Article

CHILD LANGUAGE TEACHING & THERAPY

Child Language Teaching and Therapy
1–18

DOI: 10.1177/0265659015593608
clt.sagepub.com
SAGE

Developing and testing EVALOE: A tool for assessing spoken language teaching and learning in the classroom

Marta Gràcia and Fàtima Vega
University of Barcelona, Spain

Maria José Galván-Bovaira
University of Castile-La Mancha, Spain

Abstract

Broadly speaking, the teaching of spoken language in Spanish schools has not been approached in a systematic way. Changes in school practices are needed in order to allow all children to become competent speakers and to understand and construct oral texts that are appropriate in different contexts and for different audiences both inside and outside school. Here we describe the construction of EVALOE ('Escala de valoración de la enseñanza de la lengua oral en contexto escolar' [Assessment scale of oral language teaching]), a flexible tool designed to assess the teaching and learning of spoken language in schools. The first part of the tool is an observation scale (for use in the classroom), and the second is a semi-structured interview carried out with the teacher after the observation. The instrument is designed to assess how teachers facilitate the development of spoken language. It also aims to further interdisciplinary collaboration with regard to the introduction of changes in teaching practices in order to promote the development of communication friendly classroom environments. A pilot test and expert review of the tool was carried out. Thirty-nine schools then participated in the field test, which involved 39 professionals (23 speech and language therapists, eight educational psychologists, and eight teachers) and two university students. Eighty observations, 34 in a preschool setting and 46 in primary schools, were carried out. The results suggest that EVALOE is a useful tool for assessing the teaching and learning of spoken language.

Keywords

assessment tool, communication friendly environments, classroom observation, interdisciplinary collaboration, spoken language, teaching and learning

Corresponding author:
Marta Gràcia, University of Barcelona, Facultat de Psicologia, Departament de Psicologia Evolutiva i de l'Educació, Passeig de la Vall d'Hebron, 171, Barcelona, 08035, Spain.
Email: mgraciag@ub.edu

I Introduction

Spanish legislation on education (Mallart and Sarramona, 2013; Ministerio de Educación y Ciencia, 2013) places special emphasis on linguistic competence (speaking, listening, reading and writing). Specifically, it stresses the need to teach children to produce oral texts that are appropriate to particular situations, and the need to organize the course content in a way that allows children to express their thoughts and opinions, to respond to those of others, and to expand their knowledge. These oral texts must be constructed in a context that encourages interaction and dialogue. Although since the 1990s a succession of laws passed in Spain have aimed to develop linguistic competence (Ministerio de Educación y Ciencia, 1990, 2002, 2006, 2013), the inclusion of these aims in everyday teaching practice has been an arduous process. In this context, it would be useful to have a tool for assessing the skills and strategies used by teachers to address these objectives in their normal classroom activities and for evaluating their ability to promote the development of communication friendly environments.

The oral language activities that teachers most frequently use in the classroom in Spain involve asking students to talk about previously prepared written work: for instance, stories or explanations of events in preschool contexts; oral presentations of written assignments in primary school; or sometimes participation in debates about different academic or social themes (Sánchez-Cano, 2007, 2009; Rolla, Arias, Rivadeneira, Coronado and Romero, 2012; Rolla, Coronado, Rivadeneira, Arias and Romero, 2012). Previous studies have focused on school activities designed to further the development of pragmatics in children (Blum-Kulka and Snow, 2002) both in pre-school years and at primary school. Now we need to explore whether teachers in our country think that oral texts arise naturally, or whether formal or explicit training is required in order to help children develop their pragmatic skills.

1 Conceptual perspective

This study is based on ecological and functional theories of language development in natural contexts and on the sociocultural approach to teaching and learning. The first argument is that interactive processes between children and significant adults are crucial for language development. During these processes the children's cognitive and language skills are activated, as are certain strategies on the part of the adult, which favour the acquisition and growth of linguistic competence (Tomasello, 2003, 2014). The process begins in the family context; the strategies that specifically promote oral language development at this stage have focused on the communication management through turn-taking (Kaye, 1982; Snow, 1972). Adult adaptations (Cross, 1977) to children's linguistic competence, specific strategies of over-attribution of meaning, expansions and evaluative feedback (Girolametto et al., 2004; Moerk, 1983), amongst other strategies, are relevant for children's language development. One of the strategies that has proved particularly effective for language development is to follow the child's lead and interests (Girolametto et al., 2004; Manolson, 1992). Some studies (del Rio and Gràcia, 1996; Gràcia, 2003) have proposed a three-bloc classification of these strategies and adaptations: (1) communicative context management, (2) adult language adaptation, and (3) implicit educational strategies.

School is the second natural context in which linguistic competence continues to progress and where teachers use some of the strategies that parents apply at home when interacting with their children (Justice and Ezell, 1999). During interaction at school, teachers, as the more competent participants, help children to improve their communication by providing scaffolding. As children become more independent, this help is gradually withdrawn (Wood et al., 1976, Gràcia, 2003; Gràcia and Segués, 2009). Although teachers and children are often involved in

activities and situations that resemble the family context, the main aim at school is to consolidate proficiency in academic language (Snow, 2010). The use of reciprocal teaching helps children in school to make impressive gains in reading comprehension, and the use of specific dialogic strategies by both teachers and pupils encourages more critical and elaborated contributions (Mercer, 2008).

Exploratory talk (Mercer, 2008) is an example of dialogue that can foster more effective learning in the school context, where the establishment of specific rules embodies the essential qualities of reasoned debate. As many authors have pointed out (Wells, 1999; Mortimer and Scott, 2003; Waite et al., 2003), many factors contribute to the effectiveness of classroom teaching, and talking with the teacher is a potential influence on the development of students' knowledge and understanding. However, researchers have put forward persuasive and influential arguments for the impact of the quality of teacher–student dialogue on the development of children's understanding of science and other curriculum subjects (Mercer, 2008), as in the case of oral language competence (Gràcia et al., 2015).

There is evidence that the characteristics of the educational context and the language that teachers use to address to children (de Rivera et al., 2005; Girolametto et al., 2004) partially account for children's linguistic outcomes during the school years. The conversational strategies that teachers use when interacting with children and their classroom organization seem to be among the elements that contribute to creating classroom environments that facilitate the development of spoken language (Mercer, 2008).

2 *Why develop a tool to assess oral language teaching at school?*

A review of previous work on the assessment of language (Ygual and Cervera, 1999) shows that researchers have developed and used tools to assess communication and language in natural situations, and have subsequently used the results to conduct naturalistic interventions. Other instruments have been designed to assess more general aspects of the classroom including the organization of the activity, the type of grouping and teachers' strategies (Dockrell et al., 2012a, 2012b). Tools have focused on the skills, knowledge and expertise of the school staff, the school atmosphere, and on classroom observation (Marinac et al., 2008; Marshall and Lewis, 2014; Mashburn, 2008; Smith et al., 2008). Classroom observation tools have been designed to capture emotional, social and interactive support (ECERS; Harms et al., 1998), the quality of the classroom environment such as space, activities, interaction, and program structure (CLASS; Pianta et al., 2008), and curriculum, learning environment, safety and health (Abbott-Shim et al., 1992). However, none of these tools directly assess oral language skills, and so are not sensitive to key elements in the classroom that support spoken language development. The Communication Supporting Classrooms Observation Tool (CsC; Dockrell et al., 2012a, 2012b) has been constructed with this aim in mind, as a flexible three-part tool that allows educationalists to examine the language learning environment in the classroom.

However, the CsC does not seem to have been constructed to capture the interaction between teachers and pupils: that is, how teachers try to involve pupils in the activities, or the strategies they use to encourage pupils to self-regulate their participation during conversational classroom activities. Although the first and second parts of the CsC tool are sensitive to capturing the teacher's ability to create a language learning environment ('the classroom is organized to emphasize open space or smaller group activities'), and the third part captures language learning interactions ('adults repeat what the child says and add a small amount of syntactic or semantic information' or 'Turn-taking is encouraged'), it does not capture other important elements of the instructional design of language contents during classroom activity in primary school.

In primary school, teachers must be familiar with the elements that define the pragmatics of language and must be able to incorporate strategies in their educational programme, specify the learning aims, and carry out systematic evaluations. The teaching of spoken language should not be limited solely to language classes (either mother-tongue or foreign). Spoken language is the key instrument for learning, throughout the syllabus; it is vital to make children aware of the importance of academic language, not only in order to improve their written language, but also to improve the quality and complexity of their construction of oral texts (Snow, 2010).

3 The EVALOE tool

In this article we describe the construction of EVALOE ('Escala de valoración de la enseñanza de la lengua oral en contexto escolar' [Assessment scale of oral language teaching]), a flexible tool designed to assess how oral language is taught in the classroom (Gràcia et al., 2015). EVALOE is a non-standardized tool in two parts. The first is an observation scale comprising 30 items, designed to assess three dimensions of oral language teaching and learning in the classroom: (1) Context and communication management; (2) Instructional design; and (3) Communicative functions and strategies. The second part comprises a set of questions for a semi-structured interview with the teacher who has been observed, and it assesses these aspects in greater depth. One of the tool's major strengths is that it explores the interaction between teacher and children in the classroom with regard to spoken language.

Following on from other studies (Dockrell et al., 2012a, 2012b), our aim was to create a flexible tool able to identify key classroom aspects of oral language teaching and learning in order to help school staff profile and monitor daily practices and generate reflection on these issues. An instrument of this kind can help practitioners to identify the changes needed in their classrooms to create environments that are more communication friendly and enhance language learning, and then to monitor the changes introduced over an extended time period. Classroom observations using EVALOE can also promote the introduction of new methodologies for spoken language teaching and learning (Gràcia et al., 2015).

4 Research questions

In this study we describe the creation of an instrument for examining how teachers teach spoken language at school, and assess its usefulness for improving teachers' practices in the classroom. We address two main questions:

- Is it possible to construct a tool suited to the cultural and linguistic characteristics and needs of Spanish schools which is able to assess how teachers facilitate the development of spoken language in the classroom?
- Is this tool useful for detecting aspects of the practices of school staff that might be improved in order to create more communication friendly environments and help children to develop their competence in listening and speaking?

II Method

In this section we first describe the process applied to create EVALOE and the characteristics of the first version of the tool. Second, we describe the process through which the reliability of the first version of the observation element of the tool was evaluated. Third, we describe the expert assessment and revision process carried out in order to check its content validity. Finally, we report the procedures for pilot test and field test studies.

1 Development of the EVALOE tool

The construction of EVALOE was based on a current review of the literature and on the results of our previous studies (del Rio, 1993; del Rio and Gràcia, 1996, 2003; Galván-Bovaira, 2008; Galván-Bovaira and del Rio, 2009; Gràcia, 2001; Gràcia et al., 2012; Sánchez-Cano, 1993, 2007), which focused on the analysis and improvement of the communicative aspects of interactions between mothers, practitioners/teachers and children. Before conducting some of these studies we had carried out a literature review to select definitive studies. When we embarked on the construction of EVALOE we conducted a further review using similar key words and selection criteria. This allowed us to identify studies that complemented the empirical evidence-based studies reviewed previously, including studies with a diversity of designs. We also considered policy documents from the departments of education of the Catalan and Spanish governments.

On the basis of this evidence, we then developed the two parts of the tool. The first part is an observation scale, and the second a set of questions for conducting an interview. The items devised for the observation scale were subsequently reviewed to check that they represented the key aspects of our conceptualization of the methodology of spoken language teaching by five team members who used the scale to observe videotaped classes. After discussion, decisions were taken regarding the adequacy of the items, and the three possible answers of each item. This process was repeated four times until agreement was reached on the wording of the items and the answers. The team also prepared 22 questions to be used in the semi-structured interview. Both parts of EVALOE comprised three subscales or dimensions, designed to assess:

- Context and communication management: the physical and social learning context and the strategies used to manage the communication between teacher and pupils and between pupils;
- Instructional design: the strategies that a teacher deploys during class sessions in order to share the language aims of the session with the pupils, to check previously shared knowledge and to learn linguistic content at the end of the session; and
- Communicative functions and strategies: the communicative functions (giving information, asking for information, regulating the action, and so on) that adults teach pupils, the way in which pupils use these functions, the educational strategies (expansion, positive feedback, and clarification) the teacher uses, and the pupils' response to these strategies.

The first version of the observation scale had 10 items in each of the first and second subscales and 17 items in the third subscale, all items consisting of descriptive statements of some aspect of the class with three options that the observer could select to describe degree to which each aspect was present. The first version of the guide for the semi-structured interview with the teacher after the observation comprised 22 questions, grouped under the three dimensions. A manual was prepared that included a theoretical framework, a glossary of the most important terminology, a description of the tool and guidelines for its use, correction, and interpretation.

2 Reliability of the observation scale

Qualified educational psychologists and speech and language therapists conducted an inter-rater reliability study for the first version of the observation scale. Eight classroom observations were conducted at one school. Two raters who independently scored the session observed each classroom concurrently. Inter-rater reliability for EVALOE was high, Cohen's kappa coefficient $\kappa = .78$ (Cerda and Villarroel, 2008). Individual items with low reliability were reviewed to ensure that

there were no problems with wording and no items were modified or deleted after the reliability process.

3 Expert review of EVALOE

Five experts, each with more than 20 years of experience, reviewed both parts of the tool: two researchers on language teaching and learning from different theoretical orientations; one speech and language therapist/educational psychologist who coordinates a public service for children with language and hearing disorders; and two educational psychologists from a public psycho-educational service that helps teachers to provide all children the support they need at school. The experts were requested to express their views on: (1) specific aspects of the manual; (2) the clarity, suitability, and importance of the items; (3) the type of response (in terms of importance, frequency or other criteria); (4) the length; (5) the usefulness, clarity and length of the questions included in the interview guide. Five research team members reviewed the experts' comments and suggestions, took notes and suggested changes. They then met and discussed all the suggestions generated by the experts' comments until a consensus was reached regarding changes that should be introduced in order to improve the tool. This process was considered as a test of the tool's content validity jointly with the theoretical process applied to design and construct the first version from the review of empirical evidence. The second versions of the manual, observation scale items and question guide for the interview were drawn up as a result of the discussion process. Thirty items were retained in the first part (observation tool) and there were no relevant changes in the second part (interview).

4 Pilot test

For the pilot test, participants (observed teachers and observers) were recruited from a single school in Barcelona. The lead researcher contacted the principal of the school by telephone to describe the study and explain what was required. When the school agreed to participate, the principal contacted members of staff (though not the teachers who were to be observed) in order to discuss the study. Ethical approval was obtained from the University of Barcelona Ethics Committee, and written consent was obtained from each participant. The informed consent document included only general information about the study, so as not to bias the results. Once consent had been obtained from all the participants, we gave the observers a research package containing: the manual, a copy of EVALOE, an envelope in which to return items to the University after completion, and an address where the audiotape of the interview was to be sent. Observers were asked to choose two teachers from two different levels for observation and interview.

The observed participants in this pilot study were five female teachers with more than five years of teaching experience. The observations were carried out at the following levels: Reception 3 years, Reception 4 years, Reception 5 years, Primary year 1 and Primary year 2. A speech and language therapist and an educational psychologist from the school staff participated as observers and were able to raise any queries they might have with the lead researcher at any time during the process. The observed classes lasted around 45 minutes, and the observations took place in the morning during the daily teaching activities. The observer was sited at the end of the class and did not participate in the activities. The interviews lasted around 30 minutes, and were carried out during the midday break at the school, in the teacher's room, with only the interviewer and the teacher present.

A qualitative approach was adopted to identify problems with the tool after the pilot study. We administered a questionnaire to the teachers who had participated in the pilot study to determine their perceptions of the wording and clarity of the items, and also the usefulness of the tool

(observation scale and interview). Once they had answered the questionnaire, the research team met them to discuss their answers in order to decide on the changes that needed to be introduced. However, no major changes were introduced in the manual, the observation scale items, or the interview questions.

5 Field test

a Final version of the observation scale. The final version of the first part of the tool comprised 30 items consisting of statements that were representative of the three dimensions identified previously. The items were grouped in three subscales (eight items in subscale 1, seven in subscale 2, and 15 in subscale 3). The items were rated on a scale from 1 to 3, considering the whole class session observed; there was a space for adding observations about each item if necessary. Appendix 1 shows a sample of three items from subscales 1 and 2 and 6 items from subscale 3.

b Procedure for the field test. The procedure used to recruit participants (both observers, and observed teachers) for the field test was similar to the one used in the pilot study. Team members contacted professionals at two public services, the Resource Centre for Hearing Impairment and the Psycho-Educational Service, territorial educational services, and teachers, educational psychologists, and speech and language therapists working at schools in two different regions of Spain (Catalonia and Castile-Leon). Schools were informed of the nature of the research by the researchers or by the professionals at the services mentioned, and decided which teachers would be observed and which would participate as observers and interviewers. In some cases university students from the research team were observers. Ethical approval was obtained from the University of Barcelona Ethics Committee, and written consent was obtained from all participants who agreed to participate.

The guidelines for the participants (observers and teachers being observed) were the following:

1. Observers were to observe classes at different educational levels (pre-school and primary school).
2. The observations could be done during classes in any subject (maths, language, sciences, arts and craft, and so on) or during any kind of activity (conversation, play, and so on).
3. Each observer could observe more than one class.
4. Each teacher could be observed only once.
5. Teachers had no information about the characteristics of the tool and the research aims.
6. Teachers could participate either as observed teachers or as observers.
7. The interview had to be carried out on the same day as the observation or the next day, and it had to be audiotaped.

Altogether, 80 classes were observed, and 47 teachers were interviewed. The conditions, length and time to develop the observations and interviews were the same as in the pilot study. Of the 47 interviews, 24 were audiotaped and in 23 cases the interviewer took notes. Half of the 24 audiotaped interviews were transcribed.

Thirty-nine schools with a middle–high socioeconomic status participated in the study: 24 were state schools (from pre-school to Primary school year 6) and 15 were private (from pre-school to Primary school). Thirty-nine professionals from state or private services (23 speech and language therapists, eight educational psychologists, eight teachers) and two university students participated as observers, and 80 teachers, all with more than five years of experience, were observed. Women accounted for 90% of observers and 95% of observed teachers. The proportions of teachers

Table 1. Participants in the field test.

Schools	Observed teachers	Observers			
		Speech and language therapists	Teachers	Educational psychologists	University students
39	80	23	8	8	2

Table 2. Characteristics of the schools and the class years in which the observations took place.

		Number of observations		Total
		State school	Private school	
Pre-school	Reception (3 years)	6	6	12
	Reception (4 years)	4	6	10
	Reception (5 years)	3	9	12
	Total	13	21	34
Primary school	Year 1 and 2 (5–7 years)	11	5	16
	Year 3 and 4 (7–9 years)	9	6	15
	Year 5 and 6 (9–11 years)	10	5	15
	Total	30	16	46
Grand total		43	37	80

Table 3. Characteristics of the classes in which the observations took place.

	Language	Conversation	Sciences	Maths	Arts and crafts	Total
Pre-school	11	12	10	0	1	34
Primary school	12	8	20	3	3	46
Total	23	20	30	3	4	80

Table 4. Characteristics of the observers and age groups where the observations took place.

	Teacher	Speech therapist	Educational psychologist	University student	Total
Pre-school	10	11	8	5	34
Primary school	5	23	4	14	46
Total	15	34	12	19	80

observed and observers/interviewers, the courses observed, characteristics of the schools and subjects observed are shown in Tables 1–4.

III Results

1 Psychometric characteristics of the observation scale

From the responses returned, an inter-item Pearson correlation matrix was made using the IBM SPSS statistics program. Each of the 30 items in this first part of the tool was correlated with the

Table 5. Pearson's correlations for total subscale values.

	Subscale 1	Subscale 2	Subscale 3	Total Scale
Subscale 1	1	.452**	.562**	.749**
Subscale 2	.452**	1	.535**	.755**
Subscale 3	.562**	.535**	1	.782**

Note. **Significant at $p < 0.01$.

other items on the scale in order to establish whether any item should be reviewed or removed. As none of the correlation values were lower than 0.30, no items were removed.

A further Pearson correlation matrix was made correlating the three subscales to see whether they shared common factors. Table 5 shows that although the correlations between the three subscales were significant, the values were low (from 0.45 to 0.60); correlations between each subscale and the total of the scale were slightly higher (from 0.70 to 0.80). Nevertheless, the decision was made to maintain the three subscales due to the experts' comments and the feedback from the teachers.

We also calculated the scale's reliability and the Cronbach's alpha coefficient for the total items of the scale was 0.87. We then measured the factor 'Cronbach's Alpha coefficient if some item is removed', and found no variance.

2 *Findings from the semi-structured interviews*

The responses to the 47 interviews carried out during the field test were discussed by the research team in order to evaluate the final version of the interview guide. Some useful changes to the guide were identified that would make the interview more objective and straightforward. Changes proposed for the interview manual were: (1) placing greater emphasis on the need for interviewer training; (2) improving the clarity of wording of the questions; (3) introducing tips to ask for clarification in case of ambiguous answers; (4) increasing the specificity of the recording of the answers. For example, when we asked teachers 'Do you teach students how to ask questions or how to make requests?' the answer was often another question like 'Can you repeat the question please?' or 'I'm sorry but I don't understand the question'. We decided that this specific question should be reworded to indicate to the teachers more precisely what we wanted to know. For this reason, in the last version of the interview we reformulated the question and, after providing a brief explanation of the meaning of 'communicative functions', we asked them 'What communicative functions do you teach your students?'

The other proposed changes were linked to the content of the questions: (1) increasing the clarity of the questions and (2) increasing the clarity of the examples given to facilitate the understanding of the questions. For example, when we asked teachers 'What do you usually do in your classes to help students to manage conversations?' they often thought that we wanted to know whether they tried to encourage all children to participate in all activities. Therefore, their answers missed the point of the question; in this case the problem was the misunderstanding of the concept 'managing the conversation'. Other changes were introduced in order to help the interviewer contextualize and manage the interview, such as the inclusion of an introductory paragraph in each part along with some examples of how to advise teachers if they had difficulties with certain questions. Furthermore, in the cases when the answer to the question was 'No', interviewers would be advised to tell teachers to skip the question and go on to the next one. It was proposed that the questions be grouped into three dimensions as in the observation scale. Appendix 2 shows a sample of questions of the interview guide from the proposed revised version.

IV Discussion

EVALOE, the tool described in this article, introduces a new perspective in communication and language assessment in classrooms. Two research questions were proposed for this study:

1. Is it possible to construct a tool suited to the cultural and linguistic characteristics and needs of Spanish schools which is able to assess how teachers facilitate the development of spoken language in the classroom?
2. Is this tool useful for detecting aspects of the practices of school staff that might be improved in order to create communication friendly environments and help children to develop their competence in listening and speaking?

Research question 1

The answer to our first research question is affirmative. The aim of the first part of the tool, the observation scale, was to help practitioners and school staff to observe the communicative behaviour of teachers and pupils during classroom activities. The second part of the tool, the interview with the teacher, was designed to discuss the class session observed and the conceptions, plans, and activities that could not be observed directly. After designing the tool, we carried out a field test of the reliability of the observation scale.

Our results demonstrate the reliability of the observation scale based on the results of the inter-rater reliability and Cronbach's Alpha. The results also show that all the items included in the three subscales capture specific information, and that each item contributed to the reliability of the scale, because none of the alpha values fell when we removed an item. Although correlations between the three subscales showed that they are not independent, the correlation values were low (from 0.45 to 0.60). These results suggest that the first subscale captures specific information about context and communication management, the second subscale information about instructional design, and the third subscale information about how teachers teach communicative functions, how children learn to use these functions, and about the strategies teachers use to help children improve their language. This interpretation is concordant with the experts' comments about the need to maintain the three subscales, given the clear differences between the information captured by the items included in each one.

During the tool's construction, field testing and data analysis, a number of issues arose which are relevant for practitioners and policy-makers. Based on experts' opinions of the tool (considered as the content validity in this study), in this section we discuss the results obtained in relation to the objectives of the study. First, we present evidence about the claim that the three subscales of the scale are sensitive to different aspects of the teacher's communicative behaviour. Second, we discuss our claim that the tool can help to improve school practices.

Regarding the tool's applications, the three dimensions identified and included (Context and communication management; Instructional design; and Communicative functions and strategies) are consistent with those mentioned by other authors as relevant facilitators of language development in natural contexts (Cross, 1977; del Rio and Gràcia, 1996; Marinac et al., 2008; Mercer, 2008, 2010; Moerk, 1983).

Conversational management, or turn-taking, is identified in several studies (Kaye, 1982; Manolson, 1992) as the basic skill for further academic or formal conversation (McHoul, 1978, 1990; Mercer, 2010; Snow, Lawrence and White 2009). One of the tool's main contributions is to help teachers and practitioners not only to assess the communicative activities carried out by the teacher and children in the classroom, but to focus on the teacher–child and child–child interactions at the same time, thus capturing the communicative behaviour of all the participants

simultaneously. This type of assessment is made possible by some of the items included in the first subscale. The experts who reviewed the tool before the field test and the professionals who used it during field test mentioned this as one of the scale's strengths.

The second dimension assessed by EVALOE is instructional design. During the testing of the tool in schools in Spain, observing children from 3 to 12 years old engaged in different types of activities (maths, sciences, conversation), we found that the items on this second subscale were sensitive to the efforts of teachers to make the objectives of the classes explicit and to raise the children's awareness of the linguistic contents they are learning. As we can see in the second subscale in Appendix 1, if the item is scored 3 this means that the teacher tried to explain the oral language aims clearly and in detail. In addition, although it is not included in Appendix 1, an item on the second subscale assessed whether teachers asked the students to evaluate their communicative behaviour during oral language activities in the classroom; again, a score of 3 indicates that a teacher specifically did this in class. The literature has shown that students can learn academic language in contexts other than school (with their family, friends, reading, via websites, in extracurricular activities), but numerous studies have advocated the explicit teaching of academic vocabulary crucial across content-area learning (Snow, 2010) and the assessment of the progress in academic language acquisition (Justice et al., 2009).

The third dimension is related to educational functions and strategies. It explores the teacher's ability to make students aware of the different communicative intentions that we display in our interactions with others in different contexts (e.g. asking for information, receiving information, reflecting about language). Some of the items included in subscale 3 were designed to identify the strategies (e.g. expansion, clarification, positive feedback) used by teachers to help students master the forms of expression that these intentions can take, depending on the context, the interlocutor, the content, the activity, previous shared knowledge, and so on. These strategies can help children to broaden and develop their academic language when presenting topics, making arguments, defending proposals, and synthesizing information (Snow, 2010).

Research question 2

With regard to the second research question, the first part of EVALOE is based on observation. Again the answer is affirmative, in accordance with other authors (Dockrell et al., 2012a, 2012b; Lever and Sénéchal, 2011). The observation scale provided teachers with criteria and guidelines for observing interactions with children and evaluating their strengths and needs, thus helping to create an environment that fostered listening and speaking skills. For instance, when we introduced an item that assessed whether teachers encourage children to manage conversations and respond systematically to their leads or expand their utterances at a specific time, we were providing them with criteria and guidance for promoting important behaviours with respect to what children already do and what might be improved, as well as a set of strategies to use. The tool can also be used by practitioners working from an interdisciplinary and collaborative approach (Bonals, 2007; Dockrell et al., 2012a, 2012b; Martín, 2005) in order to broaden children's opportunities to participate in academic conversations, to assess their own interventions, to be aware of their skills, and so on.

EVALOE raises professionals' awareness of aspects of the interaction that often go unnoticed. Items on which the teacher was given a score of 1 (i.e. aspect or behaviour not present) during the observation can be discussed by the observer and the teacher jointly, so that they can look for ways to facilitate language development that are more in line with the theoretical perspective of conversational methodology on which the instrument is based. This methodology encourages children to speak in class, in large or small groups, and encourages teachers and other professionals to facilitate settings to promote spoken language in daily activities (Gràcia et al., 2015).

3 Limitations and next steps

The testing of EVALOE has demonstrated its potential for assessing language and communication learning at pre-school and primary school level in state and private schools in Spain and for helping teachers to improve the quality of their interactions with children inside the classroom. Before it can be applied in other non-Spanish cultural and linguistic contexts, however, some of the items and questions may need to be adapted according to the contextual features. This adaptation should take various aspects into account:

1. the linguistic diversity of the classes and the number of languages used at school (one or more than one), and children's competency in these languages;
2. the cultural diversity at schools;
3. available professional resources such as speech and language therapists;
4. teacher training to identify skills in need of improvement in children's language development; and
5. the inclusion of group or conversational activities using oral language.

EVALOE was developed for use in classrooms in inclusive schools with children aged from 3–12 years of age. Further evaluation would be required to see if it would be useful in older children at secondary school and in special schools or social-risk schools. The 80 observations were conducted in schools with a middle–high socioeconomic status; further data is required from observations carried out in low socioeconomic status schools in order to assess the tool's effectiveness in other contexts. Finally, the observations were made during the morning to maximize observational opportunities, and so including afternoon observations in the same classroom would determine whether the time of day of the observation is a relevant variable.

The semi-structured interview has considerable theoretical and practical interest, and the revised version now needs to be tested with more teachers in order to establish its potential as a complement to the observation scale to aid teachers and school staff in reflecting on their professional practices from a collaborative perspective and to create more communicative contexts.

EVALOE has clear value as an observation scale and as a tool for gathering information from the teacher through the interview, but it must also be able to initiate interventions that promote language development and awareness on the part of both teachers and students related to the process of language. We plan to design interdisciplinary training processes to help teachers to address the speech, language, and communication needs that the use of the EVALOE has highlighted. EVALOE may then prove to be a useful tool in evaluating the effectiveness of this training by recording changes in practice within school classrooms.

Acknowledgements

We thank the professionals, services, schools and students for their participation in the study, and Dr Carles Riba, University of Barcelona for his advice regarding methodological issues.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This study was supported by the Agrupació de Recerca en Ciències de l'Educació ('Association for Research in Education Sciences'), University of Barcelona.

References

Abbott-Shim M, Sibley A, and Neel J (1992) *Assessment profile for early childhood programs research manual*. Atlanta, GA: Quality Assist.

Blum-Kulka S and Snow C (2002) *Talking to adults: The contribution of multi-party talk to language development*. Mahwah, NJ: Lawrence Erlbaum.

Bonals J (2007) El trabajo en equipo del profesorado [Teachers' teamwork]. In: Bonals J and Sánchez-Cano M (eds) *Manual de asesoramiento psicopedagógico*. Barcelona: Graó, 143–70.

Cerda J and Villarroel L (2008) Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa [Evaluation of the inter-observer agreement in pediatric research: Kappa coefficient]. *Revista Chilena de Pediatría* 79: 54–58.

Cross T (1977) Mothers' speech adjustments: The contribution of selected child listener variables. In: Snow CE and Ferguson CA (eds) *Talking to children*. Cambridge: Cambridge University Press, 151–88.

de Rivera C, Girolametto L, Greenberg J, and Weitzman E (2005) Children's responses to educators' questions in day care play groups. *American Journal of Speech-Language Pathology* 14: 14–26.

del Rio MJ (1993) *Psicopedagogía de la lengua oral: Un enfoque comunicativo* [Psychopedagogy of oral language: A communicative approachment]. Barcelona: CE12.

del Rio MJ and Gràcia M (1996) Una aproximación al análisis: de los intercambios comunicativos y lingüísticos entre niños pequeños y adultos [An approximation of the analysis of the communicative and linguistic exchanges between young children and adults]. *Infancia y Aprendizaje* 75: 3–20.

del Rio MJ and Gràcia M (2003) Psicopedagogia de la llengua oral [Psychopedagogy of oral language]. In: Del Rio MJ and Teberosky A (eds) *Psicopedagogia de la llengua oral i escrita*. Barcelona: Fundació Universitat Oberta de Catalunya, 1–88.

Dockrell JE, Bakopoulou I, Law J, Spencer S, and Lindsay G (2012a) *Developing a communication supporting classroom observation tool*. London: Department for Education.

Dockrell JE, Bakopoulou I, Law J, Spencer S, and Lindsay G (2012b) *Communication supporting classroom observation tool*. London: Better Communication Research Programme, The Communication Trust.

Galván-Bovaira MJ (2008) L'observació sistemàtica de la interacció primerenca: Els intercanvis comunicatius i lingüístics en díades mare–bebé [The systematic observation of early interaction: Communicative and linguistic exchanges in dyads mother–baby]. Unpublished doctoral dissertation, Universitat de Barcelona, Spain.

Galván-Bovaira MJ and del Rio MJ (2009) La evaluación de la interacción comunicativa y lingüística en relación a la adquisición del lenguaje infantil: Revisión de estudios observacionales y escalas de estimación [Evaluation of communicative and linguistic interaction in relation to child language acquisition: Review of observational studies and rating scales]. *Revista de Logopedia, Foniatría y Audiología* 29: 225–36.

Girolametto L, Weitzman E, and Greenberg J (2004) The effects of verbal support strategies on small-group peer interactions. *Language, Speech, and Hearing Services in Schools* 35: 254–68.

Gràcia M (2001) Intervención naturalista en la comunicación y el lenguaje en ámbito familiar: Un estudio de cuatro casos [Naturalistic intervention in communication and language in family environment: A four cases study]. *Infancia y aprendizaje* 24: 307–24.

Gràcia M (2003) *Comunicación y lenguaje en primeras edades: Intervención con familias* [Communication and language in early years: Intervention with families]. Lleida: Milenio.

Gràcia M and Segués MT (2009) *Psicologia de l'educació i de la instrucció* [Educational and instructional psychology]. Barcelona: Fundació Universitat Oberta de Catalunya (FUOC)

Gràcia M (coord.), Galván-Bovaira MJ, Sánchez-Cano M, Vega F, Vilaseca R, and Rivero M (2015) *EVALOE: Escala de valoración de la enseñanza de la lengua oral en contexto escolar. Análisis de las interacciones comunicativas entre docentes y alumnos en el aula.* [Assessment scale of oral language teaching. Analysis of communicative interactions between teachers and students in the classroom]. Barcelona: Graó.

Gràcia M, Galván-Bovaira MJ, Vilaseca R, Rivero M, and Sánchez-Cano M (2012) El cambio conceptual de dos maestras en relación con la enseñanza y el aprendizaje de la lengua oral a través de un asesoramiento: Un estudio de casos [The conceptual change of two teachers in relation to teaching and learning oral language through a counseling: A case study]. *Revista de Logopedia, Foniatría y Audiología* 32: 179–89.

Harms T, Clifford RM, and Cryer D (1998) *Early childhood environment rating scale*. New York: Teachers College Press.

Justice LM and Ezell HK (1999) Vygotskian theory and its application to language assessment: An overview for speech language pathologists. *Contemporary Issues in Communication Science and Disorders* 26: 111–18.

Justice LM, Pentimonti JM, Sofka A, and McGinty AS (2009) *Read it again! Scaffolding coding checklist*. Columbus, OH: The Ohio State University.

Kaye K (1982) *The mental and social life of babies: How parents create persons*. Chicago, IL: University of Chicago Press.

Lever R and Sénéchal M (2011) Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology* 108: 1–24.

Mallart J and Sarramona J (2013) *Competències bàsiques de l'àmbit linguistic. Llengua catalana i llengua castellana* [*Basic competences in linguistic area. Catalan and Spanish language*]. Barcelona: Servei de Comunicació i Publicacions, Generalitat de Catalunya.

Manolson A (1992) *It takes two to talk: A parent's guide to helping children communicate*. Toronto: The Hanen Centre.

Marinac JV, Woodyatt GC, and Ozanne AE (2008) Investigating adult language input and young children's responses in naturalistic environments: An observational framework. *Child Language Teaching and Therapy* 24: 265–84.

Marshall J and Lewis E (2014) 'It's the way you talk to them'. The child's environment: Early years practitioners' perceptions of its influence on speech and language development, its assessment and environment targeted interventions. *Child Language Teaching and Therapy* 30: 337–52.

Martín E (2005) Asesoramiento psicopedagógico [Psychopedagogical counseling]. *Cuadernos de Pedagogía* 348: 62–65.

Mashburn AJ (2008) Quality of social and physical environments in preschools and children's development of academic, language, and literacy skills. *Applied Developmental Science* 12: 113–27.

McHoul A (1978) The organization of turns at formal talk in the classroom. *Language in Society* 7: 183–213.

McHoul A (1990) The organization of repair in classroom talk. *Language in Society* 19: 349–77.

Mercer N (2008) Talk and the development of reasoning and understanding. *Human Development* 51: 90–100.

Mercer N (2010) The analysis of the classroom talk: Methods and methodologies. *British Journal of Educational Psychology* 1: 1–14.

Ministerio de Educación y Ciencia (1990) *Ley Orgánica 1/1990, de Ordenación General del Sistema Educativo (LOGSE)* [*Organic Law 1/1990, of General Organization of the Education System (LOGSE)*]. Madrid: BOE.

Ministerio de Educación y Ciencia (2002) *Ley Orgánica 10/2002, de 23 de diciembre, de la Calidad de la Educación (LOCE)* [*Organic Law 10/2002, 23rd December, of Education Quality (LOCE)*]. Madrid: BOE.

Ministerio de Educación y Ciencia (2006) *Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOE)* [*Organic Law 2/2006, 3rd May, of Education (LOE)*]. Madrid: BOE.

Ministerio de Educación y Ciencia (2013) *Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa (LOMCE)* [*Organic Law 8/2013, 9th December, to improve the quality of education (LOMCE)*]. Madrid: BOE.

Moerk EL (1983) A behavioral analysis of controversial topics in first language acquisition: Reinforcements, corrections, modeling, input frequencies, and the three-term contingency pattern. *Journal of Psycholinguistic Research* 12: 129–55.

Mortimer E and Scott P (2003) *Meaning making in secondary science classrooms*. New York: McGraw-Hill International.

Pianta RC, Karen M, Paro L, and Hamre BK (2008) *CLASS: Classroom Assessment Scoring System Manual*. Baltimore, MD: Paul H. Brookes.

Rolla A, Arias M, Rivadeneira M, Coronado V, and Romero S (2012) *Lenguaje en construcción 1* [*Language in construction 1*]. Madrid: EUNED Editorial.

Rolla A, Coronado V, Rivadeneira M, Arias M, and Romero S (2012) *Lenguaje en construcción 2* [*Language in construction 2*]. Madrid: EUNED Editorial.

Sánchez-Cano M (1993) L'Adquisició de llenguatge en alumnes amb NEE, al marc escolar: Un estudi interactiu [The acquisition of language in children with special education needs (SEN) in the school context: An interactive study]. Doctoral dissertation. Universitat de Barcelona, Spain.

Sánchez-Cano M (2007) El alumnado y el trabajo de la lengua oral [Students and oral language work]. In: Sarrionandia GE, Sallán JG, Pérez MF, Artiaga JF, de Diego Navalón J, Sala NR, and de León, PLM (eds) *Manual de asesoramiento psicopedagógico*. Barcelona: Graó, 771–810.

Sánchez-Cano M (2009) *La conversa en petits grups a l'aula* [*Conversation in small groups in the classroom*]. Barcelona: Graó.

Smith M, Brady J, and Anastasopoulos L (2008) *Early language and literacy classroom observation*. Baltimore, MD: Paul H. Brookes.

Snow CE (1972) Mothers' speech to children learning language. *Child Development* 43: 549–65.

Snow CE (2010) Academic language and the challenge of reading for learning about science. *Science* 328: 450–452.

Snow CE, Lawrence J, and White C (2009) Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness* 2: 325–44.

Tomasello M (2003) *Constructing a language: A usage based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello M (2014) *A natural history of human thinking*. Cambridge, MA: Harvard University Press.

Waite WM, Jackson MH, and Diwan A (2003) The conversational classroom. *Proceedings of the 34th SIGCSE technical symposium on computer science education*: 127–31.

Wells G (1999) *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.

Wood D, Bruner JS, and Ross G (1976) The role of tutoring in problem solving. *Journal of child Psychology and Psychiatry* 17: 89–100.

Ygual A and Cervera JF (1999) La intervención logopédica en los trastornos de la adquisición del lenguaje [Speech and language therapy intervention in language acquisition disorders]. *Revista de Neurología* 28: 109–18.

Appendix 1

Examples of items from the observation scale.

Tool for assessing spoken language teaching and learning in classroom (EVALOE)
Part one
Observation scale

Date: Observer:
Teacher observed:
Course: Group:
School:
Number of students:
Subject/activity:
General aim of the class, under the subject, unit or project of which it forms part:

Instructions

The observation scale is divided into three subscales: (1) Context and communication management; (2) Instructional Design; and (3) Communicative functions and strategies. Each subscale contains a different number of questions, which the observer rates from 1–3. The observer will record the total score for each subscale in the table below.

Table. Summary table of the results.

Subscales	Score
Context and communication management	__ / 24
Instructional design	__ / 21
Communicative functions and strategies	__ / 45
Total	__ / 90

Comments

Examples of items from Subscale 1: Context and communication management

1. During the work session the *arrangement of furniture and/or the students** is suitable considering the characteristics of the activity performed
    1. The arrangement of furniture and/or students is unsuitable
    2. The arrangement of furniture and/or students is not entirely suitable
    3. The arrangement of furniture and/or students is suitable

4. The teacher facilitates students to *initiate communicative interactions**
    1. The teacher does not encourage students to initiate communicative interactions
    2. The teacher encourages students to initiate communicative interactions sometimes
    3. The teacher encourages students to initiate communicative interactions systematically when appropriate

8. During class conversation, discussion and/or debate activities, teacher and students sit in a circle
    1. During these activities teacher and students do not sit in a circle
    2. During these activities teacher and students sometimes sit in a circle
    3. During these activities teacher and students almost always sit in a circle

Examples of items from Subscale 2: Instructional design

1. The teacher explains the oral language goals of the activity
    1. The teacher doesn't explain oral language goals
    2. The teacher explains oral language goals superficially or partially
    3. The teacher explains oral language goals clearly and in detail

6. The teacher asks for *self-assessments of the students' communicative behaviour** during oral language activities in the classroom
    1. The teacher doesn't ask for self-assessments from students
    2. The teacher's requests for self-assessments from students are inadequately developed and justified
    3. The teacher's requests for self-assessments from students are thoroughly developed and justified

7. The teacher assesses students' oral language competence*
    1. The teacher doesn't evaluate students' oral language competence
    2. The teacher assesses students' oral language competence, but in a non-explicit way
    3. The teacher assesses students' oral language competence explicitly

Examples from Subscale 3: Communicative functions and strategies

1. The teacher teaches giving information*
1. The teacher doesn't teach it
2. The teacher teaches it by giving some general indications
3. The teacher teaches it by giving clear, specific explanations
2. Students give information
1. Few students give information
2. Some students give information
3. Most students give information
9. The teacher teaches *'summarizing' and/or 'drawing conclusions'** after class discussions
1. The teacher doesn't teach it
2. The teacher teaches it by giving some general indications
3. The teacher teaches it by giving clear, specific explanations
10. Students synthesize and/or draw conclusions
1. A few students do it
2. Some students do it
3. Most students do it
11. The teacher expands on students' utterances
1. The teacher doesn't expand on students' utterances
2. The teacher sometimes expands on students' utterances
3. The teacher expands on students' utterances systematically when necessary
12. Students' utterances improve after the teacher's expansion
1. In only a few cases
2. In some cases
3. In most cases

Appendix 2

Examples of questions used in the revised interview guide for teachers.

Tool for assessing spoken language teaching and learning in classroom (EVALOE)
Part two
Protocol for the in depth interview

The interview provides an opportunity to discuss observed teachers' scores on the three subscales in the observation scale. This second part of the tool allows joint reflection between the interviewer and teacher regarding the way the teacher manages spoken language during daily activities with students in classroom. The joint analysis of the answers will contribute to the design of possible modifications in the teaching and learning of spoken language in all subjects on the curriculum. It can also help to refine aspects that might otherwise go unnoticed during the observation.

Context and communication management

A first important aspect when assessing the teaching of oral language in the classroom is the context: how students are distributed in the classroom and how you manage the communication (i.e. the presence of conversation rules).

1. Grouping of students

1.1. Grouping types

In your classes, how do you usually group the students?

(If the answer does not come to mind, give the interviewee an example)

2. Communication rules

2.1. Explanation of conversation rules

Do you explain conversation rules in class?

(If the answer is NO, skip the following question and go to section 3)

2.2. Types of rules

What conversation rules do you explain in class?

(If the answer does not come to mind, give the interviewee an example)

2.3. The time when the rules are explained

When do you explain the rules of conversation?

(If the answer does not come to mind, give the interviewee an example)