

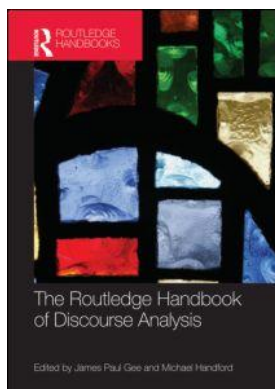
This article was downloaded by: 10.3.97.143

On: 30 Sep 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## The Routledge Handbook of Discourse Analysis

James Paul Gee, Michael Handford

### Prosody in discourse

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9780203809068.ch19>

Winnie Cheng, Phoenix Lam

**Published online on: 16 Nov 2011**

**How to cite :-** Winnie Cheng, Phoenix Lam. 16 Nov 2011, *Prosody in discourse from: The Routledge Handbook of Discourse Analysis* Routledge

Accessed on: 30 Sep 2023

<https://www.routledgehandbooks.com/doi/10.4324/9780203809068.ch19>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## Part III

# Developments in spoken discourse

---



# Prosody in discourse

*Winnie Cheng and Phoenix Lam*

---

## Introduction to discourse intonation framework

This paper aims to examine the communicative value of discourse intonation by describing the four systems of discourse intonation (Brazil, 1985, 1997): prominence, tone, key and termination. The four systems of speaker intonational choices, each of which has a general meaning that takes on a local meaning within a particular context (Brazil, 1997: xi), are moment-by-moment judgments made by speakers on the basis of their assessment of the current state of understanding operating between the speakers. The paper begins with describing Brazil's (1985, 1997) discourse intonation framework as purpose-driven, speaker controlled, interactive, co-operative, context-referenced, and context-changing, followed by the description of each of the four systems illustrated with examples from naturally occurring speech. The data analysed in this paper come from the one-million-word Hong Kong Corpus of Spoken English (HKCSE) (prosodic) (Cheng *et al.*, 2008) which is composed of the academic, business, conversation and public sub-corpora. The transcription notation used in the HKCSE (prosodic), as well as in this chapter, can be found in "Transcription conventions" at the end of this chapter.

Discourse intonation is based on the view that spontaneous speech is purpose-driven rather than sentence-oriented. It is speaker controlled, interactive, co-operative, context-referenced, and context-changing (Brazil, 1995: 26–39). Discourse intonation systems are motivated by real-time, situation-specific decisions taken by speakers to add extra layers of interpersonal meaning to words as they are spoken, and they are concerned with "the speakers' moment-by-moment context-referenced choices" (Cauldwell, 2007). The communicative value of intonation is concerned with the choices that speakers make and with their reactions to the ongoing task of making sense to their hearers in context in real-time (Cauldwell, 2002). Examining the choices of discourse intonation helps to determine the pragmatic and situated meanings of English utterances (Brazil, 1997: ix). The intonation choices that speakers make in relation to the four systems in the discourse intonation framework are independent. Altogether, thirteen intonation choices are available. Figure 19.1 summarizes the intonation choices available in the four systems of discourse intonation.

## Tone units

Brazil (1995) states that, in purpose-driven talk, intonation and syntax are considered as "being separate areas of choice," and "there is no 'normal' relationship between tone units and clauses" (Cauldwell, 2007). In fact discourse intonation moves beyond the context of the single sentence and describes the rules that govern the pitch movement beyond and between the borders of tone

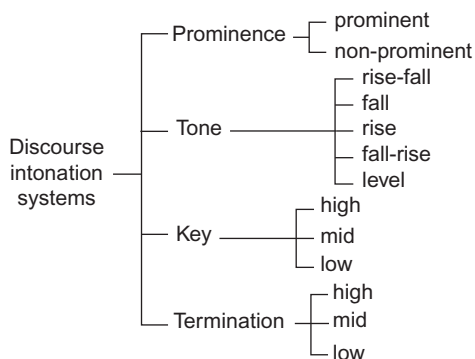


Figure 19.1 Map of the four systems of discourse intonation

units rather than sentences. Brazil (1985: 238) argues for a “need for stating the communicative value of intonation in terms of the projected contextual implications of the tone unit: only if we regard intonation as a ‘situation-creating’ device, ... can we give proper recognition to its ability to carry independent meanings.”

All of the thirteen intonation choices occur within the boundaries of a tone unit. In discourse intonation, a tone unit refers to “the stretch of language that carries the systematically-opposed features of intonation” (Brazil, 1997: 3). The internal organization of the tone unit in discourse intonation can be described in terms of three parts: non-prominent optional stretches (proclitic and enclitic segments) and the mandatory tonic segment delimited by the first and last prominent syllables, in which all the significant speaker-decisions are made (*ibid.*, 15). A tonic segment typically comprises one or two prominent syllables, any of the five tones or pitch movements (fall, rise, fall-rise, rise-fall, level) carried by the final prominent tonic syllable; the three-term pitch-level system (high, mid and low) associated with the tonic syllable; and the three-term pitch-level system (high, mid and low) associated with the onset syllable.

A tonic segment can minimally consist of only one prominent syllable, which is the tonic syllable. Example 1 shows a one-word tone unit that contains only the mandatory tonic segment, with one prominent syllable (*so*) and no non-prominent stretches:

- (1)  
 {= [< SO >]}

Alternatively, two prominent syllables, namely the onset and the tonic, may be found in the tonic segment. Example 2 shows a tone unit that contains only the mandatory tonic segment, with two prominent syllables (*so* and *have*) and no non-prominent stretches outside the tonic segment:

- (2)  
 {= [SO] we < HAVE >}

An optional segment, namely the proclitic segment, may be present in front of the tonic segment. Examples 3 and 4 illustrate tone units that contain the mandatory tonic segment as well as the proclitic segment. Example 3 is a tone unit with one prominent syllable (*why*) in its tonic segment and the proclitic segment (*so I don't know*), and Example 4 shows a tone unit which contains two prominent syllables in its tonic segment (*may* and *please*) and the proclitic segment (*so*):

- (3)  
 {\ so i don't know [< WHY >]}

- (4)  
 {= so [MAY] i have your passport < ^ PLEASE > }

Instead of appearing in front of the tonic segment, an optional segment, named the enclitic segment, may be found after the tonic segment. Examples 5 and 6 illustrate tone units that contain the mandatory tonic segment as well as the enclitic segment but not the proclitic segment. Example 5 is a tone unit with one prominent syllable (*so*) in its tonic segment and the enclitic segment (*I have to*), and Example 6 shows a tone unit that contains two prominent syllables (*so* and *checked*) in its tonic segment and the enclitic segment (*with um it with*):

- (5)  
 {= [< SO >] i have to }

- (6)  
 {= [SO] i've already < CHECKED > with um it with }

In some cases, both proclitic and enclitic segments are found in addition to the tonic segment. Examples 7 and 8 show respectively one-prominence and two-prominence tone units that contain all the three parts:

- (7)  
 {= i don't [< THINK >] so }

- (8)  
 { \ so i [ ^ MADE ] these < CHANGES > }

Table 19.1 provides a schematic summary of the examples discussed above regarding the different internal structures of a tone unit.

The composition of tone units carries important information regarding the decisions made by speakers. Specifically, it indicates whether information is to be considered integrated or distinct. Tone unit boundaries, therefore, can function as a device of disambiguation (Cheng *et al.*, 2008). When the word *so* is in a separate tone unit, for example, it is very likely to be used as a discourse particle, as in Example 9:

- (9)  
 {= [< SO >] } {= [WHERE] did you GET your < aMERican > } { \ [< ACcent >] }

On the other hand, the propositional uses of the word rarely constitute a separate tone unit on its own. In other words, when *so* is not a discourse particle, it is often in a tone unit with other items, as in Example 10:

- (10)  
 {= er [ONE] three < A > } {= [ONE] three < B > } {= [ONE] three < C > } { \ and [SO] on and < SO > forth } ...

Table 19.2 shows the frequency distribution of *so* in separate and shared tone units in the HKCSE (prosodic).

While there are more instances of *so* used as a discourse particle in the prosodic corpus regardless of whether the word is in a separate tone unit, it can be observed that there is a greater likelihood that this is the case when *so* is in a separate tone unit (Lam, 2008). Lam (2008) finds that of all the instances of *so* in a tone unit on its own, an overwhelming 98.1 percent, are discourse particles, and that only 1.9 percent of all instances of *so* in a separate tone unit express propositional content. In other words, tone unit boundaries help to disambiguate the discourse and propositional use of the

Table 19.1 Examples illustrating the three-part structure of a tone unit

(Proclitic segment)	Tonic segment	(Enclitic segment)
	{= [< SO >]}	
	{= [SO] we < HAVE >}	
{\ <b>so</b> i don't know	[< WHY >]	
{= <b>so</b>	[MAY] i have your passport < ^ PLEASE >	
	{= [< SO >]}	i have to}
	{= [SO] i've already < CHECKED >	with um it with}
{= i don't	[< THINK >]	<b>so</b>
{\ <b>so</b> i	[^ MADE] these < CHAN	ges >}

Table 19.2 The frequency distribution of *so* in separate and shared tone units in the HKCSE (prosodic)

Function of <i>so</i>	Total number of <i>so</i>	Shared	Separate
As a discourse particle	6,583 (81.6%)	4,160 (74.3%)	2,423 (98.1%)
Not as a discourse particle	1,487 (18.4%)	1,441 (25.7%)	46 (1.9%)
Total	8,070 (100%)	5,601 (100%)	2,469 (100%)

word. This is consistent with Cheng, Greaves, and Warren's (2008) observation that tone unit boundaries have an important function in disambiguation and help to indicate whether alternatives introduced by the vague expression *or something* are treated as distinct from each other or not. Further, the fact that many multi-word expressions such as phrasal verbs and idioms are often found within a tone unit provide evidence that speakers co-select words as a chunk instead of choosing individual words separately as information units (Sinclair, 1991).

## Prominence

The communicative value of the utterance is affected by intonational variations on the basis of "a small set of either/or choices," which relates to "a set of meaningful oppositions that together constitute a distinctive sub-component of the meaning-potential of English" (Brazil, 1997: 2). The finite set of oppositions, or either/or choices, available to a speaker is "a binary prominent/non-prominent choice" (Brazil, 1997: 9). A prominent syllable is one that a hearer recognizes as being in some sense more emphatic than the others in the tone unit. The first (onset) and the last (tonic) prominent syllables in the tone unit constitute sub-sets of prominent syllables. Specifically, prominence determines the beginning and end of the tonic segment.

A tone unit may contain up to four prominent syllables, though tone units with one or two prominences are considered the norm (Brazil, 1997). This is supported by quantitative findings from the HKCSE (prosodic). In the corpus, tone units with one or two prominences constitute 91.49 percent of the total, tone units that contain one prominent syllable being the most frequently occurring ones (Cheng *et al.*, 2008). Table 19.3 gives examples of tone units with different numbers of prominences.

In discourse intonation (Brazil, 1997), each prominent syllable gives prominence to a word. Prominent words, which contain prominent syllables, realize existential sense selections. A prominent word is presented as "a selection from a set of possibilities defined by the context of situation" (Brazil, 1997: 41). More correctly, a speaker's intonation projects a certain context of interaction, or projects the assumption that a particular word in the tone unit is selected, the assumption being understood as "part of the communicative value of the utterance" (p. 27). In

Table 19.3 Examples illustrating the number of prominences in a tone unit

Number of prominences in a tone unit	Examples
1	{? [< SO >] for example for}
2	{= [SO] it's < GOOD > to}
3	{\ [SO] we have a DAY < FREE >}
4	{\ [SO] you have to WEAR the rubber SHOES to < Office >}
Unclassified	{? er so it's}

other words, a speaker exploits the prominence system to project a context of interaction that suits his/her current conversational purposes.

In making a selection between prominence and non-prominence, speakers have available to them two paradigms: existential and general. The existential paradigm is “the set of possibilities that a speaker [*sic*] regard as actually available in a given situation,” and the general paradigm is the set of possibilities that is “inherent in the language system” (p. 23), the words comprising the existential paradigm being a sub-set of those comprising the general one. The selection of prominence is “what a speaker does when he chooses from an existential paradigm” (p. 45). Brazil (1997: 22–23) exemplifies the two paradigms with his well-known *queen of hearts*, said in response to *which card did you play*. In this utterance, *of* is a product of the general paradigm, because the speaker is limited in this context to this word by the language system. Conversely, *queen* and *hearts* are choices limited by the contents of the pack of cards rather than by the language system, and they are therefore part of an existential paradigm as opposed to a general paradigm. The word *queen* is a selection from an existential paradigm of thirteen members, and *hearts* of four members.

(23) Q: What heart did you play? R: // the QUEEN of hearts //

(24) Q: What queen did you play? R: // the queen of HEARTS //

(Brazil, 1997: 23)

Non-prominent words or non-selection are due to shared extralinguistic factors, which have “a very wide currency,” and to “shared experience of the immediate conversational environment of the response,” which have a circumscribed currency (p. 25). In example 11, which is taken from an academic talk, a lecturer is making a contrast between two angles. Note that in the last two tone units prominence is given to the demonstratives *this* and *that*, to highlight this contrast. In addition, the values of the angles are also made prominent (*gamma/gamma* and *five*), to emphasize the difference between measurements of the two angles. The prominence selection in these tone units reflects the deliberate choices the speaker makes out of a number of possible alternatives in the existential paradigm in order to underline words that are considered situationally informative in this local context:

(11)

...{= [< THIS >] is} {V [< FIVE >]} {= [< AND >]} {V we [KNOW] that THIS angle is < GAMma >} {\ so [THAT] angle would be GAMma minus < FIVE >}...

While the selection of prominence emphasizes words that are more important or relevant in a particular context of interaction, words may be made non-prominent for phatic reasons. When disagreement or only partial agreement is expressed, for example, words that indicate divergence of views may not be chosen for prominence, in order to tone down the difference between interlocutors, for politeness purposes (Cheng *et al.*, 2008). In example 12, which is taken from a televised interview, the talk show host (speaker b1) is asking the guest (speaker b2) a declarative



question based on the guest’s previous response. Notice the guest’s use of a non-prominent *well*, a typical marker of dispreferred response, to preface his partial agreement, followed by a prominent *yes* to highlight the convergence instead of the divergence between speakers:

- (12)
- b1: { \ so the [SEparate] rule you are TALKing < \_ aBOUT > } { \ is [< ACtually >] } { \ a [^ LOWer] < STANdard > }
  - b2: { \ well in [^ MAnY] < CAses > } { \ [< YES >] } { = er [LET] me < SHOW > you } { = [< AN >] } { \ [< eXAMple >] } { \ [< \_ HERE >] } ...

### Tone

The tone or pitch movement in the tone unit is associated with the final prominent syllable (the tonic syllable) in the tone unit, and so tone choices “attach additional meaning increments to tonic segments” (Brazil, 1997: 20). Speakers may choose from five tones: the rise, fall–rise, fall, rise–fall and level tones. Four of the tones are used to distinguish between information that is common ground, i.e. referring tones (R): rise (*r*) and fall–rise (*r*), and information that is new, i.e. proclaiming tones (P): fall (*p*) and rise–fall (*p*+). Figure 19.2 gives a graphical representation of the referring and proclaiming tone choices available to speakers.

Any spoken discourse proceeds on the basis of a considerable amount of shared knowledge between discourse participants (Brazil, 1997: 109), and it is for the speaker to decide, moment by moment, whether what he/she is saying is shared or not. Table 19.4 describes the communicative functions of the proclaiming and referring tones (pp. 82–98).

Example 13 (reproduced from example 11), which is taken from a lecture, illustrates how the selection of tone indicates whether the information presented is considered by the speaker to be shared or not:

- (13)
- ... { = [< THIS >] is } { \ v [< FIVE >] } { = [< AND >] } { \ v we [KNOW] that THIS angle is < GAMma > } { \ so [THAT] angle would be GAMma minus < FIVE > } ...

At the beginning of this extract, the lecturer makes use of the fall–rise tones to signal that the values of the angles are in the common ground, as these values are already discussed earlier. The use of the referring tone and the words *we know that* make clear to the students that this part of the discourse is not presenting new information. On the other hand, the final tone unit in example 12 has fall tone. This is because the value of the angle concerned in this tone unit is unknown to the students up to this point. By using the proclaiming tone, the speaker is presenting new information.

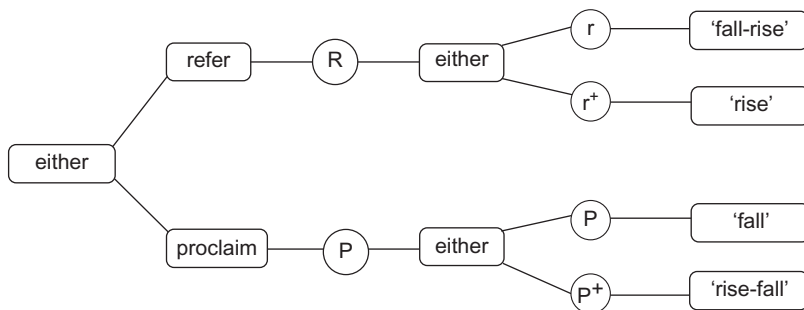


Figure 19.2 The referring and proclaiming tone choices available to speakers  
Source: Brazil, 1997: 83

Table 19.4 Functions of proclaiming and referring tones

<i>Tone</i>	<i>Functions</i>
Referring tone:rise tone (r+)	To reactivate something which is part of the common ground
Referring tone: fall–rise tone (r)	To indicate that this part of the discourse is already present in the common ground, and therefore will not enlarge the common ground assumed to exist between the participants
Proclaiming tone: fall tone (p)	To indicate that this part of the discourse is not yet present in the common ground, and so will be news and world-changing. The area of speaker–hearer convergence is being enlarged
Proclaiming tone: rise–fall tone (p+)	<ul style="list-style-type: none"> <li>• To indicate addition to the common ground and to the speaker’s own knowledge at one and the same time</li> <li>• To indicate to the hearer that no feedback of either an adjudication or concurring kind is expected</li> <li>• To indicate that the speaker intends to continue to speak and so asserts control of the progress of the discourse</li> </ul>

Its introduction enlarges the area of convergence, i.e. the knowledge in a particular academic discipline—in this context, between the lecturer and the students.

As shown by example 13 above, the proclaiming/referring opposition and the choices in the referring tone system contribute implications of common ground information or new information to the tonic segment. The selection of tones, as with other linguistic options, rests with the speaker, and the decision to present information as shared or as new is based on a subjective assessment of the state of shared knowledge between the participants, and is also open to exploitation, should the speaker choose to do so.

Tone selections are also accorded social significance, as there are tone selections that may be characterized as being participant-specific in specialized discourse types (i.e. discourse types other than conversation), and they imply a certain role relationship pertaining between the participants in a discourse (Brazil, 1997: 82–98). These participant-specific tones are the rise tone and the rise–fall tone. If a speaker selects a rise–fall (instead of a fall) in proclaiming something, or a rise (instead of a fall–rise) in referring to something, the speaker is considered to be exerting dominance and control additionally (Brazil, 1995: 243). In discourse types where one speaker is dominant in the sense of having greater responsibility for the discourse and greater freedom in making linguistic choices, that designated dominant speaker monopolizes the rise–fall/rise choice. Examples of specialized discourse types that involve an unequal power relationship between participants include teacher talk and job interviews. In academic lectures, for instance, it is the teacher who is mainly responsible for the content and process of the lesson. Accordingly, the teacher may exert his/her dominance in the discourse through a choice of tone. At the beginning of example 14 below, which comes from a lecture on count and mass nouns, a student (speaker a1) is asking the lecturer (speaker a2) whether “sheep” is considered a count noun or a mass noun on the basis of the ongoing discussion. The lecturer responds with the repeated selection of the rise tone:

(14)

- a1: {/ then [< HOW >] about the word} {\ [<; SHEEP >]} (.) {\ [<; SHEEP >]}
- a2: {/ [< ^ SHEEP >]} (.) {/ [< SO >]} (.) {/ [DO] you think < SHEEP >} {= [< IS >]}  
a) {\ [< COUNT >] noun or} {/ [MASS] < NOUN >}

The lecturer's choice of the rise tone asserts her dominance as the main speaker in the discourse and reminds the student that the answer to the question has already been established earlier in the discussion. In other words, it is perceived to be common ground between the participants.

The rise–fall tone is the least prevalent of the tones, according to Brazil (1997), and he claims that, in a discourse in which the participants are of unequal status, it tends to be the dominant speaker who alone makes this selection. In institutionalized discourse types such as classroom talk and medical consultation, for example, it is often the teacher or the doctor who uses the rise–fall tone to assert dominance and control (Cheng *et al.*, 2008). The scarcity of the rise–fall tone in naturally occurring speech is supported by findings from the HKCSE (prosodic), which show that only a negligible 0.015 percent of all tone units (49 out of 313,340) carry the rise–fall tone (Cheng *et al.*, 2008: 126). More examples of the rise–fall tone in a range of specialized discourse types are needed to investigate further the extent to which the use of this tone is confined to the dominant speaker.

In conversations, however, the selection of the rise and rise–fall tones is not restricted by the existence of institutionalized inequalities between the participants. In conversation, these tones are selected by all, some or none of the participants, depending on the moment-by-moment decisions of those involved and not, on the basis of a restrictive set of conventions. Brazil (1985: 131) argues that in conversation there is “an ongoing, albeit incipient, competition for dominance.” However, he adds that this does not necessarily imply aggressiveness or rudeness on the part of speakers; rather, when a speaker selects the rise or rise–fall tone, this can be characterized as “to remind, underline, emphasize, insist or convey forcefulness” (Brazil, 1997: 98), and so overtly the speaker assumes the status of the dominant one. The important point is that dominant speaker status is neither predetermined nor fixed in conversation and is typically interchangeable among the participants as the discourse unfolds.

In asserting dominance in discourse through the use of the rise–fall tone, speakers may also modify their view of the world at the same time. In this respect, the rise–fall tone signals to the hearer that new information has been added to the speaker's own knowledge, as well as to the common ground between the speaker and the hearer at that moment of the talk. The addition of such new information often arises from a sudden realization by the speaker of the current state of affairs or of an unexpected event, which leads to the speaker's comment (Cheng *et al.*, 2008). Example 15 below shows the use of the rise–fall tone in a conversation. In this extract the two speakers are talking about the number of universities in Hong Kong. The use of the rise–fall tone by speaker y in the first tone unit of the last utterance indicates that his view on this topic has suddenly changed and that he intends to assert his control of the talk through the continuation of speakership:

(15)

- y: {= [< \_ THERE >] is} {= [YOUR] < uniVERsity >} {= [< THIS >] one} {? and the  
[< CHInese >] university} {\ [< THAT'S >] it}
- b: {= [< WHAT >]}
- y: {= [< \_ THERE >] is} {\ [THREE] < uniVERsities >}
- b: {\ no there's [< SIX >]} {\ [< SIX >]}
- y: {\ ^ ah [< YEAH >]} {\ you [< \_ TOLD >] me} {? because i} {\ i [DON'T] <  
reMEMber >} {= the [< eXACT >] er} {= [< NUMBER >]}

The fifth tone—the level tone—is discussed in the context of the orientation or stance the speaker takes. The use of level tone projects neither a certain context of interaction nor any communicative value of the utterance. In fact this tone is used when the speaker does not intend to either proclaim or refer, and in so doing disengages from the immediate interpersonal,

interactive context of interaction. In other words, the speaker does not make “either/or” choices of any kind, and presents the language with neutral projections as to the assumption made about the current state of understanding between the speaker and a hearer (Brazil, 1997: 132).

Instead of making the binary “either/or” selection, a speaker’s choice of employing the level tone focuses on the linguistic properties or message organization of the utterance rather than on the truth of the assertion made in the utterance. Brazil (1997: 133–139) provides a detailed description of the two main contexts when speakers select the level tone. The first is when a speaker is adopting an “oblique presentation” (p. 133), or when a speaker is saying something, on paper or in the speaker’s memory, that is either pre-coded or partially coded information (pp. 136–139). Example 16 (from a public speech) illustrates such a situation, when the speaker’s continuous use of the level tone indicates that he is simply reading out pre-coded information and highlighting the words involved as an entity, which in this case is the name of a meeting. It is not until the end of the name has been reached that the speaker changes his tone choice to the fall tone:

- (16)  
 ...{/ i would [< LIKE >]} {= to [< exTEND >]} {∨ a [V Ery] warm welcome to < ALL > of you} {= who have [COME] < HERE > } {= for the [< FOURteenth >]} {= [GENERal] < MEETing > } {= [< OF >]} {= [< paCific >]} {= [< ecoNOMIC >]} {\ [coOperation] < COUNcil >}...

The second context is one in which encoding has not yet been achieved, or it presents some kind of difficulty for the speaker (p. 139)—which is likely to happen when the speaker is telling a story (p. 140) or when the speaker is talking spontaneously, as in example 17 below, which comes from a conversation, where the speaker makes use of the level tone when he has yet to formulate what to say:

- (17)  
 ...{\ i [< GUESS >]} {= [< ^ SO >] ^ er} {= i [< THINK >] er} {? [< IT >] it} {= [< IT >]} {\ [I'M] < STUdying > } {\ [< HERE >]}

Brazil (1995: 244) describes another context where the level tone is used. It is not related to encoding problems, but found when a speaker says “incremental elements” that form part of a “telling increment.” These elements are message fragments that have not yet reached the “target state” (Brazil, 1995: 165), namely the end of a discrete information unit. Typically, the incremental elements are said with level tones until the final tone unit which is said with the fall tone. Brazil states that a speaker’s selection of tone signals her/his orientation to the ongoing talk at that moment in time. In example 18 (from a conversation), for instance, the speaker repeatedly makes use of the level tone to signal the development of the incremental elements until he comes to the end of the information unit, i.e. the construction of the complete question:

- (18)  
 {= [< ER >]} {= [< ^ HOW >] < aBOUT > the} {= [ecoNOMIC] < situAtion > } {= [< \_ IN >]} {\ u [< ^ K >]}

Figure 19.3 below shows the decisions the speaker has to make for each tone unit. Direct discourse refers to the discourse in process, which is hearer-sensitive and interactive, as opposed to oblique discourse such as reading and quoting, which briefly withdraws the speaker from interacting with the hearer (Cheng *et al.*, 2008):

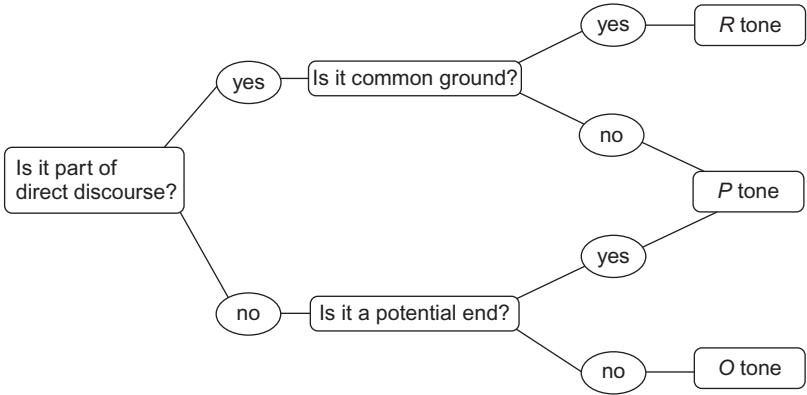


Figure 19.3 Tone choices available to speakers  
 Source: adapted from Brazil, 1997: 135–136

**Key and termination**

Key and termination refer to pitch-level choices available to speakers. Key is the pitch-level choice associated with the first prominent syllable (onset) in the tone unit. The high, mid and low pitch levels at which the onset is pitched are recognized in relation to the onset syllable of the previous tone unit. Termination is the pitch-level choice associated with the last prominent syllable in the tone unit. The high, mid and low levels at which the tonic is pitched are recognized in relation to the pitch level of the preceding prominent syllable in the same tone unit (i.e. the onset), or in the prior tone unit in the case of minimal tone units. Key and termination are therefore two systems with independent speaker choices for different meaning realizations. Key and termination choice in a particular tonic segment is “never more than one ‘level’ in the three-term system” (Cheng *et al.*, 2008: 62), namely one step above or one step below.

Brazil (1997) distinguishes between minimal and extended tonic segments, depending on whether the tonic segment contains one or more than one prominence. In tone units of both the minimal and extended types, pitch-level choices serve to determine the key and termination of the whole tonic segment. In the case of minimal tonic segments, however, it is not possible to make the selection of key and termination independently. In single prominence tone units without an onset syllable, “the first prominent syllable is also the last, so there can be no independent choices in the two systems” (p. 12), representing a simultaneous selection of key + termination. As an illustration, Table 19.5 shows some examples of tone units with different pitch-level choices of key and termination on the word *so*:

Table 19.5 Examples of key and termination pitch-level choices

Pitch-choice	Examples
high key	{\ [^ SO] at this < MOment >}
mid key	{\ [SO] this is ONE < REAson >}
low key	{\ [_ SO] they have a DIFFerent < SYStem >}
high termination	{\ [BUT] < ^ SO > far}
mid termination	{= [THANK] you < SO > much}
low termination	{\ you [DON'T] THINK < _ SO >}
high key + termination	{\ and [< ^ SO >] you know}
mid key + termination	{\ okay [< SO >]}
low key + termination	{\ [< _ SO >]}

This method of analysis needs to be justified by showing that “the meaning increments derived from the two choices are compatible and both appropriate to the situation” (ibid.). The hearer then assigns communicative value to either key or termination in the local context. Nevertheless, there are conditions where a syllable in a tone unit that is not intended to realize a sense selection can be made prominent, simply in order to achieve a choice of key or termination, since “an intonation choice can be associated with a syllable only if it is prominent” (p. 65).

Key, defined as the pitch choice on the first prominent syllable, “affects the communicative value of the whole tonic segment” (p. 50). The selection of key projects the speaker’s assumption about the hearer’s expectations as the talk unfolds. High key, for example, has contrastive implications and may show surprise, pleasure, annoyance, alarm, and so on in the local context. “Contrasting” here refers to a selection that “projects a binary opposition upon the existential paradigm and explicitly denies an alternative” (p. 45). In other words, the speaker is indicating clearly a choice made out of two options (it is “a” NOT “b”) through the use of high key. High key adds an increment of meaning to the effect that “this tone unit has a denial of expectation relations to what has preceded” (Brazil, 1997: 75–84; Cauldwell, 2007). In example 19 below (from a conversation between friends), the speaker is discussing the desserts in a foreign city that she has just visited. Notice the use of high key in the third tone unit on the word *so* to indicate a sharp contrast, and hence a surprise that the speaker feels at the variety of cakes available between the city she travelled to and her own city:

(19)  
 ... {= [< ER >]} {= [< THERE >]} {= there are [< ^ SO >]} {= many different [< TYPES >] of um} {\ [< CAKES >]}...

The selection of low key in a tonic segment projects existential equivalence to the previous topic segment. Low key has equative value, adding an increment of meaning to the effect that “[t]his tone unit has an equative relationship with what has gone before” (Brazil, 1997: 75–84; Cauldwell, 2007). In other words, low key assumes that the hearer will perceive the content as following naturally upon what has gone before, and as “being entirely in line with what the hearer would expect” (Brazil, 1995: 245). In example 20 below (from a placement interview at a hotel), the interviewer is asking the student about the department that she would like to work in. The use of low key in the final tone unit, in combination with the lexical choice *you know*, signals the alignment of expectation to what has preceded, as the beginning of one’s career is normally expected to follow one’s graduation:

(20)  
 ... {= [< ER >]} {/ [< ^ WHICH >]} {= department [THAT] < UM >} {/ you’re [parTicularly] interest < IN >} {= and would like [< TO >]} {∨ [emBARK] your < caREER >} {\ [\_ YOU] know UPON < graduAtion >}

Mid-key attributes no special expectations to the hearer. It only has additive value, “merely adding its content to what has gone before” (Brazil, 1995: 245). In example 21, which is again taken from a placement interview, the student is explaining why she is interested in working in the hotel industry and studying hotel and management subjects at her university. Her repeated selection of mid-key shows that she is simply expanding on her reasons for her study and career choice:

(21)  
 ... {= i [< THINK >] in} {∨ [THIS] < SUBject >} {= i [CAN] LEARN < aBOUT >} {= a [LOT] of er < PRACtical >} {= [< ER >]} {= [< ER >]} {\ [< KNOWledge >]} {= [SO] i THINK it is good for < ME >}...

Given the additive value associated with mid-key, it is perhaps not surprising that this key is found to be the most frequently occurring of all the pitch-level choices for key, more than 90 percent of all tone units in the HKCSE (prosodic) having been found to be produced with this key choice (Cheng *et al.*, 2008).

The selection of termination constrains the next speaker in his/her selection of key. This phenomenon is termed “pitch concord” (Brazil, 1995: 86). A speaker who conforms to pitch concord is likely to be giving a preferred response, and a speaker who does not is likely to be giving a dispreferred response (pp. 53–58).

TERMINATION	HIGH anticipates HIGH KEY response (i.e. adjudication)
	MID anticipates MID KEY response (i.e. concurrence)
	LOW sets up no particular expectations, and permits choice of high key, mid key or low key.

(Brazil, 1995: 246, 1997: 119)

Example 22 is taken from the beginning of a conversation. In this extract two friends are checking whether the recording has started.

(22)

a: {\ [< < YES > ]} {= i [THINK] it's < ER >} \* {\ [STARTED] < alREAdy >}

A: \*\* {\ [< < STARTED > ]}

\ [< < ^ YEAH > ]}

A: \* {\ [< < ^ Okay > ]} {\ [< < ^ SO > ]}

a: \*\* {\ [< < YES > ]}

At the beginning of the extract, speaker a selects high key + termination on *yeah* to seek adjudication from speaker A regarding whether the recorder is on. By providing a preferred response, which in this case is agreement with her friend's observation, speaker A also selects high key + termination on *okay* and *so*, hence achieving pitch concord.

On the other hand, example 23 shows a dispreferred response given in a televised interview:

(23)

b1: {\ [< < \_ NOW > ]} {? [< < K > ] w} {? do you} {= do you [aGREE] with < THIS >}

b2: {\ [< < ^ WELL > ]} {? [< < I > ]} {? [< < I > ]} {\ i [< < ^ THINK > ]} {= [< < ER > ]} {= i [aGREE] < WITH > er} {\ er h [\_ W] < BUT > } {= i [< < ^ THINK > ] we} {= we [NEED] to < LOOK > at the} {= [< < THIS > ] er} {= [< < ISsue > ] from a} {= from a [< < BIGger > ]} {\ [< < CONtext > ]}...

While speaker b1 selects mid-termination to seek concurrence, speaker b2 does not choose mid-key in his response. Instead he selects high key, to indicate a contrast, and suggests that he does not entirely agree with the other guest in the interview.

## Conclusion

This paper has presented an overview of the discourse intonation systems and choices and has illustrated each of the four systems through discussions of examples from the HKCSE (prosodic) to show how the systems function in local contexts so as to add communicative value to what is said. It highlights the fact that intonation is situation-specific. For instance, it can be seen from the examples above that the discourse particle *so*, among other lexical items, can be used with a range of intonational choices. Instead of rigidly tying particular lexical or grammatical elements to

intonation, the discourse intonation model thus argues that the use of intonation is context-sensitive and very much responsive to the communicative situation. While such intonational choices, consciously made by speakers, are not pre-determined, they can nonetheless be predicted to some extent through the detailed systematic observation, identification and description of general patterns. The study of the discourse intonation patterns observed in naturally occurring speech thus reveals to the analysts the intonational decisions that speakers have to make in real time interactions on a moment-by-moment basis, which reflect the rich layers of pragmatic and situated meanings expressed in speech. This is therefore an indispensable area of investigation for a deeper understanding of meanings in spoken discourse.

## Transcription conventions

### *Transcription notation used in the HKCSE (prosodic)*

<i>Symbol</i>	<i>Remarks</i>
...	parts of an utterance which have been omitted
★	onset of simultaneous speech produced by the current speaker
★★	onset of simultaneous speech produced by an interlocutor other than the current speaker
(.)	a brief, unfilled pause roughly lasts for the length of a syllable
(pause)	a unit, unfilled pause which is longer than a brief pause and normally lasts for a few seconds
(())	a non-linguistic feature such as laughter, coughing, throat clearing and applause
((inaudible))	unintelligible speech
A:	female native speaker of English
B:	male native speaker of English
a:	female Hong Kong Chinese
b:	male Hong Kong Chinese
x:	female speaker of a language other than English and Cantonese
y:	male speaker of a language other than English and Cantonese
u:	unknown speaker
{ }	tone unit boundary
/	rise tone
∨	fall-rise tone
\	fall tone
∧	rise-fall tone
=	level tone
?	unclassifiable tone
CAPS	prominent syllable
[ ]	key
< >	termination
^	high pitch level
_	low pitch level



## Further reading

Brazil, D. (1985) *The Communicative Value of Intonation*. Birmingham: English Language Research, University of Birmingham.

An important and original work on the study of discourse intonation, this book provides a detailed description of the discourse intonation framework.

Brazil, D. (1997) *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press. This is the revised edition of Brazil's (1985) seminal work.

Cheng, W., Greaves, C., and Warren, M. (2008) *A Corpus-Driven Study of Discourse Intonation: The Hong Kong Corpus of Spoken English (Prosodic)*. Amsterdam; Philadelphia: John Benjamins.

This monograph discusses the discourse intonation patterns observed in the Hong Kong Corpus of Spoken English (Prosodic), one of the largest corpora of naturally occurring speech annotated with the discourse intonation framework.

## References

Brazil, D. (1985) *The Communicative Value of Intonation*. Birmingham: English Language Research, University of Birmingham.

Brazil, D. (1995) *A Grammar of Speech*. Oxford: Oxford University Press.

Brazil, D. (1997) *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.

Cauldwell, R. T. (2002) 'The functional irrhythmicity of spontaneous speech: a discourse view of speech rhythms', *Apples*, 2 (1): 1–24.

Cauldwell, R. T. (2007) *SpeechinAction* Research Centre (SPARC). Available online at: <http://www.speechinaction.com> (accessed 16 June 2007).

Cheng, W., Greaves, C., and Warren, M. (2008) *A Corpus-Driven Study of Discourse Intonation: The Hong Kong Corpus of Spoken English (Prosodic)*. Amsterdam, Philadelphia, PA: John Benjamins.

Lam, P. (2008) *Discourse Particles in an Intercultural Corpus of Spoken English*, PhD dissertation, The Hong Kong Polytechnic University, Hong Kong.

Sinclair, J. (1991) *Corpus, Concordance, Collocations*. Oxford: Oxford University Press.