

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/125431/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

O'Grady, Gerard and Bartlett, Tom 2019. Linearity and Tone in the unfolding of information. *Acta Linguistica Hafniensia* 51 (2) , pp. 192-221. 10.1080/03740463.2019.1668621

Publishers page: <https://doi.org/10.1080/03740463.2019.1668621>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Linearity and Tone in the Unfolding of Information

Gerard O'Grady^a and Tom Bartlett^{a,b}

^a*Cardiff University, United Kingdom;* ^b*The University of Glasgow, United Kingdom*

ABSTRACT

In this article we show how speakers manage information flow in real time and signal their interactional expectations. Speech unfolds temporally, not only as a series of lexical items which are grouped into grammatical units, but also as a sequence of tone groups which have the potential to achieve an act of telling. While there is a general belief that speakers commence their discourse with information that is shared prior to telling information that updates the common ground, our analysis of a corpus of monologue and dialogue shows that matters are not so simple. Speakers' informational needs are balanced moment by moment within and between increments which are themselves shaped by the interlocutors' shifting apprehensions of communicative purpose and the extent of presumed shared information. In our analysis we combine (i) a speech functional analysis (Halliday and Matthiessen 2014), (ii) a description of a hierarchy of informational foci (Esser 1988) and (iii) prosody in order to develop a detailed description of how speakers manage information flow in real time. This enables us to show how speakers simultaneously balance informational flow while signalling their interactional expectations. Our conclusion is that speakers manage information flow by balancing textual, interpersonal and ideational choices.

KEYWORDS intonation, interaction, information, increment

CONTACT OGradyGN@cardiff.ac.uk

1. Introduction

It is a truism that speakers produce spoken language as a linear string of ordered elements in real time. Hearers perceive spoken language as meaningful action designed to update or reinforce the shared speaker/hearer common ground. Following Stalnaker (2002: 701) we define common ground as the presumed background information shared by the interlocutors. The manipulation of grammatical and prosodic elements and features within the spoken linear string signals the speakers' unfolding moment-by-moment assessments of the desirability/necessity of the updated common ground as well as projecting their attitude towards the updated common ground. The string itself is transmitted physically as a continuous stream of vibrating air molecules which is reinterpreted by hearers as a series of phonemic segments and suprasegments. In Hjelmslev's (1975: 136) view of the linguistic sign, phonemes and suprasegmental elements are situated on the expression plane of the sign. They are units that in themselves contain no meaning but at the same time Hjelmslev (1975: 99) recognised that the same units contract function to plerematic i.e. meaningful elementd, on the content plane of the linguistic sign. A unit such as the tone group, which we are concerned with in this article, realises a unit of information and is thus part of the lexico-grammatical resources of the language (Halliday 1967 and Halliday and Greaves 2008).

While it is true to say that the exact manner in how meaningless phonetic units are reinterpreted as meaningful linguistic units within a mind remains unknown – though see Cutler (2012) for ideas about how phonological structure is encoded and decoded into lexicogrammatical units – we assume that hearers as a matter of course unproblematically reinterpret phonetic structure as linguistic structure in numerous daily communicative events.

In this paper, we will introduce a lexicogrammatical/intonational unit, the increment, which we will argue is the basic unit of linear speech through which a speaker achieves a point of completion in terms of information. Such signalings of information transfer divide, in accordance with our communicative needs, into telling and asking increments (Brazil 1995, O'Grady 2010). After introducing our data in Section 2, we will examine increments in Section 3 below in order to investigate how speakers' prosodic and grammatical choices simultaneously manage information flow and signal their interactional expectations in real

time. This will enable us to examine how in extended monologues as well as conversations speakers balance their informational and interactional needs within and between increments. We will show that the management of information flow is an ongoing process which results from the balancing of choices on the textual, interpersonal and ideational levels. Once this is done in Section 4 we will discuss the limitations to Brazil's original model and show how augmenting it with a speech functional analysis and detailed consideration of prosodic systems leads to the emergence of a powerful model which can explicate how speakers in real time manage information flow. In Section 5 we will employ Labov's concept of A- and A-B events in order to illustrate how our combined analysis explicated how our speakers managed their information flow.

2. Data and method of analysis of data

In order to investigate how speakers manage information flow in real time by balancing choices on the textual, interpersonal and ideational levels we recruited nine participants: all speakers of standard southern British English, and divided our participants into three groups. One participant had to drop out at the last minute leaving us with one group of two. We recorded each group in a sound studio with only the group members and researchers present. Prior to the recording we sent a web link to a short YouTube video with sound but no talk illustrating scenes from the 2014 UK winter floods to the participants.¹ As the floods had occurred only a few months prior to the recording we anticipated that the memory of the floods would be fresh enough to (i) provoke the participants into constructing personalized and organized narratives and (ii) to foster discussion within the groups. We chose a video without talk to enable the participants to talk about their experience/reaction to the floods without being verbally primed. Finally, just before the recording began, we 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 with their cohort. Once the video was finished, we asked

¹ <http://www.youtube.com/watch?v=EjxgnpVNjJQ> [last accessed March 5 2019]

the participants to describe what they had seen, describe their reaction to it and relay any personal experiences they had had. They were asked to (i) speak in a self-arranged order for two minutes and (ii) when all group members had spoken they were asked to converse together in order to reflect on what they had heard.

In order to transcribe the corpus into tone groups containing prominent syllables and tones we employed a largely auditory analysis though one assisted by the use of Praat software (Boersma & Weenick 2013). All the sound files were played through Praat which allowed us to visualise both the F0 curve and the wave form. As a first step we segmented the corpus into tone groups. We agree with Barth-Weingarten (2016: 260) that it is frequently difficult, if not impossible, to identify tone group boundaries in natural talk. However, unlike her we do not feel that the solution is to abandon the notion of tone groups and instead argue that the fuzziness of the location of the boundaries is irrelevant to the discussion of intonational meaning. Work such as Brazil (1997), Greaves (2006) and O’Grady (2010) has shown that only the part of the tone group commencing on the first accented syllables (head or onset) and continuing until the tonic or nuclear syllable is relevant to the projection of information structure. Furthermore, by following the criteria set out in Tench (1996) we were able to identify the vast majority of tone group boundaries without due difficulty².

Tench (1996) states that tone group boundaries may be identified through a combination of factors (i) the presence of a junctural pause, (ii) a perceivable change in tempo, (iii) the presence of a tone movement. Once we had identified the tone groups, we marked up all the prominent syllables using the evidence from our ears supplemented with visual evidence from the wave form and the presence of turning points on the F0 curve (Ladd 2008). Then we transcribed the tone movement within each tone group as Fall = \ Fall-Rise = \ Rise = / Rise-Fall = / and Level = -. Finally we transcribed the relative height of the onset and tonic syllables

² In cases such as: | a \lot of houses got flooded | **and** it really damaged people’s \lives | it proved impossible to know definitely whether the *and* belonged to the first or second tone group. We decided to include such items in the following tone group as they are more naturally described as initial rather than final elements. In any case regardless of whether the item was initial or final it was by definition Given (Halliday 1967, Halliday & Greaves 2008).

as H (higher), M (mid) and L (lower) using both our auditory impressions and the Hertz values as recorded by Praat. Table 1 presents the results.³ The names of our participants have been changed and the numbers after their name indicate which group they were in. Table 1 illustrates that increments, the main focus of this article, themselves consist of one or more tone groups. In our data the number of tone groups per increment produced by the individual speakers found in monologues ranged from 3.5 to 8.2 with a mean of 4.7. The number of tone groups per increment in the conversations ranged from 1 to 9 with a mean of 6.9.

[Table 1 near here]⁴

Once we have explained the significance of the term increment as used here, we will return in Sections 4 and 5 to these data, more specifically to the different tone choices employed by speakers especially in increment final position, and their relevance to the interpretation by the hearers of the increments.

3. What is an Increment?

Brazil (1995) employed the term increment in order to describe a stretch of speech that is the minimum unit required to tell or ask.⁵ He wrote that in telling exchanges, the first speaker initiates and achieves their individual communicative purpose while the second speaker may

³ Readers will note that the number of tones and hence tone groups in the conversation detailed in Table 3 is greater than that reported for all speakers in Table 1. This is because in the conversations speakers produced a number of tone groups either as (i) backchannels or (ii) as failed attempts to grab the floor which fell outside of increment structure.

⁴ We detail precisely what increments are and how to identify them within our framework in Section 3.

⁵ The term increment is, as Auer (2007: 647) notes, a slippery beast but basically it represents an acknowledgment that speakers produce their language in a piecemeal fashion. In the Conversation Analytic framework the term refers to a stretch of speech which contains a prosodic boundary and is the grammatically structured extension of a turn completion unit (TCU) (Schegloff 1996: 59). In more recent interactional linguistics studies the increment is defined as a non-main clause continuation after a possible point of turn completion (Ford, Fox and Thompson 2002: 16). Auer (1996 and 2007) argues that a full typology of increments must incorporate syntax, semantics, prosody, semantics and action structure. His view is closest to the one proposed here, but differs in that we do not regard turn taking as a relevant factor.

acknowledge the telling. In asking exchanges, the first speaker initiates but the second speaker achieves the individual communicative purpose. The first speaker may then acknowledge. Example (1) and (2) with data drawn from the corpus under investigation illustrate what is meant by the term increment here.

	First Speaker	Second Speaker	First Speaker
(1) Telling	I don't like concrete either	<i>Uh no</i>	N/A
(2) Asking	Is it Venice that is sinking	Yeah yeah	(Silence)
	What is the name of the place that is sinking? ⁶	Venice	<i>Ok thanks</i>

Prior to producing the initial words in the increment the interlocutors are classed as being in an *Initial State* viz. the relevant background state of knowledge prior to the act of speaking assumed by the speaker to be shared between the interlocutors. Upon completion of the telling increment, the shared interlocutor background has been modified and Target State achieved. In asking increments target state is only achieved after the contribution of the second speaker. In other words the contributions italicized above in (1) and (2) are not part of increment structure. Target state, as will be detailed in the following sections, describes the modified state of circumstances realised by the completion of the increment and is identified by 3 criteria: (i) syntactic, (ii) intonational and (iii) contextual. The achieved Target State functions as the Initial State for a following exchange. Elements produced within an increment prior to Target State realise an Intermediate State, as illustrated in (3).

(3)	I	don't	like	concrete	either.
	Initial	Inter1	Inter 2	Inter3	Target
	State	State	State	State	State

⁶ As there are no example of Wh questions in the data, we have had to make one up.

Articulation of the first element *I* creates an expectation that a further element will need to be produced for a Target State to be achieved. The element *don't like* is in traditional terms transitive and requires an object (Quirk, Greenbaum, Leech and Svartik 1985: 53); the speaker is required to produce a further element in order to realise a grammatically adequate chain. While in this case the articulation of the following element *concrete* has the potential to realise a target state the speaker produces the element *either* and thereby updates the common ground; her interpersonal alignment with her hearer was required for her communicative purpose to be achieved⁷.

Brazil (1995) identified three criteria which all must be satisfied in order for an increment to be achieved. They are the grammatical, intonational and contextual criteria. This approach is radically different from other versions of real time syntax such as Auer (1996) and Haselow (2016). Unlike our approach Auer (1996) does not give prosody equal weight to syntax and so he, unlike us, would consider the string | *I think like where I –live / it was really L/bad as well* |, analysed as example (11) below, to represent a potential completion. We though because of the absence of a falling tone would consider the increment to be incomplete. Haselow (2016) focuses not on the speech signal, but on the cognitive mechanisms that enable speakers to produce speech in real time. Like Auer, he focuses on potential turn completion points. Our approach also differs from W. O'Grady (2005), whose psycholinguistic linear grammar⁸ argues that sentences are the output of an efficiency driven parser which resolve dependency relations in real time and are subject to biologically imposed constraints. Unlike our work, W. O'Grady (2005) restricts himself to the emergence of sentences and does not engage with how speakers balance informational flow and interactional needs across a discourse.

⁷ The speaker was female and in the following examples the pronoun *he* or *she* is used to refer to the actual gender of the speaker.

⁸ By linear we simply mean the ordering of words as manifested in the speech signal. However, when we refer to the concept of linearity we are referring to the assumption that speech usually emerges with informationally given elements preceding informationally new ones.

3.1. The grammatical criterion

Brazil (1995: 51) identified seven⁹ simple chains which lead to achievement of the grammatical criterion. As stated above, fulfilling the grammatical criterion is a necessary but not sufficient condition for the achievement of Target state and so speakers who produce a stretch of speech that contains a run through of the simple chains have potentially achieved Target State. Table 2 illustrates the simple chains required to achieve an act of telling in English. The chains described in Table 2 represent telling increments. Polar asking increments would commence with a sequence of VN elements. Following Brazil, N stands for Nominal, V for Verbal, A for Adverbial and E for Adjectival elements.

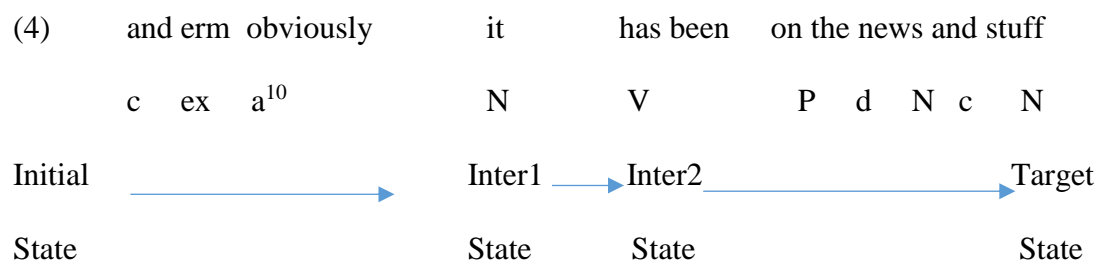
[Table 2 near here]

It can be seen that the minimal possible chain consists of between two and five obligatory elements. The articulation of the N element modifies the Initial State and creates an expectation of a following verbal element. In (i) the articulation of the intransitive V element *died* fulfils the prospected expectation and completes the grammatical chain. However, in (ii) *Susan kissed Bill* the production of the transitive V element *kissed* does not achieve a potential Target State and instead results in an Intermediate State. The potential Target State is achieved only by the articulation of the N element *Bill*. However, in (iii and iv) although the articulation of the second N element has the potential to achieve Target State, it nevertheless does not do so in either case. In (iii) the speaker's communicative need is not to tell what Susan ate but rather how she ate it. In (iv) what the speaker wants to express is not that she likes tea but at what temperature she likes her tea. Like in (ii) production of the verbal element in (v) cannot achieve Target State so that the expectation of a following – in this case E – element remains. In (vi) by contrast the V element has the potential to realise Target State, but the speaker, similar to (iii) and (iv), intends to tell how Susan danced. Finally, in (vii) while the articulation of the E

⁹ Example 11 below is an example of an 8th simple chain NVEA and hence Table 2 is not a complete list.

element results in the production of a permissible NVNE chain, the speaker's message concerned the manner of Bill's action and so the speaker produces the following A element. Hence we can see that it is impossible to discuss the adequacy of grammatical chains divorced from communicative purpose

Brazil (1995:57–68), recognising the inadequacy of simple chains such as those presented in Table 2 for describing authentic speech, proposed a number of further formal devices to augment the simple chains. The following example from our corpus illustrates the first, which is *Suspension*:



Suspension refers to the production of an element which does not alter the existing state. For instance, production of the Adverbial element *obviously* does not satisfy the speaker's communicative need, which further includes the intent to produce an N element which will result in Intermediate State 1 and prospect a further V element and so on. *Suspensive* elements have the potential to be of informational relevance as in example (4), where the a element asserts the speaker's attitude to the proposition.¹¹

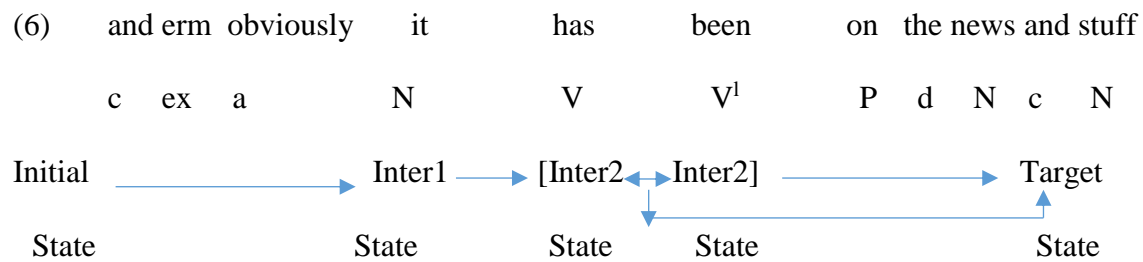
¹⁰ Suspensive elements are notated in lower case. The symbol ex stands for exclamative. Thus, in a full analysis a nominal element such as *the big light* is notated as d e N with the determiner and adjective suspending the articulation of the noun. The arrows signal that the earlier element prospects the later one or that a later element is required. In examples, such as 6, the use of double-headed arrows signals that the realisation of the later element does not fulfil the prospection.

¹¹ Had the speaker produced a suspensive Adverbial element signalling time or place such as yesterday, e.g. *yesterday it had been on the news and stuff* the suspensive a element would have signalled the angle with which the speaker intended to orient the hearer towards the unfolding proposition.

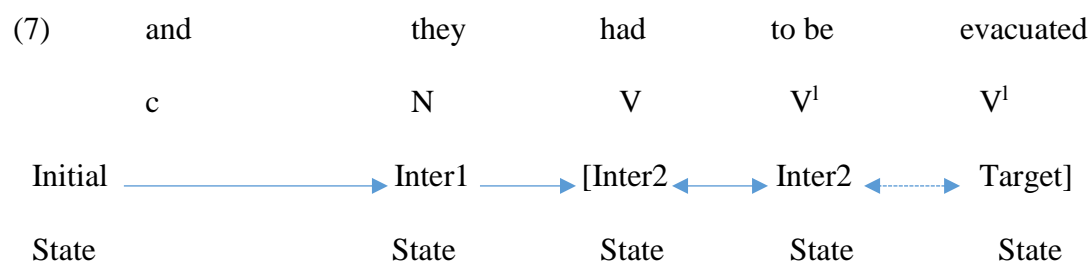
The second formal device is *Extension* which refers to cases where the speaker has exhausted all the possibilities in the chain without achieving Target State. Brazil (1995: 57) provides the following example

- (5) We wanted to search your car
 N V V¹ N

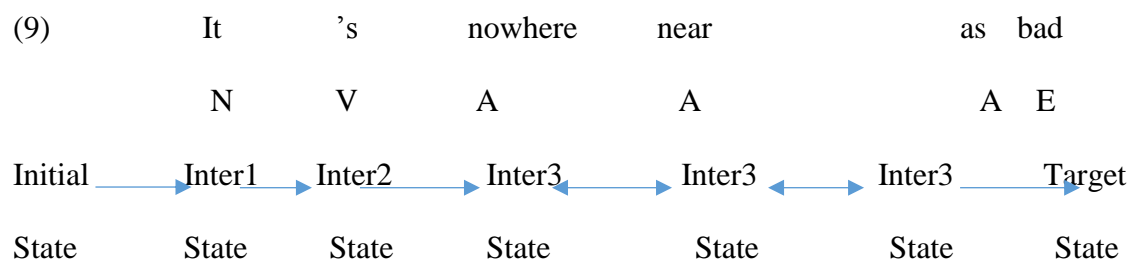
and argues that the production of the V¹ element *to search*, while in theory completing the expectations set up by the verb *want*, resets the expectations and prospects the final N element in a manner analogous to the V element in the chain *we searched your car*. The first verbal element is coded as V and the following as V¹. The same analysis applies to main verbs following auxiliaries, so that the full coding of (5) above is:



The production of the V element *has* is extended by the V¹ element *been* which does not release the speaker from the obligation to produce a following N, E or A element, in this case the following A or P/N which achieves *Target State*. As well as resetting the possibilities in the chain, the V¹ element, like any other verbal element, may on occasion achieve *Target State*.



Within the grammar there is one further formal device known as reduplication. This arises through the production of a second N or A element. In (8) the chain is extended by the production of the second N element *the floods* which itself prospects a following V element. In (9) the repeated use of A elements suspends the working of the simple NVAE chain; the expectation of the production of the final E element remains unchanged.

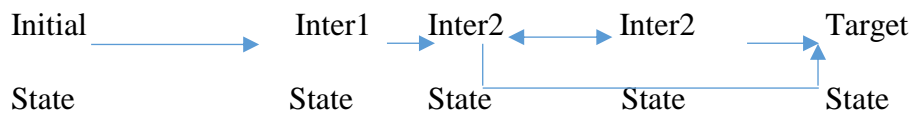


In this section we will first describe and review the intonational criteria set out in Brazil (1995) prior to suggesting some additions to the original model.

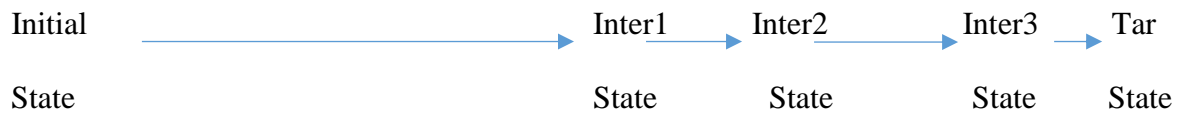
Brazil (1995) stated that no increment is complete unless it contained a falling or rise-falling tone. Falling tone movement signals completion and projects the transfer of information from

speaker to hearer (Brazil 1997, Couper-Kuhlen 1986, Cruttenden 1997, Gussenhoven 2004 and Tench 1996). A speaker could produce a falling tone (i) to coincide with the satisfaction of the grammatical criterion, (ii) after the satisfaction of the grammatical criterion or (iii) prior to the satisfaction of the grammatical criterion. The communicative significance of the misalignment of the intonational and grammatical criteria is discussed in Section 4.¹²

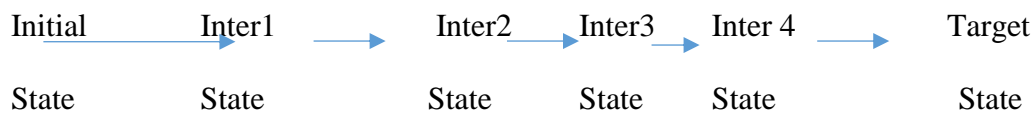
(10) | and they 're going in the \u005C\u005Cpubs |
c N V V¹ A
P d N



(11) | I think like where I L\u005C\u005Clive | it was really L/bad as well |
a a N V a E A



(12) | L\u005C\u005Cum | a \u005C\u005Cot of H\u005C\u005Cpeople | \u005C\u005Cput it \u005C\u005Cdown to | \u005C\u005Cglobal warming |
N V N A A
d N P N P e N



In (10) the falling tone is attached to the A element and signals that the speaker has in Brazil's words 'proclaimed'; the speaker projects that he has updated the hearer's world view. In (11),

¹² In Section 3.2 alone we have presented the examples with both the intonation and chains shown. We have not done so elsewhere in the paper because of space and readability issues. But all of the proposed increments presented in the paper do represent fulfilment of the grammatical criterion.

by producing an NVEA chain, the speaker has produced a successful run through of the chaining rules. However, the rising intonation can be interpreted as a signal that she has not yet updated the hearer's world view and that Target State has not yet been achieved. In (12), conversely, the speaker produces a falling tone prior to the satisfaction of the grammatical criterion. The presence of this falling tone in the second tone group ensures that the intonational criterion is satisfied: the speaker projects that her words are intended to update the hearer's world view. But the potential Target State cannot be achieved until the completion of the NVNA chain which satisfies the grammatical criterion. Hence we see, excluding for the moment the contextual criterion, that neither of the criteria is sufficient on their own to achieve a potential Target State.

3.2.2. Increments and other relevant prosodic systems

The situation is made more complex by a consideration of the comparative pitch of the initiation and termination of each increment.¹³ G. O'Grady (2010) showed that, as increments instantiate acts of information transfer from Initial to Target States, increment-initial key choices and increment-final termination choices project information about the tactic relations between increments and the anticipated hearer-response to the act of telling signalled by the achievement of Target State. An initial key – the relative intonation height of the first prominent syllable in the tone group also known as the onset – projects the speaker's expectation of how the hearer will receive the proposition realised by the ensuing Target State. A high key signals that the subsequently achieved Target State is contrastive with the expectations created by the Initial State; a mid key signals that it adds to the expectations created by the Initial State; while a low key signals that the achieved Target State is equivalent to the previous Initial State. Termination refers to the pitch height of the final tonic or nuclear syllable in the increment. A final termination choice projects the speaker's anticipation of the hearer's reaction to the achieved Target State. A high termination anticipates hearer

¹³ See Section 2 for discussion of the measuring of the relative pitch height of onset and tonic syllables.

adjudication; a mid termination expects concurrence; while a low termination projects no expectations and signals finality. See Brazil (1997: 40–66) and G. O’Grady (2010: 157–200) for a more comprehensive description of the systems of key and termination. So in summary key choices signal a speaker’s anticipation of how a hearer will respond to the updating of the common ground that is to be brought about by the achievement of the immediately following Target State. Termination choices signal a speaker’s expectation of how a hearer will respond to the updated common ground brought about by the previously realised Target State. For instance high key signals that the following content will be surprising while high termination signals that the unexpected nature of the previous content requires adjudication.

3.3. *The Contextual Factor*

Example (10), reprinted as (13) below, illustrates that the conflation of the grammatical and intonational criteria is not necessarily sufficient to achieve Target State. The example reports that people are going in pubs but did not satisfy an informational need in the context of the speaker’s narrative. It represented preliminary information. In fact the actual increment was:

(13) and they’re going in the H\pubs | still sur\living |

Target State is only achieved through the addition of a further NVaV¹ chain *still surviving*;¹⁴ which tells the hearer something about how the people have endured the storms. As they are related to the communicative purpose of speakers in context, Target States cannot be identified on purely formal criteria, either intonational or grammatical, but rather emerge as the result of negotiation between the contextually bound interlocutors. As we saw in (13), the information that *people were going into pubs* was not, in the context of a narrative about the impact of winter storms, sufficient to realise the speaker’s communicative intention. This was only

¹⁴ The NV elements (presumably *they are*) are ellipted.

realised through the addition of *still surviving*.¹⁵ In other words, (13) illustrates that, once the necessary grammatical and intonational criteria have been satisfied, the contextual factor becomes sufficient for the achievement of Target State.

3.4 What's missing?

The grammar, as described above, is capable of representing movement across a discourse from contextually bound Initial States to contextually bound Target States which themselves simultaneously represent new Initial States. The incorporation of key and termination choices, as explained in Section 3.2.2., adds to this the representation of speakers' projections of prospective and retrospective uptake. Figure 1 diagrams the unfolding of an imagined text formed out of five increments with the arrows indicating the temporal unfolding of the speech signal/.

Initial State 1 (the context shared by the interlocutors prior to the act of speaking).



Target State 1/Initial State 2 (The context shared by the interlocutors after the production of the first complete increment)



Target State 2/Initial State 3 (The context shared by the interlocutors after the production of the second complete increment)



¹⁵ In another context where the speaker had not felt that overt mention of the fortitude of the people was required, she could have realised her message simply by stating that people were continuing to go to the pub and trusted her hearer to infer that this meant that they were “still surviving”.

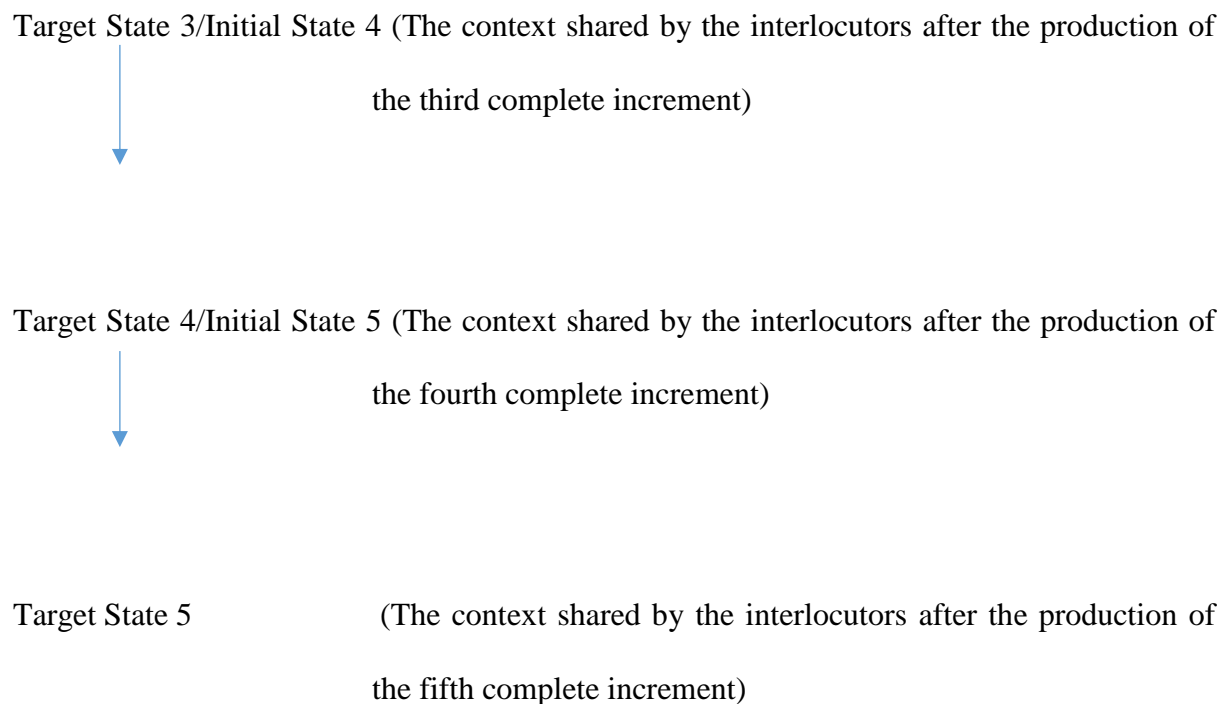


Figure 1: Increments and the unfolding of Text.

However, this description glosses over three significant points. The first is that increments are themselves formed out of sequences of tone groups which are themselves informationally significant as information units containing focal and non-focal elements. The second is that it ignores the significance of speakers' tone and key/termination choices within and between the individual increments which serve to signal an informational and tactic hierarchy within and between increments. Thirdly there is no discussion of the speech function realised by the increment.

The first point to add to Brazil's model is overt recognition that increments, as noted in Section 2, consist of tone groups whose own informational functions have to be taken into account. A tone group is a stretch of speech which contains a single major pitch movement which is anchored to a prominent syllable known as the tonic. Optionally, it may contain additional prominent syllables prior to the tonic (Cruttenden 1997, Halliday 1967, Halliday and

Greaves 2008, Tench 1996).¹⁶ There is widespread agreement that the tone group represents a single piece of information which is presented to the hearer, e.g. Chafe's (1994) idea unit, Cruttenden's (1997) presentation unit, Halliday and Greaves' (2008) information unit and what the formalist tradition calls a sense unit, (e.g. Watson & Gibson 2004). Research on speech errors has provided evidence that tone groups are handled as single behavioural acts by the central nervous system (Boomer & Laver (1968:8), Laver (1970:68), , Choe and Redford (2012: 8) and Gandour et al (2003: 155)). Within the increment the tone group is both an informationally unitary chunk and a sequence of lexical items which may prospect further lexical items. The tonic accent signals the focal element within the tone group (Halliday and Greaves 2008); its exact meaning depends, however, on a number of factors including linear position within the increment and tone choice (G. O'Grady 2016). Within a multi tone group increment, the issue of the relationship between the different foci remains unanswered and hence the description offered above does not fully map out how speakers manage information flow in real time.

The second addition to Brazil's model is complex and addresses the significance of speakers' tone and key/termination choices both within and between increments. Every tone group is the site of a tone choice. And while there is some controversy over the exact tonal inventory of English, here we adopt a five tone system: fall \, rise-fall /\, rise /, fall-rise \/ and level – (Brazil 1997, Couper-Kuhlen 1986, Cruttenden 1997, Gussenhoven 1983, Tench 1996). There is a primary distinction between tones that fall and those that rise. Figure 2 based on Brazil (1997) illustrates the main meaning contrasts coded by the different tone choices.

[Figure 2 near here]

The primary distinction visualised in the network in Figure 2 is between direct and oblique discourse. Direct discourse refers to situations where the speaker is focusing on producing a

¹⁶ Tone groups can be equated to what ToBI theorists (e.g. the chapters in Jun 2007 and 2015) label intonational phrases, see Ladd (2008).

meaningful and contextually appropriate message and is signalled by the presence of falling and rising tones. Falling tones signal a speaker expectation that the proposition encoded by the tone group is not part of the common ground, while rising tones signal that that the proposition is presumed by the speaker to be part of the common ground. A further choice is dominance/non-dominance which signals the speaker's willingness or not to assume temporary control of the situation. Oblique discourse refers to situations where the speaker signals their disengagement from the communicative situation, perhaps because of processing problems caused by searching for a word. Oblique discourse is signalled by the presence of level tone.

Esser (1988: 66–80) sketched a hierarchy of the relative importance of neighbouring tone groups in presenting the unfolding of information structure in a text. For him, tone groups containing falling tones are, all things being equal, more informationally salient than other tone groups. Tone groups with high termination¹⁷ are more intrinsically informational than those with mid, which are themselves more informationally salient than those with low termination. Example (14) illustrates this with / and \ marking tone movement and H, L and nothing termination:

(14) [H\] > [H/] > [\] > [/] > [L\] > [L/]

All high termination tone groups contain the most salient information with those containing falling tone being the most salient of all. Yet, much is missed in Esser's (1988) analysis. First of all, Esser conflates the functions of fall and rise-fall as well as rise and fall-rise. By so doing, he ignores the interactional function of tone, as we will show in Section 4. Second, he does not include key, the onset of the increment relative to the termination of the previous increment, and is therefore unable to account for how speakers signal their anticipation of how hearers

¹⁷ Esser uses the term key, though his transcriptions indicate that he is actually describing termination. In tone groups with only a single prominence key and termination are realised on the tonic syllable (Brazil 1997). G. O'Grady (2010: 157–200) shows that in such cases linear position determines which system predominates; key in increment initial position and termination in increment final position.

will respond to the updated state of affairs expressed by the immediately following Target State. Third, his sole criterion is adjacency of tone groups. This fails to take account of the fact that adjacent tone groups may belong to different semantic fields and be located within different increments. Finally, within Esser's sequence there is no place for linearity, see Firbas (1992), Fries (1995) and Halliday and Matthiessen (2014) for discussions on why the most informationally salient information is usually found towards the end of an utterance. Hence falling tones earlier in the sequence may be incorrectly classed as informationally as salient as falling tones later in the sequence.

Furthermore, Esser equates prosody solely with the presentation of information¹⁸ and ignores the multi-functional nature of prosody. Key and termination choices signal speakers' momentary apprehension of how hearers will respond to subsequent and preceding increments. Increment initial key and increment final termination selections have scope over the entire increment and do not simply project the informational importance of lexical items within individual tone groups. And, as Austin (1961) illustrated, the use of words are actions with material consequences and not merely information transferred through the aether from mind to mind.

With this in mind, we turn to the third and final point to be added to Brazil's model of increments which is consideration of its speech function (Halliday & Matthiessen 2014) i.e. the speech role it construes and the type of commodity it exchanges. The primary division in speech role is between a speaker who gives something to the hearer and a speaker who demands something from the hearer. Speakers can give/demand information or physical goods and services (Halliday and Matthiessen 2014: 135-136). Such a division leads to the following classification of speech functions.

Information	Goods and Services
Give = Statement	Give = Offer

¹⁸ Esser's data was monologic telling and hence did not require any active intervention from the hearer.

In terms of the grammar presented here, increments may map onto all four speech roles. They cannot be looked at exclusively in terms of the updating of the assumed common ground, but must also be looked at in terms of speech function which may require an overt response from the hearer. For instance, a final rising tone may not necessarily signal the informational downranking of the information contained in the final tone group, but rather it may signal the speech function of the increment. Compare:

(15a) | It is \cold in here | could you turn on the \heat |

(15b) | it is \cold in here | could you turn on the /heat |

In (15a) there are two falling tones and in Esser's hierarchy they are of equal informational weight. The increment realises a command and the hearer is projected to be in the less powerful position of having to comply. In (15b) the final rise does not seem to reduce the informational weight of the second tone group, but rather to project that the hearer, at least notionally, has the final say on whether or not to respond to the command.

To conclude, an intonational hierarchy signalled solely by tone and termination is insufficient for a number of reasons. Firstly, it is unable to capture the dynamic interplay within and between increments and their information structure. Secondly, it provides no means for exploring how speakers balance their competing informational and interactional needs. Thirdly it is unable to capture how speakers signal that they have told something which has updated the assumed common ground by achieving Target State. In order to examine more closely how speakers manage their information flow in real time by balancing the textual, interpersonal and ideational levels, we will examine in the next section the interplay of prominence, tone, key, termination, linearity, lexicogrammatical realisation and semantics in the presentation of information within the corpus.

4. Results; Dimensions of analysis and Discussion

In this section we motivate the dimensions, and reproduce the results, of our analysis of tone groups, tones, key and termination. Table 1 in section 2 above illustrates the composition of our data in terms of the speakers, the length of their monologues and conversations, and the number of increments and tone groups in these. Table 3 surveys the speakers' use of tones in their monologues and dialogues.

[Table 3 near here]

Table 3 shows that the most frequently selected tone is an end-falling movement in both the monologue (52.2%) and conversation (57.3%). Such a finding is in line with previous claims in the literature (e.g. Crystal 1969). However, it was not the most frequently selected tone for all speakers. In the monologues Phoebe, Kate and especially Mary more frequently chose a non-end falling tone. In the conversation all speakers bar Phoebe selected end-falling tone more than half the time. This, though, can largely be explained by the presence of level tones indicating some production difficulty in assembling their messages.¹⁹

The fact that end-falling is the most frequently chosen option across the corpus in both monologue and conversation means that speakers most frequently signalled that the content of their tones were assumed to expand the presumed common ground shared by the speaker and hearer (see Figure 2 in Section 3.4 above). Labov (1972: 14) classifies discourse into A- and AB-events with the former containing content known only to the speaker and the latter with content known equally to the speaker and hearer. A telling increment by definition must contain

¹⁹ For instance when Mary was describing how the images on the YouTube clip compared to her personal experiences of the flood, she produced the following stretch of 8 tone groups containing 3 level tones. The 3 dots (...) indicate non-junctural pausing and are further evidence of processing issues or in the terms we employed in Section 3, Mary was adopting an oblique orientation.

| -erm | -yeah | \and | just ... it \was quite | like \horrible weather | and it kind of does show you \that | I think like ... where I -live | it was really bad as /well |

an A-event as otherwise the speaker has not expanded the presumed common ground. More recently Heritage (2012: 32-3) has reclassified Labov's work in terms of epistemic status.²⁰ He notes that individuals' epistemic status varies from moment to moment and that speakers have a degree of freedom in how they position themselves as relatively knowing or unknowing. However, we will employ Labov's terms in the paper as we are primarily interested in examining the grammatical and prosodic traces of how speakers signal their management of common ground and not in how individual speakers position themselves as relatively knowing or unknowing.

In telling increments the speakers assume the role of the primary knower by assuming primary epistemic rights (Berry 2016 and Muntigl 2009). For instance, in (16) the hearer has no way of predicting when the speaker's friend is due to return to work and the final tone group encodes an A-event. By contrast, in (17) the speaker provides a reason for the disappearance of the coastal sand dunes in the increment final tone group. However, as the immediately prior discourse has extensively discussed winter storms and their environmental impact, this is available from the context. In (18) the final tone group contains information that is common knowledge: it is after all the purpose of barriers to protect the coastline from the sea. Thus, the final tone groups in (17) and (18) encode AB-events.

(16) like it Htook ages | to get everything back where my –friend | she supposed to \ work |

like the \ week after Christmas |

(17) | they'd like pretty much \ disappeared | just because of the really bad L/ storms |

(18) | whereas ... you could see with the \ sea | they have like \ flood | massive \ barriers |

\ yeah | that / protect that |

²⁰ Heritage's work draws upon Kamio's (1994) theory of informational territories which argues by analogy with studies of animal behaviour that if a piece of information is closer to the speaker that it belongs to the speaker's territory.

Were increments to primarily realise the textual function of signalling the expansion of the assumed common ground we could expect that there would be a tendency for falling tones to occur increment finally as, we noted in section 3.2.1, Brazil (1995) predicts. However, if as we have argued above in Section 3.4, speakers manage their information flow while balancing textual, interpersonal and ideational choices we would not expect to find such a tendency.

In order to check if there was a preference for end-falling tone – by which we mean falls and rise-falls – in increment final position, we compared the actual number of end-falling tones at increment boundaries with the expected number had there been an equal chance of any of the tones occurring in any position within the increment. Were we to find a statistical preference for end-falling tone in increment final position, this would provide some support for the notion that generally within increments new information follows given information. However, were we not to find such a preference, this would support the view presented here that increment final tone choices are motivated by the speakers' need to balance information and interaction. Table 4 presents the results with the percentages indicated in the second column showing the actual percentage of end-falling tones found in increment final position by speaker. This is compared with the predicted number of increment-final falling tones calculated on the basis that each tone had an equal chance of appearing anywhere in the increment.

[Table 4 near here]

The answer while rather mixed tends to support our hypothesis that, even in monologue, the association between increment endings and end falling tones is not automatic. The overall trend was for a slightly greater than average proportion of falling tones at increment final position, 54% versus 52%, but this was not the case for all speakers. In the monologue, Ann, Jim and Mary chose a higher proportion of falling tones at increment final position but Kate and Jane chose a lower proportion. For the other two speakers there was no real difference. In the conversation the overall trend is for a higher proportion of falling tones in increment final position. We ran a series of chi-square tests to check for significance but, as the table indicates,

bar the overall conversation data no significant results were found with a P value set at 0.05. For the overall conversation data the χ -squared = 6.7118 with a df = 1. However, the P value for the whole data set was not significant = 0.09 and hence no claims for statistical significance can be made.

Increments, as noted above, result in the achievement of Target State which itself recalibrates the context by modifying the relevant background prior to the act of subsequent speaking assumed by the speaker to be shared between the interlocutors. Such a recalibration may itself create the expectancy that the speaker will say more on a topic by the act of modifying the context or signal that the Target State achieved closes the topic. In contrast the choice of tone/termination in increment final position signals to the hearer the speaker's assumption of how they will receive the newly modified Target State. Table 6 details the tone/termination choices of all the increment endings.

[Table 5 near here]

Table 5 illustrates that all combinations of end-falling and end-rising tones and termination choices were found. However, it is clear that mid termination choices are the most frequent – occurring around 62% of the time in both monologue and conversation. Thus, in around 2 out of every 3 increments speakers neither anticipate that the achieved Target State will require hearer adjudication nor do they signal the closure of a topic. Rather they anticipate hearer concurrence with the achieved Target State. Yet, the presence of non-falling tones in increment final position illustrates that speakers are not merely passing information to their hearers. In the next section, we will examine the propositional content of the increment final tone groups in order to understand what exactly the speakers were doing on a moment by moment basis to manage their informational and interactional needs and examine how this correlates with the tonal structure of their utterances.

5. Further interpretation of results

5.1. Increments endings: tone and termination choices

Following Brazil (1995), above, in order to achieve Target State and to signal the expansion of the state of speaker/hearer shared understanding, a speaker must produce an end-falling tone. Yet, not all the propositional content is necessarily unknown to the hearers. Brazil (1997) argues that end-rising tone functions to refer to propositional content which is shared by the speaker and the hearer. If his view is correct end-rising tone in increment final position should correlate with tone groups which encode propositional information, known to both speaker and hearer while end-falling tones should correlate with propositional information known, prior to the act of speaking, to the speaker alone. Following Labov (1972: 124) we examined all increment final tone groups and classed them as either A-events (i.e. content known only to the speaker) or AB-events (i.e. content known to the speaker and hearer).²¹

Based solely on their propositional content we examined all the increment final tone groups in the context in which they were spoken and classified them as A- or AB-events. In other words, if the proposition encoded within the tone group was available in the prior context of inferable from it was classed as an AB-event. Otherwise it was classed as an A-event. Examples 19 to 32 below illustrate. As there were no differences in the distribution of tone and termination choices between the monologue and conversation we have conflated the results set out in Table 6.

[Table 6 near here]

It can clearly be seen that where the increment is completed by a tone group which realises an A-event it tends to co-occur with an end-falling tone and conversely where it realises an AB-event it tends to co-occur with an end-rising tone. Yet, 23 out of the 136 increment final tone groups with end-falling tones are found in tone groups which encode AB-events while 35 out

²¹ In the dataset there is only one example of an asking increment and as a result, the following discussion focuses only on telling increments.

of 108 tone groups with end-rising tones are found in tone groups which encode A-events. Or, to put it another way, almost one quarter of tone choices were not congruent with the informational expectations encoded in the lexicogrammar at increment endings. To make sense of our data we divided it into what we labelled congruent and non-congruent examples and examined the lexicogrammatical realisation of the focal item in the increment final tone group. By congruent we refer to increment final tone groups which contain end-falling tones and encode A-events or those which contain end-rising tones and encode AB-events. We classify all other combinations as incongruent.

Our expectation in relation to congruent A-events was that the focus, the site of the end-falling tone, would mostly be realised on freshly introduced nominal elements or on intransitive verbs occurring in final position as the lexical item that represents the culmination of information transfer across the increment. Our expectation in relation to congruent AB-events was that the focus would be realised (i) other than as a nominal or intransitive verb or that (ii) the item would itself have been previously mentioned or be available in the context.

5.2. *Congruent A-events*

As can be seen from Table 7 our prediction for congruent A-events was overwhelmingly realised with final content lexical items containing the increment final tone movement. However, it is also clear that the final focus does not necessarily correspond with a lexical item which is new to the discourse.²² Thus, even though the final tone group realises an A-event an increment is not always an accretion from recoverable lexical items to new or non-recoverable lexical items as shown in examples (19), (20) and (21).

[Table 7 near here]

²² Halliday has been consistent that focal items are not objectively new but are those which the speaker predicts as New, e.g. (Halliday 1967, Halliday & Greaves 2008, Halliday & Matthiessen 2014). Thus, we do not expect the relation between tonic accents and freshly introduced items to be 100% predictive, but we do expect that more often than not tonic items will correspond with freshly introduced items. See O’Grady (2016) for a fuller discussion.

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

(20) | there is Hstill like ^sand all over the road |

(21) | Verm | I know a few Vyears ago | muh my L /aunt | /uncle | and three of my \cousins

| \lost | all of their L\power |and L\water |for about Htwo or three H /weeks H | –erm
|because of the \flooding |

In (19), the focal element *bomb* is followed by two closed class items – a preposition and a pronoun which anaphorically refers to a previously mentioned storm damaged ice cream hut. In (20), the focus is not on the final content item *the road* which, while not overtly recoverable, is not freshly introduced but inferable from the context. The focal element *sand* is new to the discourse. The speaker in other words signals both that the focus of the tone group concerns *sand* but also that the proposition, which only she knows, does not solely discuss the existence of sand but also tells that the loose granular substance is spread all over the road. In (21), it is the whole final tone group which consists of previously introduced lexical items. The focal item *the flooding* has been previously mentioned and is in fact the very topic of the discourse. Yet the proposition contained in the increment final tone group realises an A-event as it explains the reason why the speaker's relatives lost power and water.

5.3. Congruent AB-events

Table 8 shows that some of our expectations were supported and others were not. N elements formed the majority of the increment final tonic items but the end-final tonic items more frequently occurred on items that were recoverable. Compared with A-events they were more likely to occur in non-final position.²³

²³ We conducted two chi square tests. The first checked to see if there was a significant difference between the informational status of the final tonic items in congruent A- and AB-

- (22) | \because | you know we Hbuild houses on \flood plains | and we use \concrete | –Lto
 | you’d... to create pavements and /roads | and \that kind of thing | and it doesn’t
 \saturate the water |
- (23) | so they were \like | I Hthink it \was | uhhh like at H/night | so it’s they couldn’t like
 couldn’t see out in the /garden| but they had them H/patio doors | –and | they could just
 start Hhearing this like | \water hitting | the...the Hdoors |

In (22), while the verbal element *saturate* is not recoverable from the co-text or context, the propositional knowledge that concrete, unlike a floodplain, is impermeable and will not absorb excess water is part of the common ground. Thus, although it is the final informational focus of the increment, it is contained in a tone group which expresses a proposition that does not expand the assumed common ground. In (23), the lexical item *doors* is in final position but as the tonic element is recoverable from the co-text. Little if any propositional value would have been lost had the speaker produced the following final tone group | \water hitting them |. These examples illustrate that final AB tone groups may serve the speaker’s interactional and not informational/propositional needs.

To conclude this section, we can see in the tone groups which congruently encode A- and AB- events that there is a need to bear in mind both informational prominence – the

events. The second checked to see if there was a difference in the syntactic position of the final tonic item, i.e. whether it was in final or non-final position between congruent A- and AB-events. Only the former was positive and achieved a significance value greater than $P = 0.05$. The results were that the former being χ -squared = 8.9351, $df = 1$, p -value = 0.002797 and the latter being χ -squared = 2.6845, $df = 1$, p -value = 0.1013. This suggests that there may be both significant grammatical and informational reactances in how speakers produce tone groups encoding A- and AB-events. However, to be certain we would need to confirm these tentative results by examining a far larger data set.

²⁴ One increment which ended with the discourse marked *you know* with the tonic on *know* has been excluded from the analysis. Two pronominal elements realised by *that* were coded as N elements.

recoverability of a lexical element – and propositional value signalled by the tone and termination choice. We can also see that speaker’s choices are not entirely driven by their need to expand the amount of knowledge they share with their hearer but may in fact be the product of other interactional goals. In order to further examine the relationship between prosodic choices and interaction, we examined in context all of the non-congruent examples in order to see why speakers might choose to encode an A-event but accompany it with an end-rising tone or why they might encode an AB-event but accompany it with an end-falling tone.

5.4; Non-congruent AB-events

Across the entire data set we found 23 non-congruent AB-events as set out in Table 9.

[Table 9 near here]

As can be seen in Table 10 seen the most frequent tone choice for non-congruent AB-events is a mid-falling tone. (24) illustrates.

(24) but erm | \other than that | you L-know | no one in my \family was really affected |
which ... which was \good |

An end-rising tone in final position would have invited unnecessary comment. On six occasions speakers selected a low falling tone, which projected the additional meaning of signalling the closure of a topic.

(25) | and there was loads of articles in and around about all the politicians | pointing \fingers
straightaway | Ltrying to find somebody to L\blame | <end of topic>
| and it H\kind of | it kind of Hputs the human in er us as \humans |

Here the initial high key signals a shift in topic away from the finger pointing of the politicians and towards the effect of climatic disasters on people who have themselves been preoccupied with finger pointing and what it means for humans to live divorced from nature. (26) is completed by a tone group containing a high fall.

(26) | It it just it \doesn't | allow for natural disaster recovery at H\all | \erm or to try and deplete the \amount of | Hpeople who are afected | from things like H\floods |

The high fall anticipates hearer adjudication²⁵: unlike the other non-congruent AB-events it invites a notional response: the hearer, albeit silently, is asked to make a mental note signalling their agreement or disagreement with the speaker's proposition.

5.5. *Non-congruent A-events*

We located thirty-three non-congruent A events which are detailed in Table 10. Twenty-four of them involve rising tone (examples 27–29) and 9 have fall-rising tone (examples 30–32).

[Table 10 near here]

In (27) to (29), the final rising tone signals that that speaker requires a response of some kind or other to the achieved Target State. In other words, even though the speaker produces an A-event she requires a hearer response or acknowledgement that her proposition has been received.

(27) | –and | the Hkind of scenes like my \dad would make | and maybe \put | L\online |
and like Lshare with his / L weather friends | and then he /show me²⁶ |

²⁵ Though as this example is taken from monologue there was no space for the hearer to produce an overt response.

²⁶ Despite the absence of the past tense morpheme, we have interpreted *and then he show me* as a grammatically complete structure because it is immediately followed by the

(28) | and –one of the | ideas that that was \proposed | which was quite /interesting | was um
the idea of /Hfloating cities| which \submerge | when it's really bad /Hweather | dialogue
(29)| and all the Ldirty water that's just L\ubeen | L/flooding it |

In (27), which is taken from Mary's monologue, the content expressed in the final tone group of the monologue is not known to the hearer and as such informationally, we would expect it to have been intoned with a falling tone. However, the rising tone signals to the hearers that the speaker wishes them to produce a (notional) response. By contrast, in (28), the high rise anticipates a particular kind of response: speaker adjudication. Ann asks Jim for adjudication of the feasibility of submergible floating cities. and he replies with a supportive high key back-channel 'mm'. The combination of the high termination and the rising tone signals a double response: the rise expresses that the Target State is open to speaker negotiation and the high, the fact that the speaker anticipates adjudication. Jim replies with a supportive high key back-channel 'mm' signals that perhaps contrary to Ann's expectations he is . By contrast, in (29), the low rise signals both the end of Rosa's monologic contribution and invites her hearer to produce a notional response *viz* considering the effect of the water that has been flooding people's properties. As we do not have video recordings we do not know whether her hearers responded non-verbally.

Selection of a fall-rise tone signals a contextually bound implication or as Halliday (1967) dubbed it "a but". As Table 11 shows, the fall-rise can be low, mid or high.

(30) | but \erm | yeah I think ... mostly it kind of made me glad that I'd a\uvoided | L actually
any L \flooding like that | at \all | anywhere where I \had to be |

(31) | this \one road | it was H\umore like | that was more H\specific than | the other \clips |
L\uwhere it was just | people talking about what they have \seen | and they're Hsat in
their \cars |

conjunction *but* and a further proposition which expressed her disinterest in such types of videos.

(32) | \uerm | \uand | the \uillage I \uive in is \uactually | in a \uittle /H\ualley | \uand | it \uused to be
 \umarshland /before | they L\ubuilt on it |\uerm| L\uhundreds of years ago | <end of topic>
 | so it is H\ukind of in |the ex\uact right \uplace to L\uflood | \ubadly |

In (30), selection of the increment final fall-rise generates an implication such as that the speaker was lucky that the floods did not stop her getting to the places she required. A falling tone would not have generated this implication. Similarly had she produced a rising tone she would have invited a response but not generated the implication of how fortunate she was. In (31), in addition to the implication that the action of sitting in cars watching rising water was, to put it mildly, ill advised, the high key in the final tone group creates a double incongruence in that the action of sitting in their cars is presented as contrary to expectations but as the final fall-rise signals that the tone group does not expand the state of shared speaker/hearer knowledge – the hearers are positioned as being previously aware of the stupidity of other people. In (32) we can see that the low fall-rise signals both an implication of what it means to have built hundreds of years ago on marshland and the end of the topic. The achieved Target State is immediately followed by a high key which itself signals the start of a new topic namely the effect of building on what was marshland.

6. Conclusion

Based on our novel approach we have shown that a monodimensional informational hierarchy based on adjacency, tone and termination is insufficient to show how speakers manage their information flow by balancing textual, interpersonal and ideational choices. Using a speech-functional analysis we identified final tone groups as projecting propositions which expressed A- and AB-events and using prosodic clues further subdivided these into congruent and non-congruent events. This allowed us to demonstrate that speech unfolds temporally as series of telling increments on three levels: the informational as a lexical string, the ideational as a series of tone groups and the interpersonal signalled by tone movement, key and termination. We have shown that speakers in real-time balance their informational and interactional needs and that where interaction and information are not congruent, the final tone group may not contain an expected tone movement. Lexical items are the ultimate marker of referential status but tonicity choices signal how the speaker intends the hearer to react to the presence of the referent.

The projection of textual, interpersonal and ideational information in English is a multi-layered phenomenon which operates simultaneously across and between increments, within and between tone groups and through the linear positioning of lexical items. It is signalled prosodically by key, termination, tone and tonicity choices and lexicogrammatically by the realisation of noun phrases, pronominals and ellipsis. Informational choices simultaneously function to create interactional expectations. This allows speakers to manage their hearers' expectations as they build up their message increment by increment.

Acknowledgements

Sterling work by the editors and grateful thanks to the critical comments from reviewers.

Disclosure statement

No potential conflict of interest was reported by the author

Funding

Supported by Cardiff University CURL grant AH3350E893 - RLS 1819.

Notes on contributors

Gerard O'Grady is a Reader at the Centre of Language and Communication Research at Cardiff University. He is the author and editor of 5 books including *A Grammar of Spoken English Discourse*, London: Continuum, 2010. His recent work examines differences between language as process and product and the information structure of spoken discourse.

Tom Bartlett is Reader in Functional and Applied Linguistics at Glasgow University. His research is centred around language as a social phenomenon. He specialises in Systemic Functional Linguistics (SFL) and Critical Discourse Analysis (CDA), with a particular focus on functional descriptions of Scottish Gaelic and Discourses of Sustainability.

Key of Symbols used

Intonation

\ = Falling tone / = Rising tone – = Level tone ∨ = Fall-rise tone

^ = Rise-fall tone syll = prominent syllable H = High key or high termination

L = Low key or low termination.

Grammar

N = Nominal V = Verbal V^l = Non-finite verbal A = Adverbial

E = Adjectival P = Prepositional d = determiner ex=exclamative

Lower case letters = suspension

References

- Auer, Peter. 1996. "On the Prosody and Syntax of Turn Continuations." In *Prosody in Conversation: Interactional Studies*, edited by Elizabeth Couper Kuhlen and Margret Selting, 57-100. Cambridge: Cambridge University Press.
- Auer, Peter. 2007. "Why are Increments Such Illusive Objects; an Afterthought." *Journal of Pragmatics* 17 (4): 647-658.
- Austin, John L. 1961. *How to Do Things With Words*. Oxford: Oxford University Press.
- Barth-Weingarten, Dagmar. 2016. *Intonation Units Revisited: Cesuras in Talk-in-Interaction*. Amsterdam: John Benjamins.
- Berry, Margaret. 2016. "Dynamism in Exchange Structure." *English Text Construction* 9: 33-55.
- Boersma, Paul, and David Weenick. 2013. *Praat Doing Phonetics by Computer*. Computer programme Version 5.3.52.
- Boomer, Donald, and John D.M. Laver. 1968. "Slips of the Tongue." *British Journal of Disorders of Communication* 3: 2-12.
- Brazil, David. 1995. *A Grammar of Speech*. Cambridge: Cambridge University Press.
- Brazil, David. 1997. *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.
- Chafe, Wallace. 1994. *Discourse Consciousness and Time*. Chicago: The University of Chicago Press.
- Choe, Wook Kyung & Melissa, A. Redford. 2012. The distribution of speech errors in multi-word prosodic units. *Journal for the Association of Laboratory Phonology*. Vol 3:1, 5 –26
- Couper-Kuhlen, Elizabeth. 1996. *An Introduction to English Prosody*. London: Edward Arnold.
- Croft, William. S. 1995. "Intonation Units and Grammatical Structure." *Linguistics* 71: 490-532
- Cruttenden, Alan. 1997. *Intonation*. 2nd Edition. Cambridge: Cambridge University Press.
- Cutler, Anne. 2012. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge MA. M. I. T. Press.

- Crystal, David. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Esser, Jürgen. 1988. *Comparing Reading and Spoken Intonation*. Amsterdam: Rodopi.
- Firbas, Jan. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.
- Ford, Cecilia E., Barbara A. Fox, and Sandra A. Thompson. 2002. "Constituency and the Grammar of Turn Increments." In *The Language of Turn and Sequence*, edited by Cecilia E. Ford, Barbara A. Fox and Sandra. A. Thompson, 14-38. Oxford/New York: Oxford University Press.
- Fries, Peter H. 1995. "Themes, Method of Developments and Texts." In *On Subject and Theme: A Discourse Functional Perspective*, edited by Ruqaiya Hasan and Peter Fries, 317-360. Amsterdam: John Benjamins.
- Gandour, Jack, Donald Wong, Mario Dzemidzic, Mark Lowe, Yunxia Tong, Xiaojian Li. 2003. A cross- linguistic fMRI study of perception of intonation and emotion in Chinese. *Human Brain Mapping*. Vol 18:3, 149–157.
- Greaves, William S. 2006. "Intonation in Systemic Functional Linguistics." In *Continuing Discourse on Language*. Volume 2. edited by Ruqaiya Hasan, Christian M.I.M. Matthiessen and Jonathan. J. Webster. 979-1025. London: Equinox.
- Gussenhoven, Carlos. 1983. *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Fortis
- Gussenhoven, Carlos. 2004. *The Phonology and Tone of Intonation*. Cambridge: Cambridge University Press.
- Halliday, Michael A. K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton
- Halliday, Michael A. K., and William S. Greaves. 2008. *Intonation in the Grammar of British English*. Equinox: London.
- Halliday, Michael A. K., and Christian M.I.M. Matthiessen. 2014. *An Introduction to Functional Grammar*. 4th Edition. Abingdon: Routledge.

- Haselow, Alexander. 2016. "A Processual View on Grammar: Macrogrammar and the Final Field in Spoken Syntax." In *Language Sciences* 54: 77-101..
- Heritage, John. 2012. "The Epistemic Engine; Sequence Organization and Territories of Knowledge." *Research on Language & Social Interaction* 45(1): 30–52.
- Hjelmslev, Louis N. 1975. *Resume of a Theory of Language*, edited and translated by Frances W. Whitfield. Madison WA. University of Wisconsin Press.
- Kamio, Akio. 1994. "The Theory of Information: The case of Japanese." *Journal of Pragmatics* 21: 67-100.
- Jun, Sun-Ah. 2007. ed. *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.
- Jun, Sun-Ah. 2015. ed. *Prosodic Typology II: The phonology of Intonation and Phrasing*. Oxford: Oxford University Press.
- Labov, William. 1972. "Rules for Ritual Insults." In *Studies in Social Interaction*, edited by David Sudnow. 120-170. New York: The Free Press.
- Ladd, Robert D. 2008. *Intonational Phonology*. Second Edition. Cambridge: Cambridge University Press.
- Laver, John. 1970. "The Production of Speech." In *New Horizons in Linguistics*, edited by John Lyons. 53-75. Harmondsworth: Penguin.
- Muntigl, Peter. 2009. "Knowledge Moves in Conversational Exchanges: Revisiting the Concept of Primary vs. Secondary Knowers." *Functions of Language* 16: 225-263.
- O'Grady, Gerard. 2010. *A Grammar of Spoken English Discourse: The Intonation of Increments*. London: Continuum.
- O'Grady, Gerard. 2016. "Given/New. What Do the Terms Refer to? A first (small) step." *English Text Construction*. 9-32.
- O'Grady, William. 2005. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, New Jersey: Lawrence Erlbaum.
- Quirk, Randolph., Sidney Greenbaum., Geoffrey Leech, and Jan Svartik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

- Schegloff, Emanuel. A. 1996. "Turn Organization: One Intersection of Grammar and Interaction." In: *Interaction and Grammar*, edited by Elinor Ochs, Emanuel A. Schegloff and Sandra A. Thompson. 52–133. Cambridge: Cambridge University Press.
- Stalnaker, Robert C. 2002. "Common Ground." *Linguistics and Philosophy* 25(5–6): 701–721.
- Tench, Paul. 1996. *The Intonation Systems of English*. Cassell: London.
- Watson, Duane, and Edward Gibson. 2004. "Making Sense of the Sense Unit Condition." *Linguistic Inquiry* 35(3): 508–517.

Table 1: Tone groups, increments and time by speaker in monologues and conversation.

	Monologue				Conversation			
Speaker	Time/secs	# TGs	# Incs	# TGs/ Inc	Time/secs	# TGs	# Incs	# TGs/Inc
Ann (1)	149.7	71	14	5.1	135.2	49	11	4.5
Jim (1)	137.9	77	20	3.9		16	6	2.7
Mary (2)	142.5	74	9	8.2	317.9	36	7	5.1
Kate (2)	201.8	144	23	6.3		80	22	3.6
Jane (2)	158.8	101	29	3.5		57	20	2.9
Rosa (3)	168.5	112	22	5.1	152.9	1	1	1
Phoebe (3)	207.4	103	21	4.9		36	4	9
Minnie (3)	152.9	134	33	4.1		46	10	4.6
Overall	1319.7	816	172	4.7	606	321	81	6.9

Table 2: The seven simple chains illustrated.

i	NV	Susan died
ii	NVN	Susan kissed Bill
iii	NVNA	Susan ate sushi quickly
iv	NVNE	Susan likes tea hot
v	NVE	Bill is clever
vi	NVA	Susan danced merrily
vii	NVNEA	Bill made Susan sad by accident

Table 3: The speakers' tone choices.

		Monologue						Conversation					
NAME	GROUP	\	^	/	∨	–	Tot	\	^	/	∨	–	Tot
Ann	1	37	1	14	11	8	71	19	0	15	10	8	52
Jim	1	49	1	13	11	3	77	13	1	2	3	0	19
Mary	2	23	2	32	12	5	74	27	0	10	4	4	45
Kate	2	62	5	45	17	15	144	47	3	19	14	5	88
Jane	2	50	3	26	14	8	101	41	2	6	9	4	62
Rosa	3	57	0	19	18	18	112	3	0	1	1	0	5
Phoebe	3	50	1	21	24	7	103	13	0	11	9	4	37
Minnie	3	80	5	17	22	10	134	38	0	7	5	3	53
Overall		408	18	187	129	74	816	201	6	71	55	28	361

Table 4: The correspondence of end-falling tone with increment final position.

	Monologue			Conversation		
	% \	%\ INC	P value	# \	%\ INC	P value
Ann	54	64	N/S	46	71	N/S
Jim	65	79	N/S			
Mary	34	44	N/S	62	74	N/S
Kate	47	32	N/S			
Jane	53	34	N/S			
Rosa	56	54	N/S	57	6	N/S
Phoebe	50	47	N/S			
Minnie	64	73	N/S			
Overall	52	54	N/S	57	65	.009578*

Table 5: The tone and termination choices at increment endings.

Tone + Termination	Monologue	Conversation	Overall
\ + Low Termination	13	11	24
\ + Mid Termination	61	26	87
\ + High Termination	9	8	17
/ + Low Termination	13	5	18
/ + Mid Termination	19	9	28
/ + High Termination	20	2	22
√ + Low Termination	6	1	7
√ + Mid Termination	20	10	30
√ + High Termination	2	1	3
∧ + Low Termination	1	0	1
∧ + Mid Termination	4	2	6
∧ + High Termination	1	0	1
Total	169	75	244

Table 6: Tone and termination and A and A/B events.

Tone + Termination	A-Event	AB-Event	Total
L\	19 (79.1%)	5 (20.1%)	24 (100%)
M\	71 (82.5%)	15 (17.5%)	86 (100%)
H\	16 (94.1%)	1 (5.9%)	17 (100%)
L^	0 (0%)	1 (100%)	1 (100%)
M^	5 (83.3%)	1 (16.7%)	6 (100%)
H^	1 (100%)	0 (0%)	1 (100%)
L/	4 (22.2%)	14 (77.8%)	18 (100%)
M/	9 (31%)	20 (69%)	29 (100%)
H/	12 (54.5%)	10 (45.5%)	22 (100%)
Lv	2 (28.6%)	5 (77.8%)	7 (100%)
Mv	7 (23.3%)	23 (76.7%)	30 (100%)
Hv	2 (66.7%)	1 (33.3%)	3 (100%)

Table 7: Lexicogrammatical realisation of final focal items in A-event congruent increments

	1 st mention	Previous mention	Final	Non-Final	Total
N element ²⁷	31	26	40	17	57
V element	16	10	21	5	26
E element	17	2	17	2	19
P or A element	3	4	6	1	7

²⁷ There were 4 pronominals or non-content lexical items coded as N elements: *everything*, *something*, *we* and *me*. Along with the 7 P and A elements they represent the sole instances of non-content items being make focal. We have also discounted the asking increment from this analysis as the final focal element was the exclamative *yeah*.

Table 8: Lexicogrammatical realisation of final focal items in AB-event congruent increments.

	1 st mention	Previous mention	Final	Non-Final	Total
N element	15	28	30	13	43
V element	11	11	15	7	22
E element	2	1	1	2	3
P or A element	0	6	6	4	6

Table 9: The intonation realisation of non-congruent AB-events in increment final position.

	Low	Mid	High	Total
Falling \	5	15	1	21
Rise-Falling /\	1	1	0	2
Total	6	16	1	23

Table 10: The intonation realisation of non-congruent A events in increment final position.

	Low	Mid	High	Total
Rising /	4	10	10	24
Fall-Rising \	2	5	2	9
Total	6	15	12	33

Figure 2: The tones of English in a five tone approach.

