University of Groningen

Language and science in young learners

Menninga, Astrid

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Menninga, A. (2017). *Language and science in young learners: Intervening in the balance between challenging and adapting*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 04-08-2025

Chapter 2

Language use in kindergarten science lessons

Language production and academic language during a video feedback coaching intervention for teachers

This chapter is based on:
Menninga, A., Van Dijk, M, Wetzels, A., Steenbeek, H., & Van Geert, P. (in press). Language use in kindergarten science lessons. Language production and academic language during a video feedback coaching intervention for teachers. *Educational Review and Evaluation, 23*(1-2), p 26-51. doi:10.1080/13803611.2017.1292920

2.1 Introduction

Many studies support the importance of implementing science lessons in kindergarten (e.g. Eshach & Fried, 2005; French, 2004; Greenfield, Jirout, Dominguez, Greenberg, Maier, & Fuccillo 2009). Although young children often demonstrate a spontaneous interest as well as strong intuitive insights in the scientific domain (Steenbeek, Van Geert & Van Dijk, 2011; Schwartz, 2009), they face great challenges when entering the kindergarten science classroom because of the linguistic demands of these lessons (Schleppegrell, 2004). Content-rich activities are beneficial for language development, such as the learning of conversational skills and vocabulary. Science lessons, in particular, provide a great opportunity for students to acquire complex language (Conezio & French, 2002; French, 2004; French & Peterson, 2009; Glass & Oliveira, 2014; Patrick, Mantzicopoulos, Samarapungavan, & French, 2008). At the same time, the acquisition of complex language is essential to learning science (Snow, 2014; Wellington & Osborne, 2001). So this works both ways: participation in kindergarten science activities demands and supports complex language. This complex form of language use, which is referred to as *academic language*, includes specific vocabulary and grammatical structures, as well as strategies to construct more sophisticated and precise ways of using language (Halliday, 1993; Schleppegrell, 2001; Schleppegrell, 2004).

The aim of the current study is to gain insight into the use of academic language in kindergarten science lessons, through examining the teacher-student interaction over the course of eight science lessons that are part of a video feedback coaching intervention for teachers (see Wetzels, 2015). This intervention, which is called 'Curious Minds in the Classroom' (CMC)[2], is innovative in that a coach enters the everyday teaching practices, which ensures the ecological validity of the theoretical concepts underlying the intervention. Teachers learn to observe and reflect upon classroom interaction by means of video feedback in order to improve their own teaching skills, which in the end may result in better student outcomes. These teaching skills concerned open-ended questioning strategies and the application of the empirical cycle as ways to stimulate students' reasoning. The coaching is based on video recordings of real-time interactions of the actual teaching practice in the natural science classroom setting. The way in which language use varies across different classrooms and over time will also be explored and discussed.

2.1.1 Academic language

A large body of existing research has emphasized the important role of social interaction in language learning at school (e.g., Dickinson & Porche, 2011; Powell, Diamond, Burchinal & Koehler, 2010). In this context of school, students are expected to advance in language in order later to successfully participate in academic and professional careers (Wong Fillmore & Snow,

[2] The intervention is part of the Curious Minds program, a line of research run by seven universities in the Netherlands and Belgium that focuses on stimulating science education in kindergarten and elementary school. Within the Curious Minds research, the CMC intervention was developed (Wetzels, 2015), aiming at behavioral change of the teachers.

2002). However, there are large differences between the language environment at home – and also between different homes (Scheele, Leseman, Mayo, & Elbers, 2012) – and the language expectations at school. As a result, for many students there is a discrepancy between their language proficiency and the language expectations when entering kindergarten. Therefore, teachers have an important role to bridge this gap by introducing their students to the linguistic demands of formal kindergarten classroom settings.

Over the last decades, several studies have discussed what is meant by academic language (Bailey, 2007; Chamot & O'Malley, 1994; Cummins, 1980; 1981; Halliday, 1994; Scarcella, 2003; Schleppegrell, 2001; 2004). Taken together, the commonalities of such discussions are that academic language is interpreted as a linguistic register characterized by discourse (i.e. turn-taking), lexical, and grammatical features that is used in formal settings and schools. This language of school is dynamic in nature, which means that language can be more or less academic depending on the context and interlocutors (Halliday, 1993; 1994; Snow, 2010; Snow & Uccelli, 2009). As several studies have pointed out, academic language is highly valued at school and as such the mastery of academic language has large impact on later school success (Cummins, 2000; Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Schleppegrell, 2004).

For the field of early elementary science education, then, it is important to understand the nature and development of academic language. To be clear, in this study the focus is on a usage-based approach of academic language in which language is seen as a tool for precise communication and learning during science activities (e.g. Uccelli et al., 2015). The largest part of the studies on academic language use targets adults, college students, and high school students because the language in these educational settings tends to be formal and academic. However, our knowledge is limited about academic language of students in kindergarten science classroom settings and the extent to which teachers model academic language to introduce their students to this specific linguistic register (Cassata-Widera, Kato-Jones, Duckles, Conezio, & French, 2008; Conezio & French, 2002; French, 2004; French & Peterson, 2009; Glass & Oliveira, 2014; Patrick et al., 2008). The current study is an important step in further exploring the academic language use in science activities in kindergarten, and, in particular, the development of some of the features of academic language over the course of the CMC teacher intervention.

In the context of the current study, the (indirect) impact of the intervention on language use is measured. The focus is on the extent to which the teachers model academic language (use academic language themselves) and attempt to evoke academic language from their students, and the extent to which students use academic language. The use of open-ended questions is related to elaborated student talk, diverse and sophisticated vocabulary, and the use of more causal connectives (De Rivera, Girolametto, Greenberg, & Weitzman, 2005; Lee & Kinzie, 2012; Massey, Pence, Justice, & Bowles, 2008; Oliveira, 2010; Wasik & Bond, 2001; Wasik, Bond, & Hindman, 2006). Therefore, lexical and grammatical features of academic language – or at least the most salient characteristics of it in kindergarten classrooms – can be measured in terms of speaking turns, lexical diversity and sophistication, and causal connectives. These features are simplified characteristics of academic language derived from a functional linguistic approach to academic language (e.g. Halliday, 1993; Schleppegrell, 2001), which can be used effectively to give an indication of academic language in spoken interaction (see e.g. Henrichs, 2010). The contribution of the current study is to provide more insight in the use of academic language in naturalistic kindergarten science classrooms, and to investigate the impact of a general teacher intervention on the development of academic language use during the intervention.

Turns.

Firstly, for the development of academic language students need language-learning opportunities and they need guidance in refining their language skills. Fostering extended discourse contributes to language development of students (Dickinson, 2001; Snow, 2014). For instance, open-ended questioning strategies of teachers are found to be most effective for stimulating children to talk (Oliveira, 2010). This type of questions typically generates longer responses (De Rivera et al., 2005; Wasik & Bond, 2001; Wasik et al., 2006). Therefore, focusing on and working toward verbal explanations by posing open-ended questions can be a helpful tool in stimulating complex language. However, the amount of children’s talk in science and technology lessons depends highly on teacher behavior. By working on the questioning skills of teachers, their role can shift from someone who transfers knowledge towards someone who activates students to think and participate in classroom discussions (Simon, Naylor, Keogh, Maloney, & Downing, 2008). This would result in more opportunities for the students to talk, leading to more and longer speaking turns of the students, which provide conditions that may make academic language more likely to emerge.

Lexical diversity and sophistication.

Secondly, the lexical features are one of the most salient features of academic language, as the lexical choices of an individual form the content of any message (Schleppegrell, 2001; 2004). Many studies have emphasized the impact of high quality teacher language on students’ vocabulary development (e.g. Girolametto, Weitzman, & Greenberg, 2003; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Justice, McGinty, Zucker, Cabell, & Piasta, 2013; Mayo & Leseman, 2008; Snow & Kurland, 1996; Weizman & Snow, 2001). Vocabulary development is not only influenced by quantitative factors, such as how frequently children hear a specific word in their environment (Childers & Tomasello, 2002), but also by qualitative factors such as exposing children to low-frequency, complex words or synonyms. First, as a result of the use of synonyms and more domain-specific and low frequency lexical words, academic language is characterized by high lexical diversity (Eggins, 2004; Laufer & Nation, 1995; Leseman, Mayo, Messer, Scheele & Van der Heyden, 2009; Schleppegrell, 2004; Spycher, 2009). Secondly, adults’ use of low-frequency words, which are considered to be more complex (Van Hout & Vermeer, 2007), results in a more sophisticated vocabulary of children (Weizman & Snow, 2001). The results of Henrichs (2010) show that these lexical features of academic language can be found in interactions between parents and four-to-five year old children. In addition, several interventions that explicitly targeted the use of science vocabulary of young students indicated significant gains for instance in the use of more sophisticated terms (French, 2004; Henrichs & Leseman, 2014; Hong & Diamond, 2012). A systematic literature review pointed out that science instruction promotes students' use of domain-specific words, suggesting that science instruction increases the vocabulary performance of young children (Guo, Wang, Hall, Breit-Smith, & Bush, 2016). There is also empirical evidence that open-ended questions provide opportunities for more linguistically and cognitively challenging discourse (Lee & Kinzie, 2012; Massey et al., 2008). These cognitively challenging questions have thus been linked to students' development of complex language (De Rivera et al., 2005) such as rich, sophisticated vocabulary, more on-topic responses, more content-, and task-related terms (Dickinson & Porche, 2011; French, 2004).

Causal connectives.

Thirdly, in addition to lexical features, academic language involves grammatical features such as clause-combining and complex clause-embedding structures in order to mark essential relationships between clauses (Bailey, 2007; De Temple, Wu, & Snow, 1991; Schleppegrell, 2004; Snow, 1989). The length and complexity of utterances are seen as important indicators of complex language as they are a result of clause combining structures. For instance, a study of Huttenlocher, Vasilyeva, Cymerman and Levine (2002) showed that when parents produce more complex utterances and use a more diverse vocabulary, the syntax of the children also tends to be more complex. Henrichs (2010) also indicates that Dutch children use more complex structures when adults in their environment use more clause combinations. There are several studies suggesting that prompting students to explain can improve learning (e.g., Fonseca & Chi, 2011; Legare & Lombrozo, 2014; Walker, Lombrozo, Legare, & Gopnik, 2014; Walker, Lombrozo, Williams, Rafferty, & Gopnik, 2017). For instance, when children are asked to give verbal explanations (as opposed to descriptions), they made more sophisticated causal generalizations in subsequent tasks. An implication of this can be that teachers who ask more open-ended questions relating to the how-and-why of an event might be particularly effective in promoting the cognitive functions that require academic language. Previous studies indicated that open-ended questions have been associated with more causal connectives (Peterson & French, 2008) and complex utterances (see Chapter 3; Justice et al., 2013). The intervention targets training teachers to apply the empirical cycle (see figure 1) as a mechanism that supports students' reasoning. The empirical cycle can function as a tool to learn how to formulate – for instance – adequate predictions and explanations. Predictions and explanations are linguistically more complex because they require explicit expressions of the relations between clauses. The use of causal connectives may give an indication of the students' ability to express these relations appropriately.

2.1.2 Teaching academic language during science

Many studies have emphasized that teacher language has a positive impact on the language development and production of children. This is in particular important with regard to academic language as several studies have pointed out that this language of schooling is not picked up automatically (Lemke, 1990; Vygotsky, 1987) and also not often taught explicitly (Schleppegrell, 2001; Wong Fillmore & Snow, 2002). However, teachers can play a vital role in providing children with sustained opportunities for science discourse by scaffolding students' language use during science activities (Cassata-Widera et al., 2008; Mercer, 1995; National Research Council, 2013; Rojas-Drummond & Mercer, 2003). This paper explores the impact of the CMC intervention on language use of students and teachers. The intervention is based on studies that have indicated that children's active, inquiry-based learning, self-regulated exploration, and guided education are important factors for successful science lessons (Alfieri, Brooks, Aldrich & Tenenbaum, 2011; Kirschner, Sweller, & Clark, 2006; Mayer, 2004). The main focus of the CMC intervention is on enhancing teachers' questioning skills, scaffolding strategies, and particular attention is paid to structuring the science lesson according to the empirical cycle (see figure 1). This means that the lesson is based on the continuous use of a research method by asking questions, predicting, testing, observing, and analyzing (De Groot, 1994; Dejonckheere, Van De Keere, & Mestdagh, 2009). A key aspect of the intervention is enhancing teachers' open-ended questioning strategies (Barber & Mourshed, 2007; Roth, 1996). The CMC does not explicitly target language production or academic language in particular, but it does offer a context in which an improvement in language use can occur spontaneously as language use is of great importance in early science lessons (Snow, 2014; Wellington & Osborne, 2001). Previous s have demonstrated that after the intervention teachers asked significantly more open-ended questions – such as "What do you think will happen when …?" – (p = .02, d = .77) in the course of the CMC intervention (see Wetzels, 2015). In addition, students did reach a significantly higher level of reasoning skills with a large effect (p = .02, d = .83) after the teacher intervention (see Wetzels, 2015), compared to students in the control group.

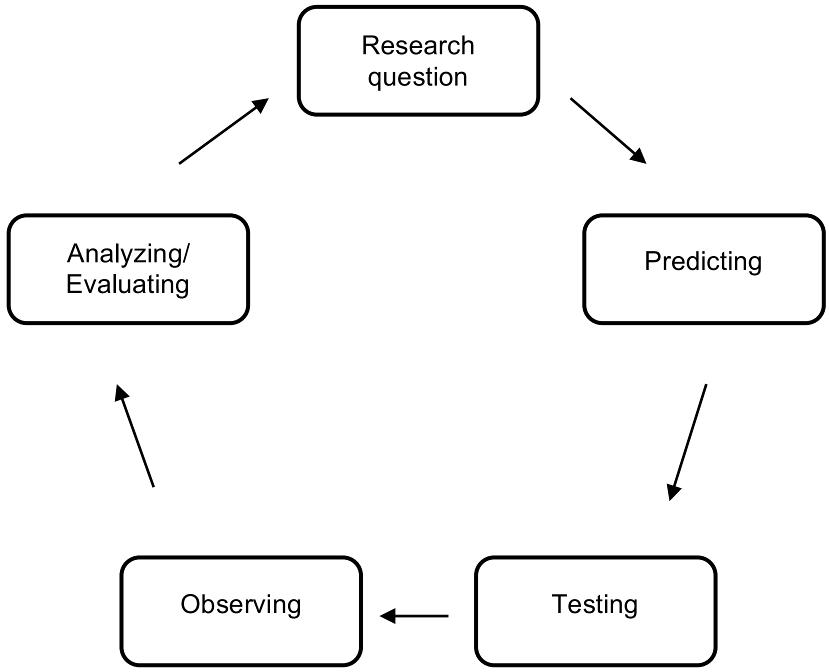


Figure 1 Schematic representation of the empirical cycle.

2.1.3 Video feedback coaching

Video feedback coaching is an effective and useful method that creates awareness of a teacher's behavior and the associated student reactions (Strathie, Strathie, & Kennedy, 2011; Van den Heijkant et al., 2006). Teaching and learning emerges in social interaction between teacher and students, in which both interaction partners constantly influence each other. Changing the existing behavior in these complex and intertwined teaching-learning processes into another more optimal form of behavior requires a lot of energy. The best way of intervening in these teaching-learning interactions is to enter the context in which these processes emerge (i.e. the classroom setting). For instance, meta-studies have shown that interventions that employ coaching in the classroom are most effective when it comes to changing the quality of interaction (Barber & Mourshed, 2007). The method of video feedback can be used to provide constructive feedback on positive teaching aspects and to select options for improvement (Fabiano et al., 2013; Noell et al., 2005; Reinke et al., 2009; Wade, 1985). Reflection by means of video feedback is a strong means to acquire new skills and improve existing skills (Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2011). For instance, Fukkink and Taveccio (2010) found that teachers became more sensitive and verbally stimulating after they received video feedback coaching. The contribution of this way of intervening is that it makes use of real-time teacher-student interaction and thus provides insight into the dynamics of interaction. This study provides a unique opportunity to advance our knowledge of academic language in kindergarten teacher-student interactions.

2.1.4 Inter- and intra-individual variability

Development and learning are idiosyncratic processes (Molenaar, 2004; Rose, Rouhani, & Fischer, 2013). This idiosyncrasy concerns variability: individuals differ from other individuals as well as individuals differ from time point to time point (Bassano & Van Geert, 2007; Van Geert & Van Dijk, 2002). This variability provides a rich source of information about learning processes. Focusing in detail on individual development across time (Molenaar & Campbell, 2009) is of more value than simply examining group averages (Barlow & Nock, 2009). Additionally, zooming in on individual development enables us to look at patterns of interaction that indicate changes on the process level. This might result in important, more detailed findings in addition to the changes on group level. As previous analyses of these data on the questioning and reasoning skills indicated, there was large individual variability between teachers and between students. This is in line with the idiosyncratic nature of developing human behavior. Although most analyses in this paper focus on the group level, we also aim to explore the individual learning curves of the individual teachers and students over time.

2.1.5 Research questions

This study aims at gaining insight into the language production of students and teachers during the CMC intervention in kindergarten science lessons. Therefore, the video recordings of the students and teachers who participated in the CMC intervention study of Wetzels (2015) were studied. Assuming an effective teacher intervention regarding the use of open-ended questions and higher reasoning levels of students (see Wetzels, 2015), it may be expected that students start using more elaborated and more academic language in the course of the intervention and that teachers reduce their talk but increase academic language. There are several areas where this study makes an original contribution to our knowledge on academic language. Firstly, this study is based on real-time interactions in a naturalistic teaching setting, which provides information on whether academic language is used in these kinds of settings in contrast to, for instance, test scores to assess the level of academic language of students. Secondly, the video feedback coaching that was used to give teachers insights in the teaching-learning processes was an innovative way of professional development for teachers as opposed to theoretical courses for teachers. Thirdly, this paper aims to investigate academic language in kindergarten classrooms, as knowledge on this topic is limited. The findings may result in practical implications on whether explicitly targeting language learning during kindergarten science activities is required, or whether academic language is already part of science lessons and greatly improves over multiple sessions. The following exploratory research questions were addressed:

1a. What are the effects of a video feedback coaching intervention within science lessons on the verbal teacher-students interaction, as measured by the proportion of student utterances and turn length in the intervention group compared to a control group?

1b. What are the patterns of change of individual trajectories of students and teachers in the intervention group with regard to proportion student utterances and turn length?

2a. What are the effects of a video feedback coaching intervention within science lessons on the academic language use of students and teachers academic language use, as measured by lexical diversity, lexical sophistication and number of causal connectives in the intervention group compared to a control group?

2b. What are the patterns of change of individual trajectories of students and teachers in the intervention group with regard to lexical diversity, lexical sophistication and number of causal connectives?

2.2 Method

2.2.1 Participants

Ten female kindergarten teachers – with a small teaching group of three to six students each – were recruited from schools in the North of the Netherlands to participate voluntarily in this study. Five teachers (M_{age}= 54 years, $M_{experience}$ = 25 years) took part in the intervention and five teachers (M_{age} = 50 years, $M_{experience}$ = 27 years) formed the control group. All teachers had very limited experience in teaching science but were willing to include this in their weekly schedule. The participating students ($N_{intervention}$ = 18, $N_{control}$ = 26) were four or five years old and were more or less evenly distributed according to gender (intervention: 50% boys / 50% girls, control: 58% boys / 42% girls). We did not include any standardized measures of the students' language development or academic performance, as we were interested in the actual language use in the real-time teacher-students interaction. All teachers and students were native speakers of the Dutch language. According to the teachers, none of the participating students had any notable developmental problems.

2.2.2 Material and measures

Transcription procedure.

The video recorded lessons were transcribed following the Codes for Human Analysis of Transcripts (CHAT) conventions (MacWhinney, 2000) by the first author of this paper. A second, trained researcher checked 32% of the transcripts (randomly chosen) for calculating reliability. Out of 4946 utterances, 256 had small differences (5%). There were no big differences between the transcripts of the different teachers (difference between transcribers for the five teachers respectively: 6%, 5%, 6%, 4%, 3%). In most cases the difference was a single word, which was not understood by the first transcriber (and was therefore marked as xx) but was understood by the second transcriber, or vice versa. The chosen unit of transcription was the utterance. Utterance boundaries were determined on the basis of turn-taking, pauses, and intonation patterns (Brown, 1973). Incorrect inflections, as well as badly articulated and hardly intelligible words or phrases, were transcribed only if the target word was entirely clear, whether by means of the context or the verbal utterance. Completely or partly unintelligible utterances were excluded from analysis, as well as interjections. Imitations and repetitions of the utterances made by students or teachers were included in the analysis because these indicate, for instance, that vocabulary is being taken over.

Preparation of transcripts.

For the linguistic analysis, separate files were created with a length of ten minutes, starting with the first content-related utterance of each transcript. There was one exception, the second post-measurement of teacher H in the control group lasted for 7.40 minutes. All transcripts were prepared for analysis using the Computerized Language Analysis software (CLAN), which is part of the CHILDES system (MacWhinney, 2000).

Linguistic analyses.

Several linguistic analyses were applied to the transcripts. The group of students as a whole was considered the interlocutor of the teacher. Therefore, the analyses were performed on the basis of the whole group of students per teacher and not considering the students individually.

Turns.

We measured the total number of utterances by use of the program CLAN[3], counting the number of utterances per participant (students or teachers). Subsequently, the proportion of student utterances was calculated by dividing the number of student utterances by the total number of utterances in the lesson. Finally, we measured the mean length of turn in words (MLT-w). Both analyses were performed using the command *mlt* in CLAN.

Lexical diversity.

Lexical diversity can be described as the number of different words relative to the total number of words in a sample, which comes down to the number of types (different words) divided by the number of tokens (total words). In this study, lexical diversity was measured by index D, which is a method that makes several adjustments to the type-token ratio (TTR). These adaptations were made in order to correct for the influence of sample size as the longer the sample, the more (repetition of the same) function words, and the lower the TTR (Richards & Malvern, 1997; McKee, Malvern & Richards, 2000). As an illustration, consider the two following sentences:

1) He puts it in there. (5 Types/ 5 Tokens = 1.00 (TTR))
2) The teacher puts the marble in the box of glass and observes that the marble sinks to the bottom. (14 Types / 19 Tokens = .74 (TTR))

These sentences vary in length with different type token ratios, but the complexity (in terms of diverse language use) is clearly demarcated between the two. The long sentence may be regarded as more academic, which does not show from the TTR. For this reason, in the current

[3] To download CLAN go to: http://childes.psy.cmu.edu/clan/. For more information on the specific procedures on calculating linguistic measures in CLAN, the webpage http://childes.psy.cmu.edu/manuals/clan.pdf provides a detailed manual.

study, this influence is also limited by taking samples that carry approximately the same length (ten minutes). The index D was calculated by the command *vocd*, which is a tool on CLAN (MacWhinney, 2000). All student and teacher samples were greater than 50 tokens – the minimum requirement for computation of D.

Lexical sophistication.

The lexical sophistication, the number of complex and low-frequency words relative to all words, was measured by means of the Measure of Lexical Richness (MLR). The MLR, developed by Van Hout and Vermeer (2007), was calculated with an online tool based on the word list by Schrooten and Vermeer (1994). This word list provides the frequency of 26,000 lemmas in a corpus of two million tokens, drawn from oral and written language input in elementary schools in the Netherlands. The corpus contains words from picture books, factual subjects and mathematics textbooks, but also from oral teacher input during instruction and other interactions in the classroom, making it an adequate corpus to use in the context of science lessons. The online tool generates an MLR index, which is a weighted score resulting from the comparison between the frequency distribution of lemmas in the submitted transcript and the distribution of the Schrooten and Vermeer (1994) corpus. Proper nouns and unintelligible words and utterances were excluded from the transcripts. Although MLR is commonly used to estimate children's vocabulary size (Van Hout & Vermeer, 2007), in the current study MLR is used for both the students and the teacher as an indicator of lexical sophistication, as it reflects the comparison between the frequency distribution of lemmas in the teacher-students interaction and the distribution of the Schrooten and Vermeer corpus (1994). When we return to the examples in the paragraph on diversity, these sentences also illustrate the difference in lexical sophistication[4]. The MLR of the first sentence would be 1.00 (which is an indication that all words are from list 1 with high(est)-frequency-words), and the MLR of the second sentence would be 3.89 (indicating that some words are from lists with low(er)-frequency words).

Number of causal connectives.

The Dutch language offers a wide variety of connectives that mark causal relations but not all of these connectives have a grammatical counterpart in English. Causal connectives in this study were defined as words that indicate causal relations, both coordinate and subordinate. The connectives that were included in the analyses are: *omdat, zodat, doordat, daardoor, hierdoor, waardoor, daarom, dus* and *want*, which are best approximated in English by variations of *because, therefore* and *so*. By means of the program FREQ in CLAN, we calculated the total number of causal connectives per transcript for the students and teachers separately.

[4] Note that for the calculation of MLR, these sentences were literally translated into Dutch (same number of words) because the program only allows Dutch input.

2.2.3 Procedure

The data were systematically collected in naturalistic classroom situations during science lessons given by the teacher to a small teaching group. The same group of students participated during the whole intervention (intervention or control group). The participating teachers, in the intervention and control groups, were recruited from schools in a regional collaboration project on science education[5]. Since all teachers participated in a regional science collaboration project, the teachers were taken to be comparable with regard to interest and experience in and frequency of teaching science. The recruitment of teachers was done in several waves over an extended period of time (around two years). Firstly, we recruited teachers for the intervention condition, and once we had enough teachers to participate in the intervention, we recruited teachers to participate in the control condition. This control group was also offered the opportunity to receive the complete video feedback coaching similar to the teachers in the intervention condition or – when preferred – a group workshop at a later time. Written consent of all parents was obtained prior to participation of the students in the study. A University's Ethical Committee of Psychology (ECP) approved to conduct this study.

The teachers were instructed to choose a small group of students (three to six), varying in age, gender and cognitive level. All lessons were recorded on video. The data collection of the intervention group consisted of two pre-measurements, four sessions that were used for coaching purposes, and two post-measurements. This resulted in eight measurements per teacher. The control group data consist of only the two pre-measurements and the two post-measurements. The teachers in the control group continued their science lessons in the intervening period since their schools were involved in a regional collaboration project on science education. Most of the data were collected over three to four months, the period during which all the coaching occurred. In the intervention group, the coaching sessions took place every one or two weeks. As some effects of the coaching would be immediately apparent in the following lessons, we filmed another two lessons four to six weeks after the coaching had stopped (post-measurements). These lessons were aimed at measuring the more long-term effects of the CMC intervention. When it comes to the control group, there was a comparable amount of time between pre- and post-measurements. Figure 2 presents an overview of the design of this study.

The teachers were instructed to give a regular science lesson on a subject of their own choice for the duration of fifteen to twenty minutes. Teachers were free to choose any topic for their science lesson, which supported their self-efficacy. The opportunities for self-efficacy are important elements that address a teacher's basic concerns, namely the concerns for competence and for autonomy (Wetzels, 2015). Most teachers chose subjects such as craftwork, designing, floating and sinking, air pressure, senses, sound and marble track. Table 1

[5] The regional collaboration project was called Bètapunt Noord. Schools could join this collaboration project when aiming to implement science into their teaching curriculum. The participation of teachers in the CMC intervention was the first step towards this implementation. Before the start of the intervention, teachers in both conditions had no access to specific information or materials regarding teaching science, other than freely available information in books or on websites.

presents an overview of the topics of the lessons per teacher in both the intervention group and the control group.

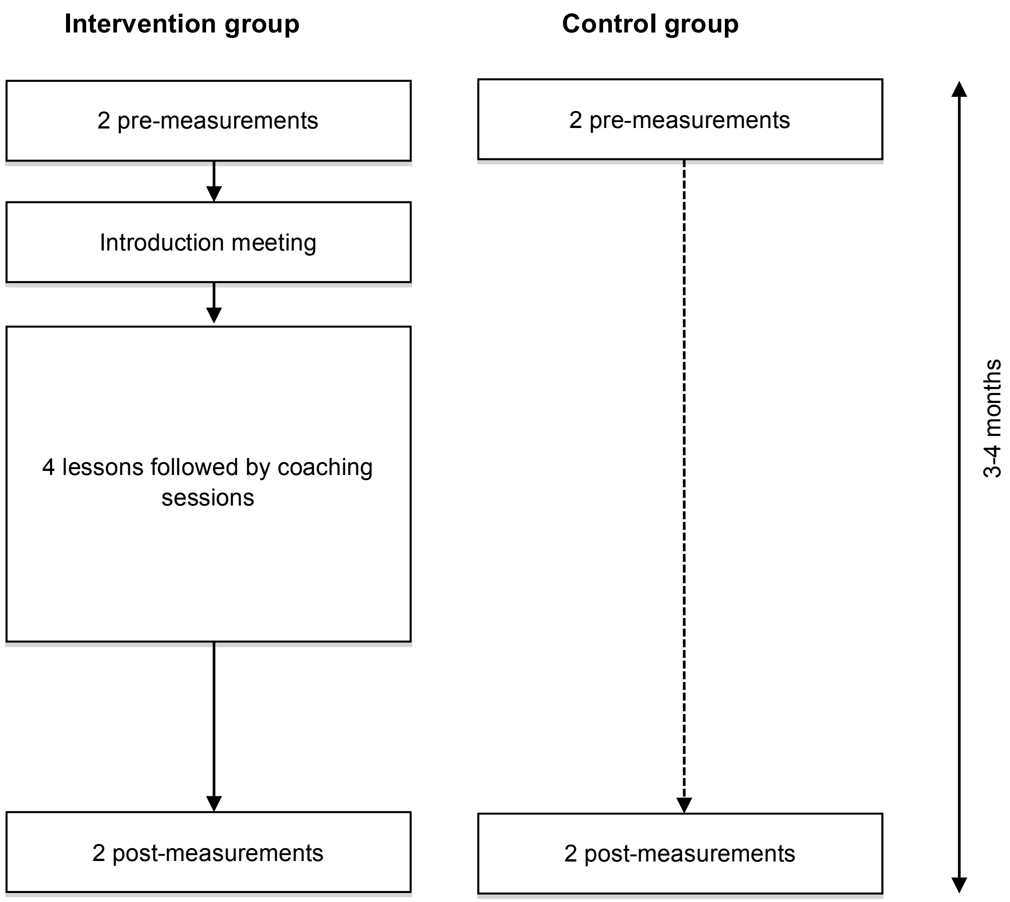


Figure 2 Design of the study

Tabel 1 Overview of (lesson) topics per teacher in the intervention group and the control group.

Les son	Intervention group					Control group				
	Teacher A	Teacher B	Teacher C	Teacher D	Teacher E	Teacher F	Teacher G	Teacher H	Teacher I	Teacher J
1	Designing	Floating and sinking	Floating and sinking	Designing	Designing	Marble track	Magnet	Mixing fluids	Constructing	Mixing fluids
2	Designing	Senses	Marble track	Designing	Designing	Floating and sinking	Weighing and gravity	Sound	Gravity	Air (pressure)
3	Floating and sinking	Sound	Mixing colors	Floating and sinking	Ramp	-	-	-	-	-
4	Air (pressure)	Sound	Flowers	Measuring and weighing	Marble track	-	-	-	-	-
5	Air (presure)	Shadow	Air (pressure)	Mixing colors	Marble track	-	-	-	-	-
6	Air (pressure)	Floating and sinking	Sound	Mirror	Air (pressure)	-	-	-	-	-
7	Measuring and weighing	Magnet	Magnet	Air (pressure)	Air (pressure)	Shapes	Constructing	Floating and sinking	Electricity	Air (pressure)
8	Senses	Floating and sinking in water/oil	Senses	Senses	Air (pressure	Building a tower	Constructing	Senses	Constructing	Air (pressure)

The students were not specifically informed of the intervention – other than we are going to conduct some science experiments – beforehand. After the first two lessons (pre-measurements), all teachers in the intervention group attended an information meeting about the general principles of teaching science as formulated in the Curious Minds program[6]. Furthermore, the teachers were instructed in recognizing and encouraging the students' talent for science, posing open-ended questions and applying the empirical cycle of De Groot (1994) in the lessons (see figure 1). Some video clips of the Curious Minds program and clips of other teachers' lessons were shown to illustrate the theory. After this, the teachers were asked to specify a personal learning goal that was used as a special point of interest for both the teacher and coach in the coaching sessions. This personal learning goal was aimed at stimulating the intrinsic motivation of the teachers and had to be in line with the coaching principles. An example of a teacher's personal learning goal was: "I want to learn how to ask questions based on the empirical cycle". In the intervention stage, lessons were immediately followed by a coaching session particularly focused on the personal learning goal, the questioning skills and the use of the empirical cycle. The small group of students in which the coaching intervention took place enabled the teachers to scaffold the students individually as well as the group of students as a whole.

For the purpose of video feedback coaching, the coach selected several moments from the lesson, based on a ratio of three moments that showed successful teacher behavior to one moment that indicated an area for development as a higher positivity ratio is needed for successful behavioral change to occur (Fredrickson, 2013). Coach and teacher discussed and reflected upon these moments to bring the teacher's behavior to a conscious level (Van den Heijkant et al., 2006). The teacher was provided with tools to enhance her skills in the following sessions. The method of video feedback coaching allowed adapting the intervention to the specific, typical classroom interactions of individual teachers. This way, the coaching provides optimal opportunities for the teachers to continue practicing the trained skills after the intervention. The intervention was extensively tested in a pilot study (Wetzels, 2015). An experienced coach and psychologist (the third author of this paper) performed the intervention using an extensive manual.

2.2.4 Analyses

Monte Carlo permutation tests.

The data were analyzed using a Monte Carlo procedure (e.g. Good, 2000). A Monte Carlo analysis is a nonparametric test, which evaluates the null hypothesis that the probability of the specific association between the variables under study is based on chance alone. To achieve

[6] The basic principles of Curious Minds are that everyone is talented, adults should learn to recognise and stimulate the natural curiosity of children, adults should become talent-experts, adults are the motor behind further development of children's talent and the best way to achieve this is case-based learning for adults.

this, data are randomly sampled multiple times in order to determine whether the empirically found value or difference (i.e., between pre- and post-measures) could be expected based only on chance. The Monte Carlo permutation test provides an estimation of the exact *p* value, which is the probability that the same or a larger difference is found if the null hypothesis is supported. The greater the number of permutations, the closer this estimation comes to the exact value (see Gigerenzer, 2004; Schneider, 2015). Monte Carlo analyses were performed in Microsoft Excel in combination with PopTools (version 3.2). This test is particularly suitable for small sample sizes with repeated measures. These analyses were performed for both the teachers and the students in the intervention group (first two lessons as pre-test and lesson seven and eight as post-test) and the control group (first two lessons as pre-test and lesson three and four as post-test). In the Monte Carlo procedure, the simulated data – based on a random shuffle of the observed data – were first shuffled (10,000 times) over the pre- and post-tests. Then, the observed data were compared with the randomly shuffled simulated samples, after which it could be determined how often the observed difference (or a bigger difference) occurred in these random samples. This procedure resulted in a p-value comprising the probability that the observed difference occurred in this distribution of 10,000 random samples. Being aware of the controversial use of a specific predetermined alpha to determine significance (see Gigerenzer, 2004; Schneider, 2015), we only report the exact *p*-values that indicate the probability that the finding is based on chance.

Effect sizes.

In addition, we computed effect sizes[7] to indicate the magnitude of the difference in relation to the variability in the sample. Effect sizes were calculated as Cohen's d, which represents the standardized mean difference between the pre-test and post-test. The effect sizes were calculated per variable, for the students and for the teachers separately. The mean scores and pooled standard deviations per variable were calculated for the two pre-tests combined and the post-tests combined. Finally, the difference between pre-test and post-test measurements was divided by the pooled standard deviation. Based on Cohen's classification (1992), effect sizes of d = .20, .50, and .80 (and negative d-values of -.20, -.50 and -.80) reflect respectively small, medium, and large effects.

Slopes.

In order to explore how changes occurred over time, we tested whether the average linguistic measures showed an increase or decrease over time for both students and teachers. The slope of the empirical data was tested against the slope of randomly shuffled data. First, the slope was calculated over the empirical data, based on an average of all students. Second, the columns and rows with empirical data of all students are randomly shuffled, after which the slope was calculated again over the shuffled data. This procedure was repeated 10,000 times

[7] The use of effect sizes – in addition to the permutation tests – was exclusively meant for descriptive purposes.

per test by using Monte Carlo permutations and results in a *p*-value, which indicates the probability that the slope of the empirical data stemmed from the distribution of slopes of shuffled data. The same procedure was followed for teacher measures.

2.3 Results

2.3.1 Effects on verbal teacher-student interaction

Regarding language use and production, we observed that the proportion of student utterances increased in the intervention group. A Monte Carlo test revealed that the probability that this difference was based on chance was very small (M_{pre} = 28 % vs. M_{post} = 38%, $p < .01$, d = 1.57). The proportion of student utterances was practically unchanged in the control group, and the Monte Carlo test showed no indications that there was a difference in the control group between pre- and post-measurement (M_{pre} = 35 % vs. M_{post} = 38%, p = .22, d = .35). These changes indicate that there was a shift in the verbal teacher-students interaction after the intervention.

The length of the students' turns, on average, increased in the intervention group in contrast to no clear increase or decrease in the control group. The Monte Carlo analysis indicated that the probability that the difference in the intervention group based on chance was relatively small (M_{pre} = 3.9 vs. M_{post} = 5.3, p = .01, d = 1.04), and that the probability that the variability in the control group based on chance was rather large (M_{pre} = 4.4 vs. M_{post} = 5.1, p = .22, d = -.29). The turn length of the teachers decreased in the intervention group (M_{pre} = 13.7 vs. M_{post} = 11.6, p = .05, d = .83) compared to no decrease in the control group (M_{pre} = 12.8 vs. M_{post} = 12.1, p = .29, d = .25). This means that the students elaborated their talk and the teachers shortened their utterances after the intervention, while the variability of the average trajectories of the students and teachers in the control group was roughly the same for the pre- and post-measure.

2.3.2 Patterns of change

Proportion of student utterances.

The average trajectory indicated a fairly gradual increase in the proportion of student utterances across the entire period (*slope* = .02, $p < .01$). A closer examination revealed that the intra-individual trajectories showed different patterns of change in the proportion of student utterances over time, as is illustrated in figure 3. Strikingly, the variation in the proportion of student utterances was relatively small in the first two lessons, while in the course of the intervention there was an increase in variability between the different classrooms. For instance, the proportion of student utterances of classroom D already changed during the first coaching lesson, whereas the students of teacher E only increased their proportion of utterances considerably after the third coaching session.

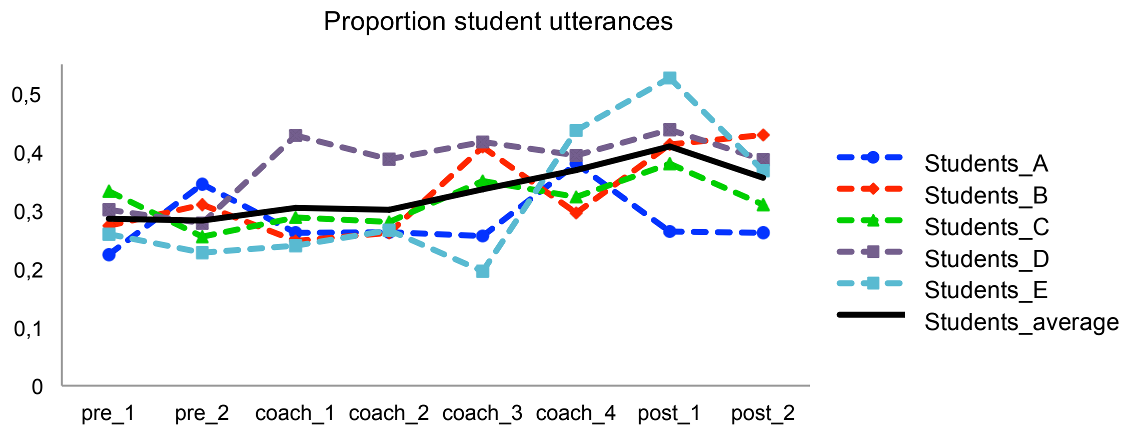


Figure 3 Proportion student utterances per teacher during the complete intervention.

Turn length.

Turn length of students increased over the eight lessons (*slope* = .23, $p < .01$). The variation between students seemed to increase. On average, teachers reduced the length of their speaking turns (*slope* = -.46, $p < .01$). Most teachers showed a similar pattern over time, with turn length decreasing after lesson 4 and stabilizing during the last four lessons (see figure 4).

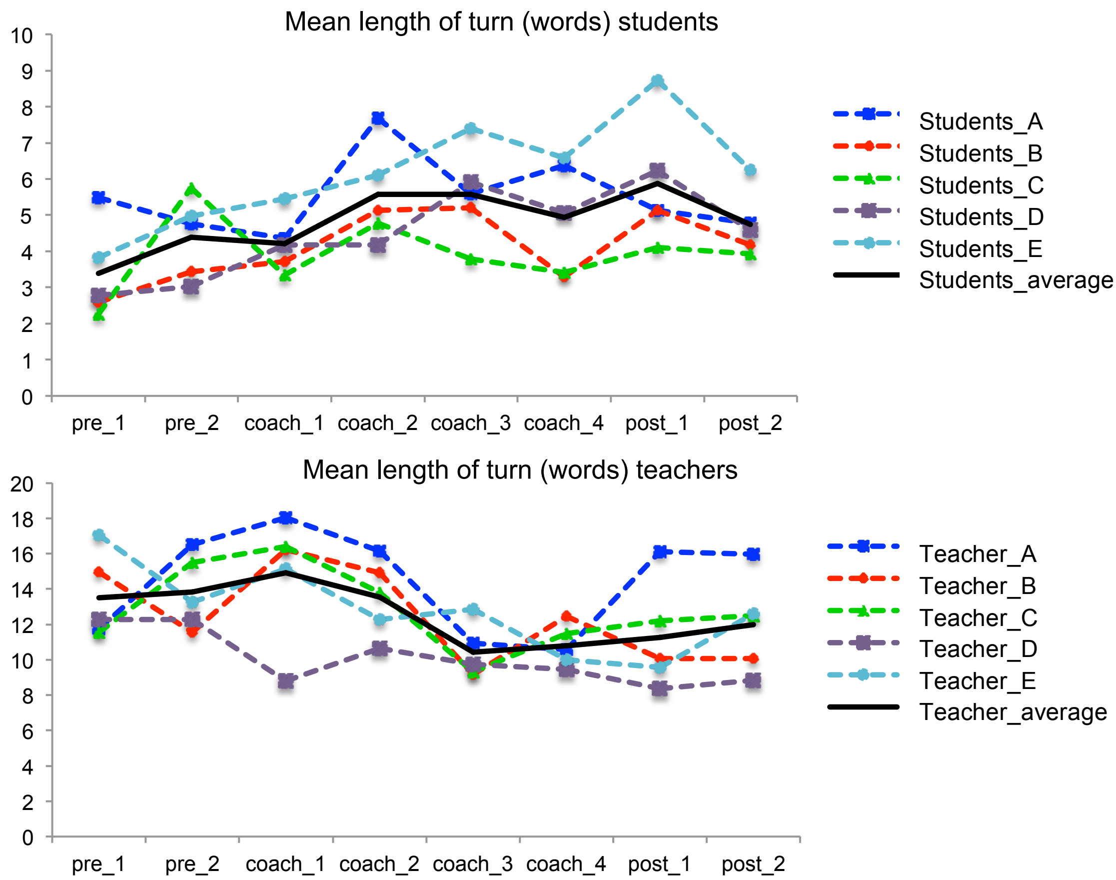


Figure 4 Mean length of turn (words) of students (upper) and teachers (lower) during the intervention

2.3.3 Effects on academic language use

Students.

The second research question aimed to describe whether academic language use of students changed after the teacher intervention. The students in the intervention group used more complex and low-frequency words in the post-measurements. As indicated by the Monte Carlo test, the probability that this difference in lexical sophistication was based on chance was relatively small (M_{pre} = 3.1 vs. M_{post} = 3.8, p = .02, d = .99). The students in the control group showed no difference (M_{pre} = 3.4 vs. M_{post} = 2.9, p = .83, d = .08). The analysis revealed that there were no indications that there was a difference in lexical diversity between the pre- and post-measurements, neither for the intervention group (M_{pre} = 49.1 vs. M_{post} = 51.6, p = .33, d = .21) nor for the control group (M_{pre} = 47.2 vs. M_{post} = 44.0, p = .71, d = .01). The students showed an increase in the number of causal connectives after the intervention. The Monte Carlo analysis indicated that there was a relatively small probability that this difference was based on chance (M_{pre} = 1.5 vs. M_{post} = 5.4, p = .01, d = 2.41). This increase in the use of causal connectives was not observed in the control group (M_{pre} = 2.4 vs. M_{post} = 2.3, p = .58, d = .38). These results indicate that the students' language included more complex words and more causal connectives, without however being more lexically diverse.

Teachers.

Analysis of the teachers' characteristics of academic language indicated an increase in lexical sophistication of the teachers after the intervention. The Monte Carlo test revealed that the probability that this difference was based on chance was relatively small and almost significant with a large effect (M_{pre} = 2.7 % vs. M_{post} = 3.2%, p = .06, d = .75). We observed no clear changes in the control group regarding lexical sophistication between pre- and post-measurements (M_{pre} = 2.6 % vs. M_{post} = 2.7%, p = .34, d = .20). The Monte Carlo analysis revealed that there were no indications that there were differences between pre- and post-measurement for lexical diversity of the teacher language (intervention: M_{pre} = 68.6 % vs. M_{post} = 66.1, p = .66, d = -.19; control: M_{pre} = 66.7 % vs. M_{post} = 66.7, p = .49, d = .01), nor for the number of causal connectives used (intervention: M_{pre} = 3.1 % vs. M_{post} = 4.5, p = .15, d = .54; control: M_{pre} = 10.10 % vs. M_{post} = 8.2, p = .80, d = .03).

2.3.4 Patterns of change

Lexical diversity.

The lexical diversity scores plotted over time, displayed in figure 5, also revealed large differences between the individual students and teachers and large moment-to-moment fluctuations among the individuals with no clear indications for an average increase or decrease ($slope_{students}$ = .15, p = .42; $slope_{teachers}$ = -.76, p = .19). For instance, students and teacher in classroom E showed an alternating pattern with peaks and lows over time. Classroom C had the widest ranges of diversity scores for both students (27.16 to 57.81) and teacher (57.61 to 82.98).

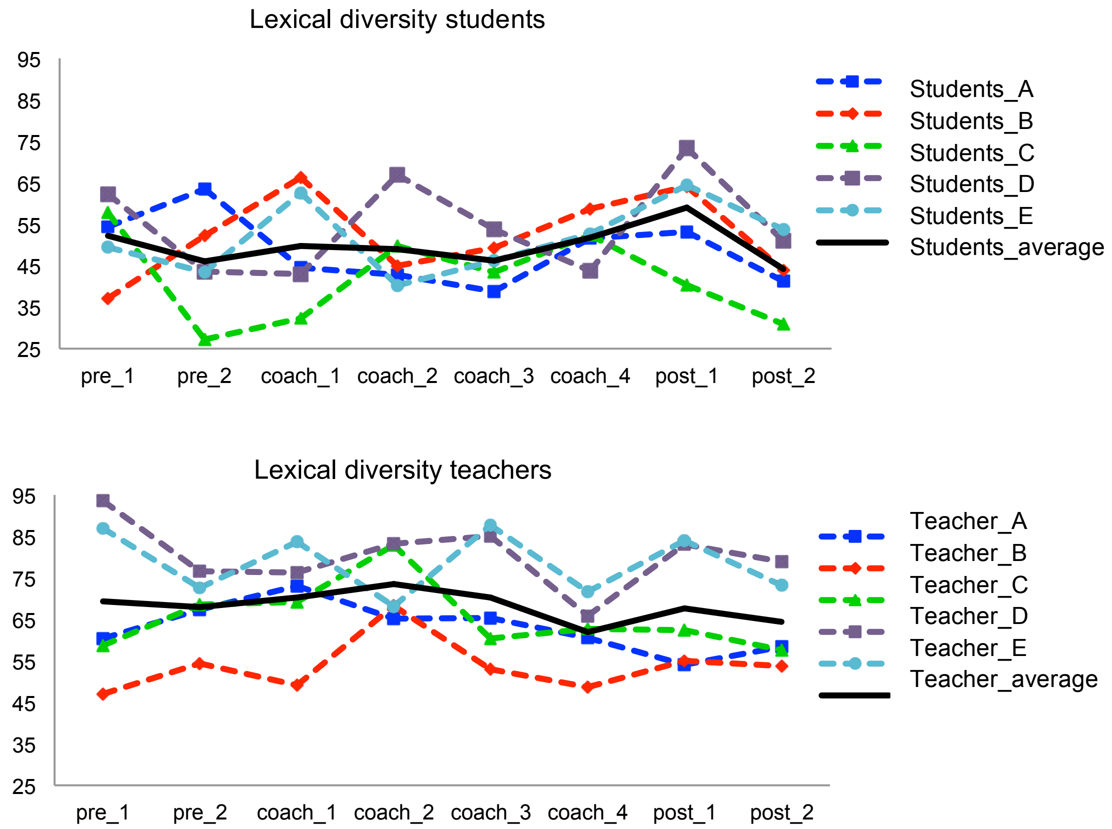
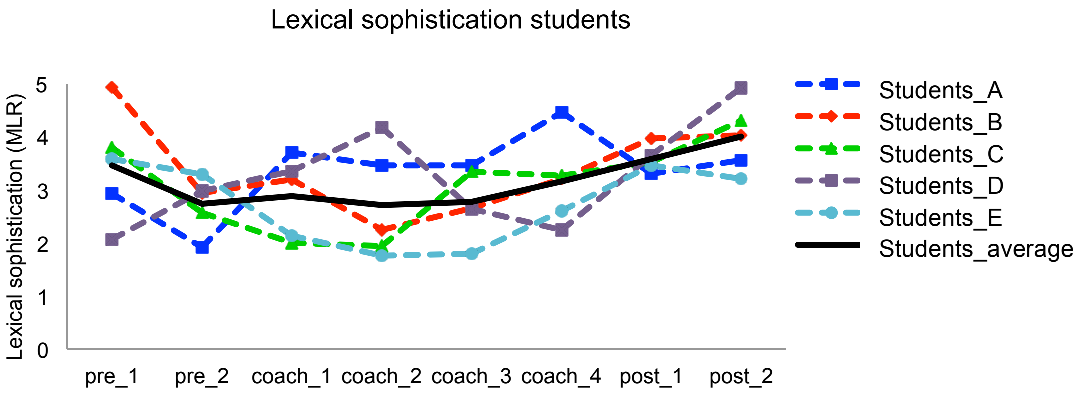


Figure 5 Lexical diversity scores (calculated by means of Index D) of the individual students and teachers in the intervention group during the complete intervention.

Lexical sophistication.

For both students and teachers, on average, lexical sophistication increased over the eight measurements ($slope_{students}$ = .07, p = .06; $slope_{teachers}$ = .11, p < .01). However, moment-to-moment variability was also observed for lexical sophistication. As shown in figure 6, the variation in sophistication was large in particular in case of students.



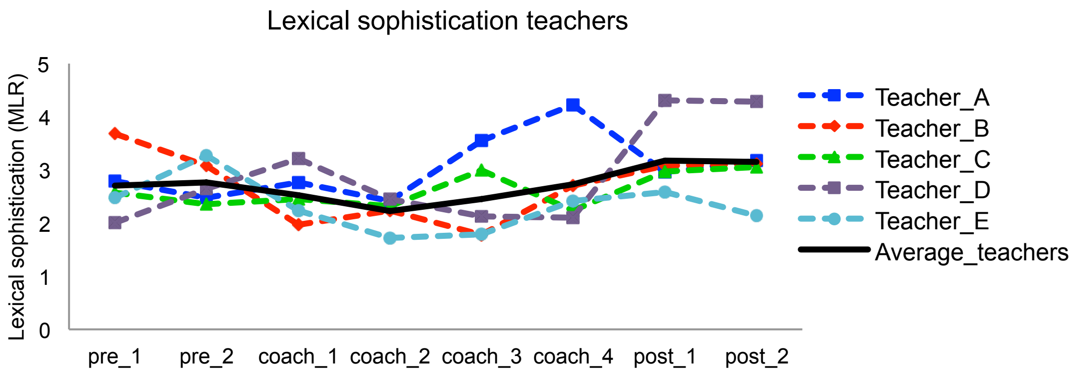


Figure 6 Lexical sophistication scores (calculated by means of MLR) of the individual students and teachers in the intervention group during the complete intervention.

Causal connectives.

The use of causal connectives indicates students' first attempts to combine clauses and to structure information. In general, students' use of causal connectives increased over time (*slope* = .54, *p* = .03) whereas teachers' use of causal connectives mainly varied over time (*slope* = .15, *p* = .25). Figure 7 reveals that classrooms, students as well as teachers, differed to a large extent in the number of causal connectives used. Some classrooms showed gradual increases in the students' number of causal connectives, for example classroom E, whereas the students of classroom B barely used causal connectives during the intervention. What also stood out was that the variation between individuals in the first two lessons was relatively small and the subsequent lessons showed much more variability.

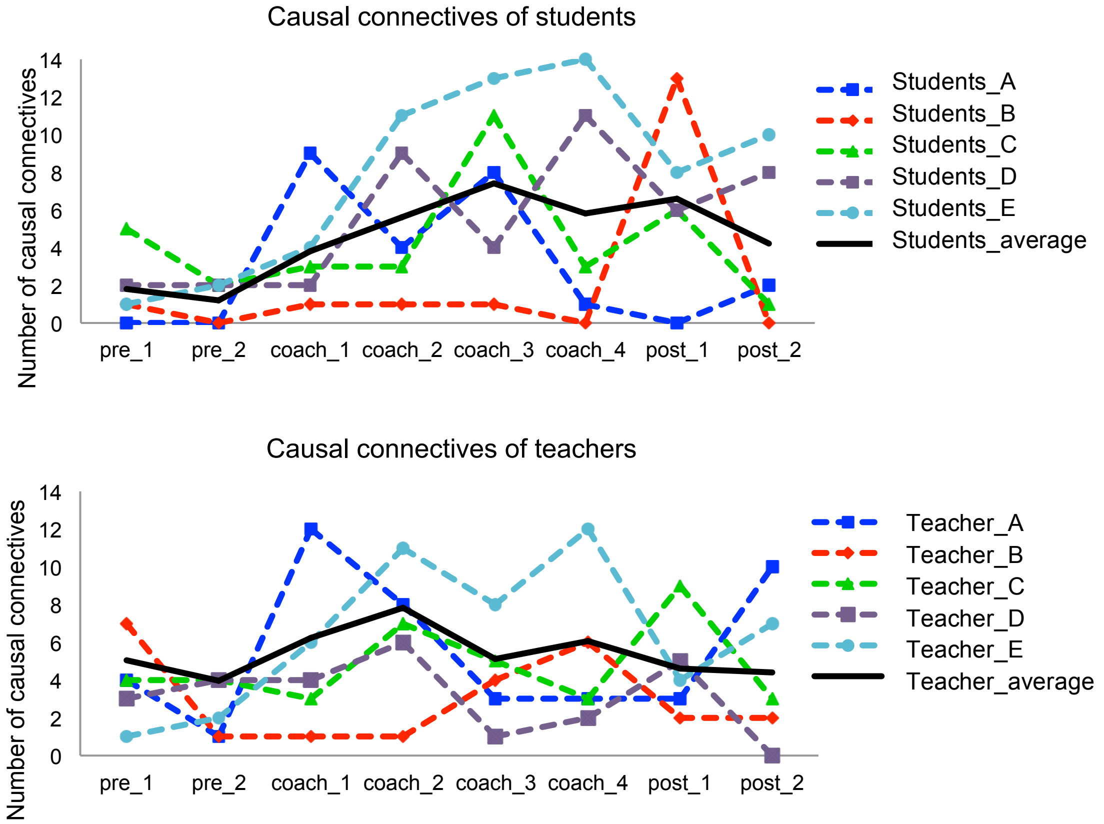


Figure 7 Total numbers of causal connectives per lesson of the individual students and teachers in the intervention group during the complete intervention.

2.4 Discussion

The aim of the current study was to gain insight into the language production of students and teachers during the CMC intervention in kindergarten science lessons. Taking data from an effective teacher intervention directed at changing teacher behavior, we found strong support for a shift in language use reflected in the proportion of student utterances, which significantly increased over time. The turns of the students also increased in length, whereas the teachers' turns decreased. This implies that the teacher invited more as well as more elaborated talk from the students after the CMC intervention. These findings are in line with previous studies (Dickinson, 2001; Oliveira, 2010; Snow, 2014). For instance, Oliveira described this as a shift in interaction from a more teacher-dominated to a more student-centred paradigm. This increase in student language production creates more language-learning opportunities, which is, thus, the first step towards refining the students' language skills to a more academic register.

2.4.1 Theoretical implications

In the current study, academic language was reflected in measures of lexical diversity, lexical sophistication, and the use of causal connectives. In line with our expectations, there were some increases in these measures over the course of the intervention. For the students in the intervention group, lexical sophistication and the use of causal connectives increased. The teachers' lexical sophistication increased as well, whereas the control group – both its students and teachers – showed no changes. This indicates that science learning and language learning, indeed, go hand in hand, even without explicitly focusing on language learning during the science activities. On the other hand, in the literature, we find that academic language is not picked up automatically by students and that teachers play an important role in acquainting students with more sophisticated language skills (Lemke, 1990; Vygotksy, 1987). Offering opportunities for language learning can be done by encouraging student to talk, elaborate, and refine their verbal contributions as well as by modelling academic language. In the context of the CMC intervention, the teachers were trained how to pose open-ended questions at the right moment in order to stimulate the reasoning of students. An increase in open-ended teacher questions – as was found after the CMC intervention – is related to more elaborated and complex student responses (De Rivera et al., 2005; Dickinson & Porche, 2011; French, 2004; Justice et al., 2013; Lee & Kinzie, 2012; Massey et al., 2008; Peterson & French, 2008; Wasik & Bond, 2001; Wasik et al., 2006). As there were no indications of changes in language use in the control group, we argue that the increases in the intervention group may, in part, be related to the better questioning skills of the teachers after the intervention. With regard to modelling academic language, there were only few indications – only weak support for increase in sophistication – that teachers used more academic language themselves. This may indicate that teaching science does not automatically imply using academic language. Wong Fillmore and Snow (2002) already emphasized that teachers have little awareness of the important role they have in acquainting students with academic language. An intervention targeting academic language in kindergarten science activities should therefore pay attention to both open-ended questioning strategies and language modelling strategies.

The changes in lexical sophistication reflect the amount of focus that is directed at content of these science lessons, in comparison to the context of a general teacher training. The fact that lexical diversity did not change leads us to conclude that an increase in lexical sophistication did not automatically cause the language to be more diverse. This result may partly be explained by the fact that, while the science lessons in this study did evoke more complex words, at the same time there was a lot of repetition of these words. Students and teachers did not use, for example, synonyms and hardly showed variation in word use, although they did use more complex words. For future kindergarten science interventions, it is thus important to instruct teachers how to be more linguistically diverse themselves and how to stimulate more diversity in students' language. Another possible explanation might be that, apart from using more complex words, the students also tended to have longer speaking turns. Longer utterances tend to include more function words, which does not add to the diversity of the language. Another factor, which we speculate to partly explain these results, might be the topic of the science lesson (Henrichs, 2010; Van Hout & Vermeer, 2007). Some topics elicit a wider variety of different and complex words, whereas others in some way impose restrictions on the vocabulary choice.

The data from the repeated measures of students and teachers over time showed large variability between individuals and between lessons (Molenaar, 2004). We found little variation between the individuals in the pre-measurements as well as between measurement one and two. However, the individuals followed their own unique path once the intervention had started. This emphasizes the idiosyncratic and nonlinear nature that is characteristic of learning processes. The individual and adaptive character of the CMC program might even strengthen the idiosyncratic character of the data since the coaching is based on the real-time interactions of particular teaching situations which are different for every teacher-student interaction (individuals) and every session (time). Focusing on the actual learning trajectory – by means of repeated measurements while the intervention is carried out instead of only comparing pre- and post-tests – reveals what it is that actually changes in the course of an intervention. The intra-individual variability can be used as a main source to describe how changes in development occur (Nesselroade & Molenaar, 2010; Nesselroade & Ram, 2004).

Although neither the inter- nor the intra-individual variability that we found is surprising, few studies have been carried out that specifically focus on interventions that take these differences into account. In recent years, however, methods and approaches that are sensitive to capturing individual differences and changes over time have become more common (e.g. Lichtwarck-Aschoff & Van Geert, 2004; Molenaar & Campbell, 2009). The results of this study support the relevance of this method for future research, whether or not it is combined with taking more global measures, such as group averages.

Beyond the promising findings of this study, there is one major limitation to point out. The results in this study are based on ten motivated teachers. It is a point of discussion whether the small number of participants limits the generalizability of the findings, and whether these results could be replicated with less motivated teachers. Ryan and Deci (2000), however, indicate that an intervention should be first performed on highly motivated participants. The next step is to perform the intervention on larger and more representative groups of teachers. The present study could be the precursor to a larger study.

2.4.1 Practical implications

The insights into academic language use in kindergarten science lessons, which were provided by this study, result in practical implications considering the role of explicitly targeting language learning when teaching science. The findings indicate that changes occurred in the verbal teacher-students interaction and academic language production after the video feedback coaching for teachers. Although this intervention was not specifically aimed at language use, we found positive changes regarding both language use as well as some aspects of academic language of students, which support the influence of video feedback coaching in professionalization trajectories (Van den Heijkant, 2006; Seidel et al., 2011). Nevertheless, in practice, awareness in teachers of the need to stimulate the use of academic language is important if the language use in science lessons has to be improved. Previous intervention studies have shown that the use of academic vocabulary can be successfully trained (French, 2004; Guo et al., 2016; Henrichs & Leseman, 2014; Hong & Diamond, 2012). Science lessons represent an appropriate context in which to acquaint students with academic language (Conezio & French, 2002; French, 2004; Peterson & French, 2008; Samarapungavan, Mantzicopoulos, Patrick, & French, 2009), and future research should focus on investigating whether using and eliciting more academic language can be an explicit goal within video feedback coaching interventions for teachers within this context, resulting in more academic language of students (Schleppegrell, 2001; Wong Fillmore & Snow, 2002). From this study, video feedback coaching for teachers has been shown to be an effective way to create opportunities for students to learn academic language.