STRUCTURE AND INTONATION OF BUSINESS TELEPHONE OPENINGS

Mark Liberman
Cynthia McLemore

University of Pennsylvania

1   Introduction

A phone call to a restaurant, a store, a lawyer's office, or similar commercial establishment produces a response that is conventionalized in textual form and in intonation. Because of this conventionalization, such utterances provide large amounts of similarly-structured material, often involving identical words, in identical discourse and interactional contexts. The set of alternatives can be exhaustively listed, above the level of details such as the establishment's name, and the various alternatives have different probabilities that can be estimated for a given class of establishments. This predictability remains true of some common kinds of exchanges that may follow the opening, such as the answers to questions like "how late are you open?" or "where are you located?"

As a result, such material should be useful in exploring (some aspects of the) correlations among intonational types, textual units, and contexts of use, since we can easily obtain hundreds of examples of a given linguistic form in a given situation. More importantly, such material should give us an opportunity to evaluate possible differences in intonational usage among speakers differing in region or gender, since we can obtain large numbers of examples that are almost entirely controlled for other situational variables.

This paper attempts to establish the validity of this method by examining the telephone openings that result from calling restaurants. We analyze this material from three points of view. First, we provide a finite-state "pragmatic grammar" that characterizes all of the responses in a large sample of such calls, and sketch the statistical structure of strings in its language. Second, we examine the range of intonational variation in this material, and the correlation of intonational types with the elements of our structural analysis of the text. Third, we examine the effect of the speaker's sex on the textual and intonational structure of our data, as a pilot example of how groups differentiated by categories such as sex, region, restaurant price range, etc. might be analyzed.

We conclude that

1. At least four different intonation patterns occur commonly in our data; the most frequent pattern is one that ends High-Mid.

2. Different textual elements of the response show different distributions of intonational patterns; especially, offers of help differ from other elements such as restaurant names.

3. The association of textual elements with intonation patterns is probabilistic, not categorical.

4. Females and males differ slightly in the length and textual structure of their responses.

5. Females and males do not differ in the intonational structure of their responses, as long as the distribution of textual material is held constant.

2   Materials

We took phone numbers from the Restaurant section of the Bell of Pennsylvania Philadelpha yellow pages, avoiding types of restaurants where the phone is likely to be answered by non-native speakers. We called each number, and for those that answered within five rings, we identified ourselves as gathering information about restaurants for a conference, and explained that the call was being recorded for future reference, as required by Pennsylvania law. If the person answering expressed willingness to answer questions, we asked about the restaurant's hours, about the price range of the menu, and about the restaurant's location.

This paper analyzes the data from the initial response portion of 134 successful calls. Judging from their voices, 69 of those answering were female, and 65 were male. One call was eliminated because a member of an after-hours cleaning crew answered. Each recorded call was digitized, transcribed, and pitch tracked. Because of the poor sound quality of many of the recordings, voicing decisions and possible octave errors were hand-corrected. There were only two cases in which F0 could not be plausibly estimated. In one other case, we judged that the speaker was non-native. These three cases were included in the textual analysis, but excluded from the intonational analysis.

About half the calls were made by each of the authors, and calls were made on various days and at various times of day. It seems reasonable to assume that some kinds of variation on the caller's side will produce variation in the responses. Thus it is quite likely that different sorts of answers would follow questions by callers differing in sex, age, race, etc. It is also likely that prosodically different questions would engender prosodically different answers. We did not attempt to explore any of these issues in this preliminary study, which in any case only analyzes the initial response of the answering party.

3   Functional structure of initial responses

Normally the first vocal move in a telephone conversation is made by the called party, who picks up the phone and says something that indicates readiness to go forward with the conversation. Schegloff (1972) calls this a "distribution rule for first utterances ... : the answerer speaks first" (p. 351). As he points out, the telephone ring constitutes a summons, and the initial response of the called party is an answer to that summons; together, the telephone ring and the response constitute a *summons-answer pair*. The summons-answer pair is the conversational unit that we investigate below; however, since the only vocal material is that provided by the answer, we will refer to it as *response* or *initial response*.

In modern America, a private individual is likely to say nothing more than "hello," "yeah" or something of the sort. However, phones that represent some institutional entity are usually answered with a more elaborate formula, typically including the name of the institution, sometimes including the name of the speaker, along with various conventional greetings and expressions of willingness to help, gratitude for patronage, etc.

In 134 successful calls to Philadelphia-area restaurants, we find that almost all of the initial responses are made up of sequences of elements drawn from a limited set:

(1) (a) Greetings, drawn from the set *hello, good morning, good afternoon, good evening* (**G**).

    (b) The name of the restaurant, sometimes preceded or followed by a place-name modifier if the establishment is a branch of a chain (**R**).

    (c) The name of the person speaking, expressed as *(this is) NAME (speaking)* (**P**).

    (d) An offer of help, expressed as *(how) may/can I/we help you (please)* (**H**).

    (e) An expression of thanks, such as *thanks/thank-you for calling RESTAURANT* (**T**).

For ease of discussion, we will refer to these five classes of elements with the single-letter symbols **G, R, P, H**, and **T** respectively, as indicated above, and will classify responses in terms of sequences of these letter symbols, as exemplified in 2 below. We have changed the personal names to protect the privacy of the speakers, but have left the names of the restaurants in their original form, since the we are dealing with speech addressed to the public at large rather than conversations with any significant private content.

(2) (a) Hello. (**G**)

    (b) Cambria. (**R**)

    (c) Good evening, Chautauqua. (**GR**)

    (d) Charley's place, can I help you? (**RH**)

    (e) Good evening, International House of Pancakes, how may I help you? (**GRH**)

    (f) Good evening, Old Original Bookbinder's, Robin speaking. (**GRP**)

    (g) Thank you for calling Smart Alex, Iris speaking, how may I help you? (**TPH**)

    (h) Hello, Ponderosa, Kelly speaking, how can I help you? (**GRPH**)

    (i) Good evening and thank you for calling Pottstown Ponderosa, how can I help you? (**GTH**)

    (j) Burger King, this is Sally, may I help you? (**RPH**)

Out of 134 responses, all but one were straightforwardly analyzable as a sequence of these elements. The orphan response

(3) mm yup?

was produced by a member of the cleaning crew at an establishment that was not open for business, and thus should perhaps be discounted; it is not tallied in the data considered below. One additional response, given as 2(i) above, included the word "and" separating the elements G and T. It is tallied as **GTPH**, ignoring the "and." Otherwise, all of the responses could be precisely characterized as sequences of one or more of the five elements, without anything additional.

Of the 153 logically possible sequences of one or more of these five elements, 15 occurred in our data:[1]

Table 1: Counts of Attested Sequences

Sequence	Overall Count	Females	Males
G	3	1	2
GR	35	21	14
GH	1	0	1
GRP	3	2	1
GRH	14	10	4
GTH	1	1	0
GRPH	7	6	1
GTPH	2	1	1
R	29	12	17
RP	2	0	2
RH	18	8	10
RPH	3	1	2
TH	7	3	4
TPH	6	2	4
RG	2	1	1
Totals:	133	69	64

As Table 2 shows, the female respondents included the **G** "greeting" element somewhat more often than the males did. A chi-square test suggests that this magnitude of difference would occur by chance about 6% of the time, which is marginal but suggestive. No other differences in unigram distributions of response elements were significantly different by sex in this data.

Table 2: Number and Percentage of Responses Containing the Various Elements.

Element	Overall	Females	Males
G	68 (51%)	43 (62%)	25 (39%)
R	113 (84%)	61 (88%)	52 (81%)
P	23 (17%)	12 (17%)	11 (17%)
H	59 (44%)	32 (46%)	27 (42%)
T	16 (12%)	7 (10%)	9 (14%)

As Table 3 and Figure 1 show, female respondents also tended to use slightly longer strings of elements.

---

[1] **R** and **T** are clearly in complementary distribution, and all of the elements obey rigid sequence constraints, as discussed in 3.1

Table 3: Sex Differences in Length of Initial Response Sequences

Length	Females	Males
1 element	13 (19%)	19 (30%)
2 elements	33 (48%)	32 (50%)
3 elements	16 (23%)	11 (17%)
4 elements	7 (10%)	2 (3%)

Figure 1: The percentage of responses with lengths of 1,2,3,4 "elements,"
plotted separately for male and female speakers

Figure 1: Distribution of Initial Response Lengths
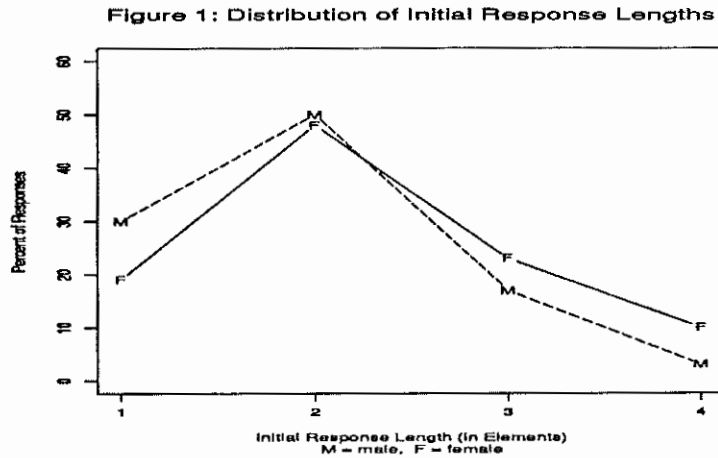
Although the indicated differences are statistically significant, it would unwise to interpret them without more information about the organization of the establishments under study, especially possible correlations between sex and job classification, or among sex, job classification and restaurant type. We report them simply as examples of how a technique of this type can be used to quantify group differences.

3.1 Nature and sources of the functional structure

We have expressed the functional structure of these initial responses in terms of a simple kind of "pragmatic grammar," which could be formalized in terms of transition probabilities among the elements we have identified.

It is not essential to our work that this analysis have any status beyond an observational one. Its categories are extensionally well defined; the categories divide the set of initial responses into sensible classes, and divide individual responses into one or more sensible chunks; the classes and

chunks enable us to make prosodic comparisons among these utterances in a way that compares comparable things.

This analysis is part of our initial description of the data, not necessarily part of a theory. Still, we find it interesting that the responses are so constrained in structure, even though their superficial form is quite varied. For instance, as we have noted, the ordering constraints among constituents are fairly rigid. Thus we recorded 23 examples in which the restaurant name **R** preceded the personal name-phrase **P**, and none that went the other way; similarly, the greeting **G** preceded the restaurant name **R** in 59 examples, and followed it in only 2. Intuitive judgments tend to agree with these observations: note the awkwardness of the **P R** and **R G** orders in (4)(b):

(4) (a) Good evening, Joe's Steak House, Quinn speaking.

(b) ? Quinn speaking, Joe's Steak House, good evening.

Other generalizations involve consistent, but logically inessential, choices of phrasing. Thus 8 of the 23 personal-name (**P**) elements were of the form "this is NAME ...," while none of the 113 restaurant-name (**R**) elements had that form, even though it seems reasonable to answer in such a way. Thus the type of answer exemplified in 5 never occurred:

(5) Hello, this is Joe's Steak House, can I help you?

Some of these these generalizations might be argued to have a functional basis. Without any doubt, the fact that the "thank you" **T** category contains the restaurant name **R** explains the fact that **T** is in complementary distribution with **R**. Some of the ordering constraints might also have a functional basis: thus the **R P** order might reflect a tendency to put the general before the particular, or the known before the unknown (since callers generally know what restaurant they're calling), or the more relevant before the less relevant. On the other hand, if the order had been the other way, we might have given some equally plausible *post hoc* explanation for that. Especially since material from other languages and cultures may well be very different, it seems wise to reserve judgment about the relative role of logic and convention in this area.

3.2   The issue of training

Many establishments no doubt instruct their personnel to answer the phone with a specific formula. However, several calls to the same establishment did not always produce the same formula, and calls to different branches of national chains likewise produced a fair amount of variation. Thus 11 calls to Burger King branches produced the patterns **G**, **R**, **GR**, **RH**, **GRH**, and **RPH**.

Therefore, we conclude that the distribution of the forms we observe are produced by tacit linguistic processes of the usual and natural kind, as much as by explicit stipulation. The intonational structure of the initial responses, to which we now turn our attention, is even less likely to be determined by explicit instruction.

4 Intonational structure of initial responses

4.1 Nature and purpose of our intonational classes

In order to look at correlations between textual and intonational categories, we have chosen to classify each subphrase according to what British intonation analysts would call its *nuclear accent*, which in the system of Pierrehumbert (1980) would include the final pitch accent, the phrase accent, and the boundary tone.[2]

However, we have not tried to use the descriptive categories given by any extant intonational theories, but instead have chosen a more "bottom-up" approach, in which we have tried to divide the tonal patterns occurring in our data into natural classes that are extensionally well-defined and likely to be intersubjectively stable, but still oriented towards the phonological distinctions that we will make later in the study.

In other words, our intonational classes are phonetic in character, just as one might distinguish the flapped, glottalized and aspirated allophones of /t/ in a speech corpus, in order to examine their distribution relative to stress and juncture at word and phrasal levels. We feel that it is particularly important to begin descriptively with categories that are as phonetically transparent as possible, with more abstract, quasi-phonological categories motivated as explanations of phonetic distributions.

Of course, like all phonetic categories, ours are subjective, in the sense that they are determined by a trained observer based on the application of explicit criteria, rather than by a completely algorithmic procedure. It is also important to note that we are basing this study on a phonetic classification of certain aspects of the utterances we examine, rather than attempting a complete transcription of all linguistically-relevant aspects of the F0 contours.

As a result of the just-described properties of our categories, we feel that they provide a relatively stable description, which other researchers could apply to similar material without a great deal of uncertainty. However, we point out that very different patterns may be found in data from very different sources—for instance, material from French or Chinese would surely motivate a different classification, as might material from different dialects of English.

4.2 Definition of our intonational categories

We attempted to assign each functionally-defined phrase of our data to one of six intonational categories, whose properties are briefly listed in 5, and discussed more extensively below.

Summary of Intonational Categories:

(6) (a) High-Mid: pitch peak on (or just after) the nuclear syllable, then a fall to a mid-to-low level. Unaccented pre-nuclear material is typically at least as low as the final value. Sometimes a slight rise at the very end. See Figures 2 and 3.

---

[2] See McLemore (1991a) pp. 36-44 for a discussion.

(b) Low-High: a rise from a relatively low pitch, with the F0 minimum variably within the nuclear syllable or earlier. Sometimes a slight fall at the very end. See Figures 4 and 5.

(c) High: a relatively high F0 value from the very beginning of the accented syllable, carried along to the end of the phrase without much of a rise or or fall. See Figure 6.

(d) High-Low-High: pitch peak on the nuclear syllable, then a fall and a rise on material that is unstressed or at least unaccented; final rise is as high or higher than the accent peak. See Figure 7.

(e) High-Low: pitch peak on (or just after) the nuclear syllable, then a fall to near the bottom of the range. See Figure 8.

(f) H-M-or-H-L: ambiguous between High-Mid and High-Low.

Sometimes two functional elements were completely run together into a single intonational phrase, so that it seems wrong to classify their intonational forms separately. Figure 2(b) exemplifies this case—the word "hello" stands as the commonly-encountered low-pitched introductory material before the peak of the high-mid pattern on "Babe's," in about the same way that the word "empty" stands to "glass" in Figure 3(a).

If (as was often the case) we were uncertain whether to merge or split the elements intonationally, we chose to split. In general, the assignment of phrase breaks in natural speech material is problematic.

4.2.1 The High-Mid pattern

Figure 2: High-Mid on phrases of the form G and G R
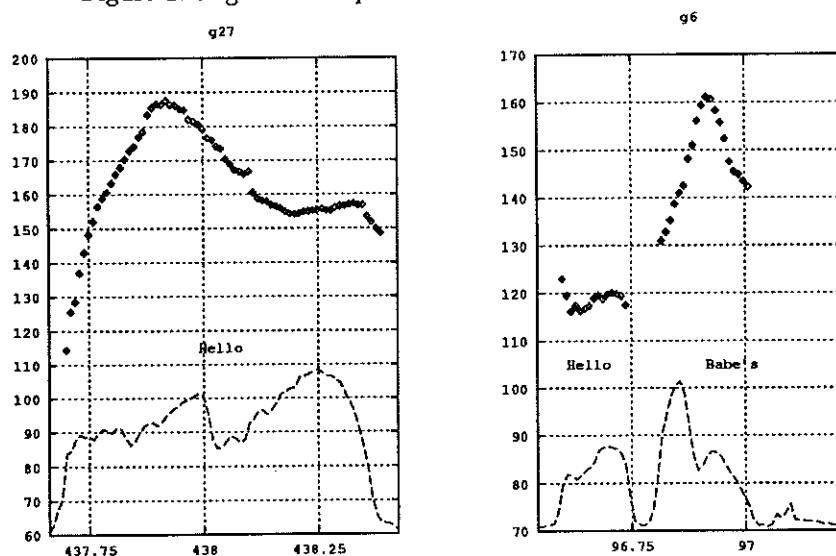
The most frequently-occurring intonational pattern in our data is exemplified in the pitch tracks shown in Figures 2 and 3.

In these figures, as in subsequent pitch tracks, the X axis shows time, the Y axis shows fundamental frequency for the function plotted with open circles, the orthographic labels are centered in the region of time that corresponds to them, and an RMS amplitude trace, plotted to an arbitrary but graphically convenient scale, is drawn as a dashed line at the bottom of the figure.

We will call this contour *High-Mid*. It is characterized by a pitch peak on (or just after) the nuclear-stressed syllable, with a subsequent fall to a mid-to-low level. If there is unaccented pre-nuclear material, it is often on a low pitch, typically at least as low as the final value. About 48% of the phrases in our initial-response data were classified as High-Mid.

Figure 3: High-Mid on responses of the form R, RH
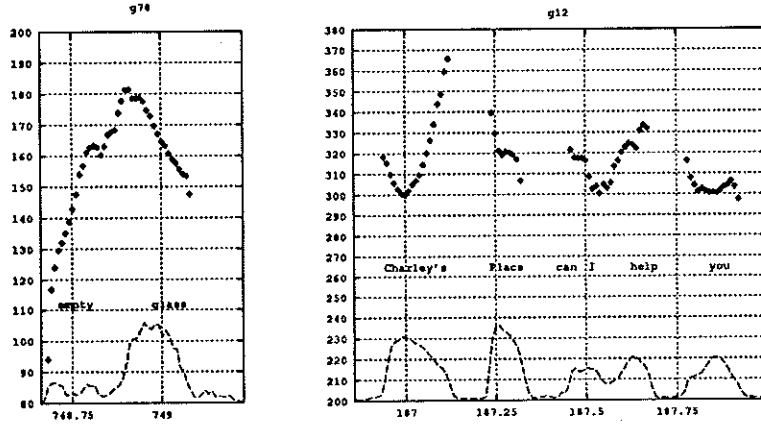

4.2.2 The Low-High pattern

Figure 4: Low-High on phrases of the form R,GR
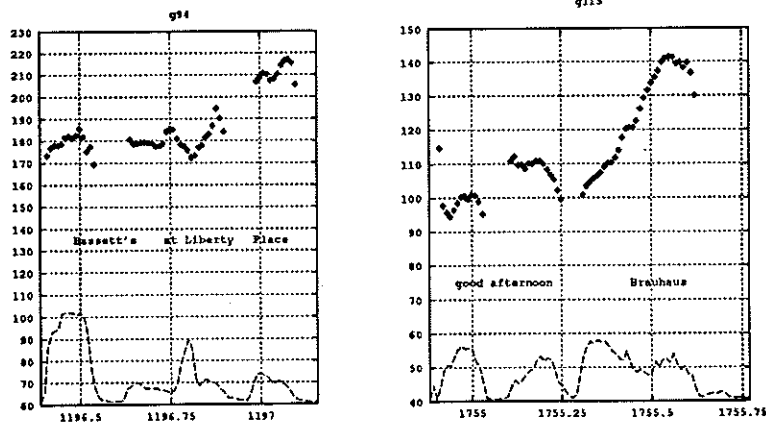
The intonational pattern shown in Figures 4 and 5 is a rising contour which we will call *Low-High*. It is characterized by a rise from a relatively low pitch. The F0 minimum value may occur somewhere within the nuclear syllable; or or it may fall on a previous word; or it may occur just before the nuclear syllable, aligned with essentially unstressed material. In any case, the F0 rises from the last accented syllable to the end of the phrase, except that there sometimes may be a slight fall at the very end as the amplitude drops off.
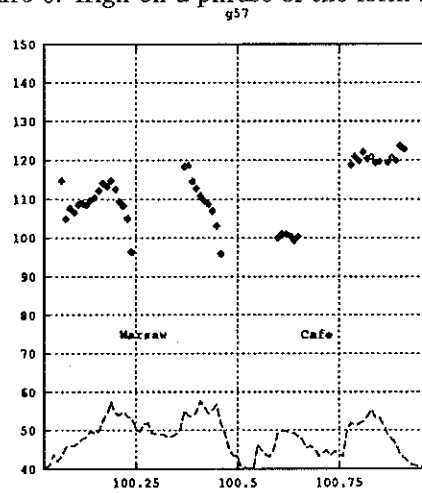
About 15% of the phrases in our initial-response data were classified as Low-High.

Figure 5: Low-High on phrase of the form (GT)PH
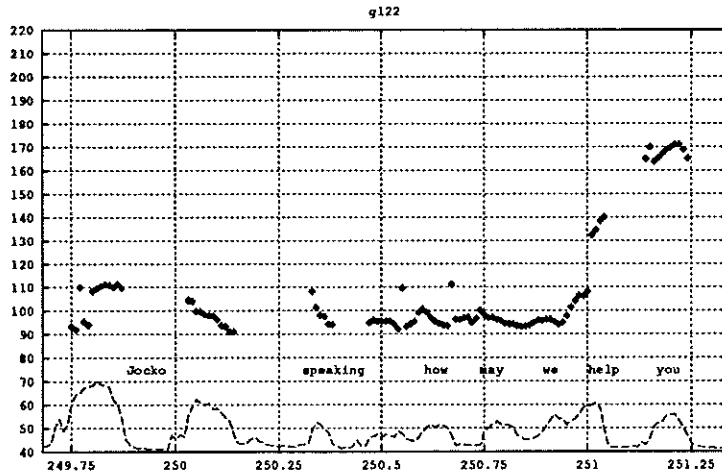
4.2.3 The High pattern

Figure 6: High on a phrase of the form R

In this category, there is a relatively high F0 value from the very beginning of the accented syllable. This F0 value is carried along without a very pronounced rise or fall to the end of the phrase. Thus, like the Low-High pattern, the High pattern ends on a fairly high pitch. Pre-nuclear material is often lower, and the overall impression is usually similar to that of the Low-High pattern.

An example is shown in Figure 6. About 11% of the initial-response phrases showed the High pattern.

4.2.4   The High-Low-High pattern

Figure 7: High-Low-High on a phrase of the form **GRPH**
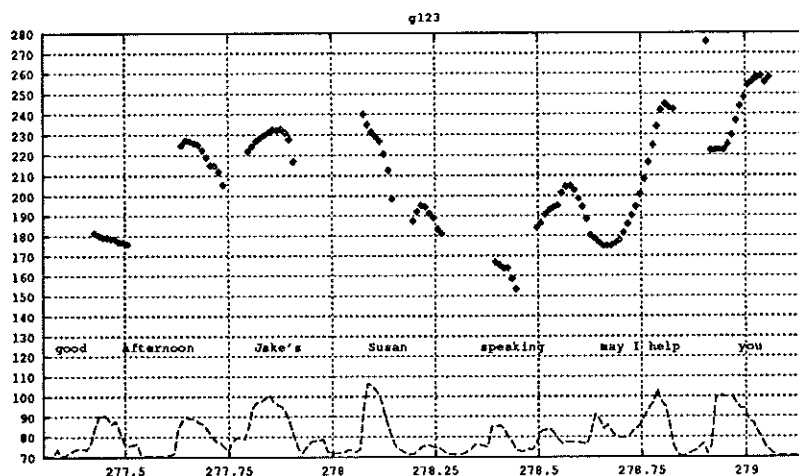
As exemplified in Figure 7, in the *High-Low-High* pattern, the nuclear syllable aligns with a relatively high F0 value, which is followed by a fall and then a rise on material that is unstressed or at least unaccented, and in which the final rise is as high or higher than the accent-aligned peak. Again, this patterns ends on a relatively high pitch, and may be a variant of the Low-High pattern. It only occurred on 1% of the initial-response phrases in our study.

4.2.5   The High-Low pattern

In this case, there is a relatively high F0 value, which may lie within the accented syllable or slightly after it if there is subsequent unstressed material, and then a following fall to a point near the bottom of the speaker's pitch range.

The High-Low pattern is exemplified in Figure 8. About 18% of the initial-response phrases were classified as High-Low, making it the second commonest pattern.

In some cases, the High-Mid contour type is difficult to distinguish from the High-Low contour type. The basis for the distinction, obviously, must be found in the post-nuclear region of the

contour, which is usually falling, although sometimes there may simply be a level region of lower pitch following a consonantal interruption of voicing.

Figure 8: High-Low on phrases of the form **G, GR, GRH**
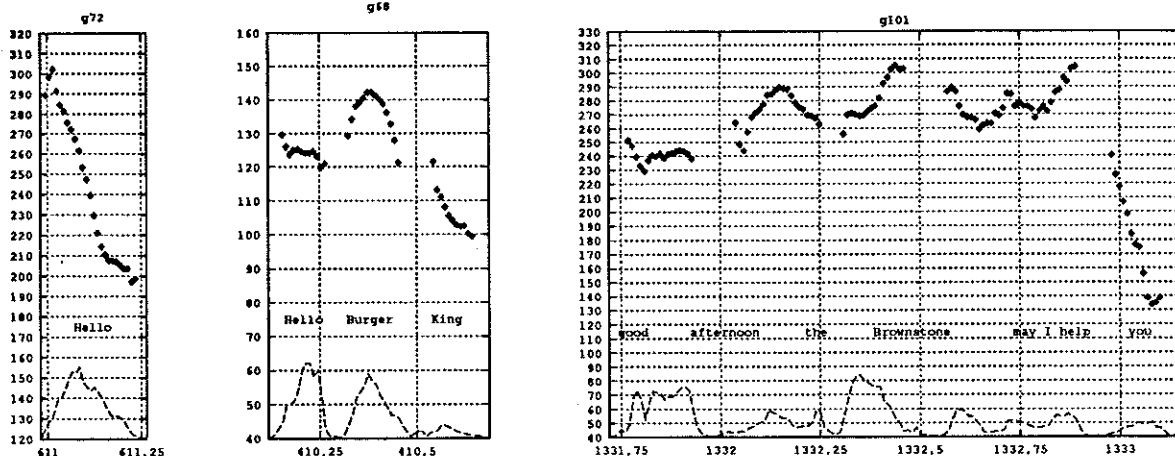
In our categorization, we chose the High-Mid rather than High-Low classification if this post-nuclear region has minimum F0 values that are as high or higher than pre-nuclear low values in the same contour; or has minimum F0 values that are clearly at or above the mid region of the speaker's normal pitch range; or is drawn out in time and remains high in amplitude; or shows a slight final rise.

If the case seemed resolutely ambiguous, we assigned it to a special category of "ambiguous between High-Mid and High-Low," which is treated separately in the tabulations below. About 6% of the initial-response phrases were classified as ambiguous in this way.

4.3   Distribution of intonational categories

130 responses were used in this phase of the analysis. Two out of the 133 responses in the first phase of the analysis were discarded due to marginal sound quality, and one due to non-native accent. There were 252 intonational phrases notated overall, distributed as in Table 4.

Perhaps the most striking thing about Table 4 is the large number of High-Mid patterns (about half of the total). Since it has been claimed that one of the uses of this pattern is "to call to people with whom the speaker is not in eye contact" (Liberman (1975)), this finding is in retrospect reasonable; but we did not anticipate it, and it does not seem to have been noted in earlier literature.

One of our findings seems to contradict a claim made in the literature. The system of into-national description presented in Pierrehumbert (1980) distinguishes two sorts of patterns within our Low-High category: what we might call a "low rise," notated **L* H H%**, and what we might call a "high rise," notated **H* H H%**. In Pierrehumbert and Hirschberg (1990), it is claimed that telephone reponses with rising intonations could only be examples of the "high rise" type, with the

"low rise" type being impossible and indeed ludicrous. This is cited as motivation for some aspects of their compositional account of the meaning of the contours in question.

We observed four Low-High patterns on response element R in final position, and one on response element R in non-final position in a case where there is clearly a following phrase boundary. Although it is sometimes hard to be certain which of the two notations is the appropriate one for a particular case (which is one reason that we didn't attempt make the distinction systematically throughout our data), we believe that four of these five cases are fairly clear instances of the "low rise" type, in that the nuclear accented syllable shows a syllable-internal F0 minimum which is also the lowest pitch value in the phrase. The classification of the remaining case is not clear. On the face of it, these observations refute P&H's distributional claim, and the argument based on it.

The overall differences in distribution between females and males shown in Table 4 are not significant at the .1 level, according to a chi-square test (in other words, there is a better than one in ten probability that the observed difference occurs by chance).

As Table 5 shows, the distribution in response-final position was not significantly different from the overall distribution, with the marginal exception of the High category (which tended to occur more often in non-final position). As before, the female-male distributions shown in Table 5 are not significantly different according to a chi-square test.

Table 4: Overall distribution of Intonational Classes

Intonational Type	Overall Count	Females	Males
High-Mid	120 (48%)	69 (48%)	51 (48%)
High-Low	46 (18%)	21 (15%)	25 (23%)
H-M or H-L	16 (6%)	9 (6%)	7 (7%)
Low-High	38 (15%)	21 (15%)	17 (16%)
High	29 (12%)	22 (15%)	7 (7%)
High-Low-High	3 (1%)	3 (2%)	0
Totals	252	145	107
L-or-M-final	182 (72%)	99 (68%)	83 (78%)
H-final	70 (28%)	46 (32%)	24 (22%)

Table 5: Distribution of Intonational Classes in Response-Final Position

Intonational Type	Overall Count	Females	Males
High-Mid	64 (49%)	38 (55%)	26 (43%)
High-Low	27 (21%)	10 (14%)	17 (28%)
H-M or H-L	9 (7%)	4 (6%)	5 (8%)
Low-High	24 (18%)	13 (19%)	11 (18%)
High	4 (3%)	2 (3%)	2 (3%)
High-Low-High	2 (1.5%)	2 (3%)	0
Totals	130	69	61
L-or-M-final	100 (77%)	52 (75%)	48 (79%)
H-final	30 (23%)	17 (25%)	13 (21%)

Although in general the sex of the respondent did not have a significant effect on the choice of intonational type in this data, the distinctions among the functional classes of response elements did sometimes relate significantly to intonational categories.

The most striking relations of this type relate to the H "help phrase" element. As an example, Table 6 summarizes the distribution of intonational categories on response-final R (restaurant name) and H (help) phrases. The R-versus-H differences are easily significant at the .001 level (in other words, the probability of these differences occuring by chance is less than one in a thousand).

Table 6: Intonation of R vs. H in Response-Final Position

Intonational Type	Overall (incl. G,P,T)	R	H
High-Mid	64 (49%)	46 (74%)	13 (22%)
High-Low	27 (21%)	3 (5%)	20 (34%)
H-M or H-L	9 (7%)	5 (8%)	4 (7%)
Low-High	24 (18%)	4 (6%)	19 (33%)
High	4 (3%)	4 (6%)	0
High-Low-High	2 (1.5%)	0	2 (3%)
Totals	130	62	58
L-or-M-final	100 (77%)	54 (87%)	37 (64%)
H-final	30 (23%)	8 (13%)	21 (36%)

We can summarize the difference between the intonation of the response-final R phrases and the response-final H phrases by saying that the R phrases showed mainly the High-Mid pattern, while the H phrases were more likely to be High-Low or Low-High. Note, however, that all of the common intonation patterns observed in our study (High-Mid, High-Low, Low-High) occur on both response-final R and response-final H phrases. The intonational differences between R and H phrases are probabilistic, not categorical.

Given the stereotyped association of yes/no questions with intonational rises, wh-questions with intonational falls, and the fact that the H phrases come in both yes/no and wh-question variants, the reader may wonder if this difference in syntactic form might explain the intonational variation in the H phrases, despite their strongly formulaic character. There were 14 wh-question H phrases and 44 yes/no question H phrases in our data, with intonational types distributed as shown in Table 7.

Table 7: Intonation of WH vs. yes/no final H-phrases

Intonational Type	Overall	WH (how)	yes/no
High-Mid	13 (22%)	4 (29%)	9 (20%)
High-Low	20 (34%)	7 (50%)	13 (30%)
H-M or H-L	4 (7%)	0	4 (9%)
Low-High	19 (33%)	3 (21%)	16 (36%)
High-Low-High	2 (3%)	0	2 (5%)
Totals	58	14	44
L-or-M-final	37 (64%)	11 (79%)	26 (59%)
H-final	21 (36%)	3 (21%)	18 (41%)

Although the differences are in the stereotypic direction, they are not significant at the .5 level by a chi-square test (that is, the probability getting such a difference by chance is better than one in two). In order to determine whether an effect of syntactic form actually exists in this case, we would need to look at more data. It is clear from the data that we have that any difference will be probabilistic, not categorical.

5   Conclusions

The data analyzed in this study show that telephone openings, as performed by Philadelphia-area restaurant employees, are highly regular in structure, both in the type of textual elements that constitute them, and in the order of those elements. Speakers use a wide range of intonational contours on these openings, but use some contours much more often than others; frequencies of contour occurrence are significantly correlated with functional type of textual element, but not with speaker sex or question syntax. The only significant difference in response structure between speakers of different sex is that females have a slight tendency to use the greeting element more often than males.

As noted, the High-Mid intonational form occurs on approximately half of the textual units analyzed in this data. This form was originally identified as a "vocative" or "calling" pattern (Leben (1976), Liberman (1975)), and was observed by Ladd (1978, 1980) to signal that some aspect of an utterance is stereotyped or predictable, i.e. "stylized." Based on observed patterns of use, McLemore (1991a) shows that speakers exploit both the iconicity of such a form and the conventionalized informational aspect of its function (e.g., given/new).

We speculate that the frequent use of High-Mid on the data reported here may well have such dual motivations. On one hand, an answer to a telephone summons establishes contact across a distance; the High-Mid intonational form iconically represents this distance, as in the vocative use of the form. On the other hand, the response to a telephone ring is expected by callers; the textual form of the response is highly regular and conventionalized, and the content of the response is also known or predictable, for the most part.

We observed above that the **R** (restaurant-name) phrases show mainly the High-Mid pattern, while the **H** (help) phrases are more likely to be High-Low or Low-High, although about a fifth of them show the High-Mid pattern. We speculate that the lower frequency of High-Mid on **H** phrases has two sources. First, **H** phrases are always in final position in the opening sequence, making the iconic representation of calling-across-a-distance with High-Mid less relevant to their use, and intonational cues to turn-taking more relevant.[3] Second, the intonational marking of help phrases as predictable, with High-Mid, would tend to impart a sense of ritualization to the offer of help, which some speakers may wish to avoid.

The fact that final **H** phrases are somewhat more likely to show rising (Low-High) intonation than final **R** phrases are might be related to the fact that the **H** phrases take the syntactic form of questions. However, the stereotypic pattern is for rises on yes/no questions and falls on wh-questions, while in our data, **H** phrases of both forms show essentially the same distribution of intonation patterns. Additional data on this issue would be useful.

---

[3]See McLemore (1991a, 1991b).

It has been noted that one stereotype of women's speech is a greater frequency of intonational rises than in men's speech (cf. Lakoff (1975); see also McConnell-Ginet (1983), Bolinger (1989)). In our data, male speakers are just as likely to use rises as female speakers, and in fact, the distribution of intonation patterns is essentially the same for males and females. Although the context of our study is a very specialized one, it still offers the opportunity for greater or lesser expression of gender identity and attitudes allegedly correlated with it, so that our failure to confirm the stereotype seems to us a significant one.

Overall, we feel that this pilot experiment is a success. The degree of apparent conventionalization in our data suggests that a comparison of telephone openings in different geographical regions would provide a good first-order indication of interactional and intonational differences across speech communities.

References

[1] D. Bolinger. *Intonation and Its Uses*. Stanford University Press, Stanford, CA, 1989.

[2] D. R. Ladd. Stylized intonation. *Language*, 54:3:517–541, 1978.

[3] D. R. Ladd. *The Structure of Intonational Meaning: Evidence from English*. Indiana University Press, Bloomington, 1980.

[4] R. Lakoff. *Language and Women's Place*. Harper and Row, New York, 1975.

[5] W. R. Leben. The tones in English intonation. *Linguistic Analysis*, 2(1):69–107, 1976.

[6] M. Liberman. *The Intonational System of English*. MIT Ph.D. Dissertation, Cambridge, MA, 1975. Published by Garland Press, New York, in 1979.

[7] S. McConnell-Ginet. Intonation in a man's world. In Kramarae Thorne and Henley, editors, *Language, Gender, and Society*, pages 69–88. Newbury House, Cambridge, 1983.

[8] C. McLemore. *The Pragmatic Interpretation of English Intonation: Sorority Speech*. University of Texas Ph.D. Dissertation, Austin, TX, 1991a.

[9] C. McLemore. The interpretation of L*H in english. In C. McLemore, editor, *Texas Linguistics Forum 32: Discourse*, pages 175–195. Department of Linguistics and the Center for Cognitive Science, University of Texas, Austin, TX, 1991b.

[10] J. Pierrehumbert. *The Phonology and Phonetics of English Intonation*. MIT Ph.D. Dissertation, Cambridge, MA, 1980.

[11] J. Pierrehumbert and J. Hirschberg. The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication*. MIT Press, Cambridge, MA, 1990.

[12] E. Schegloff. Sequencing in conversational openings. In J. Gumperz and D. Hymes, editors, *Directions in Sociolinguistics: The Ethnography of Communication*, pages 346–380. Holt, Rinehart, and Winston, New York, 1972.