View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Central Archive at the University of Reading

Early language learning: the impact of teaching and teacher factors

Article

Accepted Version

Graham, S., Courtney, L., Marinis, T. and Tonkyn, A. (2017) Early language learning: the impact of teaching and teacher factors. Language Learning, 67 (4). pp. 922-958. ISSN 0023-8333 doi: https://doi.org/10.1111/lang.12251 Available at http://centaur.reading.ac.uk/70107/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/lang.12251

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

Early language learning: The impact of teaching and teacher factors

Suzanne Graham*

Louise Courtney

Theodoros Marinis

Alan Tonkyn

*corresponding author

s.j.graham@reading.ac.uk Tel: +44(0) 118 378 2684

Institute of Education

University of Reading

Reading

RG1 5EX

UK

Abstract

This study examined the progress in lexical and grammatical knowledge among 252 learners of French in England across the last two years of primary education and into the first year of secondary school in relation to teaching and teacher factors. It compared linguistic outcomes from two different teaching approaches, one placing emphasis on oracy, the other combining literacy with attention to oracy development. It also explored the relationship between linguistic outcomes and other teaching/teacher factors: teaching time, teacher level of French proficiency, and teacher level of training in language instruction.

Learners completed a sentence repetition and a photo description task, making relatively small but statistically significant progress in both grammatical and lexical knowledge between test points. While teaching approach had little impact on such progress, other teaching and teacher factors did, particularly the French proficiency level of the primary school teacher and the amount of teaching time devoted to French.

Introduction

This study investigates the nature of the progress made by young learners of French as a foreign language in England across the last two years of primary education and the first year of secondary school. It also explores the extent to which learning outcomes are related to teaching and teacher factors within the primary setting: teacher French language proficiency, teacher level of training in language teaching, teaching time and teacher 'approach', namely the focus of instruction. Within the study we consider two 'approaches', one we term 'oracy' (where the teaching emphasis is predominantly on speaking and listening development), the other, 'literacy' (combining reading and writing with attention to oracy).

An investigation into the progress made by young learners in classroom settings is timely and important because recent years have seen a growing number of countries lowering the age at which instructed foreign language learning begins (Murphy, 2014) in the belief that an earlier start will lead to better learning outcomes. Concerns have been expressed, however, that such policy initiatives have been based on the extrapolation of findings from language learning in naturalistic settings to language learning in classroom contexts (Murphy, 2014). Such extrapolations seem to assume that both types of learning are identical and that an early start in classroom instruction will automatically lead to rapid, effortless learning. On the contrary, any age advantage reported for naturalistic L2 learning (discussed in, for example Muñoz, 2008) has not been found in rigorous, longitudinal studies in instructed contexts, with a later start (e.g. at age 11) in fact often resulting in faster, more 'efficient' learning, as was found, for example, by Muñoz (2006), for learners of English in Spain, and Myles & Mitchell (2011), for learners of French in England.

Furthermore, much variability in learning outcomes for young learners has been reported for instructed settings (see, for example, findings from Europe reported in Enever, 2011a), suggesting that factors other than age contribute to language learning for such children. Additionally, the question of age and its relevance in instructed contexts needs to be considered with reference to the ways in which young children are believed to learn a language. Dekeyser (2003) argues that they draw on Universal Grammar and implicit mechanisms, leading to better ultimate attainment as "many elements of a language are hard to learn explicitly" (p. 335). Older children and adults, by contrast, learn explicitly, drawing on "native language knowledge" (p. 334), with greater cognitive maturity enabling them to progress more quickly. Furthermore, in order to be able to draw on implicit mechanisms, young learners require vast amounts of time and input. While such amounts are available in naturalistic contexts, this is far from being the case in instructed foreign language classrooms (DeKeyser, 2003; Muñoz, 2006, 2008, 2014).

If it is accepted that young learners learn implicitly, and that this is a slow process requiring plentiful and high quality input, the amount of exposure that they receive in instructed settings is of prime importance, more important than the age at which instruction begins, and "never ceases to be a determinant factor" in learning outcomes (Muñoz, 2014: 466). This is also likely to be the case for the quality of that exposure, which may vary according to the teachers' level of foreign language proficiency, their pedagogical skills, and the teaching methods they employ. Nevertheless, while there has been plentiful research on the effects of age on language learning, much less is known about the relationship between amount and quality of language exposure on the one hand and learners' language development on the other in classroom settings. The present study seeks to address that gap and to explore how variation in teaching and teacher factors relates to learners' grammatical and lexical development.

The context of the study is the learning of French in England, a country which has followed the global trend towards an earlier start for language learning albeit more recently. A foreign language became a compulsory element of the National Curriculum at primary school level in September 2014. Learners are required to make "substantial progress" in one language during the last four years of primary education (DfE, 2013:1), that is, from Year 3 (ages 7-8) to Year 6 (ages 10-11). At the same time, annual surveys indicate a great deal of variability across schools in the amount of lesson time allocated to language learning, and in teachers' level of language proficiency and training (Tinsley & Board, 2016). Variability in teaching and teacher factors is also found both across and within other countries (Enever, 2011b, commenting on Europe), but the relative absence of the target language in the environment within England and other Anglophone contexts make it particularly important to understand how such variability affects learning outcomes.

Variability in primary language provision is also likely to have implications for learners' move to secondary

or high school, as it means that learners in England enter secondary school with widely different levels of foreign language proficiency (Tinsley & Board, 2016). This range poses a great challenge to language teachers, which they may seek to address by simply 're-teaching' what was meant to have been covered in earlier years (Office for Standards in Education, 2011). Such issues and practices occur not only in England but also elsewhere, e.g. in the USA and Australia (Lo Bianco, 2009; Pufahl & Rhodes, 2011), and indeed also occur in other curriculum areas (such as mathematics and L1 reading - Galton, Gray, & Ruddock, 1999) albeit to a lesser extent because the core nature of those areas means that primary school provision is less variable. 'Re-teaching' may account for findings across a range of contexts suggesting that learners' progress slows across the primary to secondary school transition (for example, Hill, Davies, Oldfield, & Watson, 1998, in Australia; Low, Brown, Johnstone, & Pirrie,1995 (in Blondin et al. 1998), in Scotland). By contrast, and more positively, a recent study by Courtney (2014) of 26 learners of French in England across the primary-secondary transition did find evidence of progress in breadth of vocabulary, gender assignment and verb morphology, although progress was slow and there was much individual variation. While Hill et al. (1998) and Courtney (2014) were relatively small longitudinal studies, Low et al. (1995) had a much larger sample but used cross-sectional data, making it difficult to compare findings across the studies and possibly contributing to the apparent contradictions in their findings.

There is also evidence of some plateauing at the end of primary school, as Cable et al. (2010) found for target language phonology and listening in a study of learners of French in England across the last four years of primary education (Years 3 to 6). Slower progress in Year 6 may relate back to teacher factors in primary school, particularly to teachers' subject knowledge which may be inadequate to deal with more than the beginner level . In turn, a lack of progress at the end of primary school may then form a shaky foundation for learners as they move into secondary school. Making progress across the primary to secondary transition is likely to have implications for subsequent motivation and success in language learning, both areas where

England has persistent problems (Tinsley & Board, 2016), along with other countries, for example, Australia (Lo Bianco, 2009).

The role of teacher and teaching factors

Amount of instruction

While outcomes in early language learning have often been explored from an individual differences perspective, there is also evidence of the importance of teaching and teacher-related factors. First, **teaching time** or amount of exposure learners receive, which has been an area of importance to researchers since at least the 1970s (Lightbown, 2014) and which has been investigated from a number of different angles in relation to early language learning. For example, Myles and Mitchell (2011), in a study of learners aged 5, 7 and 11 in England, established that vocabulary learning was influenced by amount of raw teacher input, i.e. how often learners met the linguistic items on which the project tests were based, for all ages. Frequency of lessons is reported to have influenced linguistic outcomes (reading, writing and listening) in a large study (N= 20,804) of Year 6 learners of English in Hungary by Nikolov (2009). Although Nikolov (2009) claims that the number of weekly lessons learners experienced was positively correlated with scores on the language tasks used, the absence of detailed results in the reporting of this study make it difficult to interpret. Furthermore, it is unclear whether the relationship between lesson frequency and outcomes was stronger for the Year 6 learners than for the Year 10 learners also involved in the study.

Looking at total amount of lesson time and length of lessons, Genelot (1997), in a study of over 1000 young French learners of English, reports a positive relationship between total amount of lesson time and learning outcomes, as well as a more positive effect of shorter, more frequent lessons compared with longer, less frequent ones. By contrast, there is also evidence that intensive periods of instruction may be more beneficial than a 'dripfeed', little and often approach. Intensive instruction may be needed at the beginning

of language study to allow learners to reach the basic level of communicative competence on which further development can be built. This is the conclusion reached by Netten and Germain (2008) in Canada, comparing 'core' French (daily lessons of 30-50 minutes) with an 'intensive' approach (a five month period of intensive exposure to French across 65%-70% of the school day, followed by a return to a more 'normal' curriculum with typically two 80 minute sessions a week). Intensive French aims to develop spontaneous oral and written communication, with literacy skills developed alongside oracy, for which, the authors argue, the longer lesson time is required. Evaluating outcomes from both core and intensive French, Netten and Germain (2008) report that learners experiencing the intensive approach developed spontaneous communication skills but core French learners did not. While it must be acknowledged that the Intensive French programme differs from the type of instruction offered in the context of the present study, there are arguably still implications that can be drawn from it with respect to the importance of time for learning.

Teaching approach: literacy and oracy

Linking oracy and literacy skills, according to Netten and Germain (2008), is central to the success of Intensive French. The introduction of reading and writing into early language learning is not universally supported, however. Indeed, in England an 'oracy first' approach to second language instruction is widely used (Cable et al,. 2010), possibly based on the premise that L2 instructed learning should follow what happens in first language development – i.e. language is firstly acquired orally and only later are grapheme-phoneme correspondences taught. Oral and aural instruction involves more implicit learning and hence might be deemed to be more appropriate for younger learners.

While literacy-based activities thus may be considered to be more suitable for older learners, it is possible that they can also benefit younger learners, however. Evidence from immersion studies suggests that oral input alone may lead to fluency but to inaccurate production, as Harley and Swain (1984) found for past

tense formation among Grade 1 English speaking learners of French in Canada. They argue that learners are less likely to notice or attend to grammatical forms from oral input and where grammatical features have lower levels of phonological salience. The authors also comment that "written input of some kind would be helpful in drawing attention to phonologically non-salient segments" (p. 295). This may be particularly the case for a language such as French, in which markers of gender and adjectival agreement are not very salient in oral input.

Vocabulary development potentially also benefits from written input rather than oral input alone, according to a range of evidence reviewed by Hu (2008) who argues that a focus on orthographical forms helps to "'fossilize' the L2 speech signals" (p.823) in the input, that is, makes them more accessible and therefore more likely to be processed effectively. Beginning learners of an L2, and especially ones with weaker phonological awareness and hence poor speech perception, may find it hard to learn from oral input alone because they have difficulty in "constructing accurate, detailed phonological representations in the process of abstracting a stable specification of the sound structure of the new word from the input" (Hu, 2008: 825). Thus presenting the written form alongside the oral form could result in a clearer and more durable memory representation, although a counterargument might be made that presenting language in more than one modality presents learners with a heavier cognitive load[1]. In Hu's own study among Grade 3 Chinese-speaking children in Taiwan, learners were taught novel English words, both with and without written forms. Half the sample (37) had higher levels of L1 phonological awareness (PA), half lower levels. Learning was found to be better in the written condition, but learners with higher PA benefited the most from the written input. Hu hypothesises that the benefit for low PA learners could be enhanced with more explicit instruction in L2 phonics, highlighting the importance of the quality of provision over and above the type of provision.

Written input can perhaps also provide more repeated encounters with lexis and grammatical structures, also potentially covering a wider range than is possible from the teacher's oral input alone (Lightbown, 2014; Porter, 2014), compensating perhaps for any shortcomings in the range and accuracy of the teacher's language proficiency. This input may then have benefits not only for learners' literacy development in the L2, but also for their oral development. Indeed, two studies indicate that oracy and literacy can develop side by side. The first, by Drew (2009), investigated the impact of an Early Years Literacy Programme on the development of learners of English (aged 8) in Norway. The programme included extensive reading of illustrated graded books that featured systematically high frequency vocabulary. An experimental group of 57 pupils received lessons combining periods of extensive and differentiated reading with oral, communication-based activities. Learners also read at least one book a week at home with the help of parents. Their performance on pre- and post-tests in listening, speaking, reading and writing was compared with that of 58 learners experiencing lessons characterised largely by whole-class teaching, choral repetition, reading dialogues aloud in groups and simple writing. The experimental group made more progress than the control group on all skills, but especially on the listening and oral tests. More recently, and in the same context as the present study, Porter (2014) conducted an action research project with primary-aged learners of French in England and found that literacy work including both reading and writing allowed learners across the attainment range to make progress orally as well as in literacy. Although both studies indicate that combining literacy with oracy can be beneficial across skills, in the first it is possible that the experimental group also received greater amounts of input than the control group from the out of class work, while the absence of a control group in the second study makes it more difficult to assess the impact of literacy work.

Teacher expertise

Teacher training

Quality of input is likely to be of as much importance as quantity and type of input, particular with regard to

the language proficiency and pedagogical skills of teachers. Neither has received extensive research attention and studies have also tended to be small and cross-sectional. For example, the impact of teacher training was considered in a study by Mihaljevíc Djigunović (2009), who looked at the relationship between attitudes and learning conditions for children in Croatia learning English. Learning conditions included teaching time, teacher training, and class size, the latter being a factor which, in a review of studies of early language learning in Europe, Blondin et al. (1998) found to have negligible impact on outcomes. Mihaljevíc Djigunović (2009) reports that 'less favourable' conditions impacted negatively on learners' attitudes towards and sense of competence in learning English, although few details are provided regarding the exact nature of the conditions in which learners were taught (for example, it is stated that in the four 'less favourable' classes "the teacher of English **might** not have had much training" (p. 78, emphasis added). Simlarly, Szpotowicz (2009) investigated vocabulary acquisition over three to five days among 67 learners of English in Poland. Although the main focus was on the learnability of different types of vocabulary, a teacher effect was also identified, with learners taught by the least qualified of four teachers achieving the lowest scores on tests of recall and recognition. This was in spite of all four teachers using the same vocabulary teaching methods prescribed by the researcher.

Teacher language proficiency

It is unclear however whether teachers in Szpotowicz (2009) also varied in English language proficiency. Teacher language competence is likely to have an impact on outcomes, with Muñoz (2006: 34) stating that not only does exposure need to be intense but also to "provide an adequate model". Teachers' linguistic skills are likely to be especially relevant to the acquisition of grammatical features, particularly within usage-based theories of language acquisition, which emphasise the importance of the linguistic environment (see Murphy, 2014, for an overview) for grammatical development. Frequency and consistency (in relation to form-function mappings) of grammatical features in the input are likely to influence how well such features are acquired (Murphy, 2014). Arguably, teachers with lower levels of proficiency in the target

language are less able to provide such frequency and consistency. The optimal level of teacher language proficiency is however more difficult to determine. Native speakers, particularly if they lack training in language teaching (Walkinshaw & Duong, 2012), may be less able to simplify their input to make it 'comprehensible' to learners (Krashen, 1982).

Overall, studies that have investigated the impact of teachers' level of language proficiency are limited in number, perhaps, as Unsworth, Persson, Prins and De Bot (2014) argue, it is a rather sensitive topic. Some relate to bilingual rather than to foreign language settings; certain tentative conclusions from them can however be drawn. For example, Bowers and Vasilyeva (2011) report that for early learners of English in a bilingual setting, the total number of words spoken by their native speaker teachers during a period of 90 minutes observation (audio recorded) was positively related to learner's receptive vocabulary knowledge, but that the average number of words per teacher utterance was negatively correlated with learning outcomes. The authors interpret this as an indication that beginning language learners need exposure to relatively uncomplicated speech, and can only take advantage of native speaker input once they are beyond the very beginner stage. It also perhaps suggests that teachers need to know how to make their input comprehensible and suitable for the learners they are teaching. Native speaker input may however sow the seeds for later development, according to Aukrust (2007), who found that in another bilingual setting amount, diversity, and discourse complexity of teacher input (video recorded and analysed using MacWhinney's 1995 Computerized Child Language Analysis) only predicted preschool learners' language outcomes two years later.

Amount and quality of input may, furthermore, interact with each other. This issue has been the subject of investigation in two studies conducted in foreign language instructed settings in the Netherlands. The most recent, by Unsworth et al. (2015), explored the impact of weekly lesson time for language learning and teacher language proficiency on the development of vocabulary and grammar skills of Dutch children

learning English in their first and second year of study. Learners constituted two groups: 168 learners in early English schools (receiving up to 220 minutes a week of English), and 26 age-matched children having more 'regular' English exposure i.e. approximately 45 minutes a week. The authors found that both teaching time and teacher oral language proficiency were very important factors in scores for grammar and vocabulary development (measured through the Test for Reception of Grammar Version 2, Bishop, 2003) and the Peabody Picture Vocabulary Test, Dunn & Dunn, 2007, respectively). This was especially the case where children received under 60 minutes a week of English instruction, and where the teachers' proficiency level was below level B on the Common European Framework of Reference (CEFR). Furthermore, a regression analysis showed teachers' language proficiency was the best predictor of outcomes, for both vocabulary and grammar, a finding which the authors interpret as evidence of the importance of lexical diversity and grammatical complexity in input gained from the teacher. Finally, teachers' language proficiency predicted grammar scores only at post-test II, suggesting that its effects take time to emerge, a conclusion that underlines the importance of longitudinal investigations of the impact of teacher variables.

Unsworth et al. (2015) note that their study did not take account of other issues that may have impacted on learning outcomes, such as teaching qualifications and the kind of instruction given, underscoring the need to take account of the full range of teaching/teacher factors that may influence learning. An earlier study in the Netherlands by Edelenbos and Suhre (1994) is one of the few to explore the teacher variables of teaching experience and spoken fluency, language teaching qualifications, teaching time, alongside teaching approach (either broadly communicative or grammar-focused courses, as ascertained through content analysis). Data on reading, writing, listening, vocabulary, grammar and spoken fluency were collected from 2116 pupils in 112 schools. Amount of lesson time was significantly and positively correlated with all scores except learners' spoken fluency. Possessing a teaching qualification was the most important teacher factor for predicting outcomes (in vocabulary, grammar, spoken fluency and listening), with teacher spoken fluency positively related only to learners' grammar scores. Once teacher and SES variables were controlled

for, however, the only difference attributable to the course followed was on the grammar test, where, perhaps unsurprisingly, learners following a course with a heavier emphasis on explicit grammar had the higher scores. Additionally, as Driscoll et al. (2004) point out, the teachers in Edelenbos and Suhre's (1994) study who used the more grammar-focused courses were also the most qualified and experienced teachers and also had more English language teaching time available to them. A further limitation of the study, however, is that it is cross-sectional and thus cannot give full insights into how teaching and teacher factors impact on learners' progress over time.

Early language learning in England

Variability in teaching and teacher factors such as those noted above is very pertinent to the context of England, because compared with what children in other European countries experience (Enever, 2011b), learners in England receive limited amounts of language instruction. In an annual survey of language teaching provision in around 600 primary schools (Tinsley & Board, 2016), 48% reported allocating 30- 45 minutes a week to language teaching, 32% between 45 minutes and an hour, and just 15% between one and two hours. Additionally, around a third of schools reported that the member of staff teaching a foreign language had only GCSE level (i.e. level A, CEFR) language competence or below. Teachers with native or degree level competence were reported by 16% and 29% of schools respectively. Nearly 60% of responding schools reported that lack of staff expertise in language pedagogy was a challenge for them. These less than ideal conditions for language learning are not unique to England, however; a survey of 142 countries by Copland, Garton, and Burns (2014) suggests that concerns about teaching and teacher factors are widespread. Nevertheless, the fact that in England the classroom is the only real contact that most learners have with the foreign language they are expected to learn (which in most schools is French – Tinsley & Board, 2016) means that teaching time, teacher language proficiency and teacher language pedagogy training are likely to be of particular relevance there.

Regarding teaching approach or focus of teaching, curriculum documents in England relating to primary languages make strong statements about the importance of including literacy in early language instruction. Thus the Key Stage 2 Framework for Languages (DfES, 2005), a non-statutory but still widely used curriculum document (Cable et al., 2010; Porter, 2014), includes both an oracy and literacy 'strand' and presents them as being of equal importance. Nevertheless, Cable et al. (2010) found that oracy received more attention than literacy in many classrooms in England.

In summary, previous research indicates that amount of teaching time, teacher language proficiency/training and teaching approach (i.e. focusing predominantly on oracy or attending to literacy as well) are likely to have an impact on the language development of young learners, but in potentially complex ways. It is not clear, for example, which factors are the most important, at which stage of learning, and whether they interact with one another. Overall, given that large proportions of young learners in England are being taught in conditions that are at odds with what research suggests are optimum for language acquisition, it is important to gain greater understanding of how learning outcomes vary with teaching and teacher factors, and with a longitudinal perspective. Furthermore, the somewhat contradictory findings regarding the nature of learners' progress across the primary to secondary transition call for further exploration of the nature of that progress.

The aims of the project

This project aimed to address the issues outlined above by using a longitudinal design to investigate the following research questions:

1. To what extent do children make progress in their knowledge of French (vocabulary and

grammar) across Years 5, 6 and 7?

2. To what extent is children's knowledge of French across Years 5, 6 and 7 related to a) teaching and teacher factors in primary school (teaching time, teacher French proficiency, teacher level of training) and b) teaching approach (oracy- vs literacy)?

Our consideration of grammatical development included three grammatical features of French which are part of the primary languages curriculum in England (DfES, 2005; DfE, 2013): article-noun agreement, adjective-noun agreement, and simple present tense verbs. Curriculum expectations are not that teachers should engage in formal, explicit grammar teaching but rather should develop learners' 'Knowledge about Language' (DfES, 2005), by drawing their attention to features such as gender class as part of more communicative, game-based activities, practices observed by Cable et al. (2010) to be common in primary classrooms in England. It might be claimed that grammatical gender is difficult for English-speaking learners of French to acquire[2]. There is however evidence that accuracy in gender assignment increases at a statistically significant level between school Years 5 and 7, as Courtney (2014) found, reaching a 69% accuracy level. We therefore considered it appropriate to include it within our broader assessment of grammatical development in relation to teaching and teacher factors. It should be noted furthermore that we do not seek in this paper to report on each grammatical feature separately.

Methods

Research design

The study was a "natural experiment" (Shadish, Cook, & Campbell, 2002:12), in which the outcomes of naturally occurring phenomena are compared. It tracked learners' linguistic development in French from the penultimate year of primary school, Year 5 (summer term, Test Point 1), to Year 6 (early summer term, Test

Point 2), and into Year 7, the first year of secondary school (half way through the autumn term, Test Point 3). At all Test Points, learners completed a Sentence Repetition Task (SR) and a Photo Description Task (PD), both described below. Stages in the data collection are shown in Table 1, along with the sample size.

To address Research Question 2, we looked at the two following teaching approaches:

1. A predominantly 'oral' approach which focuses principally on developing speaking and listening skills with very little emphasis on literacy ('Oracy');
2. A more literacy-based approach, where reading and writing activities are integrated into instruction, alongside oracy skills ('Literacy').

As the above makes clear, we are contrasting two approaches that might be termed 'oracy' and 'oracy + literacy', but for brevity and clarity we refer to the approaches as 'oracy' and 'literacy' respectively. Furthermore, we use the term 'approach' to signify the relative emphasis placed by teachers on the development of oracy and literacy skills, and the relative amount of time spent on activities involving oral/aural language or written language.

Learner participants

Data were initially collected from 254 learners across nine primary schools (for details of selection of schools, see **Literacy and oracy divisions of schools**). The data from two learners were subsequently excluded from the analysis as their English literacy scores were greater than three standard deviations below the mean, giving a Year 5 sample of 252. Numbers of learners involved in the study varied across time points, as indicated in Table 1. Attrition at the start of Year 7 arose for two main reasons: some learners transferred to a secondary school where the language taught was Spanish or German rather than French;

several secondary schools required us to test learners outside of lesson time, and we hence had to rely more on learners' willingness to attend after-school testing sessions. Not all learners attended such sessions. Those withdrawing in Year 7 had test scores at Test Points 1 and 2 that did not differ significantly from those of learners who remained in the study at Test Point 3 (as ascertained by independent-samples t-tests, with p values ranging from .16 to .99), giving us confidence that the remaining students were not untypical of the sample as a whole.

<Table 1 about here>

All learners spoke English at home (as ascertained through a parent questionnaire). One learner also spoke a language in addition to English at home (Nepali), but had levels of English literacy within three standard deviations of the sample and national average. Nepali does not mark gender or adjective agreement as French does and so the learner was not felt to be at an advantage on the measures of French used in the study. This was further supported by the fact that the student's scores on all of the measures of French were within one standard deviation of the sample mean. For these reasons it was decided to retain the learner's data for analysis. For just 18% of children was there any reported contact with French outside of school (in all cases very slight). Information on learners' English literacy levels was obtained from their primary school for reading comprehension and writing (including imaginative and factual writing, spelling and grammar). In England at the time of the study, learners' English literacy attainment was expressed through nationally applied National Curriculum 'levels' and sub-levels or descriptors (e.g. 1a, 1b, 1c), based on teacher assessment and tests, with learners in our study scoring between level 1a (lowest) and 5a (highest). For the study, sub-levels were converted to a point scale from 1 (1a) to 13 (5a). The scores thus obtained for reading and writing were added together and the mean calculated, giving each learner a combined English literacy score out of a possible 13.

Learners were all in their third year of learning French at the start of the project and in the penultimate year of primary education (Year 5, age between 9 years 10 months and 10 years 10 months). They were all in classes of under 30 learners, with most classes comprising between 25 and 30 learners.

Literacy and oracy divisions of schools

Prior to the start of the project, we distributed a questionnaire to 35 schools in the South of England, which asked teachers of Primary French about the kind of activities they personally used with Year 5 learners and the frequency with which they used them. From the replies received, nine included follow-up contact details from teachers who were interested in being part of the main study. Details of these nine schools and the one teacher per school involved in the study are given in the **Supplementary Materials**. The Key Stage 2 Framework for Languages mentioned earlier, a non-statutory document that sets out objectives and related teaching activities in the areas of oracy and literacy across Years 3 to 6 (DfES, 2005), was used to create the questionnaire items (15 literacy, 14 oracy). For example:

> *Pupils listen to a story in the language but don't see the words at the same time* (Oracy)
>
> *I read stories to learners and they follow the words on the board or from a Big Book* (Literacy)

Additionally, one question asked whether teachers drew any attention to grammatical features such as gender of words (all teachers reported doing so). In order to gain a broad, overall picture of the teaching approach adopted, we asked teachers to indicate how often they used each activity with learners in Year 5 French classes and space was provided at the end for them to list any other activities not mentioned. Finally, teachers were asked to indicate what proportion of lesson time was devoted to listening, speaking, reading and writing respectively.

The percentage of literacy and oracy activities that teachers reported using in all or most Year 5 lessons was calculated. A literacy score (1=low, 5 = high) was then allocated based on this percentage as follows:

1 = 0-20% of questionnaire literacy activities used in all or most lessons

2 = 21-30% of questionnaire literacy activities used in all or most lessons

3 = 31-40% of questionnaire literacy activities used in all or most lessons

4 = 41-50% of questionnaire literacy activities used in all or most lessons

5 = 50% + of questionnaire literacy activities used in all or most lessons

After gaining this broad view of the participating teachers' instructional focus in Year 5, we wanted to be sure that the approach learners experienced did not change when they moved into Year 6. Therefore, once the study was underway, participating teachers (one per school) were observed teaching the project learners in Year 6, once per school, using a lesson observation schedule, which also allowed for more in-depth examination of classroom activities than had been gained through the questionnaire. This confirmed that for each school the teaching approach remained constant across the two years, and gave us a rounded picture of the focus of instruction each group of children received across both Year 5and Year 6. We also requested Schemes of Work from schools in order to corroborate information gained from the questionnaire and observation, and to place the observation within a broader context. The types of literacy activities observed were recorded and analysed with reference to the Key Stage 2 Framework objectives for Literacy (DfES, 2005). We used the Framework to create a literacy scale from 1 to 5, as follows

1 = Year 3 reading activities observed but no writing activities; for example, recognition of familiar words in written form, making sound-spelling links, reading aloud simple words

2 = Year 4 reading activities observed but no writing activities; for example, reading and understanding familiar phrases, following a short written text that is read aloud, reading aloud phrases

3 = Year 3 and 4 reading activities + Year 3 writing activities observed: as (1) and (2), plus writing single words

4 = Year 5 and 6 reading activities + Year 3 and 4 writing activities observed; for example, reading short texts individually, writing words and phrases

5 = Year 5 and 6 reading + Year 5 and 6 writing activities observed; for example, reading short texts, including authentic texts, writing sentences on a range of topics using a model.

Thus a school receiving a score of 1 would only be including the type of reading activities aimed at Year 3 in the teaching of Year 6 learners. The final observation score was then added to the questionnaire score, to give each school a total literacy score out of ten. Schools with a borderline oracy/literacy score (i.e. scoring at the mean and median points, 6 and 6.5) were excluded for the analysis of the impact of teaching approaches on learning outcomes (but included for other analyses). This gave a sub-sample consisting of three schools in the oracy group, and three in the literacy group (Schools 1, 2, 3, 5, 7 and 9), with numbers of participants as shown in Table 2.

<Table 2 about here>

Schools were matched on key indicators (percentage of pupils claiming Free School Meals – FSM; percentage of pupils with Special Educational Needs; percentage of pupils with English as an Additional Language - EAL). All schools were in the Department for Education's low percentage band for EAL and FSM (the latter a measure of deprivation widely used in the UK). Information on teaching and teacher

factors (weekly amount of teaching time, teacher level of French, teacher level of training in teaching languages) was gained through a short questionnaire. Most teachers[3] were relatively highly qualified linguistically and pedagogically, compared with the national averages (Tinsley & Board, 2016) reported above. As shown in the **Supplementary Materials**, approximately two-thirds of teachers had native or degree level competence in French, and had had post-graduate training in language teaching; teaching time for French ranged from 15 minutes a week for one school to 60 minutes in two schools, with the remainder at around the national average of 30-45 minutes a week (Tinsley & Board, 2016). Once in Year 7 at secondary school, all learners received very similar amounts of French teaching per week, namely approximately two hours.

Tasks and procedures

We designed and piloted two tests to assess knowledge of grammatical features and vocabulary: a Sentence Repetition task (SR) and a Photo Description task (PD), with the same tests used at each Time Point.

Sentence Repetition Task (SR)

Sentence Repetition tasks (also known as Elicited Imitation, EI) are frequently used to assess the language development of young learners (Marinis & Armon-Lotem, 2015) because they are believed to tap into their implicit knowledge (Erlam, 2006) and have been found to be good indicators of grammatical development (e.g. Devescovi & Caselli, 2007). It is claimed (e.g. Wu & Ortega, 2013) that learners can only repeat structures that they have acquired and that EI/SR tasks thus give a 'window' into the underlying linguistic competence of the learner. As Jensen and Vinther (2003) outline, if the length of a cue sentence is such that learners could not be imitating it through mere parroting, then "we assume they imitate it after passing it 'through a filter of existing grammatical knowledge' (Eisenstein et al., 1982, p. 391)" (p. 389). As such, EI

tasks are “reconstructive” (Jessop, Suzuki, & Tomita, 2007:215):

A number of recent reviews conclude that EI tasks discriminate reliably between learners of different levels in different knowledge areas. A recent meta-analysis of 76 EI studies by Yan, Maeda, Lv, and Ginther (2016), looking at investigations across syntax, morphosyntax, lexis and phonology, found strong correlations between EI scores and other measures, leading them to conclude that “EI is a highly sensitive measure” and that there is strong evidence regarding its “construct-related validity” (p.516). Other advantages of EI/SR tasks are that they permit the examination and tracking of a range of very specific linguistic items over time. As such an SR task was felt to be well suited to the aims of our study.

In SR tasks, learners hear sentences containing the target items which they then have to repeat verbatim. Sentences need to be long and complex enough for learners to have to analyse what they are hearing rather than just ‘parroting’ what they hear, but not so long as to place a heavy burden on memory. The issues of length and complexity have been widely discussed in the literature (see Devescovi & Caselli, 2007) and are relative to the age, memory capacity, and proficiency level of the participants, with no consensus as to what is the ‘ideal’ (Yan et al., 2016). Choosing the right length and complexity for learners is often an empirical question and requires piloting to ensure that learners do not show a ceiling effect that could indicate parroting, or a floor effect suggesting that the length and complexity exceeds the participants’ processing capacity.

The SR task used in the present study (based on Marinis & Armon-Lotem, 2015) consisted of 18 sentences, with six each for the following grammatical areas: Article-noun agreement; Adjective-noun agreement; Simple present tense. Nineteen lexical items were included (see **Supplementary Materials**). Sentences had between seven and ten syllables, a range chosen based firstly on a review of the literature (e.g. Campfield &

Murphy, 2014; Erlam, 2006) and then from piloting. Piloting of sentences with different lengths with a small number of children showed that some children had ceiling effect in sentences that were shorter than seven words and floor effect in sentences that were longer than 10 words. Learners saw a picture focusing on the target item at the same time as hearing the phrase, in order to focus them on meaning, which, according to Erlam (2006), maximises the possibility of reconstructive production rather than rote repetition. Sentences were recorded with clear articulation by the second author (university degree-level proficiency in French).

It has been argued (Erlam, 2006) that the reconstructive nature of the task is also maximised and dependency on memory reduced by having a delay of around three seconds between hearing the sentence and repeating it; however, for learners of very low proficiency this places a heavy load on working memory. Piloting suggested that inserting an artificial delay meant learners could not do the task at all. Therefore in the main task no delay was inserted.

Target items were placed initially, medially, and finally in the SRT sentences, to control for order/recency effects (following Jessop et al., 2007). In order to minimise possible order effects, two sequences of the task were created and learners were randomly assigned to complete one of these sequences. Lexical items were selected by consulting commonly used Primary French teaching resources and Primary French practitioners regarding what children in Year 5 would be likely to have covered in their French lessons. The SR task was scored out of 56, with 28 points available for grammar and 28 for vocabulary (see **Supplementary Materials**).

Photo description Task (PD)

The PD aimed to gain an additional perspective on learners' grammatical and vocabulary knowledge

(focusing on the same target areas) and thus to help corroborate the insights gained from the SR task, as recommended in the EI literature (e.g. Jessop et al., 2007). It was in two parts: (1) focusing on article-noun-adjective agreements and (2) focusing on present tense verb use. For (1), pictures of coloured objects were presented and participants were asked to describe these, prompted by the question, 'Qu'est-ce que c'est'? ('What is it?'). For (2), pictures depicting actions were shown, and participants were asked to say what was happening in the picture, prompted by the question 'Qu'est-ce qu'il fait?' ('What's he doing?'). In the PD as a whole, the 17 items shown in the **Supplementary Materials** were targeted, overlapping with those in the SR task (divided between the three grammatical areas). The PD was scored out of 54, 30 points for grammar and 24 for vocabulary.

Research ethics

Permission for the learners' participation in the study was gained from school head teachers, parents/carers, and from the learners, following procedures outlined by the University in which the researchers were based, and from whom approval for the study was obtained. As a token of our appreciation for taking part in the study, learners in Year 7 were given a low-value gift token.

Procedures

Learners were tested individually, during French lesson time in Years 5 and 6, and in most cases, after school in Year 7, using a laptop in a quiet room. Pictures for each item in the SR were shown on the laptop by a researcher and the learner listened to each sentence once through headphones and gave a response (repeating the sentence) which was recorded using the free software, Audacity, and an external microphone. Similar procedures were followed for the PD, except that learners saw pictures and were asked to say in

French what they saw (see above). The researcher followed a scripted protocol and their response to the learner was limited to a simple phrase of encouragement after each attempt in the tasks.

Teaching and teacher variables

Information regarding these variables was drawn from the teacher questionnaire, in which at the start of the study primary school teachers gave information on the number of minutes a week spent on French in Years 5 and 6 and on their own French proficiency (i.e. their highest level of French qualification) and their training in language teaching (see **Supplementary Materials**). For French qualifications teachers chose from the following options: no formal French qualification; GCSE; A level (B2 level on the CEFR); degree; Native Speaker. Participating teachers' proficiency covered all these levels apart from A level. For teaching qualification, questionnaire options were: no training; some training within a generalist teacher training course; post-graduate qualification specialising in foreign languages. All of these levels were represented among the primary school teachers in the study. Teacher training and language proficiency variables remained constant across Years 5 and 6 in each primary school, as was the case for teaching approach.

Data analysis

Sentence Repetition and Photo Description task

All sentences were fully transcribed and mispronunciations were transcribed phonetically using English spelling conventions e.g. *la apa* (produced instead of *le lapin* - 'the rabbit' in the SR task). Indeterminate forms of the indefinite article were transcribed as *(if).* The sound quality was excellent and therefore inaudible words occurred very rarely and were marked with *xx* for one word or *xxx* for multiple words.

We scored only target items rather than using a more holistic scoring because our aim was to gain insights into grammatical and lexical development by tracking the same grammatical forms and lexical items over time. Scoring is outlined in detail in the **Supplementary Materials** but in summary, a 'lexis' mark was firstly awarded if the required noun or verb was produced, then 'grammar' marks for correct gender assignment, adjectival agreement, and simple present tense (as applicable).

Native-like pronunciation was not required. Scores for grammar (combining all three features) and vocabulary were calculated for each learner for each task, plus a global score (grammar plus vocabulary). Scoring was carried out firstly by the second author; the first author (also with university degree-level competence in French) then scored 10% of learners' SR and PD tasks. An inter-rater reliability rate of 98% was achieved, with differences in scores resolved through discussion. The calculation of Cronbach's alpha indicated strong reliability for both the SR task (values ranging from .81 to .84 for vocabulary, .80 to .83 for grammar and .93 to .94 for global scores) and for the PD task (values ranging from .84 to .86 for vocabulary, .90 to .91 for grammar and .93 to .94 for global scores).

Normality and homogeneity of variance assumptions were assessed by examining histograms and normality tests. Normality tests (Shapiro-Wilks) indicated that the majority of global scores from the total sample and the oracy/literacy schools sub-samples were normally distributed at each time point, except for Year 7 Sentence Repetition (total sample: S-W .97, df 160, p = .001; Literacy sub-sample: S-W .95, df 75, p = .006) and Year 5 Photo Description (total sample: S-W .97, df 160, p =.001; Literacy sub-sample: S-W .97, df 75, p = .034; Oracy sub-sample, S-W .94, df 53, p = .014). Histograms suggested however that deviations from normality were not severe and following Field (2013), it was decided that parametric tests were robust enough to cope with the slight deviations from normality for these scores (with assumptions for each individual test checked and reported separately under **Results**).

Separate grammar and vocabulary scores on both tests showed a greater level of non-normality and therefore non-parametric statistics were applied in any analyses of those scores. Unless indicated, in all cases the alpha level was set at .05. Results for Bonferroni post-hoc tests are reported using SPSS Bonferroni adjusted p-values, i.e. adjusted for multiple comparisons. Where there were any instances of violations of sphericity (Mauchly's test) degrees of freedom were corrected using Huynh-Feldt estimates of sphericity.

Results

Research Question 1: To what extent do children make progress in their knowledge of French (vocabulary and grammar) across Years 5, 6 and 7?

Across the sample as a whole, learners' scores increased steadily across the three rounds (Table 3), for global scores and separate vocabulary and grammar scores. For global scores, a one factor repeated measures ANOVA indicated that for global scores there was a significant effect for time for the SR task, F (2, 318) = 118.28, p = .001, $\eta^2 p$ .43. Post-hoc tests (Bonferroni) indicated that learners made significant progress between all three test points, (p = .001), with small effect sizes[4] (Cohen's *d*) of .34 (T1-T2) and .44 (T2-T3), and a medium T1-T3 effect size of .79. Similar results were obtained for the PD task - F (1.72, 272.77) = 96.53, p = .001, $\eta^2 p$ .38, with significant progress across all three time points (Bonferroni post-hoc tests, p = .001, with small to medium effect sizes (*d*) of .27 (T1-T2), .45 (T2-T3), and .70 (T1-T3)).

<Table 3 about here>

Examining grammar and vocabulary separately, it is clear that scores were low at all time points and especially for grammar (Table 4). Nevertheless, non-parametric analyses (a Friedman test with post-hoc

Wilcoxon matched-pairs Signed-Ranks tests) showed that learners made statistically significant progress over time in both grammar and vocabulary. While effect sizes for vocabulary growth were in the medium to large range, for grammar they were generally small to medium (Table 4).

<Table 4 about here>

Research Question 2: To what extent is children's knowledge of French across Years 5, 6 and 7 related to a) teaching and teacher factors in primary school (teaching time, teacher French proficiency, teacher level of training) and b) teaching approach?

Looking firstly at the relationship between linguistic outcomes and primary school teaching/teacher factors, a series of Spearman's rank order correlations was conducted across the total sample, for grammar and vocabulary as well as global scores. Table 5 indicates that the level of the primary school teachers' French language proficiency and their level of training were significantly related to learners' linguistic outcomes at all points for grammar and at most time points for vocabulary, although at low to moderate levels, in part attributable to the fairly narrow range of the teacher French proficiency and training ratings. Teaching time was more strongly correlated with outcomes than teacher variables, especially in Years 5 and 6, with Year 5 outcomes only weakly related to teacher factors. As learners moved into secondary school, however, the relationship between their scores and the level of French proficiency and training of the teacher who had taught them at primary school became stronger, particularly as far as grammar was concerned.

<Table 5 about here>

<u>Teacher (French) Proficiency and Teacher Training</u>

Global scores

To explore the impact of primary school teacher variables further, three groups were created for Teacher Proficiency (i.e. French language proficiency) in order to give groups of more equal size[5:] GCSE or lower; degree level; or Native Speaker (NS). For Teacher Training, two groups were created[5]: non specialist (no or limited training, within the context of a general teaching qualification); and specialist (the teacher had specialised in foreign languages for their teaching qualification). Repeated measures ANOVA tests were performed firstly on global scores with a) Teacher Proficiency, b) Teacher Training as the between subjects factor, Test Point as the within subjects factor and Teaching Time at primary school as a covariate. Figures 1 to 4 show the mean scores across teacher groups on the global scores for the SR and PD task respectively, adjusted for Teaching Time.

<Figures 1, 2, 3 and 4 about here>

Teacher Proficiency

Homogeneity of variances ($p > .05$) and covariances ($p > .05$), as assessed by Levene's test and Box's M test respectively, was established for both the SR and PD task. There was a main effect of Test Point [SR: $F(2,312) = 20.21$, $p = .001$, $\eta^2p = .12$; PD: $F(1.74, 272.09) = 28.67$, $p = .001$, $\eta^2p = .16$] but not of Teacher Proficiency group [SR: $F(2,156) = .08$, $p = .924$, $\eta^2p = .001$; PD: $F(2, 156) = 1.06$, $p = .348$, $\eta^2p = .01$]. A significant Teacher Proficiency x Test Point interaction was found for the PD task, $F(3.49, 272.09) = 3.82$, $p = .007$, $\eta^2p = .05$, but not for the SR task, $F(4, 312) = 2.13$, $p = .08$, $\eta^2p = .03$.

Post-hoc tests (Bonferroni) indicated for the PD task that while learners taught by a primary school teacher with degree level French made significant progress between all test points ($p = .001$; T1-T2 $d = .36$; T2-T3 $d = .46$, T1 – T3 $d = .80$), those in the GCSE or lower Teacher Proficiency group only made significant progress between Years 5 and 6 ($p = .044$, $d = .66$). For those taught by Native Speakers, significant progress

occurred only between Years 6 and 7 (p = .001, *d* = .69), and between 5 and 7 (p = .001, *d* = .73), but not between Years 5 and 6 (p = .168). At no test point was there any significant difference between scores across Teacher Proficiency groups.

Teacher Training

For Teacher Training, while there was a significant effect of Test Point for both tasks [SR: F(2,314) = 18.13, p=.001, η^2p = .10; PD: F(1.75, 274.90) = 31.95, p=.001, η^2p = .17] there was no significant effect of Teacher Training group [SR: F(1, 157) = .17, p = .69, η^2p = .001; PD: F(1, 157) = .13, p = .72, η2p = .001] and no significant Teacher Training x Test Point interaction for either the SR or the PD [SR: F(2,314) = 1.71, p = .18, η^2p = .01; PD: F(1.75, 274.90) = 2.66, p = .08, η^2p = .02]. It should be noted however that while the assumptions for homogeneity of variances and covariances were met for the SR scores, Levene's test was significant for PD scores in Year 5 (p = .006) and Year 7 (p = .025). The PD scores for all test points were therefore transformed using a square root transformation (so that for Levene's test p > .05) and the analysis repeated. This confirmed the results of the analysis on untransformed scores, namely a significant main effect for Test Point but not for Teacher Training group and no significant Teacher Training x Test Point interaction.

Teacher Proficiency and PD grammar

An analysis of PD grammar was then conducted, the area in which the highest correlation with Teacher Proficiency was observed, in Year 7. We explored firstly the progress of learners in each Teacher Proficiency group. After a significant Friedman test, post-hoc Wilcoxon Signed-Ranks tests (with a Bonferroni correction applied, reducing the alpha level to .006) indicated, as for global scores, that only groups taught at primary school by a teacher with degree level French made significant progress in grammar across all test points, with small to medium effect sizes (T1-T2 z = 5.31, p = .001, *r* = .31; T2-T-3 z = 5.62,

p = .001, *r* = .23; T1-T3 z = 3.22, p = .001, *r* = .39). While those taught by a teacher with GCSE or lower made no significant progress at any point, the Native Speaker group did not make significant progress between Years 5 and 6 (z = .202, p = .84) but as Figure 5 shows, they made very rapid progress at the start of Year 7, nearly doubling their Year 6 score with a large effect size (z = 4.81, p =.001, *r* = .52).

<Figure 5 about here>

Comparing PD grammar across Teacher Proficiency groups, Kruskal-Wallis tests showed that scores differed across groups at all test points (p =.01). Follow-up Mann-Whitney U tests (Table 6) showed no significant difference between the degree and GCSE group at any point. By contrast, the learners in the Native Speaker teacher group, outperforming the GCSE teacher group in Year 5 (*r* = .31), Year 6 (*r* = .32) and Year 7 (*r* = .53), significantly outperformed the degree teacher group in Years 5 (*r* =.25) and 7 (*r* = .40) but not in Year 6.

<Table 6 about here>

Teaching Time and global scores

Given that Teaching Time was the teacher/teaching variable most strongly correlated with outcomes (Table 5), we then conducted a further repeated measures ANOVA on global scores with Test Point as the within-subjects factor and Teaching Time as the between-subjects factor.

There was a significant main effect for Test Point [SR: (F (2, 312) = 68.42, p = .001, $\eta^2 p$ .31; PD: (F (1.76, 273.86) = 69.10, p = .001, $\eta^2 p$ .31], and for Teaching Time [SR: F (3, 156) = 6.73, p = .001, $\eta^2 p$.12; PD: F (3, 156) = 7.88, p = .001, $\eta^2 p$ .13]. There was also a significant Test Point x Teaching Time interaction [SR:

F (6, 312) = 3.19, p =.005, $\eta^2 p$ .06 ; PD: F (5.27, 273.86) = 4.69, p =.001 , $\eta^2 p$.08]. As homogeneity of variance was violated for the PD1 scores (Levene's test, p = .007), the ANOVA was repeated using transformed scores, confirming the results of the analysis on untransformed scores, namely a significant main effect for Test Point and for Teaching Time, with a significant Test Point x Teaching Time interaction.

Post-hoc Bonferroni tests showed that for the SR task the learners receiving 60 minutes of instruction had significantly higher scores than all other Teaching Time groups in Year 5 (p = .05, effect sizes ranging from *d* = .71 to *d* = 1.24) and Year 6 (p =.01, effect sizes ranging from *d* = .75 to *d* = 1.47). In Year 7, they still had significantly higher scores than the 15 minutes (p = .04, *d* = 1.15) and the 30 minutes group (p = .049, *d* = .55) but not the 40 minutes group (p = .34), who were catching up. During the primary school years there were no statistically significant differences between any of the groups receiving below 60 minutes of instruction. Looking at progress over time, while the Year 5-6 progress of the 60 minute group was statistically significant (p =.001, *d* = .47), between Years 6 and 7 it was not (p =.24). The 15 minutes group made very little progress between Years 5 and 6 (p =1.00), while the 30 and 40 minutes group made significant progress over all three years [30 minutes group, p = .001, effect sizes ranging from *d* = .31 to *d* = .76; 40 minutes group, p = .001, effect sizes ranging from *d* = .54 to *d* = 1.21]. Figure 6 shows progress rates for all groups in graphical form. Despite the slowing of progress for the 60 minutes group, it is also clear that at the start of Year 7 all groups receiving under an hour of instruction a week were still below, or barely reaching, the level of performance achieved by that group nearly two years earlier in Year 5.

<Figures 6, 7 about here>

Similar results were found for Bonferroni post-hoc tests conducted on the PD task. In Year 5 the 60 minutes learners again had significantly higher scores than learners in the 15 minutes (p = .001, *d* = 1.78),

30 minutes (p = .001, *d* = 1.20) and 40 minutes groups (p = .001, *d* = .90). In Year 6 differences were still significant but with smaller effect sizes (60 vs.15 minutes, p =.001, *d* = 1.61; 60 vs.30 minutes, p = .017., *d* = 64; 60 vs.40 minutes, p =.007, *d* = .71). At none of the test points were there any statistically significant differences between any of the groups receiving below 60 minutes of instruction. In terms of progress (Figure 6), the 60 minutes group had a slightly lower Year 6 score than for Year 5, while the start of secondary school saw renewed progress for them (p = .003, *d* = .41) but they did not significantly outperform any of the other groups once in secondary school. Between Years 5 and 6 the 15 minutes group made very little progress (p=.272) but moved forward significantly in Year 7 (p =.001, *d* = 1.24). Progress was made across all test points by the 30 minutes group (p = .001, effect sizes ranging from *d* = .38 to *d* = .88, 001) and the 40 minutes group (T1-T2, p = .003, *d* = .29 ; T2-T3, p = .001, *d* = .66 ; T1-T3, p = .001, *d* = .91)

Teaching Approach and global scores

The means displayed in Table 7 show little difference between the two approaches, with Oracy learners seeming to have the greatest advantage on the PD task, but the difference was small.

<Table 7 about here>

In order to further explore the impact of Teaching Approach a regression analysis was conducted in the GLM function of SPSS, permitting the inclusion of continuous and categorical factors (Hawkins, 2009). This enabled us to control for all primary school teaching variables, for which we entered the categorical variable 'School' as a fixed factor alongside Teaching Approach. The variable 'School' was used rather than Teaching Time, Teacher Proficiency and Teacher Training separately, as the separate teacher/teaching variables were highly correlated with one another, with Spearman correlations ranging from .545 to .742.

Learners' English literacy scores (which were significantly correlated with all outcomes, ranging from .532 to .632, Courtney, Graham, Tonkyn, & Marinis, 2015) were entered as a covariate. The assumptions of homogeneity, normality and linearity were checked and met (Hawkins, 2009). Table 8 shows that the model explains around 47–62 per cent of the variance in scores across Years 5–7 for the SR task, and between 51 and 57 per cent for the PD task. While Teaching Approach did not have a significant impact on outcomes, School did, explaining between 6 and 17 per cent of the variance for the SR task, and between 7 and 17 per cent for the PD task in a subsequent set of regression analyses conducted with School as the sole included variable. Thus School became a more important explanatory factor as children moved into Year 7.

<Table 8 about here>

Teaching Approach and L1 literacy

As the literature review had suggested that a more literacy-based approach might benefit learners with lower levels of L1 literacy, we also explored whether Oracy and Literacy approaches had different impacts on these learners in our sample. Hence we analysed the scores of learners with a combined Year 5 English literacy score lower than one standard deviation below the mean. This gave a very small number of learners (Table 9), which reduced further as learners moved into secondary school.

<Table 9 about here>

Mean global sores for the SR and PD tasks were calculated (Table 9), with Teaching Approach as grouping factor. Learners from both Teaching Approaches made similar progress across the three time points, although the Literacy group always outscored the Oracy group. For the PD task, however, there was clearer

evidence of greater progress between Years 6 and 7 for the learners receiving a Literacy approach in primary school compared with those receiving an Oracy approach. A Mann-Whitney U test, however, indicated that the difference was not statistically significant once a Bonferroni correction was applied (reducing the alpha level to .017) (Literacy *Mdn* = 12.50; Oracy *Mdn* = 5, U = 1.00 p = .048).

Discussion

Our findings can be summarised as follows: learners of French made statistically significant progress as they moved through Years 5, 6 and 7. The amount of progress was however modest from one year to the next, particularly for grammar. This lends weight to the argument (e.g. Muñoz, 2006) that early language learning in instructed contexts is not the rapid, effortless enterprise it is often assumed to be. The fact that learners on average did make progress across transition from primary to secondary school is however contrary to what has been found in several studies (such as those from Europe summarised in Blondin et al., 1998) but in line with more recent investigation such as Courtney (2014), arguably because we tracked development in the same linguistic items longitudinally rather than using more open-ended tasks in which progress can be more difficult to track. The factors of teaching time available for French at primary school and the primary school teachers' level of French proficiency and training in language teaching were all positively related to learning outcomes, although to differing degrees at different time points. Correlations for teaching time were generally the strongest among all teacher/teaching variables, echoing the findings of previous research (e.g. Netten & Germain, 2008; Nikolov, 2009) regarding the importance of time for learning. Teaching time correlations were strongest in the primary years, perhaps indicating that time is especially important for younger learners, to enable them to draw on the more implicit mechanisms believed to underpin early language learning (Dekeyser, 2003; Muñoz, 2014). At primary school only learners receiving 60 minutes of instruction significantly outperformed all other groups on all measures, suggesting that an hour a week is a threshold that needs to be reached before any difference in learners' progress related to amount of

instruction can be detected. The lower correlations between outcomes and teaching time at the start of Year 7 may also, however, simply reflect the fact that all learners were receiving very similar amounts of teaching time by then. Furthermore, the slowing of progress for the 60 minutes group suggests that secondary schools were concentrating on bringing all learners within a very heterogeneous intake up to a similar level of proficiency, possibly at the expense of building on the greater progress made by the 60 minute learners at primary school. By contrast, the primary school teachers' level of training and their language proficiency were more strongly related to outcomes in Year 7 than in earlier years. The relationship between teacher training and outcomes reinforces the findings of studies in other contexts (e.g. Edelenbos & Suhre, 1994; Szpotowicz, 2009) but teacher language proficiency emerged as more important in the present study than training with regard to learners' progress over time. This was particularly true for grammar on the PD task, supporting the argument that learners need an adequate model, with consistency of grammatical features in the input in order to acquire them (Murphy, 2014). While progress was most even across the three test points for learners taught by a teacher with degree level French, groups taught by Native Speakers showed little progress between Years 5 and 6 but then large amounts in Year 7. Although we did not directly measure the quality and quantity of teacher target language input in our schools, it is possible that Native Speaker teachers provided the richer input that Aukrust (2007) argues may initiate processes that develop in the longer term rather than immediately. The more even progress in groups taught by non-native speaker teachers with a degree in French may have resulted from them providing a good model combined with the more scaffolded approach that some research (e.g. Walkinshaw & Duong 2012) suggests is associated with non-native teachers.

The teaching approach learners received, Oracy or Literacy, did not seem to be a factor influencing their levels of attainment when learners of different English literacy levels were considered together. This finding reflects the conclusions of Edelenbos and Suhre (1994), and those reported in Blondin et al. (1998) for a

large body of European research, that ‘course’ is less important than other classroom factors. The only tentative benefit found for Teaching Approach was that learners with lower levels of English language literacy seemed to make more progress as they moved into secondary school and in relation to production if they had received French instruction with a stronger literacy focus at primary school. This indicates perhaps that such learners need access to the written form in order to facilitate retention of vocabulary and grammatical forms (as suggested by Hu, 2008, in relation to learners with lower levels of phonological awareness, and by Harley & Swain, 1984, commenting on the needs of young learners). It is possible that presenting language in an oral form only places a heavy burden on such learners, preventing them from retaining the language to which they are exposed or from recalling it easily.

The lack of a clearer impact of a Literacy approach on outcomes across all learners, however, suggests more strongly that quantity and quality of input were more important than focus of instruction, as indicated by the results from the GLM regression analysis in which Teaching Approach had no impact on outcomes once the combined variable of School (including teaching time, teacher training, and teacher French proficiency) was entered into the model. It is also possible however that the differences between the Oracy and Literacy approaches teachers used were not clear cut enough, even after we had excluded borderline schools. In Year 6, the amount of literacy activities increased across all schools, as teachers prepared learners for the more literacy-focused approach they would meet at secondary school. Furthermore, even in schools receiving the highest literacy score (7 and 8 out of a possible 10), there was relatively little evidence of the higher level literacy activities that may be needed for literacy to really support learning. In only one school did we see such activities, where learners read a short text and answered questions as well as writing sentences based on a model. Perhaps not coincidentally, learners in that school (School 1) had the highest mean scores when individual school scores were examined. It is possible that with more focus on higher level literacy activities a clearer advantage for Literacy schools might have been found. It should be noted, however, that

learners in School 1 also experienced the greatest amount of teaching time (60 minutes, like one other, Oracy, school), and arguably sufficient lesson time is needed in order to incorporate activities that go beyond word and sentence level. Considering this question from another angle, however, it is possible that the teacher of School 1 learners was able to include literacy activities precisely because her learners had made sufficient progress through more and higher quality teaching to enable them to cope with such activities[1].

The findings from School 1 thus suggest that teaching approaches and other teacher and teaching factors interact in complex ways. For example, the two schools which allocated the most time to the teaching of French (School 1 and School 7) both employed teachers with a degree or native speaker competence in French and who had completed specialist post-graduate teacher training in French teaching. One was in the Oracy group, the other in the Literacy group. By contrast, the teacher in the school allocating the least amount of time to French per week, 15 minutes, had had relatively little training in teaching it, although she had a degree in the subject. This suggests perhaps that primary schools where French is seen as important ensure that it is taught by a well-trained, linguistically proficient teacher, and allocate sufficient time to its teaching. The growing amount of variance explained by 'School' across our test time points suggests that it is the coming together of a number of factors that is important for the best possible outcomes for young learners as they move from primary into secondary school, rather than one single factor that makes the difference.

Limitations and conclusions

Levels of attrition in our sample size in Year 7 pose a limitation to our exploration of the impact of teaching

and teacher factors, as does the nature of the literacy-oracy division we adopted. Future research into the impact of teaching approach would most likely benefit from a more strictly experimental perspective, with the teaching in any schools involved manipulated in a tightly controlled manner, including controlling for teaching time available and teacher language proficiency/training. In the present study, we sought instead to take a more naturalistic approach, to study practice that is actually occurring in schools. While this inevitably brings with it shortcomings, arguably it also gives a more grounded picture of early language learning classrooms. Our findings suggest albeit tentatively that a literacy-based approach has the potential to help children with lower levels of L1 literacy make progress in learning another language but they are limited by the small sample of such learners that we had. Further research into provision for these learners is warranted, given that they may be excluded from language classrooms, as occurs, for example, in the USA (Sparks, 2012) and increasingly in England, especially at secondary school, on the grounds that they would find language learning too difficult and achieve limited progress (Tinsley & Board, 2016). Such research would also benefit, arguably, from a consideration of the role of working memory capacity in relation to different types of instruction, an area we were unable to explore in the present study.

It should also be acknowledged that we only assessed learners' grammatical and lexical knowledge orally, and that learners had already experienced three years of learning French prior to the start of the study. While it is probable that they received similar amounts of teaching time prior to Years 5 and 6, we cannot be certain about this, nor about the other teaching and teacher factors to which they were exposed during periods prior to the study.

While the variation in our study regarding teaching time, teacher language proficiency and training constitutes a limitation for the comparison of the two teaching approaches studied, it did allow us to gain insights into how variation in these factors is related to varying learning outcomes. The study's significance

lies in the evidence it provides from a longitudinal investigation, with a relatively large sample, regarding the importance of optimising the conditions for learning a foreign language at primary school to provide more equal opportunities for all learners, namely through teachers with sufficient pedagogical and linguistic expertise and sufficient teaching time. Such conditions are far from being guaranteed either in England (Tinsley & Board, 2016) or elsewhere (Copland et al., 2014). Thus the findings of the study are of relevance beyond the context of England, across countries implementing a lowering of the age at which the learning of another language begins. Such policies imply that language learning is a valued part of the curriculum. If this is the case, then as much if not more attention needs to be paid to providing the minimal conditions for learning as to age of learning, or as expressed by Muñoz (2011: 130), "trusting young age of learning with the burden of learning success is clearly not enough". We very much concur with that view.

Note 1: We thank an anonymous reviewer for this suggestion.
Note 2: As argued by an anonymous reviewer. We are aware that the assessment of grammatical gender can be problematic in that learners have a 50-50 chance of assigning the correct gender (masculine or feminine). To take account of the possibility of guessing we examined responses for grammatical gender separately using a one-sample t-test. At each time point, learners' performance on gender was significantly different than chance ($p < .05$).
Note 3. All of the teachers were on the primary schools' core staff and most were employed specifically to teach French. In three cases, French was taught by a 'generalist' teacher who taught them all other curriculum subjects.
Note 4. Effect size was interpreted as follows: d - small = 0.2, medium = 0 5, large =0.8 r – small = 10, medium = .30, large = .50 (Field, 2013)
Note 5: It is acknowledged that groups remained rather unequal in size.

Acknowledgement

This project was funded by the Nuffield Foundation. The views expressed are those of the authors and not necessarily those of the Foundation.

References

Aukrust, V. G. (2007). Young children acquiring second language vocabulary in preschool group-time: Does amount, diversity, and discourse complexity of teacher talk matter? *Journal of Research in Childhood Education*, 22, 17–37. doi:10.1080/02568540709594610

Blondin, C., Calendier, M., Edelenbos, P., Johnstone, R., Kubanek-German, A., & Taeschner, T. (1998).

Foreign languages in primary and pre-school education: A review of recent research within the European Union. London: CILT.

Bowers, E. P., & Vasilyeva, M. (2011). The relation between teacher input and lexical growth of pre-schoolers. *Applied Psycholinguistics, 32,* 221–41. doi: 10.1017/S0142716410000354

Cable, C., Driscoll, P., Mitchell, R., Sing, S., Cremin, T., Earl, J., Eyres, I., Holmes, B., Martin, C., & Heins, B. (2010). *Language learning at Key Stage 2: A longitudinal study.* London: Department for Children, Schools and Families (DCSF Research Reports DCSF-RR198).

Campfield, D. E., & Murphy, V.A. (2014). Elicited imitation in search of the influence of linguistic rhythm on child L2 acquisition. *System, 42,* 207–219. doi:10.1016/j.system.2013.12.002

Copland, F., Garton, S. & Burns, A. (2014) Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly, 48,* 38–762. doi: 10.1002/tesq.148

Courtney, L. (2014). *Moving from primary to secondary education: An investigation into the effect of primary to secondary transition on motivation for language learning and foreign language proficiency.* Unpublished PhD thesis, University of Southampton, UK.

Courtney, L., Graham, S., Tonkyn, A., & Marinis, T. (2015). Individual differences in early language learning: A study of English learners of French. *Applied Linguistics.* doi:10.1093/applin/amv071o0

DeKeyser, R. (2003). Implicit and explicit learning. In C.J. Doughty & M.H. Long (Eds.) *Handbook of second language acquisition* (pp. 313-348). London: Blackwell.

Department for Education (2013). *Languages programmes of study: Key stage 3. National Curriculum in England*. Available at: https://www.gov.uk/government/publications/national-curriculum-in-england-languages-progammes-of-study.

Devescovi, A., & Caselli, C. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language and Communication Disorders, 42,* 187–208. doi: 10.1080/13682820601030686.

DfES (2005). *Key stage 2 framework for languages.* Nottingham: DfES Publications.

Drew, I. (2009). Using the Early Years Literacy Programme in primary EFL Norwegian classrooms. In M. Nikolov (Ed.) *Early learning of modern foreign languages: Processes and outcomes* (pp. 108-120). Clevedon: Multilingual Matters Ltd.

Driscoll, P., Martin, C., Graham-Matheson, L., Dismore, H., & Sykes, R. (2004). *A systematic review of the characteristics of effective foreign language teaching to pupils between the ages 7 and 11*. Research Evidence in Education Library. EPPI-Centre, Social Science Research Unit, Institute of Education, London.

Edelenbos, P., & Suhre, C (1994). A comparison of courses for English in primary education. *Studies in Educational Evaluation, 20,* 513–534.

Enever, J. (Ed.) (2011a). *Early language learning in Europe*. London: British Council.

Enever, J. (2011b). Policy. In J. Enever (Ed.), *Early language learning in Europe* (pp. 23-32). London: British Council.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics, 27,* 464–491. doi: 10.1093/applin/aml001

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* London: Sage.

Galton, M., Gray, J., & Ruddock, J. (1999). *The impact of school transitions and transfers on pupil progress and attainment*. DfEE Research Report No. 131. Norwich: HMSO.

Genelot, S. (1997). L'enseignement des langues vivantes à l'école élémentaire : Eléments d'évaluation des effets au collège. *Revue Française de Pédagogie, 118,* 27-42.

Harley, B., & Swain, M. (1984). The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper, & A.P.R. Howatt (Eds.), *Interlanguage* (pp. 291-311). Edinburgh, Scotland: Edinburgh University Press.

Hawkins, D. (2009). *Biomeasurement.* Oxford: Oxford University Press.

Hill, K., Davies, A., Oldfield, J., & Watson, N. (1998). Questioning an early start: The transition from primary to secondary foreign language learning. *Melbourne Papers in Language Testing, 6 (2),* 21-36

Hu, C. F. (2008). Use orthography in L2 auditory word learning: Who benefits? *Reading and Writing, 21,* 823-841. doi: 10.1007/s11145-007-9094-6

Jensen, E.D., & Vinther, T. (2003). Exact repetition as input enhancement in second language acquisition. *Language Learning, 53,* 373–428. doi:10.1111/1467-9922.00230

Jessop, L., Suzuki, W., & Tomita, Y. (2007). Elicited imitation in second language acquisition research. *Canadian Modern Language Review, 64*, 215–220. doi: 10.3138/cmlr.64.1.215

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press

Lightbown, P.M. (2014). Making the minutes count in L2 teaching. *Language Awareness, 23,* 3-23 . doi: http://dx.doi.org/10.1080/09658416.2013.863903

Lo Bianco, J. (2009) Second languages and Australian schooling. *Australian Education Review, 54*. Available online at: http://research.acer.edu.au/aer/8/

Low, L., Brown, S., Johnstone, R., & Pirrie, A. (1995). *Foreign languages in primary schools: Evaluation*

of the Scottish pilot projects 1993-5. Final report to Scottish Office. Stirling: Scottish CILT.

Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. In: S. Armon-Lotem, J. de Jomg, & N. Meir (Eds.) *Methods for assessing multilingual children: Disentangling bilingualism from language impairment.* Clevedon: Multilingual Matters.

Mihaljevíc Djigunović, J. (2009). Impact of learning conditions on young FL learners' motivation. In M. Nikolov (Ed.) *Early learning of modern foreign languages: Processes and outcomes* (pp. 75-89). Clevedon: Multilingual Matters Ltd.

Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In C. Muñoz, C. (Ed.) *Age and the rate of foreign language learning* (pp. 1-40). Clevedon: Multilingual Matters.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics, 29,* 578–96. doi: 10.1093/applin/amm056

Muñoz, C. (2011). Is input more significant than starting age in foreign language acquisition? *International Review of Applied Linguistics, 49,* 113-133. doi: https://doi.org/10.1515/iral.2011.006.

Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics, 35,* 463–482. doi:10.1093/applin/amu024

Murphy, V. (2014). *Second language learning in the early school years: Trends and contexts*. Oxford: Oxford University Press.

Myles, F., & Mitchell, R. (2011). *Learning French from ages 5, 7 and 11: An investigation into starting ages, rates and routes of learning amongst early foreign language learners*. ESRC End of Award Report, RES-062-23-1545. Swindon: ESRC

Netten, J., & Germain, C. (2008). The future of intensive French in Canada. *The Canadian Modern*

Language Review/Revue Canadienne des Langues Vivantes, 65, 757-786. doi : 10.1353/cml.0.0077

Nikolov, M. (2009). Early modern foreign language programmes and outcomes: Factors contributing to Hungarian learners' proficiency. In M. Nikolov (Ed.) *Early learning of modern foreign languages: Processes and outcomes* (pp. 90-107). Clevedon: Multilingual Matters Ltd.

Office for Standards in Education (2011). *Modern languages. Achievement and challenge 2007-2010*. Available at: http://www.ofsted.gov.uk/resources/modern-languages-achievement-and-challenge-2007-2010

Porter, A. (2014). *An early start to French literacy: Learning the spoken and written word simultaneously in English primary schools*. Unpublished PhD thesis, University of Southampton, UK.

Pufahl, I. & Rhodes, N. C. (2011) Foreign language instruction in U.S. schools: Results of a national survey of elementary and secondary schools. *Foreign Language Annals*, 44(2), 258–288. doi: 10.1111/j.1944-9720.2011.01130.x.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Company

Sparks, R. L (2012).Individual differences in L2 learning and long-term L1–L2 relationships. *Language Learning, 62*, 5-27. doi: 10.1111/j.1467-9922.2012.00704.x

Szpotowicz, M. (2009). Factors influencing young learners' vocabulary acquisition. In M. Nikolov (Ed.) *Early learning of modern foreign languages: Processes and outcomes* (pp. 195-211). Clevedon: Multilingual Matters Ltd

Tinsley, T., & Board, K.. (2016). *Language trends 2015/16. The state of language learning in primary and secondary schools in England.* Reading: CfBT.

Available at: https://www.britishcouncil.org/sites/default/files/language_trends_survey_2015.pdf

Unsworth, S., Persson, L., Prins,T., & De Bot, K. (2015). An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics, 36,* 527-548 doi: 10.1093/applin/amt052

Walkinshaw, I., & Duong, O.T.H. (2012). Native- and non-native speaking English teachers in Vietnam: Weighing the benefits. *TESL-EJ,* 16 (3). Available at: http://www.tesl-ej.org/wordpress/issues/volume16/ej63/ej63a1/

Wu, S-L., & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals, 46,* 680–704. doi: 10.1111/flan.12063

Yan, X., Maeda, Y., Lv, J., & Ginther, A (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing, 33,* 497–528. doi: 10.1177/0265532215594643

Table 1 Data collection points and number of participants, whole sample

	Male	Female	Total
1 – Year 5, Summer	119	133	252
2 – Year 6, Spring/Summer	114	127	241
3 – Year 7, Autumn	75	89	164

Table 2 Sub-sample by teaching approach

	Oracy	Literacy	Total
1 – Year 5, Summer	73	103	176
2 – Year 6, Spring/Summer	68	100	168
3 – Year 7, Autumn	54	77	131

Table 3 Means (Standard Deviations), whole sample, for vocabulary, grammar and global scores, Years 5-7.

Time Point	Vocabulary	Range	Grammar	Range	Global score	Range
			Sentence Repetition*			
Year 5	14.40	1-26	9.08	0-24	23.48	1-50
(N = 252)	(4.86)		(5.12)		(9.58)	
Year 6	17.02	4-27	9.76	1-23	26.78	6-49
(N = 241)	(5.34)		(4.72)		(9.71)	
Year 7	19.52	6-28	11.37	1-23	30.88	8-51
(N = 164)	(4.72)		(4.59)		(8.89)	
			Photo Description**			
Year 5	9.52	0-23	4.39	0-18	13.91	0-39
(N = 252)	(4.81)		(3.65)		(8.14)	
Year 6	10.98	0-21	5.11	0-19	16.09	0-40
	(4.65)		(3.55)			

(N = 241)					(7.78)	
Year 7	12.52	0-23	7.14	0-24	19.66	4-45
(N = 164)	(4.95)		(5.08)		(8.48)	

*For the SR task, both grammar and vocabulary had a maximum possible score of 28 each.
** For the PD task, the maximum possible score for vocabulary was 24, and for grammar, 30.

Table 4 Vocabulary and grammar: comparisons across time points (Within-Subjects, Wilcoxon matched-pairs signed-ranks test)

	Sentence Repetition					
	Vocabulary			**Grammar**		
	z	**p**	**r**	**z**	**p**	**r**
Year 5 vs. Year 6	9.75	.001*	.44	3.12	.001*	.14
Year 6 vs. Year 7	7.83	001*	.39	4.55	.001*	.23
Year 5 vs. Year 7	10.22	.001*	.50	6.45	.001*	.32
	Photo Description					
Year 5 vs. Year 6	8.21	.001*	.37	4.50	.001*	.20
Year 6 vs. Year 7	5.62	.001*	.28	5.59	.001*	.28
Year 5 vs. Year 7	8.28	.001*	.41	6.60	.001*	.32

*Significant at $p < 0.008$ after Bonferroni correction.

Table 5 Task results correlated with learner variables and teaching/teacher variables (Spearman rank order correlations)

Score	Year	Teacher Proficiency	Teacher Training	Teaching Time
Sentence Repetition				
Global scores	5	.126*	.098	.239**
	6	.182**	.202**	.362**
	7	.230**	.222**	.231**
Vocabulary	5	.045	.053	.158*
	6	.179**	.181**	.389**
	7	.241**	.210**	.247**
Grammar	5	.182**	.132*	.279**
	6	.175**	.216**	.303**
	7	.192*	.211**	.211**
Photo Description				
Global scores	5	.198**	.161*	.424**
	6	.165*	.140*	.392**
	7	.239**	.183*	.235**
Vocabulary	5	.180**	.147*	.388**
	6	.130*	.083	.373**
	7	-.006	.059	.121
Grammar	5	.196**	.152*	.436**
	6	.203**	.215**	.377**
	7	.420**	.274**	.288**

*** < .001; ** < .01; * <.05 (2-tailed)

Table 6 Descriptive statistics and results of Between-Subjects, Mann-Whitney U Tests for PD grammar, by Teacher Proficiency group

*Significant at .006 after Bonferroni correction

	Medians (Min.-Max.)		
	GCSE or below	**Degree**	**Native Speaker**
Year 5	3.00 (0-12) n = 49	3.00 (0-16) n = 146	6.00 (0-18) n = 57
Year 6	4.00 (0-14) n = 46	4.00 (0-15) n = 141	6.00 (0-19) n = 54
Year 7	4.00 (0-11) n = 16	5.00 (0-16) n = 105	11.00 (2-24) n = 43
	Mann-Whitney U (p value)		
	GCSE vs. Degree	**GCSE vs. Native Speaker**	**Degree vs. Native Speaker**
Year 5	3500.00 (.82)	902.50 (.002*)	2718.50 (.001*)
Year 6	2732.00 (.107)	779.50 (.001*)	3020.50 (.025)
Year 7	619.00 (.09)	106.00 (.001*)	1096.00 (.001*)

Table 7: Mean global scores Years 5-7 by Teaching Approach, sub-sample (Standard Deviation)

Sentence Repetition

	Literacy		**Oracy**	
Time Point	**Mean (SD)**	**Range**	**Mean (SD)**	**Range**
Year 5	24.86 (10.20)	3-50	24.74 (7.51)	1-41
Year 6	28.17 (10.74)	6-49	27.85 (8.69)	7-45
Year 7	31.65 (8.99)	8-51	32.19 (8.72)	10-47

Photo Description

	Literacy		**Oracy**	
Time Point	***Mean (SD)***	***Range***	***Mean (SD)***	***Range***
Year 5	13.85 (8.14)	1-39	15.75 (9.01)	0-36
Year 6	16.04 (8.06)	2-36	17.31 (8.41)	0-40
Year 7	19.86 (8.43)	3-38	21.59 (8.53)	3-45

Table 8: GLM regression analysis for Sentence Repetition and Photo Description tasks, Years 5 to 7

	SR Y5	**SR Y6**	**SR Y7**	**PD Y5**	**PD Y6**	**PD Y7**
English	MS=2945.77	MS=3345.20	MS=2887.53	MS=2151.52	MS=2075.95	MS=2424.94
Literacy	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
	f=64.47	f=79.80	f=92.16	f=58.06	f=65.59	f=75.37
	B=2.38	B=2.69	B=2.89	B=2.03	B=2.12	B=2.65
	p=. 001**	p=.001**	p= .001**	p= .001**	p= .001**	p = .001**
	η^2p = .28	η^2p =.33	η^2p =.43	η^2p =.26	η^2p =.29	η^2p =.38
Approach	MS=23.15	MS=125.90	MS=2.28	MS=.24	MS=4.14	MS=42.21
	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
	f=.51	f=3.00	f=.07	f=.006	f=.13	f=1.31
	p=.478	p=.085	p=.788	p=.936	p=.718	p=.254
	η^2p =.003	η^2p =.02	η^2p =.001	η^2p =.001	η^2p =.001	η^2p =.01
School	MS=625.50	MS=901.15	MS=573.42	MS=614.99	MS=517.72	MS=490.54
	(1,4)	(1,4)	(1,4)	(1,4)	(1,4)	(1,4)
	f=13.69	f=21.50	f=18.30	f=16.60	f=16.36	f=15.25
	p=. 001**	p= .001**	p= .001**	p= .001**	p= .001**	p= .001**
	η^2p =.25	η^2p =.35	η^2p =.37	η^2p =.28	η^2p =.29	η^2p =.33
R^2	.474	.591	.619	.510	.547	.573

Table 9 Mean scores (Standard Deviations) for learners with low English literacy, across Teaching Approaches

Task	Year	Mean score by Approach (SD)	
		Literacy	**Oracy**
Sentence Repetition	***Y5***	14.45 (8.17) n = 11	14.00 (8.86) n = 5
	Y6	17.30 (7.75) n = 10	15.75 (7.14) n = 4
	Y7	20.67 (9.85) n = 6	18.00 (9.17) n = 3
Photo Description	***Y5***	6.19 (4.07) n = 11	4.80 (3.42) n = 5
	Y6	8.50 (6.57) n = 10	4.25 (6.33) n = 4
	Y7	13.17 (4.07) n = 6	6.33 (4.16) n = 3

FIGURES

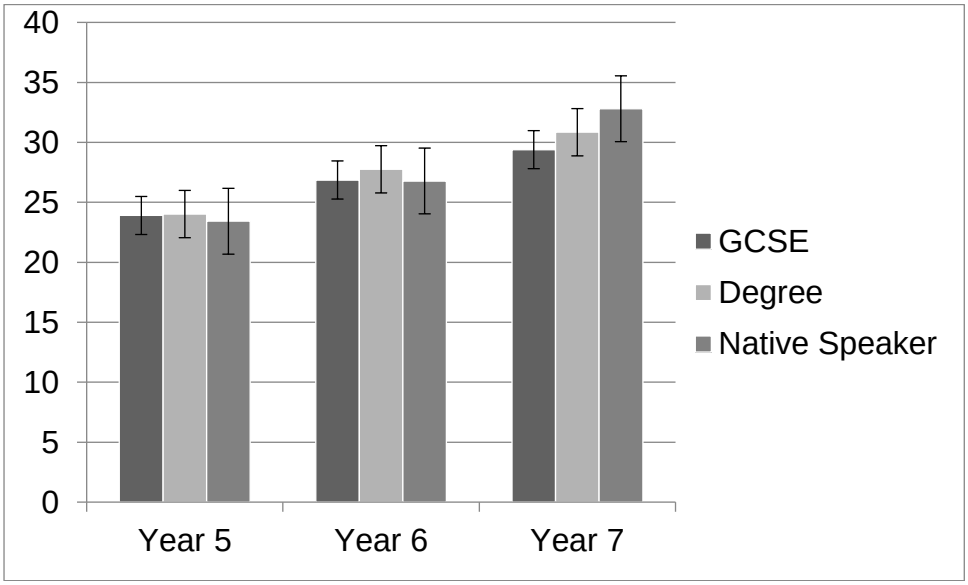

Figure 1: SR global scores: Teacher Proficiency group, adjusted means with Key Stage 2 Teaching Time as a covariate

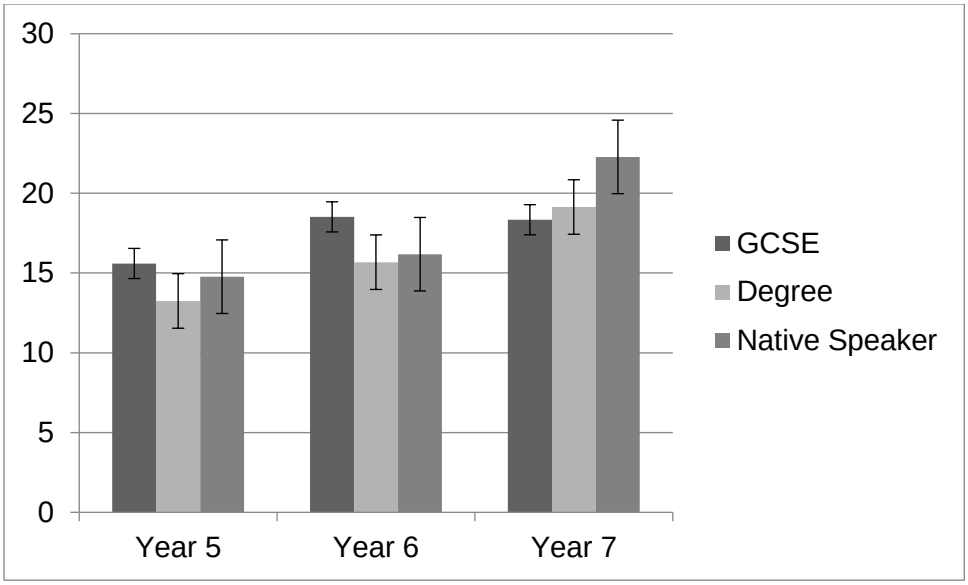

Figure 2: PD global scores: Teacher Proficiency group, adjusted means with Key Stage 2 Teaching Time as a covariate

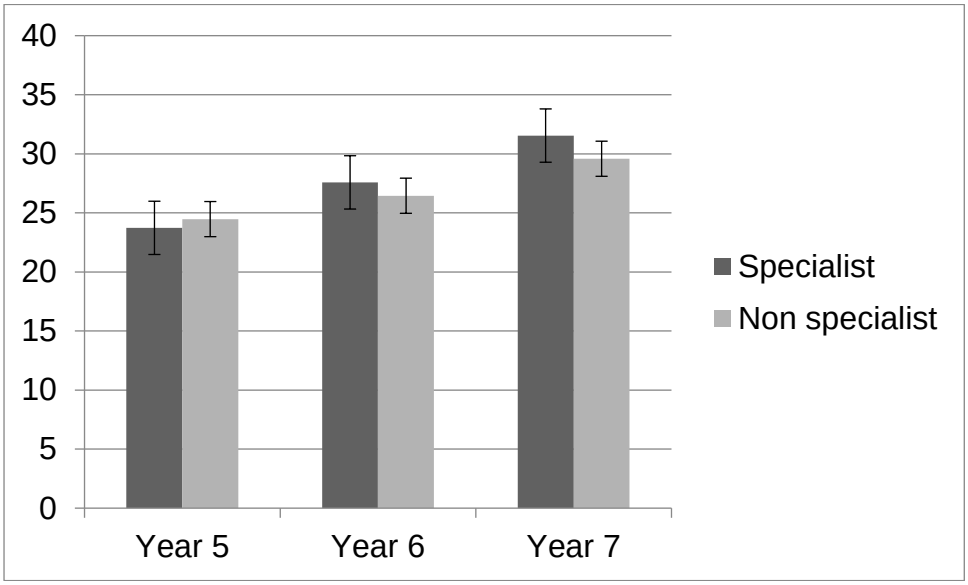

Figure 3: SR global scores: Teacher Training group, adjusted means with Key Stage 2 Teaching Time as a covariate

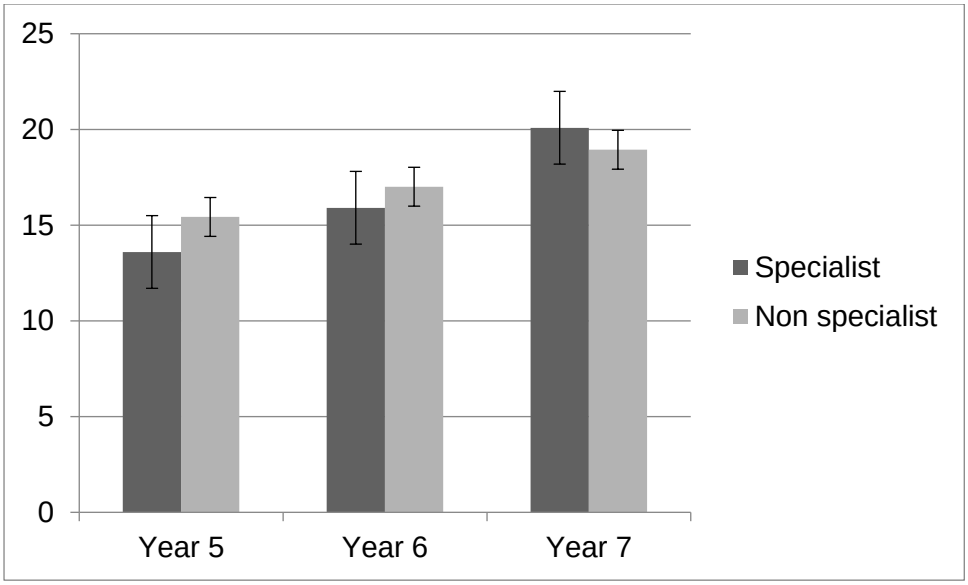

Figure 4: PD global scores: Teacher Training group, adjusted means with Key Stage 2 Teaching Time as a covariate

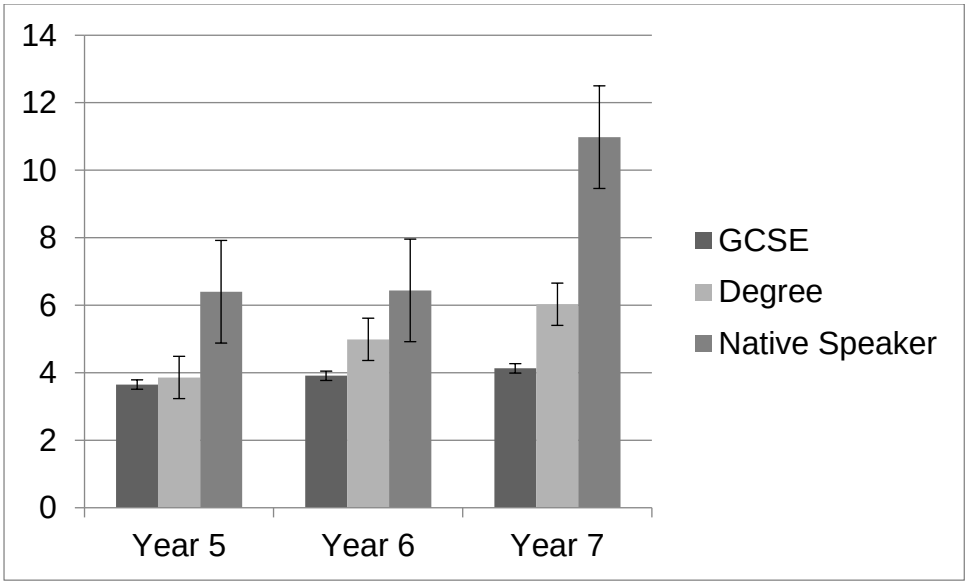

Figure 5: Mean PD grammar scores: Teacher Proficiency group

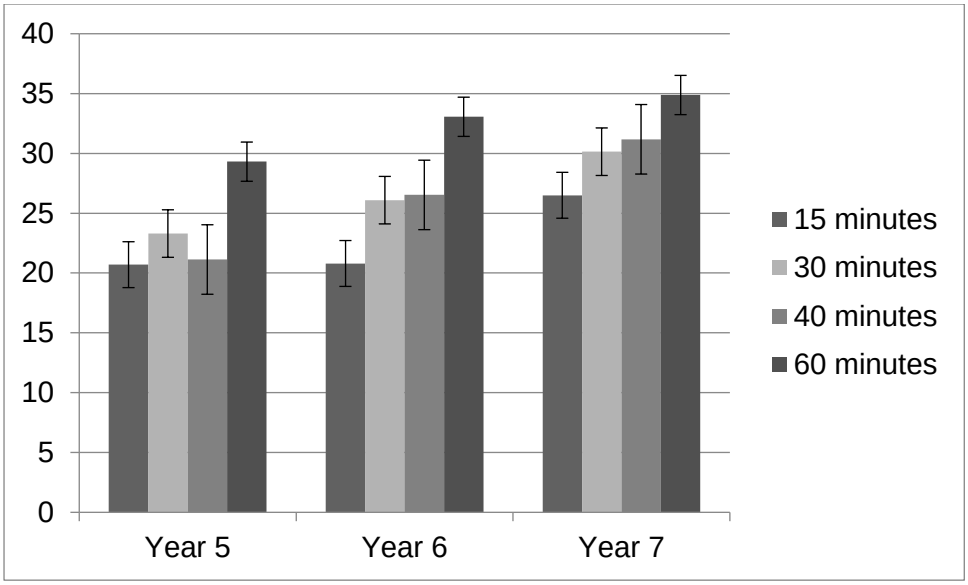

Figure 6: Mean SR global scores: Key Stage 2 Teaching Time

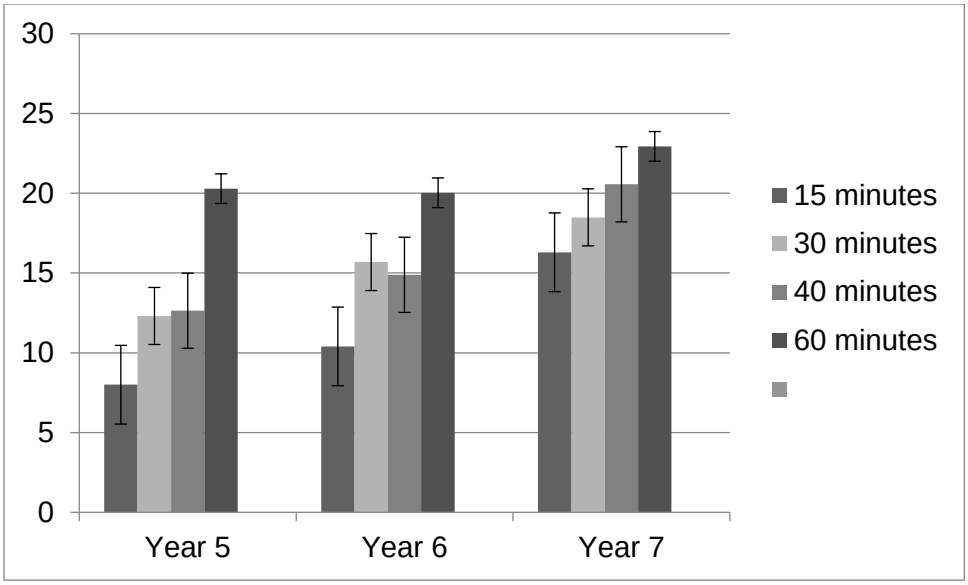

Figure 7: Mean PD global scores: Key Stage 2 Teaching Time

Supplementary Materials

Teacher and teaching factors in each primary school, Years 5 and 6, whole sample

School	Teaching time for French per week (minutes)	Teacher level of formal French qualification	Teacher level of training in language teaching	Approach (Oracy or Literacy, or borderline)
1	60	Degree	One year specialist post-graduate training	Literacy
2	40	Native Speaker	One year specialist post-graduate training	Oracy
3	30	Degree	One year specialist post-graduate training	Literacy
4	30	GCSE	None	Borderline
5	15	Degree	Some language focus within a one year general primary training programme	Oracy
6	30	None	None	Borderline
7	60	Native Speaker	One year specialist post-graduate training	Oracy
8	40	Degree	One year specialist post-graduate training	Borderline
9	30	Degree	One year specialist post-graduate training	Literacy

Tasks

Full details of the tasks used in the study can be found at http://pmlresearch.com where all tasks can be freely downloaded.

Sentence Repetition

Target Items

Nouns (10, articles omitted, grammatical gender indicated in parenthesis)

Lapin (m)
Souris (f)
Crayon (m)
Table (f)
Stylo (m)
Trousse (f)
Chien (m)
Chat (m)
Sandwich (m)
Pomme (f)

Adjectives (3)
Vert
Blanc
Brun

Verbs (infinitives shown) (6)
S'appeller
Regarder
Manger
Jouer
Faire
Ecouter

Sentence Repetition Task Sequence 1 (with target items underlined; Sequence 2 uses identical sentences in a different order)

		Possible points for lexis	Possible points for grammar
1.	Il a un lapin et une souris	2	2
2.	Il a un crayon vert dans son sac	2	2
3.	J'ai un grand frère qui s'appelle Tom	1	1
4.	Il préfère la table blanche et bleue	2	2
5.	Je regarde la télé à la maison	1	1
6.	J'ai un stylo bleu et une trousse rouge	2	2
7.	Je mange une salade avec des frites	1	1
8.	Elle a un chien blanc et noir	2	2

9. Le chat noir est sous la table	2	2
10. Après l'école il joue au football	1	1
11. Je voudrais un sandwich et une pomme	2	2
12. Il fait du sport après l'école	1	1
13. Elle veut des frites et une pomme verte	2	2
14. Le chien dit bonjour à la vache	1	1
15. Il écoute de la musique dans sa chambre	1	1
16. Le crayon rouge est sur la chaise	1	1
17. Pierre veut une trousse brune et noire	2	2
18. Le chat brun est dans ma maison	2	2
Total :	**28**	**28**

Scoring of the SR Task

For sentences assessing article+noun agreement, one mark was awarded for 'lexis' if the learner produced the noun correctly. A further mark (for grammar) was awarded if the article-noun agreement was correct. For example, where the target item was the feminine noun 'la table', a learner producing *le table* would score one point ('lexis'), another producing *la table*, one point for lexis and one point for 'grammar'.

For the sentences assessing adjective+noun agreement, the correct provision of the noun and the adjective scored two points for lexis. A further 'grammar' point was awarded for the correct agreement of the article and the noun, and a further point for the correct agreement of adjective and noun. Thus the correctly produced '*une pomme verte'* (*a green apple*) would receive the maximum of four points, two for *pomme* and *verte* as lexis, one for correct article agreement and one for correct adjective agreement (both 'grammar' points). If the learner produced '*un pomme vert'* they would be awarded three points, two for lexis and one for grammar (art.+adj agreement). While in this example the learner assigned the incorrect gender to the noun, the adjective agreement is consistent with the article they used (both masculine). In the simple present tense sentences one mark was awarded for lexis if the verb was produced but in the incorrect form (for example, in the infinitive) and a further point was awarded for 'simple present tense' if the verb was correctly conjugated with the correct pronoun. Therefore the maximum score for each of these sentences was two.

Photo Description Task

Target items

Nouns (8, articles omitted, grammatical gender indicated in parenthesis)

Pomme (f)
Stylo (m)
Chien (m)
Table (f)
Lapin (m)
Souris (f)
Trousse (f)
Sandwich (m)

Adjectives (4)
Vert
Blanc
Gris
Brun

Verbs (infinitives shown) (5)
Manger
Ecouter
Regarder
Jouer
Faire

Photo Description Task Items (with target items underlined)

Part A:

	Possible points for lexis	Possible points for grammar
1. Un stylo vert	2	2
2. Un chien blanc et noir	2	2
3. Une trousse grise	2	2
4. Une souris brune	2	2
5. Un lapin gris	2	2
6. Un crayon brun	2	2
7. Une table blanche	2	2
8. Une pomme verte	2	2

Part B:

1. Il joue avec un/le chien	2	3
2. Il écoute de la musique	1	2
3. Il mange un/le sandwich	2	3
4. Il fait du vélo	1	2

5. Il joue au football	1	2
6. Il regarde la télé	1	2
	Total : 24	**30**

Scoring of the PD Task

The scoring for the Photo Description task followed the same pattern as for the SR task, except that for elicitation of simple third person present tense forms of verbs, three points were available (one point for lexis for providing the target verb, with a possible two further grammar points for 'simple present tense' - one for providing the correct third person pronoun, and one for the correct form of the verb). Many learners produced bare finite verbs without pronouns and it was necessary to develop a coding scheme to reflect this behaviour. Two of the sentences targeting verbs also included a target noun. For article-noun agreement, we accepted either a definite or an indefinite article.