Discourse Intonation in Greek

Antonis Botinis

INTRODUCTION
This paper is about Greek discourse intonation within the framework of the KIPROS project, an ongoing research program on interactive and contrastive prosody with Swedish, French and Greek as the languages of the working material. The project is supported by the Bank of Sweden Tercentenary Foundation and is being conducted at the Phonetics Department, University of Lund, with Gösta Bruce as project leader. Swedish is represented by Gösta Bruce and Ursula Willstedt, Paul Touati is mainly responsible for French, and Antonis Botinis and recently Annika Åkerblom are mainly responsible for Greek.

Our ultimate goal in the KIPROS project is a discourse model for Swedish, French and Greek prosody through a contrastive comparison of these languages. We are working towards the construction of this model by description, examination and classification of interactive categories and the organization of topic structure in the development of spoken discourse as well as by an extensive acoustic analysis of dialogue prosody and the classification of stereotyped prosodic patterns and their relation to interactive/topic structure. Regular prosodic patterns will be tested perceptually to establish how they are associated with the corresponding discourse categories. Examples of our working methodology are found in Bruce et al. 1988 and Touati and Bruce 1989.

The prosodic structure and its phonetic realization for the three languages involved in the project have been investigated in Bruce 1977, 1983, Touati 1987 and Botinis 1989. These studies have concentrated on prosody in sentence domain under well controlled experimental conditions, with systematically selected speech material of the question ~ answer type with only one speaker at a time. An apparent advantage with this method has been the examination of different prosodic categories with minimal influence from other prosodic contributions, as well as the different interplay of the prosodic parameters across an utterance.
Nevertheless, no matter how confident and well acquainted with laboratory speech we may feel, more complicated prosodic structures which we had not encountered do appear in real life speech communication outside the laboratory as a result of discourse contribution to prosody. As a matter of fact, this was the first question we put to ourselves as soon as we undertook the laborious enterprise of analyzing interactive discourse prosody: would we come across the familiar prosodic patterns from the laboratory? As we had expected, this was a solid framework for further research.

In the present study, we have concentrated on discourse intonation, i.e. tonal relations and pitch sequences and the way they are realized in the time course with Fo variations in the organization of discourse. By this, we by no means advocate an absolute superiority of Fo over duration and intensity, the phonetic realizations of temporal and dynamic variations. We simply assume that, apart from a study of the contribution of intonation to discourse structure for its own sake, tonal variations may be accompanied by temporal/dynamic variations. In cases where temporal/dynamic variations make independent, distinct contributions to discourse we leave it to future research.

All languages show pitch variations in their spoken form. However, both structure and function of tonal variations may differ across languages. Thus, in Greek a stressed syllable may be associated with a pitch-rise forming a high plateau with several post-stressed syllables, whereas in Standard Swedish a stressed syllable may be associated with a pitch-fall, in which case the relative timing of the fall carries the distinction between the two accents. On the other hand, the focal application in Greek starts its realization mainly from the stressed syllable of the element in focus in combination with a major pitch-fall into the post-stressed syllable(s) and a regular post-focal deaccentuation forming a low plateau. In Swedish, by contrast, the focal application is mainly realized with a major pitch-rise immediately after the fall associated with the stressed syllable forming a high plateau with several post-focal syllables and a successive compression of pitch-range for post-focal accent distinctions. In French, pitch-rises may be associated with minor and/or major syntactic boundaries, whereas focal applications (accent d’insistance) are realized on the first syllable of the focal element followed by a pitch-fall and a low plateau. In other words, post-focal deaccentuation is a regular pitch-pattern under certain speech conditions in both French and Greek but not in Swedish.

Intonation is multi-functional and scholars have devoted much research to describing tonal patterns, attributing them to different functions. Thus, intonation may have a grammatical function (Jones 1975, O’Connor & Arnold 1961), an attitudinal function (Hadding-Koch 1961, Uldall 1964), and an information function (Halliday 1967, Rossi et al. 1981). In addition, intonation may have a discourse function, and discourse analysts have been particularly interested recently in interactive structure (Brazil et al. 1980, Brazil 1985), topic structure (Brown et al. 1980, Schaffer 1984) and turn interplay structure (Schaffer 1983, Cutler & Person 1986). All the above intonation contributions may interact simultaneously on the prosodic manifestation of spoken discourse along with microprosodic interference (Di Cristo 1982), in contrast to laboratory speech where we have been able to investigate the effect(s) of one particular parameter at a time.

**EXPERIMENTAL DESIGN**

**Speech material**

In the KIPROS project we started our discourse research by investigating dialogues varying both in content and nature. Although we started with different kinds of dialogues for Greek, French and Swedish, our working methodology is to cover the same type of dialogues for these languages. Our justification for this choice has been our intention of gaining an insight into as wide a range of dialogue types as possible from the very beginning in the pursuit of a common fusion of the three languages. Thus, in Greek, the first step was to record a a face-to-face interactive dialogue in the laboratory; in Swedish a spontaneous telephone conversation from a radio entertainment program; and for French a face-to-face conversation between an adult and a child (cf. Bruce et al. 1988).

In the present investigation, we have selected a telephone conversation as similar as possible to the Swedish one. It is a conversation extract from an entertainment program, xoris anasa (Breathless), in which program listeners may phone in and have a short chat with the program leader, participate in a contest with the chance of winning a record or a small present, and request their favourite music (Appendix).

The characteristics of this type of dialogue are the non-visual interaction, i.e. facial expressions and gestures are excluded, the time available is very limited, and the program leader governs the outcome of the conversation to a considerable extent. Moreover, discussion is rather restricted to the activities and popular subjects of the program, such as enigmas, anecdotes, jokes, and the like. Our expectation is that in this kind of dialogue, prosody makes a major contribution to discourse as a consequence of the program listener’s
earnest desire to speak and the program leader's desire to create a lively program.

Speech participants
The discourse participants of the investigated dialogue are a young man, the program leader, and a relatively young mother whose small daughters as well as herself are eager fans of the program. We have limited our analysis to the second part of the dialogue, when the mother takes over, after the program leader has talked with one of her two daughters. They both speak standard Athenian (in the investigator's judgement, who has grown up in central Athens), and their language style is the 'in' style of the program. Although they do not know each other, in general social terms, the program leader has a 'touch' with his listeners and his voice and speech mannerisms are very familiar to the program participant, i.e. the conversation is friendly and highly cooperative. Both speakers make an equal contribution to the development of the dialogue, but it is the program leader who has the dominant role in the control of time and, consequently, the amount of information to be exchanged.

Speech analysis
The extract from the telephone conversation (53 s) was digitized and analyzed on our VAX 11/730 computer device with the API program of the ILS package in which Fo is traced through a modified cepstral calculation. We have had considerable aid in the segmentation of speech units from the waveform along with the Fo-contours. However, no matter what sophisticated technological equipment we may use, we nevertheless face various problems in this type of speech analysis.

Firstly, the quality of the recording and the speech material do not provide constantly reliable pitch configurations throughout the conversation. Secondly, in overlapping speech, it is not an easy task to isolate each participant's pitch contribution from the overall pitch contour or to segment each participant's timing with accurate precision. Thirdly, spontaneous speech is characterized by frequent haplogies that the native speaker of the language is hardly aware of. Thus, in our Greek material segments and/or syllables, most of the ones that are highly predictable have hardly any acoustic manifestation, although the present investigator, a native speaker of Athenian Greek, 'hears' them clearly. It is, rather, that a native speaker has a 'mental' representation of a word or a combination of words into larger units and once

s/he has (provided) sufficient acoustic cues for the identification of the intended message s/he may not be aware of the various haplogies forms which are associated with different styles and tempo. However, when a precise segmentation of the spoken unit is required, then a message ~ signal relation problem arises.

In our Greek transcription, as phoneticians and with regard to our foreign language readers, we follow a broad morphophonemic transcription and not the standard Greek orthography, paying particular attention to its alignment to the acoustic representation; in cases where there is haplographic acoustic manifestation, we follow an auditive transcription which is closer to the linguistic representation, so as not to confuse our Greek readers who might be presented with 'words' which do not exist in the Greek lexicon, had we followed a narrow acoustic/phonetic transcription. Still, in our material there are hardly any segmental/syllabic haplogies in environments of major interest to us, such as stressed/emphatic and turn/topic initial prosodic manifestations.

Analysis evaluation
We shall proceed to the analysis of our speech material with some basic assumptions of its organization. First, two (or more) discourse participants involved in a conversation are likely to talk rather coherently about one or more topics. They may be centered round a particular aspect of it or, as is more probable, they may take different aspects of it, in which case the (main) topic is subdivided into several subtopics.

Second, discourse participants may proceed to the development of one or several topics by their (regular) alternation to discourse contribution in terms of turn-units. Moreover, a turn-unit may be subdivided into sub-turn-units by external and/or internal prosodic cues as well as grammar and lexis.

Third, (sub)turn-units may have a prosodic structure in relation to their own semantic/grammatical structure as well as indicating their relation to previous or following turn-units, i.e 'look backward' ~ 'look forward' discourse cues. Finally, a turn-unit may be composed of a series of stressed and unstressed syllables in accordance with its lexical/grammatical structure.

The telephone conversation analyzed in this study is opened by a greeting from the program participant (PP) and a return greeting from the program leader (PL) (underlining indicates focus):
After exchanging the customary courtesy expression as part of the greeting 'how are you?', the program participant pays a compliment to the program leader and says that she hasn't missed a single program. The program leader encourages her not to miss one, and asks her name before he continues:

The program participant compliments the program leader once again, saying that the program is so pleasant that they all follow it (at home) and the program leader responds to her compliment and continues:
The program participant interrupts the program leader to inform him that she has another daughter as well who is jealous and the program leader, before he comments on it, asks the question:

**Figure 11**

![Graph](image1)

**F0(Hz)** how old is the other daughter? Thirteen

**PL:** TU-7.2

**PP:** TU-8

The program participant remarks that there will always be another opportunity and the program leader continues, closing the topic as well as the whole conversation:

**Figure 12**

![Graph](image2)

**F0(Hz)** Thirteen you should bring her up not to be jealous

**PL:** TU-8.1

**PL:** TU-8.2

**Figure 13**

![Graph](image3)

**F0(Hz)** I'd talk to her but I can't because my time is limited

**PL:** TU-8.3
the phrase with a major tonal rise (up to 240 Hz) at the end of the sub-turn-unit, as a turn-keeping cue to indicate his intention to keep the turn. He starts his next sub-turn-unit 6.2 (Fig. 8) \textit{i korul\={a} sas ...} with a major pitch variation (up to 180 Hz) to indicate (partly) the change of sub-topic and, apart from major tonal rises at the end of his successive sub-turn-units, to ensure the turn, he keeps his tonal variations at a lower level (under 160 Hz) until he is interrupted by the program participant.

The program leader's next sub-topic change for comment is his 10th turn-unit (Fig. 14), with minor pitch variations at a high plateau (200 Hz) throughout the phrase \textit{sas xarizo to epomeno trayy\={o}l} as well as at a low plateau (100 Hz) at the exclamatory expression \textit{e...}, to indicate his intention to close the topic/conversation.

Contrary to the subtopic initial boundaries, the subtopic final boundaries for both discourse participants are rather inclined towards a fall and/or low tonal manifestation except for the program leader’s turn-unit 6.5 \textit{kerisate omos tin a\={y}ipi mu} (Fig. 10), where he is interrupted by the discourse participant into an abrupt subtopic change.

\textbf{Turn-unit organization}

The turn-units of both discourse participants vary from one-word turn-units (PP: TU-5, Fig. 2) to several syntactically complex turn-units (PL: TU-5, Fig. 2-6). Turn-units between the two discourse participants may be separated by varying degrees of pause to no pause at all. The longest pause (1.5 s) appears in turn-unit PL: TU-6.1 separating the program participant’s compliment about the program and the program leader’s response \textit{Upon afto in’efxdristo}. Shorter pauses (between 0.2 and 0.3 s) appear between turn-unit PL: TU-4.3 and PP: TU-5, and PP: TU-5 and PL: TU-5.1 (Fig. 2), where the program leader asks the program participant’s name (\textit{t’onomd sas?}); the program participant answers \textit{lili} and the program leader echoes the program participant’s name \textit{lili} before he continuous with \textit{e... madam lili}. Finally, minimal pauses (less than 0.1 s) appear between the discourse participants’ final greetings (Fig. 15) and no pause in two interruptions.

Intra turn-unit pauses also appear with a high degree of variation; the longest one is 0.2 s between the program leader’s sub-turn-unit PL: TU-5.3 and 5.4, \textit{s’afti tin ekpobi...} and \textit{pote Sen kserume...}, a shorter one (0.03 s) between the sub-turn-unit PL: TU-5.4 and 5.5, \textit{pote Sen kserume...} and \textit{kserume omos...}, and no pause between sub-turn-unit PL: TU-8.2 and 8.3 \textit{na tis m\={a}bete...} and \textit{da tis milisa...}
The intonation of the program leader’s sub-turn-units shows the following three major patterns. First, the whole sub-turn-unit appears with equal pitch variations throughout: at a low pitch-level of 100 Hz (PL: TU-5.1, Fig. 2; TU-8.1, Fig. 12; TU-10.2, Fig. 14), at a mid pitch-level of 150 Hz (PL: TU-5.2, Fig. 3; TU-8.2, Fig. 12), and at a high pitch-level of approximately 200 Hz (PL: TU-5.3, Fig. 4; TU-10.1, Fig. 14). Second, the sub-turn-unit final boundaries appear with the largest pitch variation within the sub-turn-unit, quite regularly up to 250 Hz (PL: TU-6.1, Fig. 7; TU-6.3, Fig. 9; TU-6.4 and 6.5, Fig. 10). As a matter of fact, these are the program leader’s largest pitch variations in the whole dialogue, although they are only indirectly associated with a stressed syllable. Third, the initial part of the sub-turn-unit appears with a major pitch variation up to 200 Hz (±20 Hz), associated with a stressed syllable of a word in focus, followed by a deaccented falling and flattening pitch-contour to the end of the sub-turn-unit at well under 100 Hz (PL: TU-5.4, Fig. 5; TU-5.5, Fig. 6; TU-6.2, Fig. 8; TU-8.3, Fig. 13).

It is interesting to notice that the word poté in the program leader’s sub-turn-unit s’afti tin ekpothi poté ... (PL: TU-5.3, Fig. 4) has a pitch variation up to 200 Hz, and when he repeats part of this sub-turn-unit in his next sub-turn-unit poté ben ksérume ... (PL: TU-5.4, Fig. 5) the very same word poté reaches exactly the same pitch level (200 Hz), whereas the rest of the sub-turn-unit is deaccented. Moreover, in the following sub-turn-unit ksérume ómos ... (PL: TU-5.5, Fig. 6) the pitch variation of the word ómos reaches 220 Hz, exactly the same as the two previous sub-turn-units’ word ksérume, with the rest of the sub-turn-unit deaccented as well.

From the point at which the program leader informs the program participant that his time is limited there are only minor pitch variations on stressed syllables and no focal and/or turn-keeping ones up to the end of the dialogue.

Finally, in this dialogue, there are two sub-turn-units of the program leader which elicit an answer by the program participant. The first one t’ónomá sas (PL: TU-4.3, Fig. 2), a declarative question, has minor pitch variations associated with the stressed syllables and a major pitch-rise on the final, unstressed syllable. The second one po- pósos xronón íne i ìlí korí (PL: TU-7.2, Fig. 11), a wh-word question, has a pitch-rise on the second, unstressed syllable of the wh-word pósos and falling and flattening intonation contour to the end of the sub-turn-unit.

DISCUSSION

In the present study we have investigated the discourse intonation of a short Greek dialogue. Although there are countless dialogues of the same as well as of completely different types in a language, we assume that intonation has a limited number of patterns to exploit in a great number of communication situations and, accordingly, the relation between pattern and function is likely to appear one to many. Thus, in this dialogue, we have concentrated our attention mainly on patterns which appear interesting to us either because they are familiar from laboratory speech with well documented functions or because they are new to us with, apparently, specific functions.

In our material we have found several (sub)turn-units with pitch variations throughout or only in parts of them, i.e. what has been referred to as broad vs. narrow focus (Ladd 1980, Gussenhoven 1983). With no specific element in focus, there seems justifiable phonetic evidence for three pitch-levels: low, mid and high, i.e. what has been discussed with different connotations under the term 'key' (cf. Brazil et al. 1980), dividing the speaker’s pitch-range into three approximately equal parts. Nevertheless, a more economical system of a non-high ~ high binary key has been proposed (Brown et al. 1980) which may capture a basic distinction of main ~ secondary (parenthetical) information (Bruce et al. 1988).

When the whole (sub)turn-unit is in focus, what some call 'neutral', others 'broad' focus, (i.e. when no speech item stands out as more 'informative' from the rest of the unit), quite proportional pitch variations appear associated with the majority of the stressed syllables throughout the unit. On the other hand, in narrow focus, there is a wide pitch-range on a stressed syllable in the first part of the unit, which is in accordance with the opinion that speech elements in focus with no particular syntactic constraints are usually topicalized in Greek (Botinis 1989). Similar tendencies for spontaneous speech in English are reported in Brown et al. 1980.

(Sub)turn-units with early ‘narrow’ focus utilize a postfocal deaccentuation, i.e. a typical pattern of laboratory speech, rather than enlarging the focal pitch-range in a comparison with units with no particular narrow focus. The focal item may be repeated information (Bruce et al. 1988), in which case it emphasizes the importance the speaker attaches to this item. In our material, the focal item has the largest pitch-range within the unit in which it is realized and has a backward reference, i.e. it shows the persistance of the speaker to increase the common ground of the hearer and himself (Brazil 1985) and/or modify the hearer’s belief space (Pierrehumbert...
Moreover, in our dialogue, it is mainly the dominant speaker who utilizes focal applications, asserting his privilege to decide what is important or not and let his converser agree or indirectly comment on it.

A (sub)turn-unit, apart from a major, focal pitch-change and/or pitch-range initially, may appear with a large pitch-rise during the final part of the unit, even on unstressed syllables. This pitch-rise marks the intention of the speaker to keep his turn, especially in cases where his (sub)turn-unit is of complete syntactic structure (Brown et al. 1980), in contrast to (sub)turn-unit low/pitch fall where the speaker is not so anxious to keep his turn, i.e. the hearer may take over. Similar manifestation for continuation rise has been observed for German, even for laboratory speech (Bannert 1985). Although the continuation rise clearly denotes more to come, the low/pitch-fall does not exclude this possibility, i.e. there is an asymmetry in the final (sub)turn-unit boundary; in the former case a hearer’s turn taking over would be interpreted as an interruption, whereas in the second one a take over would be permitted.

Apart from their distribution differences the two major pitch-changes, focal ~ turn-keeping, seem mutually exclusive within a (sub)turn-unit in this material. Both have an exophoric function (Halliday & Hasan 1976, Brown & Yule 1983) but whereas in the focal manifestation the speaker relates the content of the current (sub)turn-unit to the already developed discourse, in the turn-keeping manifestation the speaker emphasizes the importance of what is to follow. Moreover, their tonal manifestation is quite distinct. The focal manifestation is realized by a pitch-fall at the post-stressed syllable(s), whereas the turn-keeping manifestation by a pitch-rise.

In our previous studies (Botinis 1982, 1989; Girding et al. 1982; Bruce et al. 1988), we found the largest pitch range-variations associated with the focal elements; in the present investigation, the turn-keeping pitch-rise/maxima are constantly higher (=40Hz) than the focal maxima, reaching the speaker’s upper tonal range, as a result of the dialogue type and the limited time the discourse participants have at their disposal to exchange a certain amount of information. Pierrehumbert and Hirschberg 1987 attribute the backward ~ forward directionality of the (sub)turn-unit (intonational phrase) to the (final) boundary tones. Without disputing the significance of the final tones in relation to what follows, we think that earlier (focal) tones are equally important in relation to what precedes. However, further experimental and/or perceptual evidence is required on this issue.

Up to this point we have discussed some intonation forms and functions a (sub)turn-unit may be associated with. We have not justified its very existence and the principles on which a speaker may subdivide his turn-unit. The turn-unit is unproblematic as it is one speaker’s verbal contribution to the discourse until another speaker takes over. In principle, it is the choice of the speaker how to subdivide his turn-unit utilizing external and/or internal prosodic cues.

A major internal criterion is a wide pitch-change/range in relation to the focal ~ turn-keeping cues. An external criterion is the pause, particularly in a negative way, i.e. there is no pause within a (sub)turn-unit except in cases where the speaker has planning problems, repetitions, etc. Another external criterion is a pitch-change between complete syntactic boundaries which in combination with a pause and/or an internal pitch-change provides reliable prosodic cues to the sub-division of a turn-unit. Nevertheless, it is not seldom that there exist neither internal nor external phonetic cues, e.g. in intonational sandhi occurrences, where it seems reasonable to take syntactic/semantic factors into consideration (Cruttenden 1986).

The reader will have noticed that we have avoided taking into account the traditional ‘tone unit’ in intonation studies. This is not because of difficulties in defining and identifying it (Brown et al. 1980), since methodological inadequacies do not justify its rejection as a descriptive unit. It is rather that a tone unit presupposes at least a (stressed) syllable, the ‘tonic’ (Halliday 1967, Cruttenden 1986). We simply assume that minimal linguistic/paralinguistic verbal expressions not even reaching the level of a syllable, at least as it is traditionally described, may play a major role in the development and organization of spoken discourse in which case a speaker may have contributed with a complete turn-unit. A facial expression or a gesture may have similar effects in a face-to-face interaction. Furthermore, although we have observed some regular subtopic intonation cues, we have not discussed the issue further, awaiting a more detailed description of intonation topic/subtopic contribution in ongoing research.
The discourse material of a dialogue between the program leader (PL) and a program participant (PP) from the entertainment program xorís anása (Breathless), recorded from the local Athenian radio station Diavlos 10 in summer 1988.

PP: xá su kóstα
   Hi Kostas!

PL: xá su mamá ti kánis pos įse
   Hi mum! How are you? How do you feel?

PP: ti kánethē
   How are you?

PL: poli kálα
   Very well.

PP: tin ekpobí tin akáło ap’ton keró pú ’xi aniksi
   I’ve been listening to the program since it started.

PL: xristé mu
   Heavens!

PP: ōn éxo xási méra
   I haven’t lost a day.

PL: ōn prépi na xásume méra ſiati s’ afti tin ekpobí t’ onomá sas
   We shouldn’t lose a day, because in this program, your name?

PP: Lili
   Lili.

PL: Lili e.. madam Lili s’ afti tin ekpobí poté ōn ksérume ti b’ akásume poté
   ōn ksérume ti b’ akásume ksérume ómos t’a perásune efxráristes t’ ořes Lili, ch madam Lili, in this program we never know what we will hear, we never know what we will hear, we know however we will spend two pleasant hours.

PP: afó in’ efxráristo ſe toso poli oréa pu tin akúde ſi
   This is nice, it is so pleasant that we all hear it.

PL: lipón aftó in’ efxráristo i koríla sas móliς pro.. kérísse.. éna ãlisko.. éna kalokéri xorís anása esís ōn kerísate típota kerísate ómos tin aξípi mu
   Well, this is nice, your little daughter has just.. won.. a record .. ‘A breathless summer’, you haven’t won anything, you have however won my love.

PP: éxo ke miá áli kóri pu zílevi
   I have another daughter too, she is jealous.

PL: zílevi po- póso xronón ÿne i áli kóri
   Is she? How old is the other daughter?

PP: ŝekatrión
   Thirteen!

PL: ŝekatrion na ſiis máðete na mi zílevi t’a tis milúsa alá ōn boró ſiatí lme periorisménon apó xróno
   Thirteen, you should bring her up not to be jealous, I’d talk to her but I can’t because my time is limited.

PP: edaksi ſen pirázi
   OK, it doesn’t matter.

PL: edaksi
   All right?

PP: miá álí forá miá álí forá
   Some other time, some other time.

PL: sas xarízo to epómeno traξáκ e..
   I dedicate to you the next song, OK?

PP: efxarístο
   Thank you!

PL: okeí lipón
   All right.

PP: jía sas
   Bye!

PL: jía sas
   Bye!
BIBLIOGRAPHY


