Please cite the Published Version

Title

Difficulties assessing knowledge of grammatical terminology: implications for teacher education and teaching

Authors

Huw Bell, School of Childhood, Youth and Education Studies, Manchester Metropolitan University, UK (corresponding author)

Steph Ainsworth, School of Teacher Education and Professional Development , Manchester Metropolitan University, UK

Abstract

This paper starts by reporting on the design and trialling of an informal, low-stakes test designed to assess the knowledge of grammar terms of pre-service primary school teachers following the UK's National Curriculum, the results of which were used to inform the design of a series of optional grammar classes. The test proved surprisingly difficult to design, and the final version contains many flaws, some predictable and some less so. In this paper we describe the design process and illustrate the difficulties we encountered, using evidence from our contemporaneous discussions and notes, as well as qualitative feedback and selected results from student teachers' test scores. We speculate on the extent to which the problems of our test reflect underlying issues with our understanding of the nature of declarative grammatical knowledge; these issues are likely to be relevant in all contexts where formal grammar is taught. We also draw out specific implications for the grammatical framework used in the English National Curriculum, and the associated tests of grammar, punctuation and spelling currently in use.

Introduction

The 2014 National Curriculum for England (DfE, 2013a; henceforth NCE2014) requires that a considerable body of metalinguistic terminology be taught to primary school children. The importance of this is reinforced by mandatory testing in Year 6, when children are aged 11-12. These tests include grammar, punctuation and spelling, usually abbreviated to GPS (a term which we use henceforth). There are around 25 terms which could be considered to do with grammar rather than word formation, punctuation or text structure, making up 'a relatively

2

conservative list … based on traditional descriptive terminology' (Bell, 2015, p. 146). The Glossary for teachers attached to the Programmes of Study for Key Stages 1 and 2 – that is, Years 1-6 of primary education, for children aged 5-12 – is considerably longer and more detailed, and provides about another 20 items of grammatical metalanguage which need not be taught to children.

There is evidence that the new requirements have had considerable impact on the way teachers and children work. Safford *et al.* (2016) found teaching to the test was prioritised, and that much teaching focused on memorisation. Bell (2016) found that although some teachers are positive about teaching grammar, it is clear that many do not have adequate subject knowledge (cf. Derewianka and Jones, 2010; Jeurisson, 2012), largely because they have not studied it themselves (e.g. Kolln and Hancock, 2005). Explicit grammar teaching has become a contentious issue, with much discussion focused on the perceived difficulty and lack of utility of grammatical terminology, as well as the design of the Year 6 test and the stress it causes. In short, teaching grammar has become thoroughly politicised in the UK (although this is not new: see Cameron, 1997). These issues are relevant around the Anglophone world, where there has been a recent resurgence of interest in explicit grammar teaching and the development of children's wider knowledge of language (e.g. ACARA, 2009; CCSSI, 2010).

In this paper we start with a pressing practical issue: teacher subject knowledge. Broadly stated, the problem is that many in-service teachers or pre-service student teachers have not had a systematic education in grammatical metalanguage or concepts themselves. Our student teachers report difficulties coping with their own learning load, and uncertainty in teaching while on school placements. Most lack accurate subject knowledge and experience, and have underdeveloped understanding of the rationale behind the changes; in addition, of course, they are exposed to what might be loosely termed the 'anti-grammar' arguments, mostly without the moderating influence of a countervailing voice.

At the authors' institution we have tried to address the lack of subject knowledge amongst student teachers by providing instruction in the type of grammatical terminology required to teach KS1-2. The curriculum is crowded, however, and grammar only one small part of what must be studied. In consequence, we offer optional extra classes specifically in 'National Curriculum grammar' to groups of student teachers.

As formal instruction in grammar had largely vanished from English schools for much of the second half of the 20th century (Hudson and Walmsley, 2005), it is not realistic to expect all our student teachers to have detailed or comprehensive knowledge about grammar. However, their backgrounds are linguistically diverse, and many of them have relevant knowledge: some *have* studied English language, and others have learned English as a second language or studied other languages. Our student teachers are therefore not a uniform group in terms of what they know about grammar.

We describe more fully elsewhere (Bell and Ainsworth 2019) how these classes are designed, how they work in practice, and their outcomes. Here we need say only that the classes aim to improve terminological and conceptual knowledge; that they total 10-14 hours of input; that we have delivered the course three times with considerable revision each time; and that student teachers' perceptions of progress in subject knowledge appear higher than their test scores suggest.

This paper does not set out primarily to report on the growth and development of student knowledge of grammar terms, or to provide a detailed overview of what students do and don't know, and we have therefore deliberately limited our reporting of the test results. Rather, the paper is intended as a discussion of the difficulties of testing grammatical knowledge, to highlight those problems which appear to be inherent in such tests without a clearer understanding of what grammatical knowledge is.

Literature review

What it means to know

Our student teachers inevitably vary in spelling ability, vocabulary size and level of education in English. However, all are competent users of English in speech and writing – that is, they have good levels of procedural knowledge. The distinction between procedural and declarative knowledge will be central to our discussion, so we will make an initial attempt at clarification here.

Broadly, procedural knowledge is knowing *how to do* something, and declarative knowledge is knowing *about* something (Ryle, 1949, pp. 56-9). The key elements of Ryle's distinction are that procedural knowledge is acquired gradually and cannot usually be explicitly verbalised – it is inaccessible – whereas declarative knowledge may be obtained relatively quickly (e.g. through instruction) and can be explicitly verbalised. For example, almost all native English speakers can select past simple or past perfect verb forms as appropriate, but can describe neither how they know which one to use nor in what the differences consist: they have procedural knowledge but not declarative knowledge. Our student teachers have procedural knowledge of English grammar in that they can speak and write accurately and unhesitatingly. However, most of them do not have sufficient declarative knowledge, particularly in grammar terminology, to teach the terms required by NCE2014. We will return to these points later in our discussion.

A second pair of terms relevant here is *explicit knowledge* and *implicit knowledge*. Following Bialystok (1981) we define implicit knowledge as that which is intuitive, unanalysed, and characterised by lack of awareness. By contrast, explicit knowledge is knowledge which has been analysed, systematised, and abstracted from its context, and is therefore '"explanatory" in that its logical basis is understood independently of its application' (Bialystok, 1981, p. 34). The wider literature does not always clearly distinguish the two continua declarative-procedural and explicit-implicit, but Bialystok usefully clarifies that explicit knowledge is not

always capable of 'articulation' or formalisation – so 'we can [explicitly] know that some verbs accept objects and some do not without the linguistic notion of transitive and intransitive verbs' (Bialystok, 1981, p. 35); declarative knowledge, by contrast, can be articulated.

Much of the discussion about types of knowledge has come from the literature on second language grammar learning. This is largely because there is uncertainty over how, and how far, deliberate instruction and conscious learning can lead to second language fluency; the discussion is predicated on the relationship between procedural/implicit and declarative/explicit knowledge and the ways in which one may develop from the other.

There is less discussion about the development of metalinguistic knowledge amongst children acquiring their first language, particularly with relation to grammar. A useful exception is Gombert (1992), whose distinction between epilinguistic and metalinguistic knowledge appears to have close parallels with that between procedural and declarative knowledge, in that it outlines the relationship between unconscious and conscious knowledge. However, there appears to be no research that relates types of knowledge to learning grammatical *terminology* in children or adults. Yet we will argue later that mastering this terminology involves consideration of much the same issues about types of knowledge; we return to this point in the discussion section.

Existing tests of grammatical terminology

There are several tests of grammatical terminology from various sources, often designed as part of wider investigations into the attitudes, beliefs, confidence, and general language awareness of teachers, student teachers and undergraduates. Here we restrict ourselves to those tests which are publically available, fairly recent and in English, and which have been used to assess knowledge of pre-service or in-service teachers. In addition, our discussion is limited to those sections which directly measure participants' knowledge of grammatical terminology. Some of the terminology in these tests has been superseded by the definitions in the Glossary

to NCE2014; and their content varies according to the purpose and age of the test – Justice and Ezell (1999), for example, tested mostly terminology related to sentence patterns, clause types and phrase types. The relevant tests are summarised in Table 1.

[Table 1 near here]

Several authors in Table 1 have used the same tests. For example, Bloor's test was used by Alderson and Hudson; and Wray's test – itself based on work by Chandler, Robinson and Noyes (1988) – similarly informs that designed by Sangster, Anderson and O'Hara. Given their common lineage, many share design features. In the most common device (Format 1), participants underline or highlight an example of a named term in a sentence. For example, Bloor asks participants to identify examples of terms such as *noun* and *verb* from this sentence:

> *Materials are delivered to the factory by a supplier, who usually has no technical knowledge, but who happens to have the right contacts.*

This is similar to Format 2, except that in the latter participants create their own sentences. This gives the test-taker control over the lexical and syntactic environment; but it also requires them to do two things rather than one, and this allows them to unwittingly (or wittingly) make the test more difficult or more simple for themselves.

Format 3 asks participants to define terms, which usefully reveals something about *depth* of knowledge. However, it is time-consuming to define terms; in addition, the variation in responses means it is probably difficult to mark reliably, and the definition of some terms varies according to theoretical perspective. Asking participants to explain the difference between two terms (Format 6, used only by Williamson and Hardman) has similar problems, and is limited to a relatively small set of paired terms: it makes little sense to ask anyone to explain the difference between *clause* and *determiner*.

Format 4 requires participants to identify the class of *all* the words in a sentence. The advantage of this is that it allows terms to be tested more than once, but this is offset by the difficulty of finding or creating a satisfactory sentence which contains the required exemplars.

Finally, Williamson and Hardman asked participants to describe the function of highlighted words in a sentence. This seems unlikely to be a reliable way of eliciting terminology, if that is what was sought; the authors note drily that 'it was found helpful to ask the students to use grammatical terminology so as to avoid lengthy nontechnical descriptions of some of the items' (Williamson and Hardman, 1995, p. 119) – so this test format could be considered functionally equivalent to Format 1, although naming need not be the same as identifying a function.

One further widely-taken test, the Literacy Professional Skills Test, must be passed by all entrants to teacher training in England. However, this examines the ability to "recognise where writing fails to conform to the conventions of written Standard English" (DfE, n.d., p. 1) and reading comprehension skills: it does not test grammatical terminology and we do not discuss it further.

The Key Stage 2 test of Punctuation, Spelling and Grammar

A key difference between the Key Stage 2 test of Grammar, Punctuation and Spelling (henceforth KS2-GPS) and the tests described earlier is that the former is designed for children who are still mastering writing; to some extent it examines developing procedural knowledge as well as knowledge of terminology. The questions below are from a Key Stage 2 test (DfE, 2017):

1. Insert a **semi-colon** in the correct place in the sentence below.

   Come and see me tomorrow I will not have time to see you today.

2. The **prefix** <u>mis-</u> can be added to the word <u>read</u> to make the word <u>misread</u>. What does the word **misread** mean? Tick **one**.

   - *to read quickly*

- *to read incorrectly*
- *to read again*
- *to read before*

3. Which sentence is punctuated correctly? Tick **one**.

- *Joe went upstairs, turned on his computer took out his homework and started.*
- *Joe went upstairs turned on his computer, took out his homework, and started.*
- *Joe went, upstairs turned on his computer, took out his homework, and started.*
- *Joe went upstairs, turned on his computer, took out his homework and started.*

The questions illustrate the variety of knowledge under assessment. Q1 requires procedural knowledge of punctuation and declarative knowledge of terminology; Q2 requires only general vocabulary knowledge; and Q3 requires only procedural knowledge of punctuation.

We do not intend to provide a full-scale critique of KS2-GPS, the limitations of which have been widely noted elsewhere: evidence submitted to a Parliamentary enquiry into primary assessment described the test as 'muddled, lack[ing] coherence [and] flawed' (CRW, 2016, p. 1) and recommended that they be discontinued. In the following sections, we provide a brief overview of how we used our test with the three cohorts, and explore the design of our own test and the difficulties we encountered in measuring students' metalinguistic knowledge.

Overview of our testing context

As mentioned above, we offered our students optional extra classes in the type of grammar used in the National Curriculum. We also designed a grammar terminology test (GTT): this was originally conceived as a way of establishing how far students knew the subject knowledge they needed, as a source of data to aid in design of the optional taught sessions, and as a method of measuring their progress after the voluntary grammar classes.

We offered these optional classes three times; for each cohort of participants, all the classes took place within a 10-week period. For all cohorts, GTT was administered at the start and end of the term in which the optional classes ran, and was mandatory for all students enrolled on their programme of study whether or not they ultimately took part in the optional classes. After

it was first used, both the content and the format of GTT were modified: in this paper we refer to the different versions as GTT1 and GTT2. Table 2 shows the number of participants in each test and class and the test used.

[Table 2 near here]

Participation in voluntary classes was low, those students who did attend did so intermittently, and very few attended all classes, so we do not have enough test results from students who attended all classes to compare them with non-attenders in the same cohort, or to compare attenders between cohorts. A similar lack of directly comparable data affects confidence measurements: for Cohort 1 we did not ask students to assess their level of confidence with the grammatical terminology; for Cohort 2 we asked students to discuss their levels of confidence during each iteration of the test, with each student providing confidence ratings for each item at the same time as they did the test; and for Cohort 3 we held focus groups after all the classes had finished to provide us with a broad overview of levels of confidence. For these reasons, precise comparisons of the degree to which confidence and knowledge varied between or within groups is not obtainable. By their own account, students who did attend thought they had improved their knowledge during the lessons: a typical comment in a focus group was that 'every week I felt like I'd learned something useful.' Our impression, based on their responses in class, broadly agreed with their view.

A single example will serve to illustrate this point. Cohort 1 students who attended at least one of the classes (n=28) demonstrated a significant increase in confidence, $t(27)=3.00$, $p=.006$, from a mean of 59.9% to 69.5% out of ten. However, no significant increase in their performance scores was found, $t(28)=1.28$, $p=.21$, with pre- and post-test scores of 46.9% and 52.4% respectively. In short, the scores achieved by our test-takers appeared not to reflect their developing knowledge of grammatical terminology, and this applied to all three cohorts.

One obvious explanation for this is that our teaching was not effective. This is possible, but based on feedback from students, we think it more likely that the test was inadequate. Now, in one sense this doesn't really matter: GTT is low-stakes, cannot be passed or failed, and was mostly intended to have a diagnostic function. In another sense, however, it may matter a great deal. We constructed the test carefully, making best use of existing tests, and after extensive consideration of its purpose and our aims, yet our test was not effective. In the next section we describe in detail how GTT1 was designed, and how we changed it in response to our dissatisfaction.

Design and development of GTT

Terminology

Given that student teachers would have to teach the terminology in NCE2014, we decided not to make any reference to classifications or analyses beyond those included or implied in the Glossary, such as re-classifying certain prepositions as subordinating conjunctions (Huddleston and Pullum, 2002, p. 600). In addition, the GTT needed to be quick and reliable, with adequate coverage of the terms in NCE2014. We therefore decided to include all the grammatical metalanguage in the Programmes of Study, with a further eleven terms from the Glossary. The included terms are shown below (those in italics need not be taught; where the terms are followed by GTT1 or GTT2, this means the terms were only included in the first or second versions of the test):

Active verb form	Coordinating conjunction	Object
Adjective	Determiner	Passive verb form
Adverb	Direct speech (GTT1)	Past tense
Adverbial	Modal verb	Present tense
Article (GTT2)	Noun	Possessive pronoun
Conjunction	Noun phrase	Preposition

Pronoun	Suffix (GTT2)	*Main clause*
Relative clause	Verb	*Past participle* (GTT1)
Relative pronoun	*(Primary) auxiliary verb*	*Preposition phrase*
Subject	*Clause*	*Present participle* (GTT1)
Subordinate clause	*Complement*	*Subjunctive*
Subordinating	*Finite verb* (GTT1)	*Verb phrase*
conjunction	*Infinitive* (GTT1)	

The non-statutory terms were included because we thought teachers' knowledge should extend beyond what they had to teach. Previous experience had also suggested that teaching clause elements, for example, was more effective if *complement* was also taught alongside *subject, object, verb* and *adverbial.* Likewise, *preposition phrase* was useful in discussing both noun-preposition relationships and in exploring how phrases can include other phrases; and participle types were useful when discussing the structure of multi-word verb forms. We included *subjunctive* not because it was particularly useful but because it had occurred in KS2-GPS at least once (DfE, 2016, p. 25). We originally included *infinitive* because it represents a very common verb form, and *finite verb* for comparative purposes; we dropped these and *direct speech* from GTT2 because they added to the learning burden and were not as helpful as anticipated. In GTT2, then, 30 terms were tested twice in separate sentences, and *passive* and *active* were tested three times each. This gave 66 questions in total; we allowed 45 minutes, which given the format of the questions (see next section) was plenty.

Format

The second major issue was the design of the test mechanism. Based on our review of existing tests we did not use formats we considered too time-consuming or too difficult, or which provided results that cannot be easily compared. We considered formats which had not been used in other tests, such as asking test-takers to label highlighted parts of a sentence using an

appropriate term: trials suggested that students found this labelling considerably more difficult, presumably because it taps productive rather than receptive knowledge of terminology, and multiple answers were sometimes possible (e.g. the underlined part in 'my car is being serviced' could be *subject* or *noun phrase*).

In the first version of our test (GTT1), we chose to use exclusively Format 1 ('find an example of X'; see Table 1), using six source sentences, including one taken directly from Bloor (1986). We wanted an element of reliability, so we tested each term twice, each time in a different sentence, and awarded a mark for each correct answer, as this example shows:

GTT1 Q3   *Although my car is relatively new, it's already becoming unreliable. Jackie bought hers in 1987 and it's still working well.*

TERM	EXAMPLE
Present tense	…………….
Past tense	…………….
Possessive pronoun	…………….
Subject	…………….

This format proved acceptable for single-word answers (e.g. *past tense* 'bought') but less so where several words were required (e.g. *subject* 'my car'), as responses could be very close but include one too many or one too few words: (e.g. *subordinate clause* 'my car is relatively new'). The second version of our test, GTT2, therefore only used this format where single-word answers were required. In cases where the term referred to more than one word, we used multiple-choice:

GTT2 Q47   *Many families were from the eastern part of the city.*

- adjective phrase   □
- noun phrase   □
- preposition phrase □

- verb phrase   □

Over both versions of the test, we saw no particular reason to choose authentic over scripted text, but found it difficult to balance accessibility and length: authentic texts tended to be longer and feature distractingly complex language, but shorter scripted texts were sometimes too easy. For example, in texts such as 'They entered the house' we found most students were able to identify the subject easily, but this ability dropped away in longer or more complex sentences ('Having decided that there was no other option, they entered the house'). As the example below shows, in GTT2 we attempted to reduce the complexity of the source sentence while keeping it long enough to provide sufficient distractors.

GTT2 Q1   I love your bag – and it's really a perfect match for that lipstick!

- Underline one adjective

- Underline one adverb

- Underline one conjunction

In the next section we discuss in detail the reasons why we think our test, and by implication most similar tests, are ineffective; we then draw out the implications for teaching and testing grammar, and our understanding of what we mean by declarative and procedural knowledge of grammar.

Three limitations in tests of grammatical terminology

We divide the problems presented by our test (and many others) into two categories. The first type includes choosing the most appropriate question type and ensuring that scoring is consistent. This type of issue is extensively discussed in the literature on language tests (e.g. Bachman, 1990; Alderson, Clapham and Wall, 1995), and since it is not unique to tests of grammatical metalanguage we do not discuss it further. The second and more serious problem is the construct validity of such tests – that is, "the degree to which a test measures what it

claims, or purports, to be measuring" (Brown, 1996, p. 231) – and we will now outline three ways in which this issue surfaces in GTT and other tests like it.

Partial and developmental knowledge is not assessed

Table 3 shows students' responses when asked to identify a subordinating conjunction and a relative pronoun from GTT2 question 4: *Tara's room was next to the stairs that led to the tower, whereas mine was situated on the ground floor*.

[Table 3 near here]

The vast majority correctly identified *whereas* as a subordinating conjunction. Several answers appear unrelated to the target answer (e.g. 'next', 'to'); but the most frequent wrong answer ('that'), while not a conjunction here, does have a subordinating function and can be a conjunction in other contexts. A similar effect is seen for relative pronoun: a large number of responses correctly identified 'that', but a large minority selected 'mine'. The latter is not a *relative* pronoun, but it is at least a pronoun, and we would suggest that those who gave this wrong answer had more knowledge than those who gave other wrong answers. Again, the test cannot accommodate partial knowledge.

In both cases, it seems reasonable to suggest that some answers which must be scored as incorrect do nevertheless indicate some degree of relevant knowledge. The point this raises is that knowledge about a particular item of grammatical metalanguage is not all or nothing: it is logical to assume that knowledge of terminology grows incrementally, and that some early-stage knowledge, while 'wrong', may form an important foundation stage in developing a more accurate grasp of concepts. In fact, it seems likely that we all have only partial knowledge, no matter how expert we may appear, and that whether the limits of our knowledge are exposed depends only on how far any particular level of analysis pushes us. Tests which do not tap this partial knowledge misrepresent the knowledge of the test takers. As we discuss later on, they may also be based on an inadequate conception of the nature of grammatical metacognition.

Not all exemplars are equal

Again, we will start with an example. Over both versions of our test, students were asked to identify a preposition from source sentences. Table 4 shows the percentage of correct responses.

[Table 4 near here]

Table 4 shows considerable variation in the prepositions correctly identified (unless, as in the fourth test item, there is only one preposition). Where there is a choice, we hypothesise that test-takers feel most confident with prepositions expressing relatively straightforward functions – for example, in the first test item directional 'to' is more frequently correctly identified than agentive 'by', and in the third test item locative 'from' is more frequently identified than associative 'with'. Our tentative hypothesis is less clear in test item 2, but overall we would suggest that where a choice is available, test-takers feel most confident with prepositions which provide information about common and conceptually simple notions such as location and direction. This is possibly because (a) these are the functions most commonly used in introductory grammar reference works, and (b) these functions can be more easily pictured (compare 'on the table' and 'style of speaking').

The important point is that, the choice of preposition in such a test is likely to have a major impact on scores, particularly where prepositions are infrequent (atop), multi-partite (on top of), or in usage more commonly represent a different class (absent). We would add that this issue seems likely to affect seriously any word class (e.g. nouns, adverbs) with a wide variety of forms and functions.

Table 4 further suggests our student teachers are unaware of the checks and tests which would enable them to identify a preposition – that is, they do not know the formal properties of this word class. If so, asking them to find a preposition in *the cat sat on the mat* may amount in some cases to nothing more than a test of whether they know that 'on' is a preposition – and

knowing that does not mean they know much about what a preposition is. On the other hand, some knowledge of the formal and functional properties of a preposition ought to provide a good chance of identifying even previously unseen examples. Our point is that being able to identify one preposition as such does not equate to being able to identify prepositions as a class, and test-takers' ability to get the right answer may depend on contextual deduction or whether they have previously seen the item as an example of the term.

Archetypal exemplars cause problems

GTT1 and GTT2, like most other such tests, prefer 'archetypal' exemplars. By this we mean those examples which have commonly exemplified a word class or feature of syntax over many years, despite revisions and changes to the ways grammar is described. A good example is the subordinate clause. The archetypal subordinate clause occurs in a sentence with at least one other 'main clause' and is marked by a subordinating conjunction. The 2017 KS2 GPS test contains precisely this type of archetype: *We change places when the bell rings*. (DfE, 2017, p. 15), and virtually identical ones are used in both versions of our test. The problem is that while this class of subordinate clause is not rare, its frequency in tests comes at the expense of other structures, including some that are extremely common (e.g. non-finite clauses such as *I want to go*), and some which commonly do feature in tests but are often labelled separately (e.g. relative clause).

This archetypal representation may dominate because of a preference on the part of test designers to test basic rather than more complex forms, and this may be sensible given what we know about the knowledge base of student teachers. However, it has potentially important consequences: it may reinforce the impression that such structures are the only type available, and it may make later teaching of less archetypal examples more difficult (which may partly account for the discrepancy between our student teachers' confidence and their post-class test scores). Such examples undermine the central concept of *clause*, which is the presence of a

verb phrase, and link it instead to a formal archetype which itself only illustrates one type of clause.

Problems 2 and 3 are both related to the type of examples in GTT, but we would argue that they are nonetheless distinct. The second problem is based on the student teachers' lack of knowledge of ways to formally identify, for example, word class; the third is based on the difficulty of choosing the range of exemplars to include in the test. Both have the same result – they decrease the efficiency and accuracy of the test.

Implications

Knowing grammatical terminology is only one part of declarative, explicit knowledge about grammar, which is only one part of wider metalinguistic awareness and knowledge. And of course terminological knowledge is also not essential: grammar can be discussed without it (Chen and Myhill, 2016, p. 102; Myhill and Newman, 2016, p. 178). Further, one could argue that knowledge of grammatical terminology of the type exemplified in NCE2014 is not essential to writing development (Wyse and Torgerson, 2017) or a growing understanding of how language works, and this argument underpins some of the reaction against the terminological requirements of the National Curriculum. However, for better or worse, the English curriculum now requires that children are taught grammatical terminology, and we contend that better understanding of how knowledge of such terminology develops is therefore essential, and that knowledge of grammatical terminology is in any case likely to form part of higher-level declarative knowledge. More importantly, we think that understanding the processes involved in learning grammatical terms can illuminate how wider metalinguistic knowledge develops.

We have found it difficult to design the terminology test described above precisely because so little is known about what knowledge of grammar terms entails. Given that grammatical terminology is a specialist subset of general vocabulary, it may be worth approaching this as a

lexical issue. Nation's (2001, p. 27) account of word knowledge breaks knowledge down into three areas: form (sound, spelling and morphology), meaning (conceptual range/coverage, associated concepts) and use (word grammar, formulaic sequences, collocations, frequency). The area most relevant to knowledge of grammatical terminology is meaning. To the primary notion of denotative meaning, Nation adds range or coverage ('what is included in the concept?'), associations ('what other words does this make us think of?') and synonyms. Nation's questions can provide, within the scope of what primary school teachers might reasonably be expected to know, an initial schema of knowledge about *noun.* Below we sketch one possible outline of bare minimum knowledge of *noun* for English, based on Quirk, Greenbaum, Leech, and Svartvik (1972, pp. 127-228).

1. 'noun' is a class of words;

2. the referent of a particular noun can only sometimes be determined by the senses (concrete/abstract);

3. proper and common nouns can be distinguished by semantic and formal properties;

4. common nouns may or may not have number contrast (count/mass), and some occur with different meanings as both mass and count forms;

5. in some circumstances nouns may be replaced by pronouns, which overlap conceptually with nouns;

6. nouns are central to the noun phrase, which may also include various modifiers (e.g. by determiners, adjectives, relative clauses);

7. noun phrases have typical functions in a clause (subject, object, complement).

Of course, there is much more about nouns that *could* be known – the relevant parts of large reference grammars amount to around 200 pages (e.g. Quirk, Greenbaum, Leech and Svartvik, 1985; Huddleston and Pullum, 2002), but these are comprehensive guides. We are interested

here only in the more modest requirements of student teachers training to teach the primary National Curriculum.

The fact is that we know very little about the processes and difficulties of obtaining and controlling even the fairly basic knowledge described above. Once a learner knows that *noun* describes a class of words, what element of knowledge is most easily or logically acquired next? How do the various elements relate to each other? What other relevant terms (*mass, count, determiner* etc) are required or useful in developing *noun* knowledge? Do any of the elements described above present particular problems in learning, and what are they? How do exemplification and repetition assist learning, and how much of each is required at what stages? How long is the learning likely to take? Because grammatical terminology is 'a network of technical concepts that help to define each other' (DfE, 2013b, p. 1), effective teaching is likely to be predicated on effective mapping of the various components of knowledge against each other – and such mapping does not exist.

A consideration arising from such difficulties is the use of rules of thumb (e.g. 'a noun is a thing') in teaching. These are acknowledged as problematic in the Glossary to NCE2014, which explicitly addresses simplified definitions of, for example, 'naming words' (DfE, 2013b, p. 9), presumably because they prioritise semantic rather than formal perspectives and provide inadequate or even misleading explanations. Myhill, Jones and Watson also note that while these simple semantic definitions are 'easy to understand at a superficial level, they are rarely robust enough to stand up to the test of authentic language in use and thus ultimately generate confusion rather than understanding' (2013, p. 85).

In fact, early learning in many knowledge domains may employ overgeneralised rules based on selected or simplified examples, which are then refined as knowledge develops; and particularly in children we often expect partial knowledge at first. It seems likely that any difficulties associated with use of heuristics are therefore caused by a failure to develop

knowledge further, or an inadequate conception of where such generalisations fit into a more developed scheme of knowledge.

Tests of grammar terminology like ours, which basically ask whether a test-taker can identify any particular noun as such, are entirely inadequate to the task of deciding *how much* and *what* they actually know. A comprehensive test in these respects would probably have to break down knowledge of (for example) *noun* into its component parts and test each separately. As far as we know this has never been done, but it appears perfectly possible, and although such a test would probably be considerably longer than existing ones it would have the notable advantage of providing genuine diagnostic insight. However, such desiderata cannot be achieved without better understanding of the processes involved in learning about grammatical terminology.

A further avenue for exploration, with potential implications for both the testing and the teaching of grammar terminology, is Vygotsky's distinction between 'scientific' and 'spontaneous' concepts. Vygotsky claims that 'the concept is not possible without the word' (Vygotsky, 1987, p. 131, cited in Brooks, Swain, Lapkin and Knouzi, 2010, p. 90), and the results presented by Brooks *et al.* suggest that 'languaging' – in this sense a particular type of explicit discussion of what learners understand by particular grammatical concepts – may '[reveal] the gaps and inconsistencies between forms and meanings that the learners use and [promote] a deeper understanding of the system behind the forms' (Brooks *et al.* 2010, p. 90).[1]

Conclusion

This paper started with an account of our attempt to design a low-stakes assessment of knowledge of grammatical terminology, and moved on to an analysis of the problems we encountered. We have argued that these difficulties indicate the need for a better understanding of what is entailed by such knowledge at a level suitable for teaching children the terms required by the National Curriculum. This paper has focused on teacher education in a UK

---

[1] We are grateful to an anonymous reviewer for this point.

context, but the issue and the findings are relevant internationally in any contexts (not just English-speaking ones) where first-language grammatical terminology is taught.

The issues outlined here are also relevant to the explicit grammatical knowledge we expect children to develop during the years of instruction. Firstly, and most obviously, this is because student teachers will soon be responsible for teaching children. But there is a more important reason. We assume that most adults have better procedural knowledge of writing than children, along with better ability to concentrate and to handle abstraction. Our experience in teaching and testing even well-motivated adults, however, leads us to believe that these skills are not sufficient, and that more attention needs to be paid to precisely what content knowledge teachers need in order to teach children effectively. In addition, given especially that primary student teachers must be generalists and that they may lack extensive linguistic training, considerably more needs to be known about the timing and staging most appropriate for learning this terminology. And we further suggest that the lack of this knowledge for adults indicates a more serious lack with regard to children's knowledge: we simply have not agreed what it *means* for a child to 'know' *noun*, for example, beyond the ability to sometimes identify one in a sentence, and so we can hardly expect teachers to have well-founded routines for teaching such terms. And because these are children, we cannot assume that their procedural knowledge is an adequate base for developing declarative knowledge of certain terms.

We would argue further that one of the most serious problems in the way that grammar is currently dealt with in the UK is the Year 6 GPS test. We will leave aside the many important arguments to be had about its distorting effects on the curriculum and teaching, the stress it may cause, its design as a test and its accuracy; the point we want to make is that the Year 6 GPS test suffers from the same problem as our own low-stakes test. That is, it is not based on a principled, developed view of what constitutes explicit knowledge of a grammatical term, and we cannot therefore know precisely what it is assessing.

In this paper we have tried to avoid taking sides in the 'grammar wars' (cf. Locke, 2010), and have not discussed the evidence for or against the benefits of grammar knowledge for writing development. However, we do believe that even fairly young children can find learning about language structure just as interesting, stimulating and enjoyable as learning about other topic, and that its presence in the curriculum is not somehow inherently wrong. We do not want grammar to disappear again from the English curriculum (Hudson and Walmsley, 2005). But until we develop a more systematic understanding of what metalinguistic and terminological knowledge is, and until the curriculum and the test are explicitly based on such understandings, the charge will always be open that neither the test nor the curriculum is fit for purpose.

References

ACARA. (2009). National report on schooling in Australia 2009. Sydney: Australian Curriculum, Assessment and Reporting Authority. Retrieved from http://www.acara.edu.au/reporting/nrosia2009.

Alderson, J. C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation.* Cambridge: Cambridge University Press.

Alderson, J.C. and Hudson, R. (2013). The metalinguistic knowledge of undergraduate students of English language or linguistics. *Language Awareness* 22:4, 320-337.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bell, H. (2015). The dead butler revisited: Grammatical accuracy and clarity in the English Primary Curriculum 2013–2014. *Language and Education*, 29:2, 140-152.

Bell, H. (2016). Teacher knowledge and beliefs about grammar: a case study of an English primary school. *English in Education,* 50:2, 148-163.

Bell, H. and Ainsworth, S. (2019). Hard to Know or Hard to Say? Developing Explicit Grammar Knowledge among Primary Student Teachers. In: Wright, C., Harvey, L. and Simpson, J. (eds.) *Voices and Practices in Applied Linguistics: Diversifying a Discipline*, pp. 195–210. York: White Rose University Press. DOI: https://doi.org/10.10.22599/BAAL1.l. Licence: CC BY-NC 4.0

Bialystok, E. (1981). The role of linguistic knowledge in second language use. *Studies in Second Language Acquisition* 4(1), 31-45.

Bloor, T. (1986). What do language students know about grammar? *British Journal of Language Teaching* 24(3) 157-162.

Brooks, L., Swain, M., Lapkin, S., and Knouzi, I. (2010). Mediating between scientific and spontaneous concepts through languaging. *Language Awareness*, 19(2), 89-110.

Brown, J. D. (1996). *Testing in language programs*. New Jersey: Prentice Hall Regents.

Cajkler, W. and Hall, B. (2012). Languages in primary classrooms: A study of new teacher capability and practice. *Language Awareness* 21(1-2), 15-32.

Cameron, D. (1997). Sparing the rod: What teachers need to know about grammar. *Changing English: Studies in Culture and Education* 4(2), 229-239.

CCSSI. (2010). *Common core state standards for language arts*. Washington D.C.: wCCSSI. Retrieved from http://www.corestandards.org/ELA-Literacy.

Chandler, P., Robinson W.P. and Noyes, P. (1988). The level of linguistic knowledge and awareness amongst students training to be primary teachers. *Language and Education* 2(3), 161-173.

Chen, H. and Myhill, D. (2016). Children talking about writing: Investigating metalinguistic understanding. *Linguistics and Education*, 35, 100-108.

CRW. (2016). Written evidence submitted by the Centre for Research in Writing at the University of Exeter: PRI0106. Retrieved from http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/education-committee/primary-assessment/written/40810.html.

Derewianka, B. and Jones, P. (2010). From traditional grammar to functional grammar: Bridging the divide. NALDIC Quarterly 8(1), 6-17.

DfE. (2013a). *English programmes of study: Key stages 1 and 2 national curriculum in England.* London: Department for Education.

DfE. (2013b). *Glossary for the programmes of study for English (non-statutory)*. London: Department for Education.

DfE. (2016). *2016 key stage 2 English grammar, punctuation and spelling Paper 1: Questions.* London: Department for Education.

DfE. (2017). *2017 key stage 2 English grammar, punctuation and spelling Paper 1: Questions.* London: Department for Education.

Gombert, J. (1992). *Metalinguistic development.* Chicago: University of Chicago Press.

Harper, H., and Rennie, J. (2008). "I had to go out and get myself a book on grammar": A study of pre-service teachers' knowledge about language. *Australian Journal of Language and Literacy* 32(1) 22-37.

Huddleston, R. and Pullum, G. (2002). *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.

Hudson, R. and Walmsley, J. (2005). The English patient: English grammar and teaching in the twentieth century. *Journal of Linguistics* 41, 593–622.

Jeurisson, M. (2012). 'Perhaps I didn't really have as good a knowledge as I thought I had.' What do primary school teachers know and believe about grammar and grammar teaching? *Australian Journal of Language and Literacy* 35(3): 301-316.

Jones, P. and Chen, H. (2012). Teachers' knowledge about language: Issues of pedagogy and expertise. *Australian Journal of Language and Literacy* 35(1), 147–168.

Justice, L. M. and Ezell, H. K. (1999). Syntax and speech-language pathology graduate students: Performance and perceptions. *Contemporary Issues in Communication Science and Disorders* 26, 119-127.

Kolln, M., and Hancock, C. (2005). The story of English grammar in US schools. *English Teaching: Practice and Critique* 4(3), 11-31.

Locke, T. (Ed.). (2010). *Beyond the grammar wars: A resource for teachers and students on developing language knowledge in the English/literacy classroom.* London: Routledge.

Myhill, D. and Newman, R. (2016). Metatalk: Enabling metalinguistic discussion about writing. *International Journal of Educational Research*, 80, 177-187.

Myhill, D., Jones, S., and Watson, A. (2013). Grammar matters: How teachers' grammatical knowledge impacts on the teaching of writing. *Teaching and Teacher Education* 36, 77-91.

Nation, P. 2001. *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1972). *A grammar of contemporary English*. London: Longman.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Safford, K., Messer, D., McLachlan, J., Walker, K., Ghosh, K., Laird, S., Disney, A. and Wright, D. (2016). *Teaching grammar and testing grammar in the English primary school: The impact on teachers and teaching of the grammar element of the statutory test in spelling, punctuation and grammar (SPaG) in England.* Leicester: UKLA and The Open University.

Sangster, P., Anderson, C. and O'Hara, P. (2013). Perceived and actual levels of knowledge about language amongst primary and secondary student teachers: Do they know what they think they know? *Language Awareness* 22(4), 293-319.

Williamson, J. and Hardman, F. (1995). Time for refilling the bath? A study of primary student-teachers' grammatical knowledge, *Language and Education* 9(2), 117-134.

Wray, D. (1993). Student teachers' beliefs about knowledge about language. In N. Bennett and C. Carré (Eds.), Learning to teach (pp. 51–72). London: Routledge.

Wyse, D. and Torgerson, C. (2017). Experimental trials and 'what works?' in education: The case of grammar for writing. *British Educational Research Journal* 43(6), 1019–1047.

Table 1: existing methods of assessing knowledge of grammatical terminology*

	Format 1 Underline terms in example text(s)	Format 2 Underline terms in self-written text(s)	Format 3 Write definitions of term(s)	Format 4 Identify word class of all words in text	Format 5 Describe function of unnamed words in example text(s)	Format 6 Explain difference between term(s)
Bloor, 1986	✓(19)					
Chandler, Robinson & Noyes 1988		✓				
Wray, 1993				✓**		
Williamson & Hardman, 1995		✓(1)	✓(4)		✓ (10)	✓ (2)
Justice & Ezell, 1999	✓(11)					
Harper & Rennie, 2008	✓(3)					
Jeurisson, 2012	✓(30)		✓(30)			
Alderson & Hudson, 2013	✓(19)					
Myhill, Jones & Watson, 2013	✓(13)					
Sangster, Anderson & O'Hara, 2013			✓(50)	✓**		

* brackets show number of items tested, where available

** these contain 17 words, but 7-9 word classes depending on the classification

Table 2: number of participants and GTT version

Cohort	Number of participants			Home programme	GTT version
	Pre-test	Post-test	Optional classes		
1	127	110	8-30	BA Primary Education	GTT1
2	111	19	8-30	BA Primary Education	GTT1
3	171	69	8-30	BA Primary Education	GTT2

Table 3 correct and incorrect identifications of *subordinating conjunction* and *relative pronoun* in GTT2 (Cohort 2; pre-course test; n=111)

Subordinating conjunction	Percentage who chose this response
whereas	9.9
that	4.5
to, next, that led to the tower, whereas mine was situated on the ground floor, mine was situated on the ground floor	0.9 each
Relative pronoun	
that	37.8
mine	24.3
whereas	18.9
Tara's	8.1
on, was, next	1.8 each
tower, ground floor, situated, ground	0.9 each

Table 4 correct identifications of preposition in four test sentences

Test item	Correct responses	Percentage choosing each correct response	Test
Materials are delivered to the factory by a supplier, who usually has no technical knowledge, but who happens to have the right contacts.	to	48%	GTT1 (n=110)
	by	28%	
After leaving Oxford, Greene worked for four years as a sub-editor on The Times.	for	4%	
	as	3%	
	on	39%	
Overall, we thought the hotel was beautiful, although outside the air was thick with fumes from a nearby factory.	with	8%	GTT2 (n=88)
	from	11%	
I really love the bag – and it makes a perfect match for that lipstick!	for	45%	