

Given/New: What do the terms refer to?

A first (small) step

Gerard O'Grady

Cardiff University

Previous studies indicate that speakers signal the informational status of referents through a combination of intonation, word order and lexical realisation. In this paper, I argue for a non-binary view of information structure with referents being (1) hearer and discourse new, (2) discourse new but hearer given and (3) hearer and discourse given. Thus there can be no simple one-to-one relationship between information structure, lexical realisation and accenting. In the spoken data examined, evidence was found to substantiate a relationship between referential distance and lexical realisation but not between referential distance and tonic accenting. Tonic accents signal speakers' subjective projection of the importance of a referent but the exact informational meaning signalled by the referent depends on a combination of tonic accent, tone choice, key, linear position and lexical realisation.

Keywords: communicative dynamism, information structure, intonation, referential distance, word order

1. Introduction¹

In this study, I examine a corpus of recorded English speech in order to unpack the relationship between the opaque terms 'Given' and 'New' and their linguistic realisations. Prince (1981) argued that the meanings of terms such as old/given/known,

1. I would like to thank the participants at the *New Theoretical Perspectives above, below and across the Clause* roundtable held at Cardiff University on September 4 and 5 2014, and the attendees at the *Communicative Dynamism above, below and at clause level: An SFL approach* colloquium held as part of the 42nd ISFC at RWTH Aachen University for their comments, suggestions and questions. Thanks are also due to *Language Learning* (Grant reference number: 506536) for funding the Cardiff roundtable. Finally I would like to thank two anonymous reviewers and Alison Wray for their insightful comments and criticisms of this article. Any remaining infelicities are mine alone.

new/presupposition/shared knowledge/common knowledge are nebulous. Information can be: (a) given because it is recoverable/predictable from the context, (b) salient because it is potentially accessible or (c) presupposed as common knowledge. While a freshly introduced item is accessible and hence part of the common ground, a presupposed item is not necessarily salient, and a salient item is not necessarily recoverable.

In order to achieve my aim of unpacking the linguistic realisations of Given/New, I first provide a critical overview of previous work, specifically focusing in Section 3 on the prosodic systems used to signal informational status. While it has been generally believed that accenting signals New information and deaccenting signals Given information, studies such as Cruttenden (2006) and Baumann & Grice (2006) indicate that the picture is much more complicated. In Section 4, I introduce the spoken corpus and examine how the co-occurrence of prosody, referential indefinite and type identifiable lexical items (Gundel et al. 1993), and referential distance (Givón 1983) contribute to the projection of the informational status of a referent as either Given or New. Finally in Section 5 I summarise the findings and indicate some of the remaining questions.

2. Given and New: A review of the concepts

In this section, I review studies which have demonstrated the importance of word order and lexical realisation to the signalling of information structure. While I will mention prosody in passing, detailed discussion of the role of prosody is deferred until the following section.

2.1 Linear ordering into Theme and Rheme

One of the first scholars to give priority to the structuring of text was the 19th century philologist Henri Weil, who, in his comparison of ancient and modern languages, said that speakers' word order mirrored their order of ideas. Sentences contain a point of departure equally present to speaker and hearer and culminate with a goal (Weil 1887:29). The opening word(s) of a sentence serve as a link between what has been said and what is to be said, and between what is known and not known. Firbas (1974:12) reported that Weil was hugely influential in Mathesius' formation of the notion of functional sentence perspective (FSP). Mathesius (1975:81ff) argued that sentences contained thematic elements which he defined as both the point of departure and the basis or foundation. More recently Vallduvi (1990:57) schematizes Weil's view as a structure consisting of a focus and a non-focus, with some non-focal elements potentially functioning as a link with the

prior co-text. McGregor (1997: Section 7.2.2) similarly notes that theme fulfils a linking role.

While thematic and rhematic elements are defined somewhat differently by scholars working within the FSP tradition – see Hajičová & Sgall (1987) who define thematic elements in terms of contextual boundness – the approach developed by Firbas (1992) defines thematic elements as those that have the lowest communicative dynamism (CD) or contribute the least to the development of the message. Firbas (1992) asserted that the CD of spoken elements depended on the interplay of a number of factors: the contextual factor, linear modification, the semantic factor and prosody. He argued that tonic accents occur on elements containing the highest degree of CD within the sentence. Daneš (1972: 229) disagreed and claimed that prosody is reflective of different underlying topic comment structures. He provided the following minimal pair:²

- (1) He also visited Prague (Others visited Prague and so did he)
- (2) He also visited Prague (He visited Prague and some other places)

His claim is that (1) has an underlying topic comment structure where the fact that others visited Prague is already known. Conversely (2) has the topic comment structure where *the fact that he visited other places* is already known. At most Daneš is prepared to concede that intonation adds an emotional colouring to utterances. Lambrecht (1994: 144), like Daneš, argues for a distinction in the informational patterning between ‘topic comment’ sentences and ‘event-reporting’ sentences. He claims that prosodic accenting reflects rather than determines the underlying formal difference in information status between topic comment sentences and event-reporting sentences (1994: 221).

The answers to the questions in Examples (3) and (4) below are examples of topic comment and event-reporting sentences respectively. In English the topic comment sentence is indicated by the presence of the pronoun and the tonic accent on the verb while the event-reporting sentence is indicated by tonic accenting on the subject noun phrase. The equivalent examples from Japanese and Greek illustrate that different languages employ different resources to differentiate topic comment structures from event-reporting structures. In Japanese it is the presence of the particle *wa* or the ellipsis or the subject plus *wa* which indicates a topic comment sentence while the presence of *ga* indicates an event-reporting sentence. In Greek the ellipsis of the subject noun phrase indicates a topic comment sentence while the overt realisation of the subject noun phrase indicates an event-reporting sentence.

2. Underlined syllables are tonic.

- (3) *What happened to your car? It broke down*
 (Anata no) kuruma ni nani ga okatta? (Kuruma wa)³ kowareta
 your of car to what broken? car broken
 Ti sinevi sto autokinito sas? (To autokinito mou) halase
 What happened car your? car mine broken
- (4) *What happened? My car broke down*
 Nani ga okatta? Kuruma ga kowareta
 What happened? Car broken
 Ti sinevi? To autokinito mou halase⁴
 What happened? Car mine broken

It is also, however, clear that in topic comment sentences in English, Japanese and Greek the initial lexical element may be backgrounded either by being elided or pronominalized.⁵ While I know of no corpus evidence in English detailing the proportions of topic comment sentences versus event-reporting ones, there seems little reason to suspect that Lambrecht is anything but correct to claim that topic-comment sentences are syntactically (and prosodically) unmarked (1994: 122).

Halliday, recognising the significance of clause initial position at least in English, popularised the concept of Theme in the analysis of the English clause. In so doing he expressly separated the point of departure (Theme) from the basis/foundation (Given) (see Davidse (1987)). For Halliday Theme is defined as culminating with the first element functioning in Transitivity (Halliday & Matthiessen 2014: 91). Theme in declarative mood is marked if the subject is not found within the Theme e.g.

- (5) [Floods] [you can't avoid them]
 THEME RHEME

Marked Themes are significant in that they signal that the point of departure for the clause is not what it would have been had all things been equal. The perspective being developed in the clause is approached from an unusual angle, in Example (5) the event rather than the person. Halliday & Matthiessen (2014: 120)

3. Brackets indicate ellipsis. In Japanese the *ga* and *wa* are topic and subject markers respectively with *ga* indicating a new referent while *wa* indicates a given referent. In Greek *sto* and *to* are accusative and nominative neutral case markers.

4. Greek, like English, has malleable tonic accenting but unlike English always places the tonic on the final rightmost element which is in focus (Baltazani 2003).

5. English, Japanese and Greek are examples of subject first languages. Dryer (2013) reports that 76.5% of 1377 surveyed languages are subject first ones.

claim that Theme and Given very frequently coincide. They will, however, vary if a clause has a point of departure which is not recoverable,⁶ or if a clause is not coterminous with a tone unit.

2.2 How long can a referent remain recoverable?

Prince (1992: 309) argued that once a lexical entity has been introduced into a discourse it remains recoverable until the completion of a discourse. However, it is by no means clear how she defined the beginning or end of a discourse. She stated that freshly introduced items may be unanchored, e.g. *a man*, or anchored, e.g. *a man I spoke with*. Halliday similarly suggested that once an element had been entered into a discourse that it might remain recoverable. However he is clear that speakers' tonic accents present lexical elements as if they were not recoverable irrespective of referential distance. Chafe (1994), who classes items as given if they are accessible, disagrees and argues that once a lexical item has left short-term memory it can be considered to be non-active and potentially non-recoverable. Presumably, as working memory is widely believed to hold 7 items, a referential distance of more than 7 clauses between mentions would entail that the item is no longer given (1994: 183). Yet, this is not entirely clear. Chafe's (1994: 6) one idea hypothesis suggests that the tone unit would be a better measure of referential distance. Svoboda (1981: 89) studied an Old English text and concluded that elements typically remained retrievable for a span of about 7 clauses. Firbas (1995: 40) examined a corpus of modern English literary texts and similarly argued that while irretrievability is inherently gradable, referents typically remain retrievable for seven sentences. If this were to be the case more generally then we could assume that recoverability in a text would tend to conflate with accessibility. We could further presume that referential distance would correlate with activation cost. The greater the referential distance the higher the activation cost for the hearer.

Geluykens (1989) notes that factors such as a change of topic or speaker, referential distance and dysfluency impact upon whether interlocutors treat a lexical item as recoverable or non-recoverable. Givón (1990) suggests that more memorable information may be more efficiently memorized and more easily retrieved. He proposes an upper limit of 20 clauses as the referential distance beyond which a referent is no longer recoverable from the context (1983: 13). I will examine the issue of referential distance in Section 4.

6. This is likely to be signalled by marked Theme and/or high key, e.g., see Example 10 this article.

2.3 Given/New: Discrete or gradient?

Prince (1981) classified salient references as being those which, while not previously mentioned, are inferable by hearers with appropriate cultural knowledge. For instance, one of the speakers in my small recorded corpus said:

- (6) with their Christmas tree and ... their um like presents literally just floating

The referent *their Christmas tree* activates a cultural schema of gift giving. While the referent *presents* is freshly introduced it is very much inferable: presents are after all what one would expect to find under a Christmas tree in England. Chafe (1987:29) includes previously active (deactivated) referents as those which belong to a cognitive frame introduced by the discourse. He distinguishes between what is inaccessible (non-recoverable), active (recoverable) and inactive (salient) based on the degree of mental processing involved. The final type of givenness presupposition is text-external. The speaker can assume that even when a cultural frame has not been previously introduced to the discourse that the speaker can infer the existence of the referent. Consider, in this regard, Example (7):

- (7) the water just started pouring into their carpet

The lexical item *their carpet* was freshly introduced into the discourse but its existence was presupposed by hearers' familiarity with English housing. A comparison of Examples (6) and (7) indicates that the boundary between given (salience) and given (presupposition) is at best porous. In both examples, the existence of the referent is implied either by the previous context or co-text. While more mental processing is required to recover *carpets* than the primed item *presents* (see Branigan et al. (2008:75) on the potential interaction between priming and the accessibility of a referent) both referents are 'potentially recoverable' or hearer given (Prince 1992:6). Loock (2013) argues that a further subcategory 'hearer indeterminacy' is needed to complete the picture. However, I do not include this category as I follow Halliday (1967) in arguing that speakers project lexical items as Given or New irrespective of their hearer's actual knowledge.

Gundel et al. (1993) have produced a givenness hierarchy which situates lexical items along a continuum of more or less givenness. Figure 1 reinterprets their work in terms of three discrete categories, recoverable (discourse given), potentially recoverable (hearer given), and non-recoverable (new). It can be seen that there is no one-to-one relationship between lexicogrammatical realisation and information status. However, pronouns and ellipsis signal recoverable information while '*a*' plus noun phrase and '*this*' plus noun phrase (referential indefinite) signal discourse new and hearer new information. The remaining structures may signal non-recoverable or potentially recoverable information.

The givenness hierarchy does not incorporate prosody, but as New items attract tonic accents, type identifiable and referential indefinite referents should contain tonic accents. The presence of tonic accents on lexical items which are either potentially recoverable or recoverable would provide some support for the view that prosody re-evaluates rather than reflects information structure in English (see Firbas 1992). A positive relationship between greater referential distance and tonic accenting would provide support for the view that intonation is reflective rather than determinative of information structure, while a negative relationship would provide support for the opposite view.

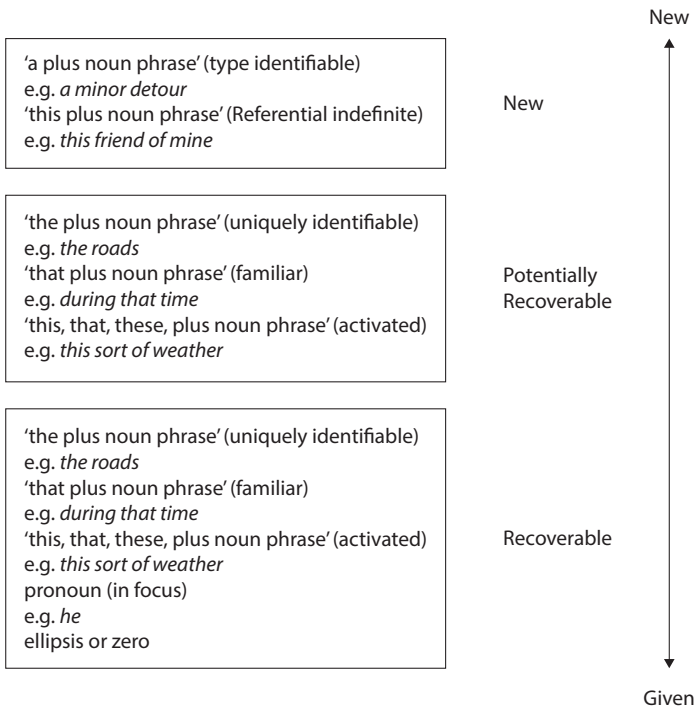


Figure 1. A reinterpretation of the givenness hierarchy

Lambrecht (1994: 109) provides a slightly more delicate schema with two primary options: identifiable (Given) and non-identifiable (New).⁷ In all there are seven

7. Lambrecht (1994) analyses information structure as a combination of the cognitive categories of knowledge and activation. But as he provides clear though defeasible grammatical and prosodic realisation rules for Given and New referents I have re-interpreted his cognitive categories as textual ones.

options which are numbered in Figure 2. New referents are either Unanchored, e.g. (1) *a woman*, or Anchored, e.g. (2) *a woman friend of Johnny's*. They are normally realised as either '*a* plus noun phrase' or as a referential indefinite '*this* plus noun phrase' and are prosodically prominent. Given referents are themselves divided into Unused (3), Potentially Recoverable (4 to 6) and Recoverable (7) elements. Unused referents are typically coded as a definite noun phrase or as a generic noun and receive some prosodic prominence. Referents may be potentially recoverable. This category consists of three subcategories, textual (4), situational (5) and inferential (6) which do not have unique structural realisations. For instance a definite noun phrase may refer to a textually prior item, an item present in the situation, or one that speakers project that hearers can infer. All 3 subcategories of potentially recoverable referents are discourse new and hearer given. Referents may be signalled as recoverable (7), i.e., discourse and hearer given, by the use of pronouns or zero and are typically unaccented.

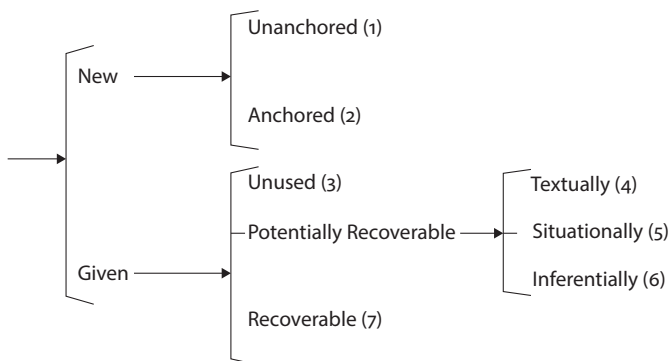


Figure 2. A schema for recoverable and non-recoverable information

Lambrecht's typology is clearly a significant advance, in that it incorporates prosody. However, it is by no means complete. For instance its discussion of the accenting of unused referents says that they receive some prosodic prominence. It is not clear if the qualification *some* signals a reference to a pre-tonic accent or to a tonic accent which is contained in the Theme. In short, more needs to be done to explore the relationship between the lexical signalling of the informational status of a referent, its prosodic realisation and whether it occurs in Theme or Rheme position.

3. Intonation and the signalling of givenness and newness

Within the Systemic Functional Linguistic tradition Halliday (1967 and 1970) and Halliday & Greaves (2008) argue that speakers' tonic placements project their

assessment of the information structure. Full lexical items when first introduced into a discourse are likely to be presented as tonic but speakers are free to project them as Given. Tonic lexical items signal the focus of the tone unit and represent the culmination of the New within the tone group. Pretonic items may be Given or New while post-tonic items are Given (Halliday & Greaves 2008:57). In many cases the beginning of the New is indicated by a pretonic prominence. For instance in (8) *it* is post-tonic and therefore Given. The pretonic prominence on *the insurance* signals the start of the New. The verbal group *wouldn't cover* is the tonic (Focal New).

- (8) | and the insurance wouldn't cover it |
 Given New Given
 ↔ ←—————→

Taglicht (1984:34) points out a problem with Halliday's treatment of information structure. In his example, presented below as (9), which is a response to the question *what happened*, the labelling of *Peter* as recoverable seems motivated only by Halliday's theoretical presupposition that a non-prosodically prominent content lexical item, in this case *phoned*, in the second tone unit must represent the boundary between the Given and the New. Despite *Peter* being unmentioned and prosodically prominent, it cannot be included as part of the New.

- (9) | John phoned Mary | and Peter phoned Jane |
 New Given New
 ←—————→ ←—————→

Lambrecht (1994:263) argues that while newly introduced referents are accented, not all accented referents are necessarily new. Pretonic accents may or may not signal that referents are new, while tonic accents signal newness or newsworthiness (Lambrecht 1994:325; see also Ladd 2008). Such views are very much in accord with Halliday's except that they remove the awkward anomaly of assuming that New information within a tone unit must not contain any intervening Given material.

To illustrate, in Figure 3 the syllable *gar* is accented in consecutive tone units.⁸ Its lexicogrammatical realisation coupled with the fact that it is freshly introduced into the discourse suggests that it is presented as discourse new and hearer new in the first tone unit. But the issue remains as to why the speaker accented it in

8. I used Praat software (Boersma & Weenick: n.d.) both to transcribe and visualize the data. Praat is open source speech analysis software available from <http://www.fon.hum.uva.nl/praat/>. H* refers to higher pitched accent, the % indicates a boundary tone with H% and L% normally indicating a final rise and fall respectively, see Gussenhoven (2004) and Ladd (2008) for a fuller description of the notation.

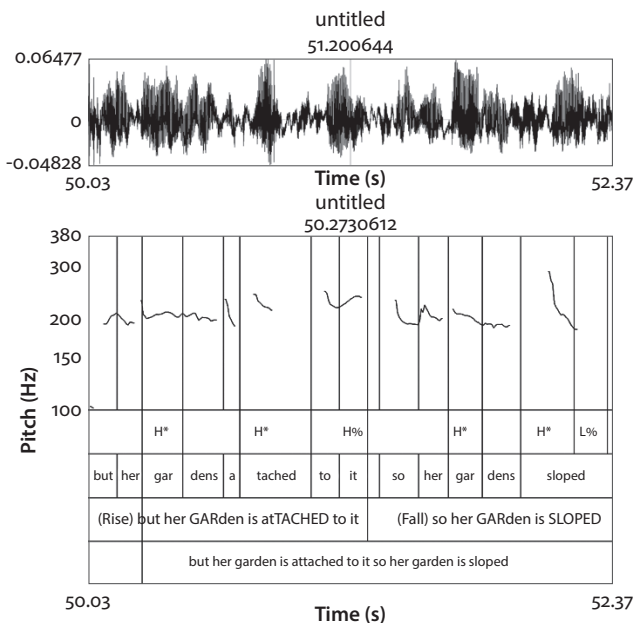


Figure 3. A spectrograph of a repeated referent

the second mention, as it is clearly recoverable from the cotext.⁹ Lambrecht's suggestion is that the second mention of the referent *her garden* is not new. Instead it may be a topic accent signalling the relation of the referent to the proposition (1994: 275).

Speakers signal their expectations of how hearers will receive the message contained in the following tone unit through their selection of Key (Brazil 1997). Key refers to the pitch height of the onset accent and is coded as high, mid or low compared with the prior onset. High key presents the following content as contrary to the previously generated expectations. Low key signals that the following content is equivalent to previously generated expectations. Mid key signals neither contrast nor equivalence. In Figure 3 the mid key on the second mention of *gar*

9. Alison Wray usefully pointed out that had the speaker used a pronoun in the second tone unit, the pronoun would have had ambiguous reference. However, the desire to disambiguate does not explain the accenting on *gar* in the second tone unit. It is possible though that this accent is a rhythmic artefact (see Kaltenböck 2009). Working in the FSP tradition Svoboda (1983) labels the second mention of 'garden' a diatheme. Firbas (1989: 48) notes that diathemes frequently co-occur with prosodic prominences which signal their status as the elements containing the highest degree of CD within the Theme.

signals that the following information is of a type that a hearer is presumed to be easily able to accommodate to their worldview. In other words the motivation for accenting may extend beyond signalling that a lexical item is in and of itself being highlighted. Miller (2006: 129) in a review of the literature claims that focused elements, while frequently new, are not necessarily so; the highlighting may instead signal a contrast, the reintroduction of an entity into the discourse, or a shift of attention to a new entity. Indeed Lehman (1977: 218) found that 23.5% of tonic accents were on items that were discourse given. So it seems that prosodic prominence neither entails newness nor non-recoverability.

Halliday (1967) claims that tone units are typically coextensive with clauses and that a tonic syllable typically occurs within the final lexical item in a tone unit. Yet, the corpus evidence in support of this claim is rather mixed. Halliday & Greaves (2008: 101) report that in around 60% of cases, ranking clauses are realised as single tone units in conversation. Croft (1995), while not disagreeing, nevertheless recognises that syntactically complex clauses which exceed the limits of short-term memory are likely to be chunked into more than one tone unit. Similarly elements such as adjuncts which are not central to clause structure tend to be chunked into a separate tone unit. Furthermore speakers tend to place parallel structures into separate tone units. Without a one to one mapping between clause and tone unit there can be no expectation of Theme and Rheme consistently conflating with Given and New respectively.

Speakers' tone choices signal how they manage information. Brazil (1997) and Gussenhoven (2004) posit that speakers produce falling tones to signal their presumption that they are adding information to the speaker/hearer common ground. Fall-rise/rising tones presume that the speakers are referring to the pre-existing speaker/hearer common ground. Cruttenden (1997) dubs end-falling tones as closed or speaker centred while end-rising tones are open and directed towards the hearers. Pierrehumbert & Hirschberg (1990) and Tench (1996) express a slightly different view. For them falling tones project major and potentially complete information while rising tones preceding falling tones signal incomplete information. Rising tones following falls modify or gloss the preceding information. Regardless, for an act of telling to occur a speaker must produce a falling tone.¹⁰ In Figure 3 the initial rise signals the first tone unit's informational dependency on the content of the following tone unit(s).

10. This statement may not be strictly true for all varieties of English such as Ulster English (see e.g. Wells & Peppé 1996) and perhaps most famously for the so-called high-rising tone or uptalk in Australian English (see e.g. Fletcher et al. 2002).

In the next section I describe the corpus which I used to investigate how the interplay of tonic accent, referential distance, lexicogrammatical realisation, word order and tone movement contributes to the presentation of information as Given and New.

4. The corpus: Description and discussion

The corpus consists of recordings made in a sound studio of eight speakers of standard southern British English. They were divided into 3 groups consisting of 2, 3, and 3 members respectively. Originally 9 participants had been recruited but 1 dropped out at the last minute. Prior to the recording I gave each participant a web link to a short video with sound but no talk illustrating scenes from the 2014 UK winter floods.¹¹ As the floods had occurred only a few months prior to the recording, I anticipated that the memory of the floods would be fresh enough to provoke the participants into constructing personalized and organized narratives. I chose a video without talk to enable the participants to articulate their experience/reaction to the floods without being primed by the talk on the video. Finally just before the recording began I played the video in the recording studio to each group. This was done to enable the groups to share the experience, so that each participant could know how much context they shared within their cohort. Once the video was finished I asked the participants to describe what they had seen, tell their reaction to it and relay any personal experiences they had. They were asked to speak in a self-arranged order for 2 minutes and once all group members had spoken they were asked to converse together in order to reflect on what they had heard. Table 1, with the participants' names changed, details the results.

Using Praat software I transcribed the corpus into tone units with boundaries signalled by |. Accented syllables were underlined with tone movement signalled as: Fall = \ Fall-Rise = \ / Rise = / Rise-Fall = / \ and Level = -. High and low key were transcribed respectively as H and L. Once I had identified all the tonic syllables I made a list of all first mentions of noun phrases containing tonic syllables on the head and classified them based on their lexicogrammatical realisations as new, potentially recoverable or recoverable. In so doing I treated each individual recording as a separate speech event. This had the advantage of allowing me to easily calculate referential distance within recordings but also meant that previous mentions of a referent within the cohort had to be discounted. The results are presented in Table 2. The numbers in brackets indicate the total number of times that

11. <http://www.youtube.com/watch?v=EjxgnpVNjJQ> (Last accessed on 22 March 2015).

Table 1. Description of the corpus

	Number of tone units	Number of seconds
Group 1		
Ann	77	149.9
Jim	78	137.9
Ann/Jim	74	135.2
Group 2		
Mary	76	142.5
Kate	132	201.8
Jane	101	158.8
Mary/Kate/Jane	206	317.9
Group 3		
Rosa	109	181
Phoebe	108	168.5
Minnie	134	207.4
Rosa/Phoebe/Minnie	97	152.9
Total	1192	1953.8

an individual speaker produced a referent in the monologue and conversation. For instance, Ann produced 21 tonic first mentions of referents of which 12 were in the monologue and 9 in the conversation.

Table 2. The number of New, potentially recoverable (Pr) and Given referents

	New	Pr	Given
Ann	12 (21)	7 (7)	0 (1)
Jim	13 (14)	7 (11)	0
Ann/Jim	10	4	1
Mary	6 (12)	3 (4)	0
Kate	20 (35)	9 (11)	3
Jane	19 (33)	13 (16)	0
Mary/Kate/Jane	35	6	0
Rosa	12 (12)	10 (10)	1
Phoebe	13 (20)	5 (6)	0 (2)
Minnie	25 (29)	15 (22)	1 (3)
Rosa/Phoebe/Minnie	11	8	4
Total	176	87	10

Unsurprisingly within the corpus all the speakers more frequently made tonic the first mentions of referents which were signalled by the lexicogrammar as new than they did referents signalled as either potentially recoverable or given. Yet as Table 2 indicates, around a third of all freshly introduced noun phrases were not lexically signalled as new. The lexical realisation of first mentions of referents does not necessarily entail newness. Indeed on 10 occasions the speakers made tonic referents which were marked by the lexicogrammar as discourse given. By so doing they presented the given referent as focal New. Table 3 lists all examples by speaker, details the tone movement and indicates whether the referent was Thematic or Rhematic.

Table 3. Tonic lexically given items listed by speaker and position

Speaker	Example	Theme/Rheme	Tone
Ann	even if it's ...yu ... in this \case it was you know	Theme	Fall
Kate	where \I was from	Theme	Fall
Kate	where \we were	Rheme	Fall-Rise
Kate	like ten \minutes from /me	Rheme	Rise
Rosa	um I think \we had	Theme	Fall
Minnie	and \having my sister come pick \me up	Theme	Fall
Phoebe	while /we're	Theme	Rise
Phoebe	while we're going through \this	Rheme	Fall
Phoebe	on at L /you	Rheme	Rise
Phoebe	and uh \we	Theme	Fall

As can be seen, six of the given referents were Thematic with five containing falling tone. By making the given/Thematic referent focal New the speaker draws attention to its status not only as the focus of an independent tone unit but also as the starting point for what is to come. One of Kate's examples illustrates this:

- (10) | where H\I was from | it was affected by a lot of the /flooding |

In (10) the anaphoric unaccented *it* in the second tone unit is co-referential with the content expressed in the first tone unit. The presence of the falling tone coupled with the high key, which signalled a topic shift, forces the hearer's attention onto the speaker and her predicament. Had Kate produced the utterance *where I was from was affected by a lot of the flooding* as a single tone unit she would have produced a slightly different meaning. The Theme would have been *where I was from* and a pre-tonic accent on *I* would have presented the pronoun referring to herself as highlighted but not focal. The following tone unit which largely consists

of Rheme is backgrounded by the presence of the rising tone. Of the four examples of a tonic given item in the Rheme, three do not have falling tone, e.g. (11):

- (11) |but for H\longer |because of the \time |it takes for that \car | and all the water is /spraying |on at /you

Informationally the presence of the falling tones in the first three tone units, a single clause, indicates that the speaker, Phoebe, presented these three tone units as information which expanded the speaker/hearer knowledge. This is highly marked as the first two tone units consist solely of Thematic matter. Her tonic marking of the indefinite pronoun *you* as focal New coupled with the rising tone signals that it is information which is already part of the shared knowledge. The Rhematic is in other words backgrounded. To conclude, there is some evidence that tonic accents in the Theme are likely to be accompanied by falling tone, while tonic accents in the Rheme are likely to be accompanied by a non-falling tone. This has the effect of respectively foregrounding the Thematic tonics as focal New while backgrounding the Rhematic tonics as reminders of pre-existing shared knowledge. The presence of the tonic on a lexically signalled given referent presents it as recoverable and focal New. Table 4 details the 176 tonic new referents by position within the clause and accompanying tone.

Table 4. Tone and Theme/Rheme choices of all tonic New items

Speaker	Theme					Rheme				
	\	/	∨	∧	–	\	/	∨	∧	–
Ann	0	0	1	0	0	5	11	3	0	0
Jim	2	0	0	0	0	8	3	2	0	0
Mary	0	2	1	0	0	3	3	3	0	0
Kate	3	1	3	0	1	13	10	2	2	0
Jane	2	0	0	0	0	19	5	5	2	0
Rosa	3	0	0	0	0	3	4	2	0	0
Phoebe	2	2	0	0	0	4	8	4	0	0
Minnie	4	0	2	1	0	15	2	4	0	0
Total	16	5	7	1	1	70	46	25	4	0

Most but not all first mentions of new referents are Rhematic, though around 18% occurred in the Theme. Examples (12) to (14) detail some of the possible meanings. In (12) Minnie introduces a referent *this guy in his car* into the discourse.

- (12) | and this H\guy | in his H ∨car in the middle | and he H \got out |

Both the noun phrase *this guy* and the prepositional phrase *in his car* are focal and contained within the Theme. The three high keys present the following tone units as contrary to expectations. *This guy in his car* is both newsworthy and the starting point of Minnie's message. In the corpus as a whole, out of the 30 instances of first mentions of lexically signalled new referents found within the Theme, 8 co-occurred with high key while 3 occurred with low key, e.g. (13):

(13) | ɪ\fancy houses | just like ...(.36) \marooned |

The noun phrase *fancy houses* is simultaneously presented as lexically new, focal and Thematic. However, the presence of the low key projects that the proposition expresses a meaning equivalent to the pre-existing shared understanding. In other words while the referent is newly introduced to the text, the referent itself is one that was hearer given in the sense that it was available to the hearer prior to first mention. Most lexically signalled new referents in Theme position occurred with mid key, e.g. (14):

(14) | pre in\dustrial men just kinda going okay |

Jim introduces the referent in Theme position, the starting point of his proposition, and makes it focal. The referent describes the advantages of a nomadic lifestyle in avoiding flooding. The mid key presents the proposition as neither contrary to nor equivalent with prior expectations.

The vast majority of first mentions of lexically signalled new referents were in Rheme position, e.g. (15) and (16).

(15) | and then we found like a World War two \bomb underneath it |

(16) | and my granddad lives in like a little \village |

The lexically signalled new referents *a world war two bomb* and *a little village* are focal and presented by Jane and Rosa respectively as newsworthy. In (15) it is what was found that matters, not where it was found, and in (16) the type of habitation her grandfather lives in is the newsworthy element. To conclude the discussion of tonic lexically signalled new referents, unless they correspond with low key, they are focal and signal newsworthiness. The function of what is signalled by newsworthiness depends on whether or not the focal referent is contained within the Theme or Rheme, the key of the tone unit and the tone movement.

The discussion to date may have given the erroneous impression that first mentions of lexically signalled new referents are automatically made tonic but this is in fact not the case. Analysis of the corpus revealed that 20 (10%) of the lexically signalled new referents were not the site of tonic accents.

Table 5. Number of New Lexical items not tonic

	Pre-tonic prominence	Not accented
Ann	3	0
Jim	5	0
Mary	1	0
Kate	2	0
Jane	1	0
Rosa	1	0
Phoebe	3	1
Minnie	4	0
Total	19	1

Example (17) is an apparent example of the first mention of a lexically signalled new referent being deaccented. However, this may be an artefact of my decision to consider referents as first mentions only within individual discourses and not within cohorts.

- (17) | I think you are /right | about people – being | sort of \not | really thinking
about what they're /doing |

Phoebe introduced the referent *people* at the opening of the conversation. However, in her earlier monologue she had previously mentioned the referent while her interlocutors Rosa and Minnie had produced the referent on 4 occasions. Thus, Phoebe's deaccenting of *people* may have been motivated by its previous mentions.

Ann's utterance, (18), is a more typical example of first mentions of lexically signalled new elements. It received a pretonic accent, was in Theme position and the site of a key selection. In other words, it functions in a manner analogous to Example 12, except that it is informationally backgrounded by not being the focus of an independent tone unit:

- (18) | and you'd see Hkids like /kayaking |

The final issue to examine is the interaction of referential distance, the lexical form of following mentions and their prosodic realisation. Of the 176 first mentions of the lexically signalled new referents 26 (13%) were re-introduced into the same discourse as full lexical items,¹² see Table 6 for details. The underlining of a

12. This figure under-reports the amount of repetition of referents. Out of a total of 253 noun phrases 71 (28%) were repeated with either tonic or pre-tonic accents.

speaker's name in the conversation indicates that the speaker produced the referent in the dialogue. To illustrate, the referent *people* was spoken 15 times, of which 5 mentions were tonic, 4 accented and 6 unaccented. Its first mention was tonic and it was produced by all three speakers.

Table 6. Repeated mention of lexically signalled items and prosodic prominence

NAME	REFERENT	# MENTIONS	# TONICS	# ACCENT	1ST MENTION
Jim	disaster	3	1	1	Accented
<u>Ann</u> /Jim	climate change	2	1	1	Accented
<u>Ann</u> /Jim	flooding	5	3	0	Tonic
Kate	friend	3	2	0	Tonic
Kate	power (cut)	5	4	1	Accented
Kate	christmas	5	3	1	Tonic
Jane	town	3	1	1	Accented
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	people	15	5	4	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	river	10	4	2	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	flood defence	2	2	0	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	Debenhams	2	1	0	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	London	2	0	0	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	friend	2	1	1	Tonic
<u>Mary</u> / <u>Kate</u> / <u>Jane</u>	house	3	2	0	Tonic
Rosa	weather	2	2	0	Tonic
Phoebe	Wiltshire	2	2	0	Tonic
Phoebe	flooding	2	2	0	Tonic
Minnie	Portland	3	1	0	Tonic
Minnie	road	7	4	1	Tonic
Minnie	village	2	1	1	Tonic
Minnie	car	3	3	0	Tonic
Minnie	mum	2	1	1	Tonic

As expected, the majority of first mentions were tonic but the relationship between a first mention and tonic accent is not automatic. There were 4 occasions when speakers' first mention of a referent is accompanied by a pre-tonic accent, e.g. (19):

(19) | they were talking about climate change and flooding |

Ann's first mention of the referent *climate change* was produced with a pre-tonic accent. The coordinate referent *flooding* is the focus or most newsworthy item in

the tone unit. Climate change, by contrast, is non-focal and so less newsworthy. In Prague school terms it is not the element with the highest degree of CD.¹³ To conclude, it seems that the typical pattern is that first mentions of referents are tonic and subsequent mentions may be tonic, accented or unaccented. Inductively it would seem that the smaller the referential distance the more likely that the second mention will be deaccented while the larger the referential distance the more likely it will be that the second mention will be accented. In order to calculate whether this was in fact the case I measured the distance in tone units between any two mentions of a referent. However, this did not prove to be straightforward as a number of tone units consisted solely of filler material such as filled pauses, conventions such as *yeah* and false starts/repetitions. As these stretches of speech did not progress the speakers' message I discounted them from the calculation of referential distance. Table 7 summarises the results.

Table 7. Referential distance and accenting of subsequent mention

	Number	Longest	Shortest	Average
Tonic ... Tonic	16	70	1	12.3
Tonic ... Accent	8	30	1	11
Tonic ... Deaccent	13	53	1	17.2
Accent ... Tonic	6	44	7	21.7
Accent ... Accent	0	N/A	N/A	N/A
Accent ... Deaccent	5	44	1	13.2
Deaccent ... Tonic	6	42	3	15.2
Deaccent ... Accent	5	23	2	8.4
Deaccent ... Deaccent	4	20	5	13.5

In the corpus there were 51 instances where a tonic mention was followed immediately by a further tonic mention. The longest referential distance between consecutive tonic mentions was 135 tone units and the shortest 1 tone unit. The average referential distance between consecutive tonics was 24.6 tone units with a standard deviation of 29. It is clear that referential distance in the examples studied did not correspond with the presence or absence of prosodic prominence. While the small size of the corpus necessitates caution, this suggests that referential distance

13. Jim had mentioned the referent *climate change* in his monologue so it could be argued that Ann's mention in the dialogue is not a first mention. However, even if this were the case, Kate's mention of *power cuts* was the first such in her group.

is more likely to correspond with the speaker's decision to foreground or background an individual referent as newsworthy, e.g. (20):

- (20) 1 | a few flooded \roads and things | **Newsworthy** ... (13 Tone units)
 2 | we had to then drive through \country roads | **Focus on type of road and not the generic category** ... (26 tone units) 3 | blockade of \water on the road | **Deaccented, focus on water, lexically signalled as recoverable** ... (10 tone units). 4 | and all of the roads around the \town | **Accented, focus on the town, lexically signalled as recoverable** ... (19 tone units)
 5 | one of the roads | **Tonic, focus on the roads, lexically signalled as recoverable** ... (4 tone units) 6 | They opened up the - roads | **Tonic, focus on the roads, lexically signalled as recoverable** ... (1 tone unit) 7 | H\ one of the roads | **Deaccented, focus on quantity, lexically signalled as recoverable**.

Only the first mention of *roads* is realised as lexically new. All the following referents are presented lexically as recoverable. This suggests that referential distance between the mentions was not large enough for the speakers to feel the need to reintroduce the referent. In order to check whether this may in fact be the case I examined the lexical realisation of repeated tonic referents where the referential distance was 20 tone units or more. Table 8 details the results.

Table 8. Referential distance and lexicogrammatical form

Referent	Earlier mention	Later mention	Referential distance
Weather	New	New	70
Village	New	PR	30
Flooding	New	PR	30
Friend	New	PR	39
House	PR	New	53
Power cut	PR	New	44
People	New	New	44
Disaster	PR	New	42
Video	New	New	135
Water	PR	New	31
River	New	New	91

There appears to be a relationship between the length of referential distance and the lexical status of the following referent. With one exception, when the referential distance was larger than 31 tone units its subsequent mention was lexically

signalled as New. Though this apparent correlation may be no more than an artefact of the sample size, it suggests, in line with the Givenness hierarchy, that referential distance is more likely to correlate with lexicogrammatical realisation than prosodic prominence. Full lexicogrammatical realisation may be necessary to disambiguate an individual referent with a distant earlier mention from intervening referents. Tonic accents conversely appear to signal speakers' subjective projection of the importance of a referent to the interpretation of their individual goal irrespective of whether or not the referent is ambiguous.

5. Conclusion

To summarise, the study of the corpus provides some support for the informational hierarchies proposed by Gundel et al. (1993) and Lambrecht (1994). The distance between two mentions of a referent corresponded with the lexical realisation of the second mention. While this study measured referential distance in terms of tone units and not clauses, it is broadly supportive of Givón's (1983) claim that a referential distance of 20 clauses or more entails that the second mention of a referent cannot be anything other than new. However, it is important to state that the examination of second mentions was restricted to cases where either the earlier or later referent was lexically signalled as new, so further investigations are needed in order to provide more definite evidence for the link between referential distance and the lexical realisation of the following referents. Furthermore this study has focused solely on referents (noun phrases) and has not examined the information structuring role of repeated mentions of other content lexical items.

Tonic accents did not correspond with referential distance (see Table 7). While freshly introduced referents tended to receive tonic accents, the presence of a freshly introduced referent did not entail the presence of a tonic accent. This provides some support for Halliday's (1967) view that speakers make tonic the items they signal as the most newsworthy regardless of whether the item is recoverable or non-recoverable. The lack of a positive relationship between larger referential distance and the presence of tonic accents offers some support for Firbas' (1992) claim that in spoken language tonic accenting re-evaluates rather than reflects information structure, and is necessarily *contra* Daneš (1972). In FSP terms the prosodic factor may override the contextual factor, signalled by the lexical realisation of an item, in a clause. Yet tonic accents do not necessarily entail that all focal elements carry the highest CD. This is because speakers may realise tone units as other than a single clause (see Croft 1995),

and the presence of low key signals that the referent, while focal, is not in and of itself newsworthy.

To conclude, tonic accents represent the speaker's projected assessment of whether or not the focal items are newsworthy in the context in which they were expressed. Tonic accents, while typically signalling freshly introduced referents, do not always do so. The recoverability of a referent is instead determined by its lexical realisation and falls broadly into three categories: (1) discourse new and hearer new, (2) discourse new but hearer given/inferable, discourse distant and hearer given/inferable (3) discourse and hearer given. More work is needed to (1) disambiguate the status of the two subcategories of potentially recoverable items, (2) examine the possible informational structuring effect of pretonic accents and (3) the role of linear modification in signalling the most newsworthy (high CD) item in clauses and clause complexes.

References

- Baltazani, Mary. 2003. Broad focus across sentence types in Greek. Paper presented at the *8th European Conference on Speech Communication and Technology (EUROSPEECH 2003 – INTERSPEECH 2003)*, Geneva, Switzerland, 1–4 September 2003. (http://www.isca-speech.org/archive/eurospeech_2003) (Last accessed on 28 December 2015).
- Baumann, Stefan & Martine Grice. 2006. The intonation of accessibility. *Journal of Pragmatics* 38: 1636–1657. doi:10.1016/j.pragma.2005.03.017
- Boersma, Paul & David Weenick. n.d. *Praat doing Phonetics by Computer*. Computer programme Version 5.3.52.
- Branigan, Holly P., Martin J. Pickering & Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118: 172–189. doi:10.1016/j.lingua.2007.02.003
- Brazil, David. 1997. *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.
- Chafe, Wallace. 1987. Cognitive constraints on information flow. In *Coherence and Grounding in Discourse*, Russell Tomlin (ed.). Amsterdam and Philadelphia: John Benjamins, 21–51. doi:10.1075/tsl.11.03cha
- Chafe, Wallace. 1994. *Discourse, Consciousness and Time*. Chicago: The University of Chicago Press.
- Croft, William. S. 1995. Intonation units and grammatical structure. *Linguistics* 71: 490–532. doi:10.1515/ling.1995.33.5.839
- Cruttenden, Alan. 1997. *Intonation*, 2nd edition. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166973
- Cruttenden, Alan. 2006. The de-accenting of old information: A cognitive universal? In *Pragmatic Organisation in the Languages of Europe*, Giuliano Bernini & Marcia L. Schwartz (eds). Berlin: Walter de Gruyter, 311–358. doi:10.1515/9783110892222.311
- Daneš, František. 1972. Order of elements and sentence intonation. In *Intonation*, Dwight Bolinger (ed.). London: Penguin, 216–232.

- Davidse, Kristin. 1987. M. A. K. Halliday's Functional Grammar and the Prague School. In *Functionalism in Linguistics*, René Dirven & Vilém Fried. (eds). Amsterdam and Philadelphia: John Benjamins, 39–79. doi:10.1075/llsee.20.04dav
- Dryer, Matthew S. 2013. Order of subject, object and verb. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info/chapter/81>) (Last accessed on 9 November 2015).
- Firbas, Jan. 1974. Some aspects of the Czechoslovak approach to the problems of functional sentence perspective. In *Papers on FSP*, František Daneš (ed.). Prague: Academia Press, 11–37. doi:10.1515/9783111676524.11
- Firbas, Jan. 1989. Degrees of communicative dynamism and degrees of prosodic prominence weight. *Brno Studies in English* 18: 21–66.
- Firbas, Jan. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511597817
- Firbas, Jan. 1995. Retrievability span in Functional Sentence Perspective. *Brno Studies in English* 21: 17–45.
- Fletcher, Janet, Lesley Stirling, Ilana Muhi & Roger Wales. 2002. Intonational rises and dialog acts in the Australian English map task. *Language and Speech* 45 (3): 226–253. doi:10.1177/00238309020450030201
- Geluykens, Ronald. 1989. Information structure in English conversation: The Given-New distinction revisited. *Occasional Papers in Systemic Linguistics* 3: 129–147.
- Givón, Talmy. 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam and Philadelphia: John Benjamins. doi:10.1075/tsl3
- Givón, Talmy. 1990. *Syntax: An Introduction*, Vol. 2. Amsterdam and Philadelphia: John Benjamins.
- Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69 (2): 274–307. doi:10.2307/416535
- Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511616983
- Hajičová, Eva & Peter Sgall. 1987. The ordering principle. *Journal of Pragmatics* 11: 435–454. doi:10.1016/0378-2166(87)90088-9
- Halliday, M. A. K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton. doi:10.1515/9783111357447
- Halliday, M. A. K. 1970. *A Course in Spoken English: Intonation*. Oxford: Oxford University Press.
- Halliday, M. A. K. & William S. Greaves. 2008. *Intonation in the Grammar of British English*. Equinox: London.
- Halliday, M. A. K. & Christian M. I. M. Matthiessen. 2014. *An Introduction to Functional Grammar*, 4th edition. Abingdon: Routledge.
- Kaltenböck, Gunther. 2009. Initial *I Think*: Main or Comment Clause. *Discourse and Interaction* 2 (1): 49–70.
- Ladd, D. Robert. 2008. *Intonation Phonology*, 2nd edition. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511808814
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620607
- Lehman, Christina. 1977. A re-analysis of Given-ness: Stress in discourse. *Proceedings of the Chicago Linguistic Society* 13: 316–324.

- Loock, Rudy. 2013. Extending further and refining Prince's taxonomy of given/new information: A case study of non-restrictive, relevance-oriented structures. *Pragmatics* 23 (1): 69–91. doi:10.1075/prag.23.1.04loo
- Mathesius, Vilém. 1975. *A Functional Analysis of Present Day English on a General Linguistic Basis*. The Hague: Mouton. doi:10.1515/9783110813296
- McGregor, William. 1997. *Semiotic Grammar*. Oxford: Clarendon Press.
- Miller, Jim. 2006. Focus in the languages of Europe. In *Pragmatic Organisation of Discourse in the Languages of Europe*, Giuliano Bernini & Marcia L. Schwartz (eds). Berlin: Walter de Gruyter, 121–124. doi:10.1515/9783110892222.121
- Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in Communication*, Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds). Cambridge: MIT Press, 271–311.
- Prince, Ellen. 1981. Toward a taxonomy of given–new information. In *Radical Pragmatics*, Peter Cole, (ed.). New York: Academic Press, 223–255.
- Prince, Ellen. 1992. Subjects, definiteness and information status. In *Discourse Description: Diverse Linguistic Analyses of a Fund-raising Text*, William. C. Mann & Sandra A. Thompson (eds). Amsterdam and Philadelphia: John Benjamins, 295–325.
- Svoboda, Aleš. 1981. Two chapters on scene. *Brno studies in English* 14: 81–92.
- Svoboda, Aleš. 1983. Thematic elements. *Brno studies in English* 32: 49–85.
- Taglicht, Joseph. 1984. *Message and Emphasis: On Focus and Scope in English*. London: Longman.
- Tench, Paul. 1996. *The Intonation Systems of English*. London: Cassell.
- Vallduvi, Enric. 1990. The informational component. Ph.D. dissertation, University of Pennsylvania. Published 1993. New York: Garland Press.
- Weil, Henri. 1887. *The Order of Words in Ancient Languages Compared with that of Modern Languages*. Translated by Charles W Super. Boston: Ginn & Co.
- Wells, Bill & Sue Peppé. 1996. Ending up in Ulster: Prosody and turn taking in English dialects. In *Prosody in Conversation*, Elisabeth Couper-Kuhlen & Margret Selting (eds). Cambridge: Cambridge University Press, 101–130. doi:10.1017/CBO9780511597862.005

Author's address

Gerard O'Grady
Centre for Language and Communication Research
Cardiff University
John Percival Building
Colum Drive
Cardiff CF10 3EU
United Kingdom

OGradyGN@cardiff.ac.uk