



O'Grady, G. and Bartlett, T. (2019) Linearity and tone in the unfolding of information. *Acta Linguistica Hafniensia*, 51(2), pp. 192-221.
(doi:[10.1080/03740463.2019.1668621](https://doi.org/10.1080/03740463.2019.1668621))

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/197264/>

Deposited on: 23 September 2019

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

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 by moving from an Initial State to a Target State. 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 speech functional analysis of a corpus of monologue and dialogue shows that matters are not so simple. Speakers may complete increments with lexical elements which aim to manage the spoken interaction or they may use non-congruent prosodic cues to manage their interaction or to upgrade or downgrade the salience of information within the increment and thus breach the expected given before new order. Our paper is of significance in that it incorporates prosody into a detailed description of how speakers manage information flow in real time and secondly it shows how in extended monologues speakers simultaneously manage informational flow and signal their interactional expectations. In other words, turns are no longer privileged as the sole sites where speakers balance informational and interactional needs; instead such 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.

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. The manipulation of grammatical and prosodic elements and features within the spoken linear string signals the speakers' 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 hearers somehow have the ability to reinterpret as a series of phonemic segments and suprasegments. Hjelmslev (1975: 136) dubbed the elements on what he called the physical or expression plane *cenemes*, but at the same time he recognised that the same units functioned *plerematically* to express lexico-grammatical meaning on the content plane (ibid: 99). *Cenemes*, (Nöth 1995: 70-71) says are empty and in themselves contain no meaning. Nöth equates *cenemes* with phonemes but suprasegmental elements such as syllables and feet are in and of themselves devoid of meaning and so we use *ceneme* as a superordinate term in this article. *Pleremes* by contrast are elements which in and of themselves contain meaning.¹ We have classed an element on the expression plane, the tone group, as being semiotically meaningful as it 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 *cenemes* are reinterpreted as *pleremes* within a mind remains unknown – though see Cutler (2012) for ideas about how phonological structure is encoded and decoded into lexico-grammatical units – we assume that people as a matter of course unproblematically reinterpret *cenemes* as *pleremes* in numerous daily communicative events. We also assume that the fact that words follow words in speech masks the real fact that language is stratified and compositional. Halliday and Matthiessen (2014) propose the following rank hierarchies which we take as our starting point.

¹ We thank Kristin Davidse for bringing Nöth's work to our attention.

| | | |
|---|---|---|
| 1 | (a) Expression plane TONE GROUP Foot Syllable Phoneme | (b) Content plane CLAUSE COMPLEX ² CLAUSE PHRASE/GROUP WORD MORPHEME |
|---|---|---|

It can be seen that it is not easy to directly equate units on the expression plane with those on the content plane. Halliday (1967) claims that the sole point of convergence is, that all things being equal, a clause is realised as a single tone group.³ It can further be seen that there is no equivalent of a clause complex on the expression plane in Halliday’s framework. While we will not propose a phonological unit as an equivalent of the clause complex, we will introduce a lexico-grammatical-intonational unit, the increment, which we will argue is the minimum unit required to tell or to ask (Brazil 1995, O’Grady 2010). We will examine increments— see Section 2 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. Thus we will show that turns are not the sole sites where speakers balance informational and interactional needs.

2 What is an Increment?

The concept of an increment, the minimum unit required to tell or ask, was introduced in Brazil (1995) in order to describe a stretch of speech that realises either a telling or asking exchange. He wrote that in telling exchanges, the first speaker initiates and achieves their individual communicative purpose while the second speaker may 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 2 with data drawn from the corpus under investigation, see Section 3 for full details, illustrates:

| | | | |
|---------|---|---------------------|-------------------------|
| 2 | First Speaker | Second Speaker | First Speaker |
| Telling | I don’t like concrete either | Uh no | N/A |
| Asking | Is it Venice that is sinking *What is the name of the place that is sinking? ⁴ | Yeah yeah Venice | (Silence) Ok thanks. |

Prior to producing the initial words in the exchange the interlocutors are classed as being in an *Initial State*; the relevant background state of knowledge prior to the act of speaking assumed

² The presence of small capitals indicates that the element is a plereme.

³ Croft (1995) reports the results of a small corpus study which provided support for Halliday’s view. But nonetheless he found that that (i) syntactically complex clauses which exceeded the limits of short-term memory were chunked into more than one tone group, (ii) speakers tended to place adjuncts and (iii) parallel structures into separate tone groups.

⁴ As there are no example of Wh questions in the data, we have had to make one up.

by the speaker to be shared between the interlocutors. Upon completion of the telling increment, the speaker has achieved Target State: the state assumed by the speaker after the articulation of the increment. The achieved Target State may itself serve as a modification of the shared interlocutor background and function as the Initial State for a following exchange. Elements produced within an increment prior to Target State realise an Intermediate State. Thus:

| | | | | |
|---|---------|----------------|----------------|---------|
| 3 | I | don’t like | concrete | either. |
| | Initial | Intermediate 1 | Intermediate 2 | Target |
| | State | State | State | State |

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 expected to produce a further element. While in this case the articulation of the following element *concrete* has the potential to realise a target state the speaker decided that the element *either* was required to achieve the desired *target state*; her interpersonal alignment with her⁵ hearer was required for her communicative purpose to be achieved.

Brazil (1995) identified 3 criteria which 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 that proposed by those such as: (i) Auer (1996) and Haselow (2016), whose versions of real time syntax consider potential completion solely in terms of turn-taking. Example 11 below illustrates a sequence which represents a potential completion for Auer but not for the approach followed here; and (ii) W. O’Grady (2005), whose psycholinguistic linear grammar 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.

2.1 The grammatical criterion

Brazil (1995: 51) identified seven⁶ simple chains which led to achievement of target state. Table 1 illustrates.

Table 1: The seven simple chains illustrated⁷

| | | |
|-----|-----------|---------------------------------|
| | | |
| i | N V | Susan died |
| ii | N V N | Susan kissed Bill |
| iii | N V N A | Susan ate sushi quickly |
| iv | N V N E | Susan likes tea hot |
| v | N V E | Bill is clever |
| vi | N V A | Susan danced merrily |
| vii | N V N E A | Bill made Susan sad by accident |

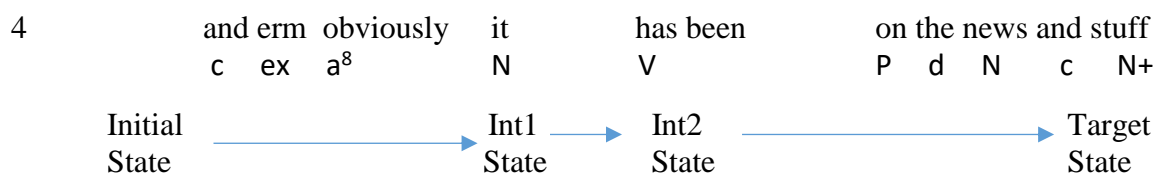
⁵ The speaker was female and in the following examples the pronoun *he* or *she* is used to refer to the actual gender of the speaker.

⁶ Example 11 below is an example of an 8th simple chain NVEA.

⁷ The chains described in Figure 1 represent telling increments. Asking increments would commence with a sequence of V N elements. Following Brazil, N stands for Nominal, V for Verbal, A for Adverbial and E for Adjectival elements.

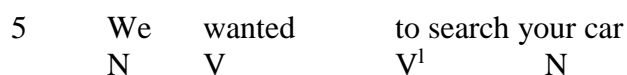
It can be seen that the minimal possible chain consists of between two and five obligatory elements. The articulation of the N element modifies the Initial State and creates an expectation of a following verbal element. In (i) the articulation of the intransitive V element *died* fulfils the prospected 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 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 does not do so because the speaker's communicative need is not to tell what Susan ate but how she ate it, and in (iv) it is not that she likes tea but the temperature of the tea which the speaker wants to tell. In (v) as in (ii) production of the verbal element cannot achieve Target State and the expectation of a following – in this case E – element remains. In (vi) by contrast the V element could achieve Target State, but the speaker as in (iii) and (iv) intends to tell how Susan danced. Finally, in (vii) while the articulation of the E element results in the production of a permissible N V N E chain, the speaker's message concerned the manner of Bill's action and so the speaker produces the following A element.

Brazil (1995:57–68), recognising the inadequacy of simple chains such as those presented in Table 1 for describing authentic speech, proposed a number of further formal devices to augment the simple chains. The following example from the 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 relieve the speaker of the expectation that she will produce an N element which will result in Intermediate State 1 and prospect a further V element and so on. *Suspensive* elements are of informational relevance as in example 4, where the A element asserts the speaker's attitude to the 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:



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 agnate 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 4 above is:

⁸ Suspensive elements are notated in lower case. 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.

pubs 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 N V a V¹ chain *still surviving*;⁹
which tells the hearer something about how the people have endured the storms.

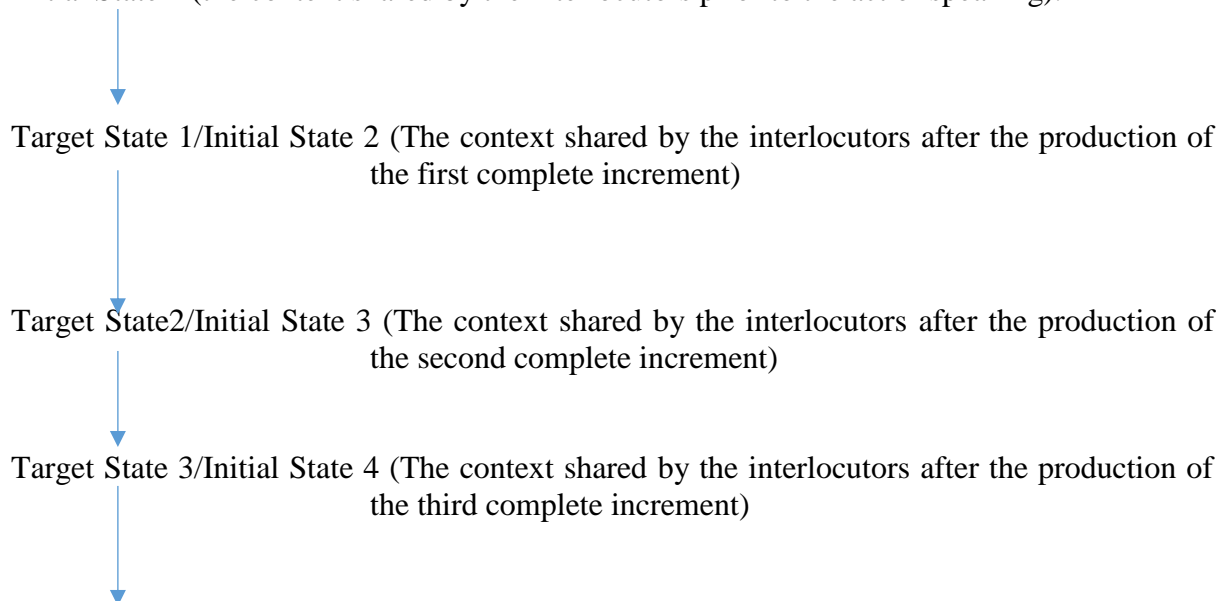
2.3 The Contextual Factor

As they are related to the communicative purpose of speakers in context, Target States cannot be identified on purely formal criteria, either intonational or syntactic, 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 realised through the addition of *still surviving*. 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. However, this is potentially problematic for designers of descriptive grammars as, while replicable criteria can be established for identifying where *potential* Target States may be found, the grammarian was not part of the context of the original talk and hence may not always be able to accurately locate *actual* Target States. Yet, assuming the speaker is a competent speaker of the language, we can see what the speaker signalled and how the hearer responded and use this to identify Target States. But recognising that linguistic communication is not 100% reliable, we argue that an indeterminate descriptive grammar is as good as it gets. Similarly, we can hope for no more than a less than perfect coding.

2.4 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 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.

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



⁹ The N V elements (presumably *they are*) are ellipted.

Target State 4/Initial State 2 (The context shared by the interlocutors after the production of the fourth complete increment)



Target State 5 (The context shared by the interlocutors after the production of the fifth complete increment)

.....

Figure 1: Increments and the unfolding of Text.

However, this description glosses over a number of 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' prosodic choices within and between the individual increments which serve to signal (i) an informational and tactic hierarchy within and between increments and (ii) the speech function realised by the increment.

On the first point tone groups are stretches of speech which contain a single major pitch movement which is anchored to a prominent syllable known as the tonic. Optionally they may contain additional prominent syllables prior to the tonic (Cruttenden 1997, Halliday 1967, Halliday and Greaves 2008, Tench 1996).¹⁰ There is widespread agreement that the tone group represents a single piece of information which is presented to the hearer, e.g. Chafe (1994) as an idea unit, Cruttenden (1997) as a presentation unit, Halliday and Greaves(2008) as an information unit and in the formalist tradition as a sense unit, 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 and Laver (1968: 8) and Laver (1970: 68)). 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 though depends 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.

On the second point every tone group is also 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 v 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 adopted from Brazil (1997) illustrates.

In direct discourse, where the speaker is focusing on producing a meaningful and contextually appropriate message a falling or rise-falling tone projects that the speaker has enlarged speaker/hearer common ground. Hence, within the increment, tone groups which contain falling or end-falling tones are projected as more dynamic in the sense that they contain propositions which move the message on by signalling the expansion of speaker/hearer common ground.

¹⁰ Tone groups can be equated to what ToBI theorists (e.g the chapters in Jun 2007 and 2015) label intonational phrases, see Ladd (2008).

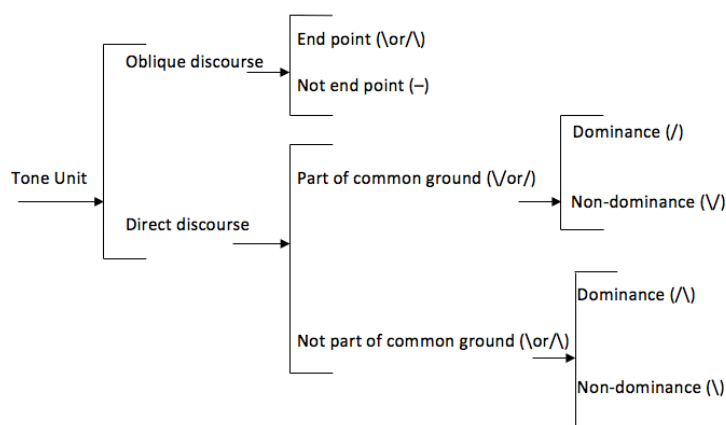


Figure 2: The Tones of English in a five tone approach

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 termination¹¹ are more intrinsically informational than those with mid which are themselves more informationally salient than those with low termination. Example 14 illustrates:

14 [H] > [H] > [V] > [V] > [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. First of all, Esser conflates the functions of fall and rise-fall as well as rise and fall-rise. By so doing, as will be shown later, he ignores the interactional function of tone. Second, he does not include key and is therefore unable to account for how speakers signal their anticipation of how hearers will respond to their words. Third and more seriously 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 so within the 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.

Furthermore Esser's equation of prosody solely with the presentation of information¹² ignores the multi-functional nature of prosody. Key and termination choices signal speakers' momentary apprehension of how hearers will respond to their words. 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 words are actions with material consequences and not merely information transferred through the aether from mind to mind. They realise the speech functions of offer and command, statement and question. The primary division in speech role

¹¹ 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 linearity determines which system predominates; key in increment initial position and termination in increment final position.

¹² Esser's data was monologic telling and hence did not require any active intervention from the hearer.

is between a speaker who gives something to the hearer or one who demands something from the hearer. Speakers can give/demand information or physical goods and services (Halliday and Matthiessen (2014: 135-136).

In terms of the grammar presented here, telling and asking increments may map onto all four speech roles. Increments cannot be looked at exclusively in terms of information 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 dominant 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. It does not capture the dynamic interplay within and between increments. Prosody alone neither signals movement across a discourse from contextually bound Initial States to contextually bound Target States nor highlights the informationally most salient elements. Much work remains to be done. Here we will examine the interplay of prominence, tone, key, termination, linearity, lexicogrammatical realisation and semantics in the presentation of information within the corpus to be described in Section 3.

3 Data

In order to investigate how informational importance is signalled within increments, we recruited 9 participants: all speakers of standard southern British English, and divided our participants into 3 groups. One participant had to drop out at the last minute leaving us with one group of 2. 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.¹³ 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 provoke the participants into constructing personalized and organized narratives. 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 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 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. With the help of Praat (Boersma & Weenick 2013) we transcribed the

¹³ <http://www.youtube.com/watch?v=EjxgnpVNjJQ> [last accessed March 22 2015]

data prosodically and then, using the criteria detailed above, into increments – see G. O’Grady (2016) for details of the prosodic coding. Tables 2 and 3 present the results.¹⁴

Table 2: The data presented as tone groups and increments

| Speaker | Monologue | | | | Conversation | | | |
|------------|-----------|-------|--------|---------------|--------------|-------|--------|-----------|
| | 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 |

As can be seen all the speakers spoke for more than 2 minutes when asked to describe the YouTube clip. The contributions to the conversation were unequal with one participant, Rosa, producing a minimal contribution. 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, for Phoebe and Kate can largely be explained by the presence of level tones indicating some production difficulty in assembling their messages. However, Table 3 does not tell us where the falling tones were found within the increments.

Table 3: The Speaker’s 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 |

¹⁴ Readers will note that the number of tones and hence tone groups in the conversation detailed in Table 2 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.

In order to check if there was a preference for end-falling tone 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 linear nature of 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.

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

| | Monologue | | | Conversation | | |
|---------|-----------|--------|---------|--------------|--------|----------|
| | % \ | %\ INC | P value | # \ | %\ INC | P value |
| Ann | 54 | 64 | .6554 | 46 | 71 | .1573 |
| Jim | 65 | 79 | .3703 | | | |
| Mary | 34 | 44 | .7924 | 62 | 74 | .07265 |
| Kate | 47 | 32 | .288 | | | |
| Jane | 53 | 34 | .1341 | | | |
| Rosa | 56 | 54 | .9362 | 57 | 6 | 1 |
| Phoebe | 50 | 47 | 1 | | | |
| Minnie | 64 | 73 | .4235 | | | |
| Overall | 52 | 54 | .7611 | 57 | 65 | .009578* |

For whole data set χ -squared = 2.8673, df = 1, p-value = 0.0904

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 choose 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.

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 or 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 5 details the tone/termination choices of all the increment endings.

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 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 how this correlates with the tonal structure of their utterances.

4 Discussion

Following Brazil, 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 hearer alone. Following Labov (1972: 124), therefore, we examined all increment final tone groups and classed them as either A events (content known only to the speaker) or AB events (content known to both speakers). While a telling increment as a whole must contain an A-event, it may contain tone groups which encode AB-events. A tone group was classed as an A event when it contained a proposition that was neither available from the context nor commonly known information.¹⁵

In telling increments the speakers assume the role of the primary knower, (Berry 2016 and Muntigl 2009). They assume primary epistemic rights Muntigl (ibid: 260). In 16 the hearer has no way of predicting when the speaker's friend is due to return to work. 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

¹⁵ In the dataset there is only one example of an asking increment and as a result, the following discussion focuses only on telling increments.

4.1 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. 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: Lexicogrammatical realisation of final focal items in A event congruent increments

| | 1 st mention | Previous mention | Final | Non-Final | Total |
|-------------------------|-------------------------|------------------|-------|-----------|-------|
| N element ¹⁶ | 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 |

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 | ^erm | I know a few ^years ago | muh my L /aunt | /uncle | and three of my \cousins
| \lost | all of their L\power |and L\water |for about H two or three H /weeks H | -erm
|because of the \flooding |

In 19 the focal element *bomb* is followed by two non-content items – the final one of 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.

4.2 Congruent A/B events

Table 8 shows that some of our expectations were supported and others 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.¹⁷

¹⁶ 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 exclamation *yeah*.

¹⁷ 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 A/B 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 A/B events. Both results were positive with the former being $\chi^2 = 8.9351$, $df = 1$, $p\text{-value} = 0.002797$ and the latter being $\chi^2 = 2.6845$, $df = 1$, $p\text{-value} = 0.1013$. This suggests that there

Table 8: Lexicogrammatical realisation of final focal items in AB event congruent increments¹⁸

| | 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 |

- 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 | uuhh like at /nightH | so it’s they couldn’t like couldn’t see out in the /garden| but they had them H/patio doorsH | –and | they could just start H\hearing this like | \water hitting | the...the /doorsH |

In 22 while the verbal element *saturate* is not recoverable from the co-text or context, the propositional knowledge that concrete is an impermeable artificial building material is common knowledge in the UK. 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 hearers’ common knowledge. In 23 the lexical item *doors* is in final position but as the tonic element is recoverable from the co-text *doors* does not simply function informationally. The final A/B tone group may as a result be the product of 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 A/B events that there is a need to bear in mind both informational prominence – the 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 their interactional goals. In order to examine further 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 A/B event but accompany with an end-falling tone.

4.3 Non-congruent A/B events

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

Table 9: The intonation realisation of non-congruent A/B 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 |

may be both significant grammatical and informational reactances in how speakers produce tone groups encoding A and A/B events. However, to be certain we would need to confirm these tentative results by examining a far larger data set.

¹⁸ 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.

As can be seen the most frequent tone choice for non-congruent A/B 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 |

Had the speaker chosen an end-rising tone in final position she would have invited comment, which would appear unnecessary. On 6 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 Lblame | <end of topic> | and it Hkind of | it kind of Hputs the human in er us as \humans |

Here the initial high key signals the introduction of a shift in topic away from the finger pointing of the politicians and towards the effect of climatic disasters on people 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 Hfloods |

The high fall anticipates hearer adjudication; in other words unlike the other non-congruent A/B events it invites a response: the hearer albeit silently is asked to make a mental note signalling their agreement or disagreement with the speaker's proposition.

4.4 Non-congruent A events

We located 33 non-congruent A events which are detailed in Table 10. 24 of them involve rising tone (examples 28–30) and 9 have fall-rising tone (examples 31–33).

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 |

In 29 to 31 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 me |

28 | and -one of the | ideas that that was \proposed | which was quite /interesting | was um the idea of /floating citiesH | which \submerge | when it's really bad /weatherH |

29 | and all the Ldirty water that's just Lbeen | Lflooding it |

In 27 the content expressed in the final tone group 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. The combination of the high termination and the rising tone signals a double response: the rise that the Target State is open to speaker negotiation and the high the fact that the speaker anticipates adjudication. By contrast in 29 the low rise signals both the end of Rosa's contribution and invites her hearer to produce a notional response: considering the effect of the water that has been flooding people's properties.

Selection of a fall-rise tone signals a contextually bound implication or as Halliday (1967) dubbed it "a but". As Table 9 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 \avoided | L actually
any L \flooding like that | at \all | anywhere where I \had to be |
- 31 | this \one road | it was H\more like | that was more H\specific than | the other \clips |
L\where it was just | people talking about what they have \seen | and they're Hsat in
their \cars |
- 32 | \erm | and | the village I live in is \actually | in a little /valleyH | \and | it used to be
marshland /before | they L\built on it |\erm | L\hundreds of years ago | <end of topic>
| so it is H\kind of in |the exact right place to L\ flood | \badly |

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. Had she produced a falling tone no such implication would have been generated. 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 signal 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.

5 Conclusion

Based on our novel approach we have shown that a monodimensional informational hierarchy based on adjacency, tone and termination is insufficient. Using a speech-functional analysis we identified final tone groups as projecting propositions which expressed A and A/B 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 propositional as a series of tone groups and the interactional 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 lexical and propositional 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. Hearers operating under real time constraints and within the confines of a shifting speaker/hearer negotiated context select the first or most relevant interpretation of the cues available to them in the context of the discourse in inferring the function of a cue based upon their individual apprehension of the state of convergence shared between the interlocutors. This allows speakers to manage their hearers' expectations as they build up their message increment by increment but also entails that verbal communication is open to misinterpretation and does not contain a single meaning.

References

- Auer, P. 1996. On the prosody and syntax of turn continuations. In E Couper Kuhlen and M Selting (eds.) *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press, 57 – 100.
- Austin, J. 1961. *How to do things with words*. Oxford: Oxford University Press.
- Berry, M. 2016. Dynamism in exchange structure. *English Text Construction* 9: 33-55.
- Boersma, P & D. Weenick. 2013. *Praat doing phonetics by computer*. Computer programme Version 5.3.52.
- Boomer, D. & Laver, J. 1968. "Slips of the Tongue." *British Journal of Disorders of Communication*, 3 (pp. 2-12).
- Brazil, D. 1995. *A grammar of Speech*. Cambridge: Cambridge University Press.
- Brazil, D. (1997). *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press
- Chafe, W. 1994. *Discourse Consciousness and Time*. Chicago: The University of Chicago Press.
- Couper-Kuhlen, E. 1996. *An introduction to English prosody*. London: Edward Arnold.
- Croft, W. S. 1995. Intonation units and grammatical structure. *Linguistics*, 71, 490–532
- Cruttenden, A. 1997. *Intonation*. 2nd Edition Cambridge: Cambridge University Press.
- Culter, A. 2012. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge MA. M. I. T. Press.
- Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Esser, J. 1988. *Comparing reading and spoken intonation*. Amsterdam: Rodopi.
- Firbas, J. 1992. *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press
- Fries, Peter. 1995. Themes, Method of Developments and Texts. In R Hasan and P Fries (Eds.) *On Subject and Theme: A discourse functional perspective*. Amsterdam: John Benjamins
- Gussenhoven, C. 1983. *On the grammar and semantics of sentence accents*. Dordrecht: Fortis
- Gussenhoven, C. 2004. *The Phonology and Tone of Intonation*. Cambridge: CUP.
- Halliday, M. A. K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton
- Halliday, M. A. K., & Greaves W. S. 2008. *Intonation in the Grammar of British English*. Equinox: London.
- Halliday, M. A. K., & Matthiessen, C. I. M. 2014. *An Introduction to Functional Grammar*. 4th Edition. Abingdon: Routledge.

- Haselow, A. 2016. A processual view on grammar: macrogrammar and the final field in spoken syntax. In *Language Sciences* 54, p 77-101
- Hjelmslev, L. 1975. *Resume of a theory of language*. Frances W Whitfield (Ed and Trans). Madison WA. University of Wisconsin Press.
- Jun, S-A (2007 ed.) *Prosodic Typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, S-A. (2015 ed.) *Prosodic Typology 11: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Labov, W. 1972. Rules for ritual insults. In D Sudnow (ed.) *Studies in Social Interaction*. New York: The Free Press, 120–170.
- Ladd, D. R. 2008. *Intonational Phonology* 2nd Edition. Cambridge: Cambridge University Press.
- Laver, J. 1970. The production of Speech. In J Lyons (ed.) *New Horizons in Linguistics*. Harmondsworth: Penguin, 53 – 75.
- Muntigl, P. 2009. Knowledge moves in conversational exchanges: Revisiting the concept of primary vs. secondary knowers. *Functions of Language* 16: 225-263.
- Nöth, W. 1995. *Handbook of Semiotics*: Bloomington & Indianapolis: University of Indiana Press.)
- O'Grady, G. 2010. *A Grammar of Spoken English Discourse: The intonation of Increments*. London: Continuum.
- O'Grady, G 2016. Given/New. What do the terms refer to? A first (small) step. *English Text Construction*. 9–32.
- O'Grady, W. 2005. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, N.J.: Erlbaum,
- Tench, P. 1996. *The Intonation Systems of English*. Cassell: London.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartik (1985). *A comprehensive grammar of the English Language*. London: Longman.
- Watson Duane & Edward Gibson (2004). Making sense of the Sense Unit Condition. *Linguistic Inquiry*. Vol 35: 3, 508 – 517.