

Visualizing Spoken Discourse: Prosodic Form and Discourse Functions of Interruptions

Li-chiung Yang

CREST, Japan Science and Technology & Information Sciences Division, ATR
Seika-cho Soraku-gun, Kyoto, Japan 619-0288

yang@isd.atr.co.jp

Abstract

In this paper we show that interruptions are important elements in the interactive character of discourse and in the resolution of issues of cognitive uncertainty and planning. By representing discourse graphically, we also show that interruptions are part of the local and global coherence that is brought about through the systematic phrase-to-phrase prosodic patterns of discourse. The specific pitch height of the interruption varies with the expression of emotion, signals of attention-getting, and signals of competitiveness. These prosodic forms are potentially usable in spoken dialogue systems to provide intelligent responding systems that are responsive to human motivations in dialogues.

1 Introduction: Interruptions and Dialogue

One characteristic of human conversation is that it's highly interactive, spontaneous and mutual information building, and the demands of the ongoing mutual negotiation process often cause imbalances in informational adequacy and desired topic direction. Interruptions play a key role in signaling and resolving these imbalances and in bringing about a mutually satisfactory accommodation of the interests and knowledge states of participants.

Because interruptions act to mediate the content and redirection of a conversational exchange, and are informationally packed with respect to these communicative pivot points, it is important to understand how interruptions are used in human communication, and determine which elements of interruptions can be accommodated in building a more flexible, more efficient spoken dialogue system.

1 Research Goals and Procedures

In this study, our goal is to look at the distribution of interruptive occurrences in natural speech, and investigate their respective functions and characteristics. Several questions that we address are the following: What are the different types of interruptions present in dialogues, and to what extent are prosodic-acoustic features significant in distinguishing between these different types of interruptions? What are some of the underlying factors that cause interruptions to occur, and how can such information on the prosodic features be utilized in spoken language systems in detecting interruptions and in constructing appropriate response strategies in human-computer interactions?

Our data for this research consists of fifteen dialogue segments extracted from a corpus of 2 hours of spontaneous conversation. The speech data were digitized and annotated for discourse relations, topic structure, interruptions, and speaker turns. The acoustic measurements of f0, amplitude and duration were correlated with the specific characteristics of the interruptions in the data. In this paper we concentrate mostly on pitch, but make reference to amplitude and duration where appropriate.

In our analysis we take a multi-level approach. In order to capture the different domains at which prosodic patterns are manifested, we analyzed the data at the within-phrase as well as the inter-phrase level. An additional level of our analysis focuses on how discourse evolves over extended stretches of conversation. As a way of representing the prosodic structures in spoken dialogue, we plotted the highest pitch points of 600 continuous utterances, about 20 minutes conversation, for each speaker and color-coded the interruptions by speaker and type. This

technique allows us to visualize and track important discourse events and points of interest easily, and allows us to form more appropriate generalizations accordingly.

2 Prosodic Characteristics of Interruptions

2.1 Types of Interruptions

What constitutes an interruption? Interruptions can be seen as situations in which one person intends to continue speaking, but is forced by the other person to stop speaking, at least temporarily, or the continuity or regularity of that person's speech is disrupted. This can happen when the interruption causes the main speaker to become hesitant in speech, even while continuing on the intended path or when the speaker continues speaking, but the interruption causes the speaker's topic direction to be modified. Interruptions, therefore, can be seen as consisting of three essential ingredients: intention of the main speaker to continue, entrance of the other person into the conversation, and disruption or stopping of the main speaker, at least temporarily.

In general, interruptions can be of two types: competitive vs. cooperative. Competitive interruptions occur when one speaker attempts to take the floor by making his or her own remarks a higher priority over the main speaker's speech when the main speaker intends to continue. This competitiveness can be on two different levels. Speakers can compete for speech space and they can also compete for topic or idea. In either of these competitive cases, interruption acts to take the attention away from the main speaker, at least temporarily, and focus it on the interrupter's speech.

Cooperative interruptions occur when one speaker wants to support or reinforce the main speaker's point without disrupting the main speaker's continuation. These types of supportive remarks are often in the form of short commentaries or clarifying questions. Such clarifying questions often support the continuing flow of the main speaker by keeping both speakers in synchrony on the topic development.

Both types of interruptions may or may not involve overlapping speech, since overlapping speech is not necessarily an indication that speakers are in conflict over which speaker has the right to the floor at that moment. For

example, in conversation, speaker's feedback utterances or back channel signals such *umhum* or *yeah* often overlap with the main speaker's speech, but they are not interruptions as they do not interrupt the main speaker's flow. In fact, they often contribute to the smooth flow of the main topic because of their supportive nature. Conversely, competitive interruptions can happen even when there is no overt overlapping in speakers' speech. This can occur when one speaker is not completely finished and intends to go on, but is at the end of an episode or a possible turn completion point. The other speaker may not be aware of the main speaker's plan, and may think the current speaker is completely finished, so starts to take the floor and creates an unintentional interruption, i.e. *mistiming*. Therefore, it is the degree of disruption to the intended continuation of the main speaker which is the critical element, and the degree of competitiveness or cooperation is determined by the actions and intentions of both speakers.

2.2 Competitive Interruptions

Analysis of our discourse corpus shows that competitive interruptions are typically high in pitch and amplitude. In spontaneous discourse, speakers often compete to gain control and dominance in the conversation. In competitive situations participants need a strong immediate signal to attract the attention away from the ongoing speech. In general, the more audible the signal is, the more forceful and effective it will be in overcoming the current focus and successfully taking the floor. Prosodically, this competitiveness and need for a strong signal are iconically reflected in the vocal cues of high pitch and high amplitude.

Competitive interruptions are often closely tied to topic development and reflect relevance, urgency, degree of importance, and interest in the current topic. In conversation, speakers often feel the need to express something which is emotionally significant to them. Speakers often encounter moments of uncertainty and have an urgent demand for information and immediate attention at a critical moment. This urgency and immediacy are a key characteristic of interruptions and are directly related to the relevance of the current topic. Speakers often grab the opportunity while the current topic is hot to clarify something, add a pertinent fact, or express

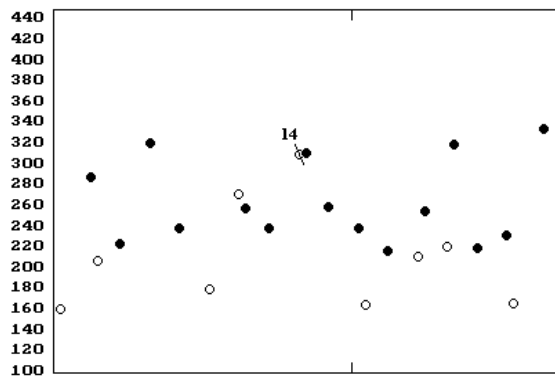


Figure 1: A competitive interruption of Speaker A for demanding new information at phrase unit 14. Circles mark Speaker A's utterances whereas filled dots mark Speaker B's utterances.

an immediate opinion. Often the high pitch and loud amplitude in competitive interruptions are caused by the emotions motivating these situations.

2.1.1 Demand for new information or clarification

The first example illustrates a typical case of competitive interruptions: (interruptions are indicated by an arrow in the transcript)

- (1)
 13 A: So this one's better then
 B: Umhum
 14 It's better than the regular ones.
 → A: | What brand is this?
 15 B: This one's -
 it's called - Marantz
 16 A: Oh.
 B: Marantz.
 Umhum.
 17 A: Five hundred?.
 B: Close to five hundred.

In this example the interrupter (A) comes in when the main speaker is at a slight pause in the middle of a response and at a low pitch level. The interrupter interrupts with a direct question 'What brand is this?' at U14 (utterance 14) using high pitch, loud amplitude and at a fast speed (see Figure 1). These prosodic features are direct results of the immediacy and urgency of the interrupter's demand for additional information for her interest.

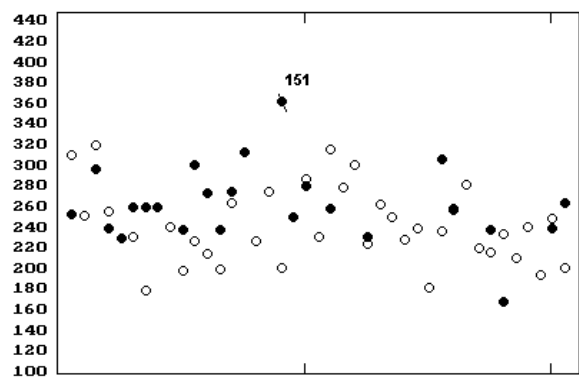


Figure 2: A case of an interruption of Speaker B at phrase unit 151 expressing an important point at a critical moment.

2.2.2 Expressing Strong Opinions

- (2)
 149 A: It's just - hmmm
 150 It's just to say that the one who speaks
 151 it's just that you - you - (pause)
 → B: But you have to speak slowly, right?
 152 It has to be very clear.
 A: Be | cause every -
 153 A: Right.
 154 Because everyone's pronunciation is
 different
 B: Umhum Right

Competitive interruptions within an ongoing topic also occur when a speaker wishes to express a strong opinion or disagreement. In the beginning section of this fragment shown in example 2, the main speaker (speaker A) is talking and speaker B mainly provides feedback. Speaker B's interruption at U151 occurs at a point where the main speaker is hesitant and pausing. Anticipating the main speaker's point, speaker B takes this opportunity to express her strong opinion on that point, and the forcefulness of her disagreement is reflected in the high amplitude and high pitch of the interruption. Comparing with the peak points for the utterances in this section (see Figure 2), we can clearly see that this interruption has a sudden pitch jump to 360Hz, and is an abrupt isolated point by comparison to the rest of the pitch points in this area, about 50Hz higher than the other points in this region. Note that the intention of the interrupter here is not to take the floor for the long term, but to *make an important point at*

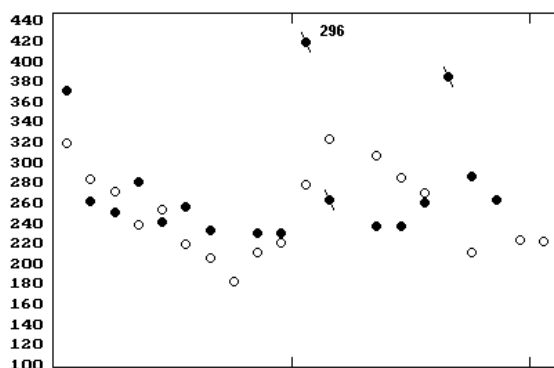


Figure 3: A very high-pitched interruption of Speaker B at U296 to shift topic.

a critical moment, and this intention is indicated by the brevity of her remark and her supportive feedback thereafter.

2.2.3 Shifting Topic

The critical moment urgency of many interruptions is shown in the above example. Interruptions often occur in the normal give-and-take of conversation as participants negotiate their own interests in the conversations. Therefore one key motivation for competitive interruptions is to change topic direction. This can happen when one speaker has a topic of greater interest, wants to avoid a topic, or wants to return to an old topic. Such interruptions often occur in the form of questions, as questions obviously are a natural way to attract attention, to demand information, and to direct or guide a speaker's speech and the direction of discourse.

- (3)
- 287 A: Then it's just - it's just teamwork
 288 it's not just
 B: | Right Umhum
 289 A: just they are doing the work
 290 B: Umhum umhum umhum
 291 A: They are also doing teamwork.
 B: Umhum
 292 A: Some people are responsible for the
 linguistic analysis,
 B: Umhum
 293 A: some people are responsible for the
 software design.
 B: Umhum umhum umhum.

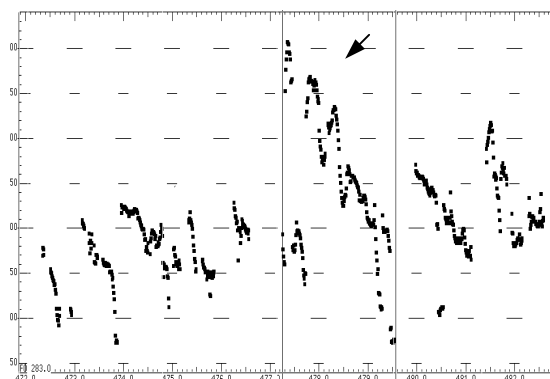


Figure 4: Pitch track of U296 with preceding and following utterances.

- 294 A: It's just - it's just teamwork.
 B: Umhum
 295 A: It also needs to be done like this,
 B: Umhum
 A: in order to do a good job.
 296 → B: | Then then then the conference at
 Central Research Institute, was that
 good?
 A: Cen - tral - Re - search - Insti - tute -
 B: Umhum
 297 A: Be cause that -
 → B: | Do you remember? M.

In example 3, the main speaker (Speaker A) is finishing up her topic, and her intention to conclude can be inferred by her repetition of the phrase 'It's just - it's just teamwork.' in U294 to tie the topic back to her beginning statement at U287. Her pitch level is getting low here. Anticipating Speaker A's completion, Speaker B comes in to shift the topic back to a previous topic. Her pitch level for this utterance (U296) is very high at 420Hz as seen in Figure 3 and Figure 4, in fact, it is one of the highest points for this speaker in the discourse. We can see that there is a dramatic and abrupt rise in pitch level. This is clearly indicated by the sharp increase of approximately 190Hz from Speaker B's previous utterance at 230Hz. Her amplitude is also loud and forceful. This interruption is followed by another lower-pitched and soft prompting interruptive question 'Do you remember?' to reinforce the intended turn in topic direction.

2.2.4 Degree of Topic Relatedness

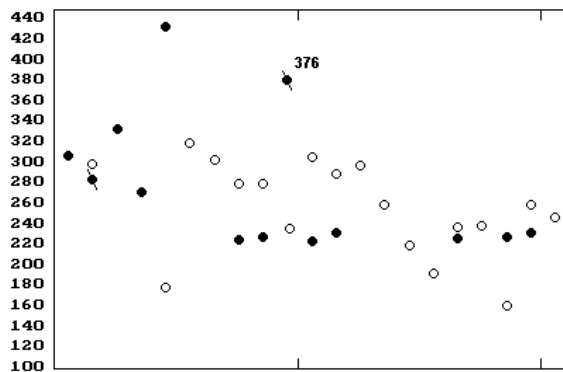


Figure 5: A high-pitched competitive interruption of Speaker B at U376 to shift to a related topic.

- (4)
- 371 A: Later that Computational Linguistics Conference
B: Umhum
- 372 A: that one's pretty good, too
373 that one's also pretty good.
B: Umhum
- 374 A: That one's just -
Oh, it's just the emphasis is more on computational linguistics,
B: Umhum
- 375 A: It's just the scope was narrower.
- 376 → B: | Does that have anything to do with what you are working on?
A: Ah...partially.

The pitch height of an interruption is closely related to the abruptness of topic shift and the intensity of expression. By contrast, for a later interruption in example 4 at U376, speaker B's pitch level is high at about 380Hz (Figure 5), but is about 40Hz lower than the interruption to shift the topic in the previous example of U296. The reason for the higher pitch level of the previous interruption may be explained on two grounds. One is that the *interest level* involved, i.e. the intensity of the emotion of the speaker, is different. In the previous interruption, the speaker is bringing in a topic in which she has great interest, whereas in the current case, the interruption is just a leading question to provide an opportunity for a further topic. The second reason concerns the *degree of relatedness of topics*. A greater cognitive effort is involved has not been present or active for some time, when a shift is made to return to a topic which therefore

requiring a stronger prosodic signal to flip back to the previous topic world, and to bring it back into the current memory of participants. In the example here, the topic shift is just one step away from the current topic which is in the participants' active memory, and thus this requires less cognitive effort, hence a less strong intonational signal.

2.2.5 Resistance to Topic Shift

Interruptions are an important element in the interactive character of discourse. This interaction comes about because of the mutual negotiating to satisfy each participant's needs in the conversation. In the above examples, topic shifts were viewed from the perspective of the interrupter, but the perspective of the main speaker also needs to be considered. When encountering interruption, a speaker may respond by yielding, by ignoring the interruption, or by continuing through forceful prosodic counter-measures. The particular response used is determined to a large degree by the existing balance of floor rights at that point in the conversation.

Whether a main speaker yields or not is decided by the degree of competitiveness and urgency of the interrupter, and how related the interruption is to the ongoing topic. In resisting interruptions, the main speaker often reacts by using both loud amplitude and a high pitch level, the principle being to first grab back the floor, then proceed with content. The prosodic give-and-take of interruptions really expresses the interactively established understanding of current floor rights and participants' intensity.

- (5)
- 75 B: Then then at the next step
76 it's already digitized,
A: Umhum
- 77 B: It's already stored on the computer.
A: Umhum
- 78 → A: This time when I went back ---
B: | Then
then you just -
- 79 A: m m
B: you can just do ---
- 80 you just have your own file, huh,
81 A: Umhum
B: then at that point you can just look at-

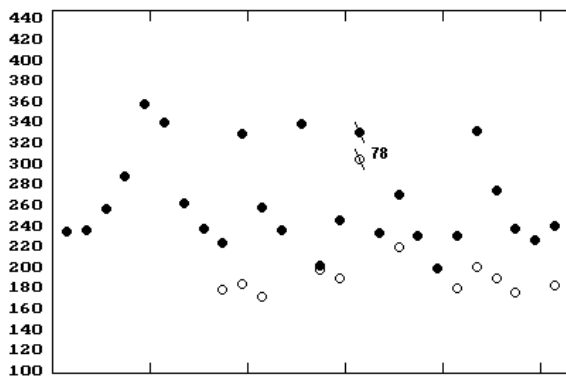


Figure 6: A case of a competitive interruption of Speaker A and Speaker B's resistance response.

82 all of the words from beginning to
end
A: Umhum

In this example, as the main speaker (speaker B) is coming down to the end of a subsection as signalled by the descending pitch level to a low pitch level at U77, speaker A comes in to initiate a topic of her own with a high pitch level of 310Hz. Speaker B immediately counters this interruption with a high pitch and loud amplitude. Once the threat to the floor rights is over however, speaker B immediately returns to a more normal pitch and amplitude level to resume her topic, as we can see by the dropoff in pitch and amp here. The raised pitch and differential in pitch level may be in proportion to the degree of competitiveness involved.

2.3 Cooperative Interruptions

The examples presented so far have mostly illustrated the discourse reasoning and the prosodic characteristics of competitive interruptions. In general, competitive interruptions are marked by a high pitch level, and a loud amplitude, expressing the participants' competition for the focus of attention. By contrast, cooperative interruptions are more supportive of the main speaker's floor rights, and the intention is to keep the attention on the main speaker's point. This difference in cooperativeness has a corresponding influence on the prosodic patterns of such supportive interruptions. Because of their non-disruptive nature, they often occur at low or medium pitch levels, and even when they are high for emotional reasons, they are generally lower in

pitch than competitive interruptions. The amplitude of cooperative interruptions can vary. In our data, the amplitude is generally low in cases of acknowledging and prompting, but often high when an interruption is used to express strong opinion or emphasis. These characteristics can be seen in the following examples:

2.3.1 Expressing Supportive Agreement

- (6)
398 A: Then there are some Taiwanese
399 graduate students
B: Uhhuh
399 A: they also had -
400 Some of them also presented their
papers, right?
401 They also went up there on the stage
to present
B: Umhum
402 A: They also presented in English.
B: Umhum umhum umhum
403 A: I really think that they did a great job.
B: Umhum umhum
404 A: They did really well.
B: Umm
405 A: So I looked at them and say ...(laugh)
406 these Taiwanese graduate students
407 they are really good
at this International Conference they
408 → B: | They
did really well
409 A: Emm B: Umhum
410 A: They did really well. (laugh)

The non-disruptive nature of cooperative interruption can be seen in this example. From the pitch plot (Figure 7) we can see that speaker A is very excited in this segment, and is speaking at a very high pitch in her range. The very excited and involved state of speaker A is clearly evidenced in the fact that her pitch level is the highest point in the entire conversation, among all her utterances. This excited state is also indicated by the abrupt 105Hz pitch elevation from her previous utterance in U406 at 325Hz. By contrast, speaker B's supportive and agreeing interruption comment at U408 is said at a relatively low pitch level of about 260Hz and at a moderately low amplitude, in agreement with the implicit goal of avoiding disruption to the main speaker's progress.

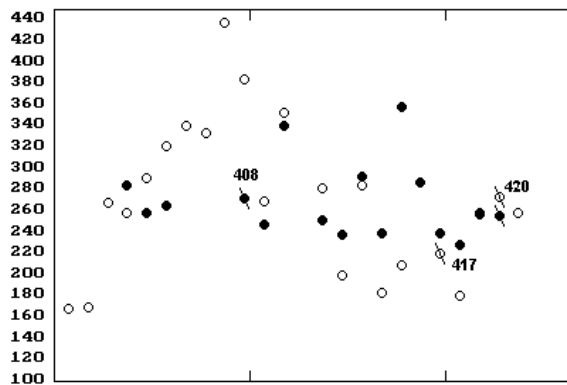


Figure 7: Three cases of low-pitched supportive interruptions at U408, U417, and U420.

2.3.2 Completing An Anticipated Point

Cooperative interruptions frequently occur when the main speaker is in the middle of completing a point and the other speaker already anticipates that point and is in agreement with the main speaker. In such cases, the other speaker often comes in to finish that point for the main speaker, presenting the interruption as a prompt. In our data, these instances typically occur at relatively low pitch levels, because of the certainty and confidence of the interrupter, and at a relatively high amplitude, reflecting increased emphasis.

- (7)
- 413 A: Take a look at what other people are doing, right?
- 414 um
- 414 B: Right right right right. Umhum
- 414 B: Now at this stage it becomes very important, I feel
- A: That's right.
- 416 B: Because the things we learn at school
- 417 I mean when you get to a certain level
- A: | There's a limit
- 418 B: Right.
- A: Mm.
- 419 B: It's just that you need to know more about the outside world.
- A: Right.
- 420 B: At the same time, it's also --- (pause)
- A: | different,
- 421 differ - ent -
- 422 at different places, huh, B: Umhum
- 423 A: what types of things they're working on
- B: Umhum umhum umhum umhum

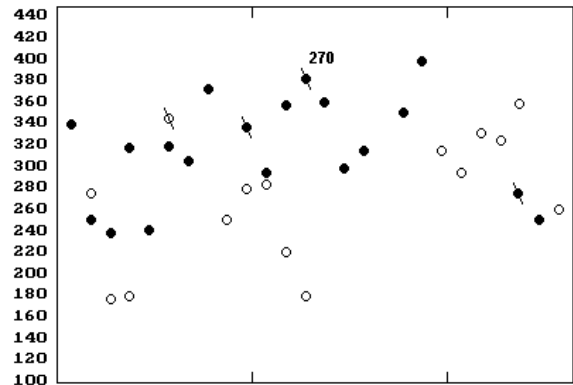


Figure 8: A high-pitched cooperative interruption at U270 for expressing strong agreement.

In this example, speaker A interrupts at U417 to finish for speaker B, expressing the point in an emphatic way: 'There's a limit'. Speaker A's pitch is at a relatively low level of 220Hz, and that reinforces the expression of unanimity or agreement with the main speaker. The loud amplitude on this phrase signals the strong opinion and emphasis that speaker A is expressing (Figure 7). At U420, speaker A again anticipates speaker B and comes in to cooperatively develop B's point at a moderate pitch level of 270Hz and a loud amplitude, signaling the joint cooperative nature of these interruptions.

2.3.3 Variations in Intensity

One complication is that cooperative interruptions are also affected by the intensity of the accompanying emotion, and therefore may also occur at high pitch levels, as seen in the following two examples:

- (8)
- 268 A: If both sides can cooperate
- B: Umhum right
- 269 A: that will be really good.
- 270 → B: | That's exactly right! You just have to cooperate. Right!
- A: Mm.
- (9)
- 298 A: That one's an int'l conference,
- B: Uhhuh
- 299 A: The conference location was at the Central Research Institute,
- 300 but | the people who came to present
- B: (pause)| Uhhuh

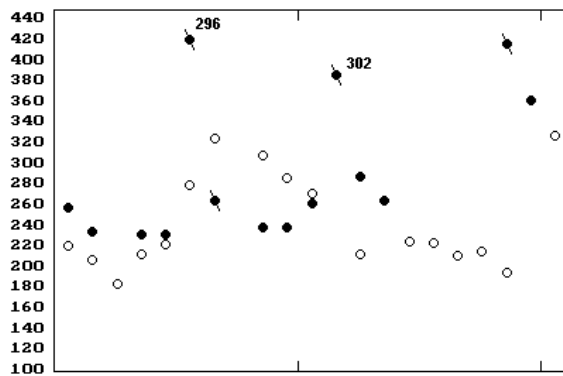


Figure 9: A supportive high-pitched interruption with salient new information at U302.

- 301 A: these speakers were from all over the world.
 302 → B: Uhhuh uhhuh | Lots of famous people.
 303 A: Right. Lots - lots of famous people.
 B: Right umhum umhum
 305 A: Lots of famous people.
 B: Umhum

In the first example, example (8), the interrupter is giving a strong expression of support and enthusiastic agreement, and this is evident in the semantic content 'That's exactly right. You just have to cooperate. Right!'. Because of the strong emotion involved, speaker B's pitch level here is very high at 385Hz, as seen in Figure 8, and the amplitude is also loud. The interruption in example (9) is also a strong interruption to support the main speaker, but adds new salient information to the ongoing topic by explicitly bringing up a notable fact 'lots of famous people'. The pitch level of this phrase (Figure 9) is high at about 380Hz and the amplitude is also great. The supportive intention of this cooperative interruption is further indicated by the continuing feedback speaker B provides, and this intention is recognized and appreciated greatly by the main speaker, as shown by her repeated echoing of speaker B's remark in U303 and U305.

2.4 Interruptions & the Resolution of Uncertainty

2.4.1 *The Integration of Discourse Elements and the Variations of Pitch Height*

Our data show that the complexity of interruptions increases with the complexity of the discourse relationships. Interruptions are

complex discourse phenomena. They are informationally packed, as they mediate the differing interests and knowledge states of participants in a conversation. The specific nature of each interruption is a reflection of the underlying motivation of the interrupter. The content and timing of interruptions are directly linked to the interrupter's urgent and intense emotional need for an immediate resolution. That is, it is the urgency of the emotion that is causing the interrupter to express the need to address a particular salient topic immediately at this particular time.

Another factor that contributes to this complexity is that competitiveness and cooperativeness are not polar opposite characteristics of interruptions, but occur as a gradient process. The degree of competitiveness arises from the intensity of the emotions underlying the interruption. Speaker intensity is also closely linked to the degree of certainty and uncertainty inherent in the ongoing topic progression. The forcefulness of the expression also affects how the main speaker responds. An intense expression often creates a critical need for an immediate response, and speakers are more prone to stop and address the issue raised by the interrupter, hence such interruptions are more competitive.

The degree of competitiveness or cooperativeness is also influenced by how related the interruption is to the ongoing topic, i.e. the degree of relatedness of topic, and the knowledge states of the participants, and by how long the interrupter intends to take the floor for. A short interruption for a clarification on the current topic is more cooperative than an interruption to change both the topic and the floor. The specific strength of signal needed to adequately overcome the ongoing topic may vary by the changing interruptability or resistance level of the topic. Because of the intentions of participants, in spontaneous discourse interruptions occur to varying degrees of intensity and varying degrees of competitiveness and cooperativeness.

Interruptions thus are a complex combination of expressions of emotion, signals of attention-getting, and signals of competitiveness, and their prosodic manifestations are directly linked to these motivations. Our data show that the pitch level of interruptions can occur at varying heights;

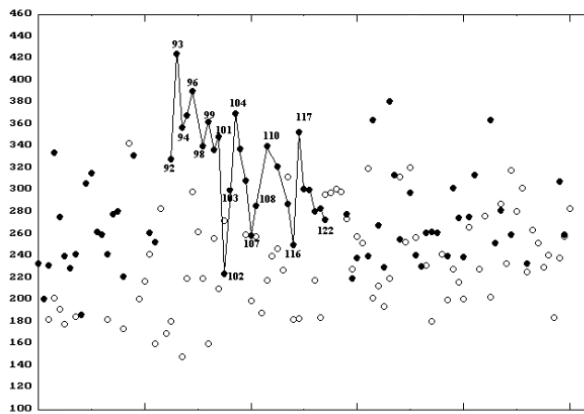


Figure 10: Gradual resolution of uncertainty of speaker B expressed in descending increments of pitch heights.

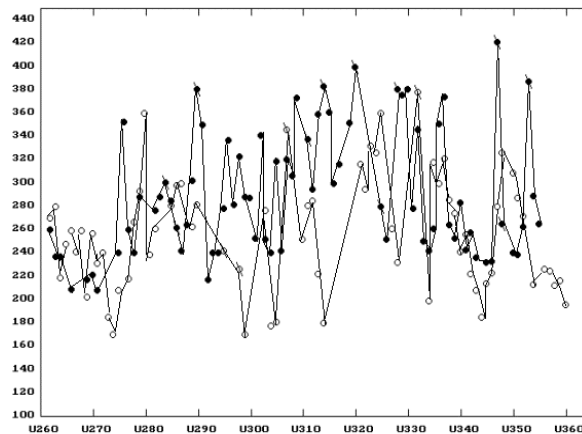


Figure 11: Interruptions in context of matching rise-fall phrase patterns of both speakers over a 100-phrase dialogue section.

the higher the intensity, the higher the pitch level. The specific pitch height of the interruption is determined jointly by the need to attract attention, the intensity of the emotion present, and the strength of signal needed to overcome the attention and focus on the current topic.

2.4.2 Extended Episodes

Example 1

The prosodic patterns for this segment of 100 utterances (see Figure 10) are very revealing of the complex emotional and discourse forces at work, and also illustrate the point that intensity and degree of uncertainty and certainty are significant determinants of topic direction and prosodic structure. In this part of the conversation, speaker A is talking about a conference she attended previously and Speaker

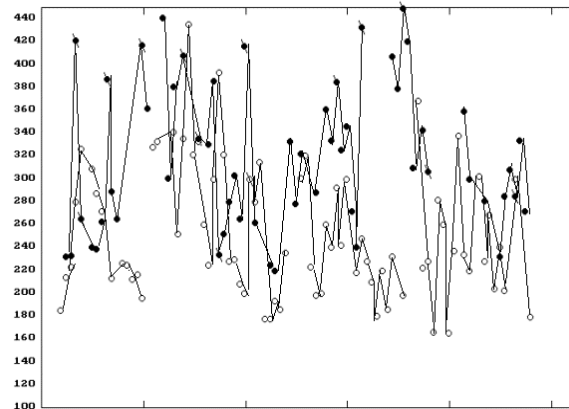


Figure 12: Speaker involvement patterns with intensification and resolution in a topic over an extended episode.

B mistook it to be the conference that she was interested in, so she initiates a series of short questions (in the form of interruptions) to confirm and clarify the information. The general cognitive pattern seen here is that the interrupter encounters an initial high unsettled state of uncertainty and gradually progresses to a more settled and certain state. This is clearly expressed in the overall downward trend in the pitch level for these utterances in the peak pitch chart. The pattern of alternating doubt and certainty is very revealing itself. At each interruption that expresses doubt and a need for clarification, there is a local rise in pitch. Those interruptions which express acknowledgement and certainty are locally lower in pitch. The specific strength of signal needed varies systematically with the resolution of the differing interests and knowledge states of participants.

The overall prosodic structure of this example also provides a vivid illustration of the importance of the process of intensification and normalization in discourse. The very high pitch at U93 reflects the abrupt climax of emotional intensity and uncertainty, and as this emotional uncertainty is expressed and cognitively resolved through the sequence of interruptions, normalization in the intensity of the cognitive state and the pitch level is then achieved.

Example 2

Taking a more extended view of our data shows that pitch movements of interruptions also vary according to overall patterns of topic development and intensity of speaker

involvement. Analysis of the discourse text shows that the rise-fall arc seen in Figure 11 also coincides with the development of a major subtopic that both speakers actively contribute to. This involvement is signaled by the large amount of dots at varying heights of both speakers. Both speakers' involvement reaches a peak of excitement roughly at the U320-U330 section, and then gradually descends as speaker A gives more specific details in concluding the topic. As shown, the pitch levels of the interruptions of both speakers also converge and follow the same rise-fall pattern as interest in the topic increases and then is resolved. This supports the view that interruptions are a part of an overall systematic prosodic structure that integrates topic progression and speaker involvement through a process of climax and resolution.

Example 3

What is happening in the conversation in Figure 12 is that one speaker (speaker B) begins to develop a topic that she is interested in but that had not been successfully communicated, and the interest level and the speaker's involvement are intensified as she attempts to overcome the mismatch as indicated by many very abrupt high pitched interruption points, whereas speaker A's pitch movements are expressed in more uniform overall descending pattern. The descending part of the curve also coincides with the resolution of an issue that speaker B had been very uncertain about throughout that section of dialogue. This reinforces our conclusion that at each level of analysis, prosody links speaker interaction, topic progression and expression of cognitive state.

4 Implications for Dialogue Systems

How can we use the above information to help build an intelligent spoken dialogue system? We can focus on 2 related aspects: Detection and Response. For example, a high pitch and amplitude would be detected as a competitive interruption of higher urgency and indicate a possible mismatch of the current state with the user's desired state. The system would respond by searching the possible topic space, adding the lexical-semantic content of the interruption to prior information to aid in the search.

Ongoing monitoring of the prosody can also provide important information on the direction

the dialogue is taking. For example, if user responses or interruptions follow an increasing pitch pattern, then the system can interpret this as increasing uncertainty, and modify the topic search direction. Conversely, decreasing pitch pattern can indicate that the user's certainty and progress toward a desired goal are increasing in a satisfactory way.

As spoken dialogue systems become more receptive to natural human speech, the disfluencies and prosody of human speech can provide critical information to guide progress along interactively developed system paths, mirroring aspects of human-human conversational speech. Further work on adapting prosody detection to dialogue systems provides a foundation for systems which are truly interactive, taking advantage of active inputs by the user, adapting to the knowledge base of users, and providing clarifications according to search strategies that account for information presented in more natural ways, and ultimately making systems more intelligent by adapting to human motivation.

Conclusion

In this paper we have shown that interruptions are important elements in the interactive character of discourse and in the resolution of issues of cognitive uncertainty and planning. We have also shown that interruptions are part of the local and global coherence that is brought about through the systematic phrase-to-phrase prosodic patterns of discourse, and are an important component for speech understanding and intelligent dialogue systems.

References

- Peter Heeman and James Allen (1999) Speech Repairs, Intonational Phrases, Discourse Markers: Modeling Speaker's Utterances in Spoken Dialogue. *Computational Linguistics*, Vol. 25-4.
- Barbara Grosz and Candace Sidner (1986) Attention, Intentions, and the Structure of Discourse. *Computational Linguistics*, 12(3):175-204.
- Marilyn Walker and John D. Moore (1997) Empirical Studies in Discourse. *Computational Linguistics*, 23(1):1-12.
- Li-chiung Yang (1995) *Intonational Structures of Mandarin Discourse*. Ph.D. dissertation, Georgetown University.