

PG/12/02

What are the main differences between spoken and written English grammar? To what extent and in what circumstances is it appropriate to teach 'spoken grammar'?

1 Introduction

The grammar of spoken English is seen by many people as being a somewhat defective version of the 'correct' grammar of written English. One reason for this lies in the origin of linguistics itself, which was only 'discovered' once language had been written down (Halliday, 1989: 36). The codifying of language into a written form allowed scholars to examine it in order to identify and analyse the constituent parts. Spoken grammar was just an unconscious element of everyday life.

Comparisons with written grammar were inevitable, especially as technology advanced to the point where researchers were able to record speech and conversation. The differences between transcribed speech and a carefully written paragraph are striking: speech appears disjointed, repetitive, poorly planned and unfocused. But conversely, if one were to read a scholarly article aloud, it would seem burdensomely dense and overly complicated. In fact, even when a paper is presented at a conference, to the most informed and interested audience possible, a successful delivery requires numerous adjustments to the structure of the message.

These changes from the written form are not haphazard or random, and the analysis of speech over the last thirty years has revealed many of the unique systematic elements of spoken grammar. This paper will describe several of these structures and the implications of these findings for language teachers.

2 Different modes of English grammar

Spoken and written grammars share many features and skilled language users will manipulate the different stylistic choices available depending on the desired effect. Written texts can adopt features of spoken English in order to sound more friendly and casual (in email messages or advertisements) while speakers can choose forms more often associated

with writing in order to sound more authoritative or professional (in public speeches or presentations). The implication is that the grammars are not defined by the media in which they are used, but the mode the user is trying to convey. A text can be as 'speakerly' or as 'writerly' as is required (McCarthy, 2001).

The major difference between the two modes is that writing is a refined product while speaking is an unconscious process (Halliday, 1985: xxiv). A piece of writing may have been revised several times before it is presented to the reader whereas the participants in oral communication are under constant pressure to respond to the many changes that occur in their environment while they are speaking. The theme and information must be negotiated even as the verbal and non-verbal responses from the other participants add to the communicative demands. 'The complexity of the written language is its density of substance, solid like that of a diamond formed under pressure. By contrast, the complexity of spoken language is its intricacy of movement, liquid like that of a rapidly running river' (Halliday, 1989: 87). Spoken grammar is the means of navigating this river. It enables greater fluency and provides a variety of methods of monitoring appropriacy during the talk (Mumford, 2008). These two categories can be used to classify the important features of spoken grammar.

It is vital when defining these features to identify only those which occur in natural speech across a range of users and contexts. Carter and McCarthy, in their *Cambridge Grammar of English* (2006), have based their findings on the CANCODE corpus (Cambridge and Nottingham Corpus of Discourse in English). Biber et al, in the *Longman Grammar of Spoken and Written English* (1999), have made use of the LSWE corpus (Longman Spoken and Written English). Factors such as age, gender, location and socio-economic class have been cross-referenced to verify the occurrence of the features across the spectrum of language groups. Examples appearing in the following sections have been taken from these corpora.

2.1 Fluency-enabling features of spoken grammar

Biber et al (ibid.: 1067) describe 'online construction of spoken English grammar' in terms of the need to keep talking, with limited planning time, but with the ability to elaborate on what has been said after the fact if necessary. The processing demands on the participants are

high and native speakers use three techniques to alleviate this: phrasal chains, flexible word order and certain non-canonical forms.

2.1.1 Phrasal chains

The speaker's primary goal is to be understood, so he or she transmits the information in small, easy-to-process units, progressing from what is known or relevant to what is new. This structure of speech, resembling short clauses in succession, is described by Carter and McCarthy as phrasal chains (op. cit.: 168). These phrasal chains enable other features which are common in spoken grammar: a preference for coordinating (rather than subordinating) structures; a preference for simple, post-modified clauses; and a high frequency of lexical chunks.

2.1.1a Simple, coordinating structures

Subordinate clauses generally require more mental processing (for both the speaker and the listener) than is available during real-time communication (ibid.: 170). According to the data, there is a much stronger preference for coordinating conjunctions (*and, but, or*) which permit an easier flow from one piece of information to the next.

*I was driving along talking to Sue **and** we'd, like, stopped at some traffic lights **and** then - bang - there was this almighty crash **and** we got pushed forward all of a sudden.* (ibid.: 170)

Simple subordinating conjunctions (*so, because/'cos'*) are less frequently found in speech when compared to written texts and often function like coordinating conjunctions in speech, linking subjects rather than nesting them.

*A: I bet they paid her for Sunday and not paid her for the Bank Holiday, Friday and Monday. **Cos** that would make your nine hours wouldn't it.
B: Yeah. **Cos** she's got the Saturday down same as I did **cos** we all did the Saturday.* (ibid.: 170)

2.1.1b Simple noun phrases

Noun phrases are also much simpler in speech. While it is common to find nouns extensively modified in written English, 85% of the noun phrases found in conversation in the LSWE corpus had no modification at all (Biber et al, op. cit.: 578). Pronouns occur with much higher

frequency in oral communication. They are even simpler than noun phrases and the shared context of much spoken interaction allows for the use of pronouns until there is a need for clarification (Carter and McCarthy, op. cit.: 168). When speakers do modify a noun phrase in speech, in the majority of cases only one adjective is used.

*It's a **large house**, lovely, just right.* (ibid.: 168)

2.1.1c Lexical chunks

A 'chunk' is a cluster of two or more words which are regularly found together without necessarily being grammatically complete: *you know, I mean, sort of, have to, I don't know, do you think*. Some of these chunks can occur more frequently than many everyday words on their own (ibid.: 828).

These language chunks show the repetitiveness and the interpersonal meanings inherent in the spoken mode of communication. By utilising formulaic utterances to perform various common functions in conversation, speakers can give themselves more processing space to devote to the content of the message rather than the structures.

2.1.2 Flexible word order

Brazil (1995: 47-48) states that speakers select the phrasal chains in their talk in order to progress from an initial state to a target state. The initial state is a body of knowledge assumed to be shared by the participants. The target state is reached when the speaker feels they have sufficiently informed their audience. Any time a speaker realises they are not going to reach their desired target state, they must adjust their message to compensate. The attention of the audience is focused and re-focused with the insertion of phrases, ideas and evaluations as necessary.

2.1.2a Headers

A header is an item of information presented at the start of an utterance to highlight something which will then be described more completely. The header is usually a noun phrase and is usually referred to in the main utterance with a pronoun. This allows the listener to more easily separate the identification of the object from the meaning or action the speaker wishes to attach to it.

That leather coat, it looks really nice on you. (Carter and McCarthy, op. cit.: 192-3)

Headers may also contain more complex constructions, particularly chains of noun phrases. These typically move from known information to new information and they help orient the listener to the topic of the remark and the relationship to the currently shared information. (ibid.: 193)

His cousin in London, her boyfriend, his parents bought him a Ford Escort.

Madge, one of the secretaries at work, her daughter got married last week.
(ibid.: 193)

2.1.2b Tails

Tails similarly occur outside the subject-verb-x clause structure but are found at the end of the utterance. They are most often noun phrases which clarify or elaborate an item in the main clause, which in most cases has been given as a pronoun. Sometimes the tail may also consist of prepositional phrases or complex noun phrases (ibid.: 194-195).

They're incredibly nice, our neighbours.

I put it there, on the fridge.

It never occurred to me, the danger I was in. (ibid.: 195)

2.1.3 Non-canonical forms

The value placed on efficiency and a light processing load in the transmission of ideas sometimes leads to the violation of certain written grammar rules in the process of speech. *There's* can often be found followed by plural noun phrases in conversational English (Biber et al, op. cit.: 186).

Gary, there's apples if you want one. (ibid.: 186)

Biber et al propose that the contraction has become a single lexical unit when processing speech. They have found that this error very rarely occurs in the past tense form which is not

contracted and requires two syllables to pronounce. A similar tendency with *here's*, *where's* and *how's* has also been identified (ibid.: 186).

Carter and McCarthy have noticed an increased use of *less* with plural countable nouns, in contrast to the traditional use of *fewer* (op. cit.: 103). Other systematic alterations of written grammar which regularly occur in speech can be found in ellipsis and response tokens.

2.1.3a Ellipsis

Ellipsis takes a number of forms across spoken English. Items which are frequently ellipped include pronouns, auxiliary verbs, pronouns and auxiliary verbs together, existential *there*, conditional *if*, determiners, and prepositions. The patterns of ellipsis are not haphazard: in general, speakers only omit things which, because of the shared context or an inferred understanding, do not need to be made explicit. (ibid.: 181)

A: Don't know what's gone wrong here.

B: Oh. Need any help? (ibid.: 181)

Must have been half a million people. (ibid.: 186)

You want anything else, just help yourself (ibid.: 187)

A: Where do you want this?

B: Bottom shelf, please. (ibid.: 186)

Rather than thinking of ellipsis as a form of omission, it may be more appropriate to say that 'writing and formal speech typically need to elaborate more for the sake of readers/listeners and so "add" items that might otherwise be unnecessary in everyday informal speech' (ibid.: 181).

There are also a great number of lexical chunks and response tokens into which ellipsis is already incorporated.

2.1.3b Response tokens

Good. Brilliant. Okay, well... (ibid.: 191)

This sequence of words would be quite out of place in an academic paper, but in speech,

sequences of single words like these are quite common. While these words are traditionally identified as adjectives and adverbs, this particular function of theirs in spoken grammar is as a feedback tool: evaluating, affirming or negating a previous speaker's utterance (ibid.: 189-90). When clustered together as in the example above, they can mark boundaries between topics. Response tokens are also commonly clustered and exchanged between speakers in formulaic greeting, thanking, closing and pre-closing rituals such as telephone conversations:

A: **Great. Yes.**
B: **Fabulous.**
A: **All right.** [laughs]
B: **Okay.** Thanks for that.
A: **Okay Len.**
B: **Cheers.**
A: **Bye.** (ibid.: 191)

2.2 Appropriacy monitoring features

The interpersonal aspect of speech is another area which requires a large amount of attention and mental processing from the participants. Not only is the information within the interaction being constantly negotiated, but opinions and attitudes will need to be expressed while being sensitive to the other parties involved. Some methods of performing this in spoken English include the use of vague language, hedges, the contracted form of *will* and certain constructions such as two-step questions.

2.2.1 Vague language

Vague language is not the result of carelessness or laziness, skilled speakers use it to make their messages less assertive and authoritative as well as easier to process (ibid.: 202). *Kind of, sort of, like, stuff, thing, or something, or anything, and so on* are phrases which can be used to remove precision so that listeners will not be required to devote unnecessary attention to relatively unimportant details of the interaction.

*I was down in er a place called erm, down in the Urals as well, erm Katherineburg. It's **kind of** directly east of Moscow.* (ibid.: 202)

Other times it may not be practical to list all of the possibilities in a given situation and vague language allows for other options.

A: *She doesn't like coffee.*

B: *Well, she can have an orange juice **or something**.* (ibid.: 202)

2.2.2 Hedges

Many markers of vague language are also used as hedges, another device of distancing speakers from their messages and making them sound less blunt. *Apparently, just (about), I think, kind of, like, maybe, perhaps, probably and surely* all can be used to perform this function (ibid.: 223).

*There's crocuses around the base of the trees but they're **just about gone I think**.*
(ibid.: 223)

2.2.3 Independent 'll

Carter and McCarthy have described the contracted form of will as an independent form (ibid.: 632). They state that the full form of will is used when a speaker wishes to make their intentions clear, the contracted form is less assertive and is used to relate instantaneous decisions and arrangements.

*There's a garage. **We'll** just stop and get some petrol.*

*Then **we'll** get a bus to Bridgford, the boys'**ll** get off near Asda and **we'll** all walk to the school.* (ibid.: 632)

2.2.4 Two-step questions

Another method of indirectness in speech is the two-step question. Speakers may feel their request is too general and create a context for it, or they may feel their request is too personal and so ask a yes-no question as a preface to their main inquiry (ibid.: 201)

A: ***Are you going to the match tonight?***

B: *Yeah, I am.*

A: ***Do you mind if I tag along?*** (ibid.: 201)

3 Teaching implications

The job of the teacher has been made only slightly easier by the identification and analysis of the features of spoken English grammar. A number of issues must be addressed when

discussing how to teach these features: for example, which features should be taught at particular levels and the best way to teach them. There is, however, a larger sociolinguistic concern: whether the native-speaker model is appropriate or even necessary for the millions of people who are learning English as a lingua franca.

3.1 The appropriacy of the native-speaker model

Native-like proficiency is not the goal for the majority of English learners. A large number of these learners only want enough English to allow them to communicate effectively with people from different language backgrounds. Writers such as Seidlhofer (2001) and Jenkins (2002) have argued that these learners can achieve everything they need without making use of overly complicated formulaic structures or being faithfully accurate in their production. As non-native speakers now outnumber native speakers of English, Seidlhofer proposes that they should no longer necessarily be measured against the standards of English as a native language but that English as a lingua franca should be recognised as a dialect in its own right.

Teachers should certainly be sensitive to issues of cultural superiority, but there are very pragmatic reasons for students to learn the patterns of native speech. The first is that very few students want to study an incomplete model. Students are free to use whichever forms they feel are most useful in their situation, but any 'deviant' structures should only be acceptable as an end product rather than as a starting point (Kuo, 2007). Kuo (2006) found that her students recognise their own limitations and look to the native speaker model as their target form, even in their interactions with each other. If international communication is the aim of the learner, eventually they will encounter native speakers as part of the global community. Learners will be at a severe disadvantage if they have never encountered this language group. Finally, increased fluency is often a stated goal at the start of courses so the methods devised by native speakers to achieve this should be valuable tools for learners.

3.2 When to teach spoken grammar

Any means of improving the ease of communication and increasing student production should be encouraged. Many of the fluency-enhancing features of spoken grammar will be very beneficial to low-level learners. An emphasis on phrasal chaining and simple structures will reduce the pressure of having to speak in complete sentences.

Learners should be encouraged to learn chunks, and shown how often they are used by native speakers. These phrases lighten the processing load by letting students produce more language with less effort. An error which is frequently heard in the classroom is the phrase: *Teacher, what means _____?* But with a few lexical chunks, learners can easily request the same information in an acceptable manner: *I'm sorry, 'chunks'? What does that mean?* This allows the learner to separate the processing of the form and function. The lexical chunks do not change, so they can be learned as units with the one variable word taking a slot which requires no grammatical processing. When students become more comfortable with the phrases, they will realise that *that* can simply be replaced with the unknown lexical item. This should encourage them to look for other chunks of language in everyday expressions.

Vague language is another feature of spoken English which can help lower level learners. Not knowing what to say or not wanting to be precise are common feelings for everyone, and giving students the means of managing this in English should give them some relief. This provides another means of lightening the processing load and allows speakers to choose where to focus their attention.

At intermediate and higher levels, it becomes appropriate to introduce the more affective aspects of spoken grammar such as hedges and two-step questions. These features do not directly ease the transmission of information but they do allow the speaker to sound more polite which ultimately makes communicating easier. Students at higher levels should have experience with unscripted dialogue and so be able to recognise examples of these features if they are presented in a lesson. Raising learners' awareness (discussed further in section 3.3.3) will help them see these forms in their daily encounters and eventually start using them.

3.3 How to teach spoken grammar

A number of people can learn acceptable forms of spoken English without ever setting foot in a classroom. The challenge for the teacher is how to re-create these results in an environment which is often less than fully immersive. One advantage the non-classroom learner has is the large amount of high-quality meaningful input. Teachers should attempt to include as much authentic spoken material as possible in their courses, encourage their

students to interact with it and develop their own understanding of the meanings of the forms.

3.3.1 Authentic input

Learning spoken grammar is exceptionally difficult without exposure to it in a meaningful context. But the speed of delivery, the unpredictable vocabulary, the repetition, the overlapping talk of the participants and the interpersonal relationships involved during the talk can make it very difficult for learners to understand. The key to avoiding motivation loss is making the material comprehensible, and an important tool in achieving this is by providing a transcript.

The internet provides easy access to thousands of transcribed speeches, interviews and conversations. The websites *TED.com* (TED, 2013) and *ello.org* (Beuckens, 2013) each have several hundred recordings and transcriptions, searchable by topic and, in the case of *ello.org*, by level. For specific purposes, teachers can make recordings themselves by interviewing friends or making what Thornbury (2005) describes as a semi-scripted recording: give the speakers one or two main points and let them improvise the rest. Newer coursebook series, like *English File Third Edition* (Latham-Koenig and Oxenden, 2013) and *English Unlimited* (Doff and Goldstein, 2011), include natural interviews with both native and non-native speakers.

These recordings and their transcripts are an excellent way to introduce students to the 'messiness' of spoken English embedded in a meaningful context. They can be used to help students notice the different structures that are common in the different modes of communication. The *Cambridge Grammar of English* recognises forms which are frequent in both the written and spoken corpus (including those which may be proscribed by prescriptive grammars), forms which are frequently found in one mode but not the other, and forms which are regionally or culturally specific and not very widespread (Carter and McCarthy, op. cit.: 168). By presenting spoken grammar alongside written grammar, teachers can encourage their students to see these forms as part of a system and search for other features in speech rather than simply disregarding them as 'errors'.

A further benefit of *TED.com* and coursebooks adopting new technology is the ability to use video in the classroom. The addition of non-verbal information vastly improves the

comprehension ability of students. Other than telephone conversations, there are very few real-life occasions when communication happens 'blind'. The use of gesture and facial expression reinforces salient points in the conversation and also provides another way of interpreting the dialogue if students cannot follow the verbal element. Students can analyse the relationship between the participants and their interaction with each other and then compare their findings with the words and structures used.

3.3.2 Authentic output

Students need to turn this input into meaningful output. For many learners, their main opportunity for interaction with native speakers is with their teachers. But often, even if 'teacher-talk' is avoided, there is a noticeable difference between conversations between native speakers (NS-NS) and talk between a native speaker and a non-native speaker (NS-NNS). Larsen-Freeman and Long (1991) have highlighted a number of simplifications which are common in NS-NNS speech: native speakers tend to control the discourse by using closed (yes-no) rather than open questions; lexical range is limited and intonation is exaggerated to emphasise key points; NNS errors can be overlooked by the NS and the topics of discussion are more focused on the present environment instead of more distant or difficult to describe situations and feelings.

There are certainly pragmatic reasons for these features, and so they should not be wholly discarded. But it is important to let students control the conversations as well. Larsen-Freeman and Long (ibid.) claim that NNS-NNS interaction is just as valuable as NS-NNS, because the participants will have to take responsibility for understanding the communication. The forms may not be correct, but the experience of managing conversations will help them notice things they need to improve.

3.3.3 Induction and automaticity

Noticing is one of the key stages towards the cognitivist concept of awareness (Thornbury, 2005). Attention is the first stage and consists of a state of readiness and curiosity so that a target feature may then be noticed. Noticing is the second stage and involves a focus on an item. The item may be unexpected, marked, salient or may have been previously pointed out. Understanding is the final stage of awareness and is achieved when a pattern is paired with a meaning. This usually requires a number of examples of the form in order for the learners to

integrate it into their system of knowledge. Once students are aware of a form, they can begin to manipulate it and eventually incorporate it into their own linguistic repertoire.

McCarthy and Carter (1995) describe a methodology of 'Illustration-Interaction-Induction'. 'Illustration' means the use of real data, embedded as much as possible in its original context. 'Interaction' is the means of awareness-raising and takes place through the observation and discussion of the choices made by the speakers in that context. 'Induction' is a result of learners drawing their own conclusions about the communicative functions of different lexicogrammatical features and developing an ability to notice these features as they occur in everyday interactions.

Both of these systems place a strong emphasis on the importance of authentic material and a keen analysis of the decisions made by the speakers in meaningful situations. Learners must then use these examples to generalise their own personal rules for use which can then be experimented with and evaluated to make further generalisations.

4 Conclusion

There remains a great deal to be explored in relation to the explicit teaching of the peculiarities of spoken English grammar. One of the main risks is the transference of appropriate spoken forms to writerly modes. Teachers are right to be cautious of this especially as many English examinations still remain which place a high value on the written grammar of English and formal styles of speech.

However, a number of features of spoken grammar which occur across age, gender and regional boundaries have now been clearly identified. These features have been developed and incorporated by native speakers over hundreds of years to improve the fluency of their communication and to mediate the directness of their language. Learners should be highly motivated by these goals and so be willing to invest the time and energy necessary to understand and use the necessary forms.

These forms can only be noticed if learners are exposed to authentic language and encouraged to use the language authentically themselves. Teachers must strive to provide

tools and guidance to make these experiences productive and fruitful when analysed. Learners need to accept the variation which occurs across the different modes of language so they will be more successful in their interactions with native and non-native speakers alike.

5 References

- Beuckens, T. (2013) **English Listening Lesson Library Online** [online]. <http://www.elllo.org/> [Accessed 31 May 2013]
- Biber, D., Johansson, S., Leech, G., et al. (1999) **Longman Grammar of Spoken and Written English**. Essex, England: Longman
- Brazil, D. (1995) **A Grammar of Speech**. Oxford: Oxford Univ. Press
- Carter, R. and McCarthy, M. (2006) **Cambridge Grammar of English**. Cambridge: Cambridge University Press
- Doff, A. and Goldstein, B. (2011) **English Unlimited**. Cambridge: Cambridge University Press
- Halliday, M.A.K. (1985) **An Introduction to Functional Grammar**. Suffolk: Edward Arnold
- Halliday, M.A.K. (1989) **Spoken and Written Language**. Oxford: Oxford University Press
- Jenkins, J. (2002) A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language. **Applied Linguistics**, 23 (1): 83–103
- Kuo, I.-C. (2006) Addressing the issue of teaching English as a lingua franca. **ELT Journal**, 60 (3): 213–221
- Kuo, I.-C. (2007) A response to Cem Alptekin. **ELT Journal**, 61 (3): 269–271
- Latham-Koenig, C. and Oxenden, C. (2013) **English File**. Third. Oxford: Oxford University Press
- Larsen-Freeman, D. and Long, M. H. (1991). **An introduction to second language acquisition research**. London: Longman.
- McCarthy, M. (2001) "Discourse." In Nunan, D. and Carter, R. (eds.) **The Cambridge Guide**

to Teaching English to Speakers of Other Languages. Cambridge: Cambridge University Press

McCarthy, M. and Carter, R. (1995) Spoken grammar : what is it and how can we teach it ? **ELT Journal**, 49 (July): 207–218

Mumford, S. (2008) An analysis of spoken grammar: the case for production. **ELT Journal**, 63 (2): 137–144

Seidlhofer, B. (2001) Closing a conceptual gap: The case for a description of English as a lingua franca. **International Journal of Applied Linguistics**, 11 (2): 133–158

TED. (2013) **TED: Ideas worth spreading** [online]. <http://www.ted.com/> [Accessed 31 May 2013]

Thornbury, S. (2005) **How to Teach Speaking**. Essex, England: Pearson Longman