
8	x	x	x	x	x	x			x		x	x		x		x			x		
9	x	x	x	x	x	x				x	x	x	x	x	x	x			x		
Sum	7	6	6	5	5	6	2	0	1	3	4	5	1	5	4	3	3	2	3	2	0

Second measurement Blue = similar teacher per class in the three target language lessons

Table 11: Scores teaching method components to stimulate language use in the three language lessons T2

Results show that there was an increase of the use of pedagogic tasks, group work, and demonstration in all three language lessons, and little decrease in the use of individual seatwork in all three target language lessons. Role-play was used in a few Dutch and in one English language lesson, but not used in Frisian lessons. Moreover, there was a little increase of the use of whole group discussion in the English lessons' during-phase and a larger increase in Frisian lessons. In both measurements, demonstration was most observed and individual seatwork was the least observed in English lessons.

To summarise, in T1 the use of teaching methods to stimulate student language use was most observed in Dutch and English lessons. However, these observations show that the least variance was observed in English lessons. In T2 the use of teaching methods to stimulate student target language use was most observed in English lessons and the least in Dutch lessons. In all three target language lessons an increase of scores was observed for the use of teaching methods, but the most in English lessons and the least in Dutch lessons.

Of the teaching methods, the use of pedagogic tasks was most frequently observed in both measurements and in all three target language lessons. In T2, also group work was quite often observed in the three target language lessons. Predominant use of demonstration was observed in only T2 of Frisian and English lessons. Role-play was observed in just three lessons over T1 and T2.

4.4 Relation between the use of teaching methods and students' language development (RQ2)

For vocabulary development, table 12 shows no clear patterns for the way the use of teaching methods and language development in the three languages align. In Dutch lessons, the highest mean score (120.3) for Dutch vocabulary was observed in a class in which a pedagogic method score of 1 was observed. For Frisian vocabulary, the highest mean score (26.7) was observed in classes with a pedagogic method score of 3. The mean score of 24.3 is observed in one class of only 7 students, which is a rather small number.

	Dutch				**Frisian**				**English**			
	VOC		RC		VOC		RC		VOC		RC	
	Mean (n)	*SD*	*Mean (n)*	*SD*	*Mean (n)*	*SD*	*Mean (n)*	*SD*	*Mean (n)*	*SD*	*Mean (n)*	*SD*
score 0					24.3 (7)	4.79			84.4 (24)	15.84	71.1 (24)	14.52
score 1	120.3 (15)	10.46	-0.69 (17)	1.40								
score 2	112.6 (44)	26.41	0.15 (43)	0.65	21.0 (46)	6.57	48.3 (22)	5.73	82.1 (7)	15.00	66.7 (7)	17.63
score 3	113.3 (65)	24.75	0.22 (65)	0.90	26.7 (53)	5.00	45.8 (43)	8.48	88.6 (77)	16.00	74.7 (77)	19.08
score 4									83.5 (20)	17.74	72.3 (20)	19.32
total	113.9 (124)	24.10	0.07 (125)	0.95	24.1 (106)	6.29	46.7 (65)	7.70	89.6 (128)	16.19	73.2 (128)	18.20

Table 12: Mean scores and standard deviations language development in the three languages related to teaching methods

Confusing patterns were observed for the relationship between the scores of teaching methods and reading comprehension development in the three target language lessons. For Dutch, table 12 shows that higher scores for reading comprehension coincide with higher scores for teaching methods. Contrarily, table 14 shows that higher scores for Frisian reading comprehension correspond with lower scores for the use of teaching methods in Frisian lessons. No pattern can be observed for the relation between English reading comprehension development and the use of teaching methods in English lessons. To test the data presented in Table 12 for significance, analyses of co-variance were conducted, the results of which are presented in Table 13.

	F	*p*	*df2*	*df2*	**observed power**
Dutch Vocabulary Development	0.45	.64	2	121	.42
Frisian Vocabulary Development	3.12	.05	2	103	.59
English Vocabulary Development	0.75	.52	3	124	.21
Dutch Reading Comprehension Development	9.99	.00	2	122	.98
Frisian Reading Comprehension Development	3.00	.09	1	63	.40
English Reading Comprehension Development	0.20	.90	3	124	.09

Table 13: Relation between teaching methods and language development in the three languages

Table 13 shows a relation between the teaching method scores in Frisian lessons and Frisian vocabulary development (p = .05), indicating that higher scores for teaching methods are related to higher scores for Frisian vocabulary development. Furthermore, for Dutch reading comprehension development significant differences in teaching methods scores in Frisian lessons (p = .00) are found as well. This finding means that higher scores for teaching methods and higher scores for Dutch reading comprehension development are related. In Frisian target language lessons however, an indication was found that, conversely, higher scores for Frisian reading comprehension development were related to lower pedagogic method scores in Frisian lessons.

To summarise, unclear patterns were observed for the relationship between the scores of teaching methods and reading comprehension development in the three target language lessons. In Dutch lessons, higher scores for reading comprehension coincide with higher scores for teaching methods. Contrarily, in Frisian lessons higher scores for Frisian reading comprehension align with lower scores for the use of teaching methods. No pattern was observed for the way English reading comprehension development and the use of teaching methods in English lessons coincide. A relation was also found between Frisian vocabulary development and teaching methods in Frisian lessons. This indicates that higher scores for teaching methods and higher scores for Frisian vocabulary development are related. Secondly, a relation was found between Dutch reading comprehension development and the variety of the use of teaching methods Dutch lessons. This finding means that higher scores for teaching methods and higher scores for Dutch reading comprehension development are related. In contrast, in Frisian target language lessons an indication is found that higher scores for Frisian reading comprehension development are related to lower variety of the use of teaching methods scores in Frisian lessons.

5. Discussion and conclusion

Focussing on CLIL classes in 9 Frisian trilingual primary schools, this paper aimed at exploring the strategies aimed at focus on meaning observed in the content lessons in the three target languages and examined whether these strategies were related to student language development in these languages. In addition, it focused on identifying teaching methods to stimulate target language use in CLIL classes and how these are related to student language development.

Regarding the first research question on the use of the strategies aimed at FoM in the three target languages, differences were observed between measurements, between classes, and between target language lessons. We registered an increase of the use of FoM strategies in all three language lessons, but mostly in Frisian lessons. In this sample, Dutch is the dominant language in peer contact outside school. Therefore, in school the use of Frisian needs to be more stimulated in authentic and meaningful tasks. This is in line with the CLIL principles focusing on engaging students in meaning making activities to stimulate the use of non-dominant languages. As there is more exposure to both Dutch (as the dominant national language) and English (see Günther-Van der Mei, 2018), it can be hypothesised that these meaning-making activities are less needed for the development of communicative language use in these two languages than for the Frisian language.

With respect to the different FoM components, both the use of real-world tasks and of demonstration were the least observed in Frisian lessons. The use of real-world tasks can be considered as the core of FoM, since the use of real-world tasks can be perceived as a motivation booster for student engagement in meaning-making activities (Van Oers, 2012) for which language is used. Demonstration, by which students present the outcomes of the during-phase task in the target language, was most observed in English lessons. Conversely, attention to pre-knowledge and reflection on content (to complete content learning) were least observed in English lessons. Thus, although meaning-making activities to stimulate communicative language use were found (Ellis, 2009b), it is questionable whether the pupils are systematic supported in the construction of knowledge in the English lessons.

Pedagogic tasks were used in all three target language lessons in both T1 and T2. In the second measurement, group work was observed in all three target language lessons as well. However, role-play was observed in just three lessons, whilst that format is highly recommended by Llinares and Dalton-Puffer (2015) to stimulate target language use in a CLIL setting. Furthermore, predominant use of demonstration was observed during T2 in Frisian and English lessons only. Demonstration can not only be related to the formats listed by Llinares and Dalton-Puffer

(2015) but is more importantly part of Skehan's three-phase model (1996, 2009), which has been argued to have a positive effect on communicative language use in the during-phase (Skehan, 2016). Regarding the relation between teaching methods and language development, in this study relations were only found for Frisian vocabulary development and Dutch reading comprehension development. These relations indicate that better performance on Frisian vocabulary development is related to the use of a broader variety of teaching methods in Frisian lessons, and that higher scores for Dutch reading comprehension development are related to the use of a broader variety of teaching methods in Dutch lessons. This is in line with Llinares and Dalton-Puffer's (2015) findings for the use of these CLIL teaching methods. However, our study shows a relation between for the use of a variety of these CLIL teaching methods/formats in meaning-making activities.

That no relations were found between the use of teaching methods in the three CLIL target language settings and Dutch vocabulary, Frisian reading comprehension, and English language development respectively, can possibly be explained by the small group sizes in the testing for significance. However, that no relations were found for Dutch vocabulary development and English language development can also be related to the exposure to Dutch and English. As argued, the effects of CLIL are related to exposure outside the school context (Sylvén, 2013; Günther-Van der Mei, 2018), and predominantly concern communicative competence (Pérez-Cañado, 2012) and reading comprehension skills (Admiraal et al., 2006).

These new insights can possibly help CLIL educators to improve the stimulation of language use in meaningful content-based language lessons, in particularly in regions in which one of the target languages enjoys less exposure outside of school. Further research could work towards overcoming some of the limitations of our study, such as the frequencies of the teaching practices and the amount of out-of school language use by media (Günther-Van der Mei, 2018). Moreover, further research could include specific moments in the interrelating development process (De Bot et al., 2007), as well as also tests to evaluate the complexity, accuracy, and fluency of oral language use (Ellis, 2000; Robinson, 2011; Skehan, 2009).

Our study suggests that to stimulate better vocabulary development in minority languages, it is therefore recommended to focus on meaning and to plan teaching formats that stimulate language use. For that, the use of a model including a post-phase seems beneficial to provide the pupils with a need to use the target language presenting the task-outcomes of the during-phase, and so reflection on content is equally stimulated (Tjepkema, 2021). This study furthermore

adds new perspectives to the field of studies on focussing on analysing longitudinal CLIL practices in relation to language development, as it made use of observation data to map teacher practices in three language lessons and assessed their relationship to student language development.

References

Admiraal, W., Westhoff, G., & De Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation 12*(1), 75-93.

Alberts, N., Egberink, M., Feddema, M., & Van Zuijlen, N. (2006/2007). Me2! Engels groep 7 Handleiding. Cito.

Alberts, N., Egberink, M., Feddema, M., & Van Zuijlen, N. (2007). Me2! Engels groep 8 Handleiding. Cito.

Bruton, A. (2005). Task-based language teaching: For the state secondary FL classroom? *The Language Learning Journal, 31*(1), 55-68.

Bygate, M. (2016a). Sources, developments and directions of task-based language teaching. *The Language Learning Journal 44*(4), 381-400.

Bygate, M. (2016b). TBLT through the lens of applied linguistics. ITL-International Journal of Applied Linguistics, 167(1), 3-15.

Candlin, C. (2009). Towards task-based language learning. In K. Van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based language teaching: A reader.*, (Vol. 1, pp. 21–40). John Benja mins.

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical Analyses of CLIL: Taking Stock and Looking Forward. *Applied Linguistics, 35*(3), 243–262.

Cenoz, J. (2015). Content-based instruction and content and language integrated learning: the same or different? *Language, Culture and Curriculum 28*(1), 8-24.

Coyle, D. (2002). Relevance of CLIL to the European Commission language learning objectives. In CLIL/EMILE European Dimension: Actions, Trends and Foresight Potential. European Commission, Public Services Contract DG 3406/001-001.

Dalton-Puffer, C. (2007). Discourse in content and language integrated learning (CLIL) classrooms. Discourse in Content and Language Integrated Learning (CLIL) Class rooms, 1-342.

Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From Practice to Principles?. *Annual Review of Applied Linguistics, 31*, 182-204.

Dalton-Puffer, C. (2017). CLIL in Europe: Opportunities and Challenges in Plurilingual and Multicultural Education. In Language Education Policy and Practice in East Asia (pp. 101-116).

Dalton-Puffer, C., Hüttner, J., & Llinares, A. (2022). CLIL in the 21st Century: Retrospective and prospective challenges and opportunities. Journal of Immersion and Content-Based Language Education, 10(2), 182-206.

Dalton-Puffer, C., Llinares, A., Lorenzo, F., & Nikula, T. (2014). "You can stand under my umbrella": Immersion, CLIL and Bilingual Education. A Response to Cenoz, Genesee & Gorter (2013). *Applied Linguistics 35*(2), 213-218.

Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). Charting policies, premises and research on content and language integrated learning. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (Vol. 7, pp. 1-19). John Benjamins Publishing.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism, 10* (1), 7 – 21

De Bot, K. (2016). Multi-competence and dynamic/complex systems. In V. Cook, & L. Wei (Eds.), *The Cambridge handbook of linguistic multi-competence* (pp. 125-141). Cambridge University Press.

De Jong, W., Osinga, N., Aalbersberg, H., Huistra, A., & De Vos, T. (2002). DLE-test begrijpend lezen. Uitgeverij Eduforce/GCO Fryslân.

Duarte, J., & Günther-Van der Meij, M. (2018). A holistic model for multilingualism in education. EuroAmerican *Journal of Applied Linguistics and Languages, 5*(2), 24-43.

Ellis, N. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The modern language journal, 92*(2), 232-249.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*, 193- 220.

Ellis, R. (2009a). Task-based language teaching: Sorting out the misunderstandings. *International journal of applied linguistics, 19*(3), 221-246.

Ellis, R. (2009b). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Multilingual Matters.

Ellis, R. (2014). Taking the critics to task: The case for task-based teaching. Proceedings of the Sixth cls International Conference clasic 2014, (pp. 103-117). Singapore.

Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: The effects on accuracy, complexity, and fluency of second language performance. *Canadian modern language review, 69(*3), 249-273.

Gorter, D. (2005). Three languages of instruction in Fryslân. *International journal of the sociology of language, 171,* 57-73.

Günther-van der Meij, M. (2018). The impact of degree of bilingualism on L3 development: English language development in early and later bilinguals in the Frisian context. Rijksuniversiteit Groningen.

Hemker, B., & Jongen, I. (2011). Begripend lêzen Frysk groep 8. Cito.

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research 19*(1), 70-88.

Jonkman, R. (2017). Wêr hast it wei? Evaluaasjesysteem Frysk (Grip-project 2014-2016)

[Where did you get that from? Asessment system Frisian (Grip project 2014-2016)]. NHL hogeschool.

Klinkenberg, E., Jonkman, R., & Stefan, N. (2018). Taal yn Fryslân. De folgjende generaasje [Language in Fryslân. De next generation]. Fryske Akademy.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied linguistics, 22*(1), 1-26.

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. The Open Applied Linguistics Journal 1(1), 30-41. Lin, A. M. (2015). Conceptualising the potential role of L1 in CLIL. Language, *Culture and Curriculum, 28*(1), 74-89.

Lasagabaster, D., & Doiz, A. (2017). CLIL students' perceptions of their multilingualism. *International Journal of Bilingual Education and Bilingualism, 20*(6), 653-667.

Llinares, A. (2015). Integration in CLIL: A proposal to inform research and successful pedagogy. *Language, Culture and Curriculum, 28*(1), 58-73.

Llinares, A., & Dalton-Puffer, C. (2015). The role of different tasks in CLIL students' use of evaluative language. *System 54*, 69-79.

Llinares, A., & Pastrana, A. (2013). CLIL students' communicative functions across activities and educational levels. *Journal of Pragmatics 59*, 81-92.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, C. Kramsch, & R. Ginsberg, *Foreign language research in cross-cultural perspective, 2*(1) (pp. 39-52). John Benjamins Publishing.

Long, M., Granena, G., & Yilmaz, Y. (2016). Major research issues in SLA. Brill Research Perspectives in *Multilingualism and Second Language Acquisition, 1*(1). Koninklijke Brill.

Lorenzo, F., & Casal, S. (2019). Content and Language Integrated Learning in multilingual educational contexts: the CALIE project. *International Journal of Bilingual Education and Bilingualism, 22*(7), 802-815.

Mahan, K. R. (2022). The comprehending teacher: Scaffolding in content and language integrated learning (CLIL). The Language Learning Journal, 50(1), 74-88.

Marsh, D. (2008). Language awareness and CLIL. *Encyclopedia of language and education, 6*, 233-246.

Ortega, L. (2015). Researching CLIL and TBLT interfaces. System, 54, 103-109.

Pérez-Cañado, M. (2012). CLIL research in Europe: Past, present, and future. International *Journal of Bilingual Education and Bilingualism, 15*(3), 315-341.

Pica, T. (2005). Class room learning, teaching, and research: A task-based perspective. *The modern language journal, 89*(3), 339-352.

Riemersma, A., & de Vries, S. (2011). Trilingual primary education in Fryslân. Trilingual primary education in Europe, 46-67. Fryske Akademy

Robinson, P. (2011). Task-Based Language Learning: A Review of Issues. *Language Learning 61*(1), 1-36.

Ross, S., & Kasper, G. (2013). Assessing second language pragmatics. Palgrave Macmillan.

Ruiz De Zarobe, Y. (2013). CLIL implementation: From policy-makers to individual initiatives. *International Journal of Bilingual Education and Bilingualism, 16*(3), 231-243.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied linguistics, 17*(1), 38-62.

Skehan, P. (2009). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied linguistics, 30*(4), 510-532.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics, 36*, 34-49.

Smit, U. (2013). Content and language integrated learning (CLIL) in the multilingual class room. *European Journal of Education, 48*(3), 295-308.

Sylvén, L. (2013). CLIL in Sweden–why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism, 16*(3), 301-320.

Tedick, D., Christian, D., & Williams Fortune, T. (2011). *Immersion education: Practices, policies, possibilities* (Vol. 83). Multilingual Matters.

Tjepkema, E. (2021). Exploring content-based language teaching practices to stimulate language use in grades 7 and 8 of Frisian trilingual primary education. Rijksuniversiteit Groningen.

Van den Branden, K. (2016). The role of teachers in task-based language education. *Annual Review of Applied Linguistics, 36*, 164-181.

Van Oers, B. (2012). Meaningful cultural learning by imitative participation: The case of abstract thinking in primary school. *Human Development, 55*(3), 136-158.

Van Ruijven, E., & Ytsma, J. (2008). Trijetalige skoalle yn Fryslân [Trilingual school in Fryslân]. Fryske Akademy.

Verspoor, M., & Edelenbos, P. (2009). Tweetalig onderwijs: beter geschoolde leerlingen in 2024. In R. De Graaff, & D. Tuin, De toekomst van het talenonderwijs: Nodig? Anders? Beter? (pp. 147-163). NaB-MVT/IVLOS Universiteit Utrecht.