An Investigation of the Empirical Basis for

Discourse Intonation as a Model of American English Intonation

Mark McAndrews

Northern Arizona University

Abstract

Discourse Intonation (DI) is a model of intonational forms and meanings that was developed to describe British varieties of English (Brazil, 1997; Brazil, Coulthard & Johns, 1980). It has subsequently been applied to American English for linguistic research and pedagogical purposes (e.g., Chun, 2002; Gorsuch, 2013; Gorsuch, Meyers, Pickering & Griffee, 2012; Kang, Rubin & Pickering, 2010). However, there is a lack of published empirical evidence that supports the suitability of DI as a description of American English intonation. Furthermore, DI-based investigations assume that observed differences in productive intonation between English as a Second Language (ESL) learners and native speakers reflect learners' lack of skill in matching appropriate pitch movements with desired meanings, ignoring the possibility that these differences are instead due to mismatched sociopragmatic norms (e.g., Pickering, 2001). To investigate these shortcomings, the current study employed a sequential exploratory mixed methods design: In Phase 1, native speakers of American English and Chinese learners of English completed an open-ended questionnaire meant to elicit sociopragmatic expectations related to two meanings that are central to DI: friendliness and interactional control; in Phase 2, participants matched carefully controlled audio clips with speakers in hypothetical interactive situations that were informed by Phase 1 responses. Quantitative results provided some preliminary support for DI as a model of American English, as well as the proposal that learner difficulties come from both linguistic and sociopragmatic sources.

An Investigation of the Empirical Basis for

Discourse Intonation as a Model of American English Intonation

Background

Prosody plays a crucial role in spoken communication at various levels of linguistic

analysis: lexical, syntactic, semantic, discoursal, and pragmatic. Therefore, justifiably, prosodic

features of language have received a good deal of attention in second language studies. One

model of English intonation, known as *Discourse Intonation* (Brazil, 1997; Brazil, Coulthard &

Johns, 1980), has become a popular framework for pedagogical and research purposes. The

current study aims to address some of the potential weaknesses of applications of this theory,

namely its lack of empirical support, and the need to distinguish linguistic and sociopragmatic

sources of observed differences among speakers.

Discourse Intonation

The model of English intonation known as *Discourse Intonation* (henceforth DI) was

established by British linguists at the University of Birmingham in the 1970s and 1980s. Like

other theories of intonation, it attempts to describe the ways in which suprasegmental

characteristics of speech are systematically manipulated by speakers to communicate meaning.

What is more or less unique about DI is the set of meanings it associates with various types of

pitch movement. These distinctive pitch contours are known in DI as *tones*. The inventory of

tones in DI consists of: fall, fall-rise, rise-fall, rise, and level.

The most fundamental meaning distinction proposed by DI is between the fall and fall-

rise tones. It is proposed that fall-rise tones mark the information in the containing tone unit as

previously being among the shared knowledge of speaker and listener. In other words, propositions associated with fall-rise tones are already part of the common ground that exists between interlocutors. For this reason, the fall-rise tone is part of a category known as *referring* tones. In contrast, fall tones are associated with information that the speaker wishes to add to the common ground, and are part of the *proclaiming* tone category. Importantly, DI contends that the information structuring function of referring tones also carries sociopragmatic meaning. It is proposed that, by singling out information that is already shared between speaker and listener, referring tones are also expressions of solidarity. Brazil (1997) describes this phenomenon as a "verbal hand-on-your-shoulder gesture" (p. 80).

A second key claim made by DI concerns the change in meaning it associates with a speaker's choice of rise-fall instead of fall, and rise instead of fall-rise. DI considers rise-fall and fall tones as both members of the proclaiming category, and rise and fall-rise as members of the referring category. There is an incremental change in meaning accomplished by choosing the rise-fall or rise instead of the associated unmarked tones described above. DI suggests that the function fulfilled by this choice is related to the sociopragmatic meaning of *controlling* or *dominating* an interaction. While a speaker who produces a rise tone is still singling out a proposition as already existing in the common ground, by choosing this tone instead of a fall-rise, he or she, it is claimed, is asserting control over the interaction. Brazil (1994) provides an example of a chairperson at the opening of an academic meeting. The chairperson follows implicit guidelines about how meetings begin, including greeting members, making announcements, and introducing other speakers. This person is "in charge" (p. 56), and would be expected to assert this control through the use of rise tones. Table 1 summarizes the key claims of DI described thus far.

Table 1

Tones and Their Associated Meanings in the Theory of Discourse Intonation

Tone	Primary Meaning	Secondary Meaning
fall	proclaim new information	-
fall-rise	refer to information in the common ground	convey solidarity
rise-fall[A]	proclaim new information	assert control
rise	refer to information in the common ground	assert control

[A] This tone is infrequent and its use is restricted to a small number of specific contexts.

Research Questions

Two problems emerge from the application of DI to second language teaching and research. First, there is little or no evidence that the intonational knowledge of native speakers of American English conforms to the DI model. Second, previous research has failed to disentangle the potential sources of observed differences between native and non-native speakers, namely linguistic and sociopragmatic knowledge. The current study takes a step towards addressing these issues by posing the following research questions:

1. Do native speakers of American English and Chinese learners of American English associate:

    a) fall-rise tones with friendliness?

    b) rise tones with interactional control?

2. Are there differences between native speakers of American English and Chinese learners of American English in terms of the interactive situations with which they associate expressions of:

    a) friendliness?

    b) interactional control?

Methods

The current study adopted a sequential exploratory mixed methods design (Creswell, 2013), and was comprised of two phases. The goal of Phase 1 was to identify interactive situations with which Chinese and native-speakers of American English associated the sociopragmatic meanings of friendliness, and control. This goal was pursued through open-ended data collection and qualitative analysis in order to capture participant-generated interactive situations, instead of imposing restrictive, a priori classifications. Crucially, the patterns that emerged from the qualitative analysis were used to inform the design of items for the Phase 2 instrument, a speech judgment task. Thus the goals of Phase 2 were two-fold: first, to test whether the cross-cultural patterns that emerged in Phase 1 were generalizable to different samples of the populations, and second, to obtain a quantitative measurement of the extent to which participants matched tones and meanings in the DI framework.

Materials

Phase 1: Open-ended questionnaire. Participants were asked to describe situations that occur on campus in which they would expect one person to express friendliness, or control. By limiting the range of possible interactive situations to only those that frequently occur on campus, we intended to elicit responses that were constrained to that domain, and thereby increase overlap between participants and strengthen potential patterns in the qualitative data. See Appendix A.

Phase 2: Speech judgment task. In the second phase of the study, a speech judgment task was used to investigate participants' sensitivity to the tonal distinctions being studied (i.e., fall versus fall-rise, and fall-rise versus rise). The instrument was delivered online using the Qualtrics platform. It consisted of 20 questions in total, eight of which were devoted to the

friendliness construct, and eight to the control construct. Both sets of eight questions were further

divided into those that were based on interactive situations which emerged from the Chinese and

native-speaker groups' responses to the questionnaire in Phase 1. As a result, all participants in

the speech judgment task, whether Chinese or native-speaker, answered some items that

originated from within their cultural group, and some items that originated from the other group.

Four additional questions consisted of non-experimental items which were designed to be

answered unambiguously: participants had to identify the sounds of common animals, like dogs

and cats (these are referred to as *quality control items* in table 2). An incorrect answer to one of

these items suggested that a participant was just clicking randomly, which triggered the removal

of their data from the study.

 Individual items in the Phase 2 task consisted of a brief written description of an

interactive situation, and below that, two embedded sound clips, labeled "A" and "B."

Participants were instructed that each item involved a situation in which one person was

speaking, and that the situations all occurred on campus. They were asked to read the description

first, and imagine what the person in the situation might sound like. The following example

situation was provided to familiarize participants with the format of the activity: *Sean is a*

freshman student. Sean's roommate has been under a lot of stress lately, due to school and work

commitments. Sean has just heard that his roommate failed a midterm exam. Try to imagine what

Sean's voice sounds like as he speaks to his roommate. Once participants had read the situation,

and imagined how the target individual might sound in that context, they were asked to listen to

the two short sound clips as many times as they wanted. Finally, participants had to select the

sounds clip, "A" or "B" that they felt was more reminiscent of speech of the person in the

described situation.

Audio stimuli. The sound clips used in the 16 experimental questions were comprised of a total of six unique recordings. These six recordings were cross-spliced together from shorter segments taken from the audio companion to Brazil (1994), and were thus originally designed to be illustrative of the DI principles being investigated in the current study. Each short recording was roughly four seconds long, and consisted of three "tone units" separated by 250 milliseconds of silence. Each tone unit was composed of between three and six words. Crucially, each tone unit was associated with one tone (i.e., meaningful pitch movement). The identities of the three tones in a single clip were manipulated by cross-splicing different tone units together. The identities of the first two tones in the sequence were varied systematically. The final tone did not vary: it was always falling. Audio stimuli therefore followed one of three tonal patterns: *fall+fall+fall*; *fall-rise+fall-rise+fall*; or *rise+rise+fall*.

Two distinct male voices produced the stimuli (these are referred to as *voice X* and *voice Y* in Table 2). In any given item, both stimuli came from the same speaker. Each interactive situation appeared twice in the questionnaire, once with the stimuli from each speaker.

Table 2

Structure of Item Distribution in Speech Judgment Task

20 total items																
8 items targeting *friendliness* construct								8 items targeting *control* construct								4 quality control items
Situation 1[A]		Situation 2		Situation 3		Situation 4[B]		Situation 5		Situation 6		Situation 7		Situation 8		
X[C]	Y[D]	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	

[A] Black cells denote situations that originated from a participant's own culture group (i.e., Chinese or native speaker)
[B] Grey cells denote situations that originated from outside a participant's own culture
[C] refers to male voice "X"
[D] refers to male voice "Y"

Table 2 summarizes the structure of the items in the speech judgment task. Items were presented to participants in random order. All stimuli recordings were subjected to a low-pass filter which removed acoustic information above 500 Hz, resulting in obscured segmental information (i.e., words were unintelligible), but which left suprasegmental features intact. See Appendix B for a complete list of items.

Participants and Procedure

Phase 1. Participants who completed the open-ended questionnaire were 9 Chinese learners of English studying at an intensive language program, and 15 native speakers of American English who were undergraduate students enrolled in a linguistics course at an American university. In terms of English proficiency, the Chinese participants were assessed at between 32 and 69 on a TOEFL-equivalent scale by the language program where they were studying. Chinese participants were given the option of responding to the questionnaire in Chinese. All participants completed the questionnaire online in a location of their choosing.

Phase 2. Ten Chinese learners of English and 15 native speakers of American English completed the speech judgment task. Crucially, the pool of participants in Phase 2 did not include any of the Phase 1 participants. In terms of English proficiency, Phase 2 Chinese participants were between 45 and 56 on the TOEFL-equivalent scale administered by the language program where they were studying.

Analysis

Phase 1: Qualitative analysis. Responses to the open-ended questionnaire were coded along three dimensions: the action performed during the interaction, the relationship between interactants in terms of social distance and level of power, and the level of any imposition that was associated with the interaction. Based on iterative rounds of reviewing the coded data,

coherent categories emerged among the responses of Chinese and native-speaker participants. The frequency with which each category was manifested in the data was recorded. Frequent categories were taken to be representative of potentially generalizable characteristics of the populations being investigated.

Phase 2: Quantitative analysis. To test research question one, 1-sample t-tests were performed to determine whether participants successfully matched tone sequences with within-culture sociopragmatic meanings at a rate that was significantly greater than chance. To test research question 2, 2-sample t-tests were conducted to compare participants' scores on within-culture items to outside-culture items.

Reliability. Reliability of different parts of the quantitative instrument were calculated according to the recommendations put forward in Brown (2005). Kappa coefficient estimates were obtained by first calculating appropriate standardized cut-point scores and K-R20 values, then matching these figures with equivalent coefficient estimates in the table provided (p. 205; adapted from Subkoviak, 1988). The raw cut score was set at 66.6%.

Results

Phase 1: Qualitative Results

The results of the qualitative data analysis show that the most frequent category for both friendliness and control was the same for both groups. Responses are summarized in Tables 3 and 4.

Table 3

Most Frequent Categories of *Friendliness* Situations

Native-Speaker Group	Chinese Group
1. Offering "small" help (various relationships)	1. Offering "small" help (various relationships)
2. Comforting (various relationships)	2. Offering service (university employees to students)
3. Seeking "medium" benefit (various relationships)	3. Comforting (various relationships)

Table 4

Most Frequent Categories of *Control* Situations

Native-Speaker Group	Chinese Group
1. Giving information (expert to novice)	1. Giving information (expert to novice)
2. Enforcing rules	2. Giving information (among peers and intimates)
3. Managing peers	3. Managing peers

Phase 2: Quantitative Results

Research question 1. In terms of matching tone sequences with sociopragmatic

meanings within the DI framework, both Chinese and native-speaker groups performed at above-

chance levels for both constructs. For *friendliness*, the overall mean score for Chinese

participants for within-culture situations was 53%. The overall mean of native-speakers for

within-culture *friendliness* items was 58%. Within the *control* construct, Chinese participants

had a mean score of 60% for within-culture items. The figure for native-speakers was 59%. None

of the 1-sample t-tests were significant at alpha = 0.05.

Research question 2. To see if the sociopragmatic differences identified in Phase 1 were generalizable to another sample of the populations being investigated, comparisons were made between within-culture and outside-culture items scores for both native-speaker and Chinese groups. For *friendliness*, the Chinese group scored 45% on outside-culture items, 8% lower than on within-culture items. The native-speaker group scored 50%, also 8% lower than they did on within-culture questions. For *control*, the Chinese group scored 50% on outside-culture items, which is 10% lower than their mean score for within-culture items. The native-speaker group scored 73%, which is 14% higher than their within-culture mean score. Thus in three out of four cases, participants performed worse on items that were based on situations that emerged from outside their own cultural group. However, a closer inspection of the outside-culture items revealed that they were generally low-performing items for both Chinese and native-speaker groups. The one exception was the *offering service (university employees to students)* category. Chinese participants scored 65% on these items, and native-speakers only 50%, as mentioned above.

Reliability. Only within-culture items were included in calculating reliability, and the Cohen's kappa estimates were generally low for the different sections of the speech judgment instrument. For *friendliness* items, the kappa estimates were 0.33 and 0.06 for native-speakers and Chinese participants, respectively. For *control* items, the kappa estimates were 0.19 and 0.33 for native-speakers and Chinese participants, respectively. Generally, the kappa estimates reflect the raw scores on the different sections of the instrument, except for the native-speaker group's performance on *control* items. The 0.19 kappa value is relatively low given that the mean score for native-speaker participants on this section was 59%. This could be due to a specific problematic item.

Relevance to PIE and Second Language Learning

The picture painted by the quantitative Phase 2 results is decidedly unclear. In regards to the current study's first research question, some evidence was found to support the proposal that native-speakers of American English associate fall-rise tones with friendliness, as suggested by DI. For Chinese participants, however, their average score on items that tested this form-meaning pairing was only 53%. Furthermore, the estimated reliability of this section of the speech judgment instrument was only 0.06 for Chinese participants. These results seem to support the conclusion that, in general, Chinese learners do not have this association in their knowledge of English (Pickering, 2001). For the *control* construct, both groups associated rising tones with interactional control, as predicted by DI. These results can only be treated as preliminary and must be interpreted with caution, due to a lack of statistical significance and generally low reliability of the instrument.

In regards to the current study's second research question, only one sociopragmatic difference that emerged in Phrase 1 seemed to generalize to the new groups of participants in Phase 2: the situation in which university employees are friendly towards students. Chinese learners tended to expect university employees to express friendliness to students, while native speakers did not.

This study found only weak evidence to support DI as a valid model of American English intonation. Lack of statistical significance and reliability could be due to methodological shortcomings. On the other hand, the current study did not find evidence against DI in the American context. In terms of implications for PIE and English learning in general, this study supports the claim made in Pickering (2002) that Chinese learners tend to lack the linguistic knowledge to map fall-rise tones with friendliness. This could be addressed in a pronunciation

lesson. Finally, the current study highlights the importance of including instruction on American sociopragmatic norms, since a lack of expertise in this domain could be a source of interactional difficulties.

References

Brazil, D. (1994). *Pronunciation for advanced learners of English*. Cambridge: CUP.

Brazil, D. (1997). *The communicative value of intonation in English book*. Cambridge, UK: Cambridge University Press.

Brazil, D., Coulthard, R. M., & Johns, C. (1980). *Discourse intonation and language teaching*. Harlow, UK: Longman.

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. New York, NY: McGraw-Hill College.

Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Amsterdam: John Benjamins.

Creswell, J. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Gorsuch, G. J. (2013). Helping international teaching assistants acquire discourse intonation: Explicit and implicit L2 knowledge. *Journal of Teaching English for Specific and Academic Purposes*, *1*(2), 67-92.

Gorsuch, G., Meyers, C., Pickering, L., & Griffee, D.T. (2020) *English communication for international teaching assistants*. Long Grove, IL: Waveland Press.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*, 554-566.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*, 233-255.

Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement*, *25*(1), 47-55.

Appendix A – Open-ended Questionnaire

Thank you for your time! Please read the following instructions carefully:

In this questionnaire, you will be asked to describe situations in campus life where people

verbally interact with each other in TWO specific ways (Question 1 & Question 2). You will

be asked to describe TWO different situations for each question.

Before you move on to Question 1 & Question 2, please take a few minutes to read the

following EXAMPLE question, and TWO possible responses. This example question and

possible responses will give you an idea of how to answer Question 1 & Question 2.

Example Question:

Describe TWO situations that often occur in campus life, where one person verbally expresses

sympathy.

Possible Responses:

SYMPATHY *Situation 1:*

Who *are the people involved in the interaction, and what is/are their relationship(s)?*

A student and his roommate. They are good friends.

Who *expresses sympathy?*

The roommate.

Why *does one person express sympathy?*

The student feels overwhelmed with assignments, and work and family commitments. The student describes this situation to his roommate and explains that he is feeling stressed. The roommate wants the student to feel supported.

When *does one person express sympathy?*

After the student has explained his situation to the roommate, the roommate takes a minute to express sympathy.

Where does the interaction take place?

In a dorm room.

Thank you for reading the example question and possible responses. Now, please answer Question 1 and Question 2. Please describe TWO situations for each question. Remember, these situations could take place anywhere on campus, and they involve people speaking (not just using body language). Try to think of situations that are very common. Answer the questions Who, Why, When, and Where to create a detailed description of the interaction.

Question 1:

Describe **two** *situations that often occur in campus life, where one person verbally expresses friendliness.*

FRIENDLINESS Situation 1:

Who *are the people involved in the interaction, and what is/are their relationship(s)?*

Who *expresses* **friendliness***?*

Why *does one person express* **friendliness***?*

When *does one person express* **friendliness***?*

Where *does the interaction take place?*

Question 2)

Describe **two** *situations that often occur in campus life, where one person* **controls the flow of speaking***. (The person in control would feel comfortable shifting to a new topic, and ending the interaction. The person without control would not feel comfortable interrupting, changing topics, or ending the interaction.)*

CONTROL Situation 1:

Who *are the people involved in the interaction, and what is/are their relationship(s)?*

Who *expresses* **control***?*

Why *does one person express* **control***?*

When *does one person express* **control***?*

Where *does the interaction take place?*

Thank you so much for your time!

Appendix B – Speech Judgment Items

Friendliness Construct Items
1. Rashad is an undergraduate student. It is the beginning of the semester. Another student, who Rashad has never met, approaches Rashad for directions to the Biology building. Try to imagine Rashad's voice as he speaks to the other student.

2. Sam is a Teaching Assistant. One student in Sam's class will make a presentation during today's class. Before the class starts, Sam notices that this student seems very nervous. Sam walks over to the student. Try to imagine Sam's voice as he speaks to the student.

3. Mo is an undergraduate student. He is attending a professor's office hours to ask some questions, and hopes to receive extra help. Try to imagine Mo's voice as he speaks to the professor.

4. Bill works at the front desk of a residence hall. A freshman student living in that residence approaches Bill to notify him of a problem with his sink. Try to imagine Bill's voice as he responds to the student.

Control Construct Items
5. Dr. Jackson is a professor of Chemistry. Try to imagine his voice as he delivers a lecture in an undergraduate class.

6. Ben is an undergraduate student. He is working on a group project with two classmates. Ben is sure that he has great ideas for the project, and wants the rest of the group to follow them. Try to imagine Ben's voice as he describes his plan to the other group members.

7. Stewart is a Residence Assistant. A student resident on Stewart's floor has broken a rule about playing loud music late at night. Try to imagine Stewart's voice as he speaks to the student resident.

8. Sal is an undergraduate student. He is playing a video game with his roommate. Sal has more experience with the game's strategies, and is a better player than his roommate. Try to imagine Sal's voice as he talks to his roommate about game strategy.