



The Pragmatics of Intonational Meaning

Julia Hirschberg

AT&T Labs – Research
180 Park Avenue, Florham Park, NJ 07932-0971, USA
julia@research.att.com

Abstract

Starting from Carlos Gussenhoven's proposal that intonational meaning be understood in terms of three biological codes, this paper suggests an augmentation of that proposal to connect human experience of these biological phenomena to the process of spoken communication. Under this scenario, the codes give rise to a set of Conversational Implicatures, similar to those defined by H. Paul Grice in his description of Cooperative Conversation. Certain additions to Grice's Maxims of Cooperative Conversation are suggested, to capture the communicative effect of features of Gussenhoven's Frequency, Effort, and Production Codes.

1. Introduction

In his thought-provoking paper, "Intonation and interpretation: phonetics and phonology" [11], Carlos Gussenhoven outlines a broad account of universal intonational meaning in terms of three biologically determined codes, which are exploited as speakers control the phonetic implementation of utterance production to convey different kinds of intonational meaning. These interpretations may be either affective, conveying attributes of the speaker, or informational, conveying attributes of the message in nature. Furthermore, they may be signaled not simply by the physiological condition defining the code, but by making reference indirectly to that condition using SUBSTITUTE FEATURES, phonetic forms that the hearer can associate with the primary form. Universal intonational meanings may be grammaticalized in particular languages such that universal meanings are more likely to be perceived and interpreted by speakers of the languages than by others.

For example, Ohala's [16] FREQUENCY CODE, which Gussenhoven adopts as his first biological code, derives from the fact that larynxes vary in size, leading to differences in the speech of adults and children, males and females. Traditional cultural dominance exercised by adult males has led to an association of lower pitch with dominance and higher pitch with submission. So, the use of lower or higher pitch by any speaker may convey affective information associated with dominance (e.g., confidence, aggressiveness) or submission (e.g., politeness, friendliness) [19, 13]. The Frequency Code may be used to convey informational interpretations also: Gussenhoven proposes that the uncertainty and questioning interpretation of certain intonational contours derive from the high pitch or rising pitch associated with some interrogative contours vs. the lower or falling pitch associated with assertions [14, 12]. Interpretations derived from speaker and hearer's knowledge of the Frequency Code may be conveyed not simply by an increase or decrease in overall pitch, but by various substitute features, such as delayed peak, in some languages. The prime example of grammaticalization of the informational use of the Frequency

Code is the common encoding of rising contours as questions, although not all rising contours function as such [3]. And some languages indeed have falling interrogatives and rising declaratives [8].

Gussenhoven also identifies two other biological codes which function similarly as reference points for intonational meaning — the EFFORT CODE and the PRODUCTION CODE. The Effort Code associates the increased effort expended on speech production with the increased precision of articulation and a wider overall pitch range. Gussenhoven notes that a general interpretation of such expanded range is that the speaker intends the item or proposition associated with the speech to be seen as of greater importance than other items. Affective meanings derived from the Effort Code may be obligingness, surprise, or agitation [19, 12]; the most widely attested informational interpretation is 'emphasis', when listeners interpret higher peaks associated with a mentioned item as conveying greater informational prominence as well [20]. The meaning of intonational prominence is often grammaticalized in the interpretation of prominence as intonational focus (*John only introduced MARY to Sue* vs. *John only introduced Mary to SUE*). Late peak can also function as a substitute for peak height in conveying prominence for this code, as when complex pitch accents appear to be interpreted as conveying narrow focus in several languages.

The Production Code is defined by Gussenhoven as speakers' expenditure of increased effort on the beginnings of phrases, where subglottal pressure is higher, than at the end. So, there will be a gradual drop over the phrase in both intensity and f_0 , known as DECLINATION. Gussenhoven claims that the information-bearing aspects of declination are not associated with the slope of this decline but with the relative highs and lows of the edges — the beginning and end of the phrase. There is considerable evidence showing that high beginnings signal changes in topic structure, high endings indicate continuations of topic, and low endings indicate topic endings [7, 15, 5]. Delayed peak in the first accent of an intonational phrase and high register can both substitute for wide pitch span for the Production Code [20, 6].

The account Gussenhoven outlines, by which universal experience of physiological codes can be exploited by speakers to convey particular meanings — even while the physical conditions which naturally give rise to these codes may not obtain in the particular conversation — is an intriguing one. However, it remains vague as to how universal experience of the physical aspects of speech might come to be linked to assumptions about speaker meaning. A small step in addressing that issue might be taken by attempting to make these communicative assumptions explicit. A possible model for linking conventional knowledge of human behavior — albeit in the social realm — with communicative effect, can be found in the Maxims of Cooperative

Conversation described by H. Paul Grice 75.

2. The Gricean Framework

In his account of Cooperative Conversation, Grice posited that knowledge of certain conventions of communicative behavior, shared by speakers and hearers, can be seen as licensing certain interpretations of utterances beyond their simple semantics, which he termed *IMPLICATURES*. For example, each of the following utterances may, in some contexts, convey more than a speaker actually *says*:

- (1) a. Some people left early.
b. Mary got married and had a baby.
c. George has three children.

Truth-functional semantics cannot capture the additional meaning which may be understood from the utterance of (1a) that *not everyone left early* — the fact that some left early is true even if everyone left. Nor can it explain the likely inference from (1b) that there is an ordering among the conjuncts: Mary first got married and then had a child. And how can we explain why (1c) may be interpreted as an exhaustive count of George's children? George having three children is perfectly true even if he has, in fact, five. The context-dependence of each of these examples is illustrated by the *DEFEASIBILITY* of the inferences just suggested, as shown in their counterparts in (2):

- (2) a. Some people left early, but not everyone.
b. Mary got married and had a baby, but not in that order.
c. George has three children, and, in fact, he has five.

In each case, the second clause is said to *CANCEL* the implicature licensed by the first, in the sense that the implicature simply does not arise in this context.

Grice distinguished those (truth-functional) inferences, that logically follow from an utterance, by applying standard rules of deduction upon the utterance's semantic representation ("what is said") from the non-truth-functional and context-dependent meanings described above ("what is implicated"). He termed these latter meanings *CONVERSATIONAL IMPLICATURES*.¹

Grice explained how conversational implicatures are licensed by speakers and understood by hearers by proposing that participants in conversation share knowledge of certain underlying universal conversational goals, subsumed under his *COOPERATIVE PRINCIPLE (CP)* — "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." [9, page 45] Because the CP is shared knowledge, speakers can communicate inferences beyond the conventional force of an utterance, by comparing 'what is said' to 'what might be said' in the exchange. That is, by interpreting what a speaker says in the context of the shared goals of the conversation, hearers may infer that nothing important to the current conversation has **not** been stated. For example, if a speaker utters (1a) when asked "Did everyone leave the party early?", the questioner is entitled to understand the implicature

¹Grice also identified an intermediate form of meaning, non-truth-functional meanings but context-independent, which he termed *CONVENTIONAL IMPLICATURE*. These are exemplified by the meaning conveyed by a conjunct like *but*, which is typically represented in the same way as *and* in truth-functional semantics. *He's a New Yorker but I like him* appears to convey something different from *He's a New Yorker and I like him*, for example.

Not all people left earlier, if the questioner assumes that the speaker is behaving cooperatively. In another context (say, the context is "I hear Jones' direct reports were told they must all stay till 6 p.m.", where the extent of early departures might be less important to the conversation than any early departures at all, a hearer might *not* draw the same inference.

This context-dependence of implicature is critical to an extension of Gussenhoven's account of intonational meaning. As is well known, emphasis may convey focus, it does not always do so; while increased pitch may signal a new topic, this is not always the case; and so on.

To codify the conduct embodied in the CP more specifically, and to explain how particular implicatures arise, Grice identified a number of maxims of cooperative conversation, four of which are usually treated as core:

Maxim of Quality:

Try to make your contribution one that is true.

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Maxim of Quantity:

a) Make your contribution as informative as is required (for the current purposes of the exchange).

b) Do not make your contribution more informative than is required.

Maxim of Relation

Be relevant.

Maxim of Manner:

Be perspicuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

The Maxim of Quality enjoins speakers to be truthful. The Maxim of Quantity enjoins them to say as much but only as much as is relevant to the exchange. The Maxim of Relation captures the notion that hearers expect what speakers say to be relevant to the purpose of the conversational exchange, and the Maxim of Manner requires that speakers provide information in a form appropriate to the hearer and to the purpose of the exchange. Thus the implicatures licensed by (1a) and (1c), in the absence of a mitigating context, can be derived from the Maxims of Quality and Quantity: if the speaker has truthfully said all that is relevant to the exchange in these cases, the hearer can conclude from the utterance of (1a) that others did **not** leave and from the utterance of (1c) that George has **only** three children.

Critically, for our adoption of a similar account of intonational meaning, the Gricean program does not assume that conversational participants always obey his Maxims of Cooperative Conversation; the knowledge that these conventions are shared by the larger community is sufficient to account for how implicatures arise.

The Gricean maxims may also be *FLOUTED*, or ostensibly violated, to communicate some additional meaning. Grice's classic example of this is the case of a philosophy professor who appears to violate the Maxim of Quantity in writing the following letter of recommendation for a pupil applying for an academic position:

Dear Sir,

Mr. X's command of English is excellent, and his attendance at tutorials has been regular.

Yours, etc.

Grice explains that of course no maxim is actually violated by this letter. The very fact that in such a situation a writer would generally be expected to make some reference to the pupil's aptitude in philosophy — but that this writer does not — conveys to his communicative partner that the writer has said as much as he truthfully can say (obeying both the Maxims of Quantity and Quality). Mr. X has no credentials which fit him for the position.

We might hypothesize that the biological codes outlined by Gussenhoven form the basis for similar conversational maxims, similarly understood by speakers and hearers, which, while not always followed, nonetheless represent “norms” of speech production. The shared knowledge of these norms, then, may form the basis for certain additional meanings which can be conveyed via intonational variation. To flesh out this proposal a bit more, we will propose a few maxims which might connect Gussenhoven's biological codes to communicative conventions.

3. Intonational Meaning and Conversational Implicature

To take a Gricean approach to intonational meaning, we might propose some additional Maxims of Cooperative Conversation derived from the Biological Codes identified by Gussenhoven. Knowledge of the Frequency, Effort, and Production Codes might be encapsulated in these additional communicative conventions, such that a cooperative conversational partner will employ linguistic cues based upon these codes to signal associated meanings.

The Frequency Code might give rise to a Maxim of Pitch. This might be specified as: “Try to match the rise or fall in the pitch of your utterances to the degree of confidence you wish to convey. Let your pitch rise to convey uncertainty and fall to convey certainty.” A classic example of the exploitation of this maxim might be the old Russian emigre joke about a staunch Bolshevik forced to confess publicly, who reads the following sentences — each with rising intonation: “I was wrong? And Stalin was right? I should apologize?”

Different languages may conventionally choose to convey uncertainty under different circumstances, accounting for the fact that not all languages have rising contours for *yes-no*-questions. But since, as Bolinger has noted, *wh*-questions in most languages are stereotypically produced with falling contours [3], even while *yes-no*-questions are produced with rises, and there is no reason to think that speakers of one form of question must be less certain than those of another.

Viewing the meaning conveyed by intonational variation as a case of conversational implicature, arising in some cases from the Maxim of Pitch, also allows us to account for cases in which rising pitch does *not* result in the impression of speaker uncertainty. Like other conversational implicatures, intonational meaning appears to be both non-truth-functional and context-dependent. For example, not every rising contour conveys speaker uncertainty. A speaker may obey the Maxim of Pitch when they are truly uncertain, but they may also exploit the shared knowledge of the maxim to different effect, e.g. by using a rising contour to convey irony or to produce a rhetorical question. So, *Are we disturbing you, Mr. Smith?* said to a student asleep in class will *not* convey any genuine uncertainty on the part of Mr. Smith's professor.

Similarly, from the Effort Code, we might derive a Maxim of Emphasis, such as “Try to make informationally important portions of your speech intonationally prominent.” Speakers

or language groups may implement this prominence differently from the higher pitch range, loudness and higher peaks common in Germanic and other languages. So, in languages like English, focus may be realized by emphasis placed on the linguistic realization of the focussed item, as in *John asked MARY to talk to Sue*. However, the importance such emphasis might attach to *Mary* is context dependent. If Sue has previously been talked about in the discourse, emphasis on *Mary* may only reflect the deaccenting of Sue, as GIVEN, or “old” information, as in (3a).

- (3) a. A: Sue is being so unreasonable. I think someone should talk to her.
B: John asked MARY to talk to Sue.
- b. A: Did John ask Rita or Mary to talk to Sue?
B: John asked MARY to talk to Sue.

Alternatively, in (3b), the prominence of *Mary* may be reasonably interpreted as increased importance, since she is being selected from a set of potentially relevant discourse entities. The context dependence of intonational prominence can also be seen in cases of structurally ambiguous narrow focus, as in (4a) and (4b):

- (4) a. A: Is she the girl in the red skirt?
B: She's the girl in the red DRESS.
- b. A: Which is your friend's cousin?
B: She's the girl in the red DRESS.

In (4a) B's reply can be interpreted as narrow focus, in a contrastive context. But in (4b) it is more likely to be interpreted as focussing broadly on the entire NP.

Finally, from the Production Code, we might derive a Maxim of Range: “Let the width of your pitch range reflect the location of your utterance in the topic structure of the discourse. Increase your range to start new topics. Decrease your range to end old ones.” Clearly this maxim is not always followed by speakers, especially in more casual speech. However, there is considerable empirical evidence that over larger spans of speech this maxim does hold true for a variety of languages [4, 17, 1, 2, 10, 18, 15].

A second maxim, also related to production, might capture the observation that speakers tend to “chunk” their speech into meaningful units, either syntactically or semantically, although they may not always observe this regularity either. So, a Maxim of Phrasing might be formulated as: “Phrase your utterance so that it is divided into meaningful portions of speech.” Again, patterns of behavior found across large corpora of speech suggest that, while this maxim is not always obeyed, it may be viewed as something of a norm. A speaker who said *He takes the nuts — and bolts approach*. would probably be viewed as flouting the Maxim for comic effect — or suffering some production failure.

4. Discussion

This paper suggests an augmentation to Gussenhoven's proposal that speaker and hearer's shared knowledge of three biological codes can be seen as giving rise to a variety of meanings associated with intonational variation. This addition is based upon an additional hypothesis, that much of intonational meaning can be viewed as an instance of Gricean conversational implicature. That is, that these meanings are context-dependent and defeasible.

While students of intonational meaning generally look for the regularities in intonational interpretation, such as “Increased prominence is interpreted as focus” or “Phrase boundaries occur

at syntactic boundaries”, there are too many counter-examples in normal speech production to conclude that particular intonational behavior maps simply to clear interpretations. Even when empirical studies have found regular associations between phenomena such as increased pitch and new topics, or intonational prominence and perceived focus, these studies also find many occasions when the commonly accepted “meanings” of intonational features do not seem to hold. The context-dependence of intonational meaning then, appears to justify its classification as a form of conversational implicature. Some additional maxims particular to cooperative **spoken** conversation are proposed, in order to link Gussenhoven’s biological codes to the context-dependent interpretations that hearers appear to understand.

The set of maxims outlined here is intended to suggest rather than to define the pragmatics of intonational meaning. Indeed, even if one accepts as a working hypothesis that intonational meaning can be characterized as a form of conversational implicature, one might still wish to modify the set of conversational maxims that best encapsulates these conventions of spoken discourse. One can imagine, for example, a corollary to the Maxim of Emphasis, capturing the frequent but far from universal tendency of given information to be deaccented, while new information is accented. The Maxim of Pitch might be augmented as well to encompass contour intonation beyond simply the meaning of rising contours. Whatever the optimal set of such maxims may be, the nature of intonational meaning as non-truthfunctional, context-dependence, and defeasible remains a strong claim and one which should be tested.

5. References

- [1] Cinzia Avesani and Mario Vayra. Discorso, segmenti di discorso e un’ ipotesi sull’ intonazione. In *Corso di stampa negli Atti del Convegno Internazionale "Sull'Interpunzione"*, pages 8–53, Vallecchi, Firenze, 1988.
- [2] Gayle M. Ayers. Discourse functions of pitch range in spontaneous and read speech. Presented at the Linguistic Society of America Annual Meeting, 1992.
- [3] D. Bolinger. Yes-no questions are not alternative questions. In H. Hiz, editor, *Questions*, pages 87–105. Reidel, Dordrecht (Neth), 1978.
- [4] G. Brown, K. Currie, and J. Kenworthy. *Questions of Intonation*. University Park Press, Baltimore, 1980.
- [5] Johanneke Caspers. Who’s next? the melodic marking of question vs. continuation in dutch. *Language and Speech: Special Issue on Prosody and Conversation*, 41(3-4), 1998.
- [6] A. Chen, C. Gussenhoven, and T. Rietveld. Language-specific uses of the effort code. In *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 2002.
- [7] Ronald Geluykens and Marc Swerts. Prosodic cues to discourse boundaries in experimental dialogues. *Speech Communication*, 15:69–77, 1994.
- [8] Matthew K. Gordon. The intonational structure of Chickasaw. In *Proceedings of the XIV International Congress of Phonetic Sciences*, pages 1993–1996, San Francisco, 1999.
- [9] H. Paul Grice. Logic and conversation. In *Syntax and Semantics*, volume 3. The Academic Press, New York, 1975. From 1967 lectures.
- [10] Barbara Grosz and Julia Hirschberg. Some intonational characteristics of discourse structure. In *Proceedings of ICSLP-92*, Banff, October 1992.
- [11] C. Gussenhoven. Intonation and interpretation: Phonetics and phonology. In *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 2002.
- [12] C. and T. Rietveld Gussenhoven. The behavior of H* and L* under variations in pitch range in Dutch rising contours. *Language and Speech*, 43:183–203, 2000.
- [13] J. Haan, L. Heijmans, T. Tietveld, and C. Gussenhoven. The morphological structure of dutch rising contours. Submitted for publication.
- [14] Kerstin Hadding-Koch and Michael Studdert-Kennedy. An experimental study of some intonation contours. *Phonetica*, 11:175–185, 1964.
- [15] Julia Hirschberg and Christine Nakatani. A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th Annual Meeting*, Santa Cruz, 1996. Association for Computational Linguistics.
- [16] J. J. Ohala. Cross-language use of pitch: An ethological view. *Phonetica*, 40:1–18, 1983.
- [17] K. Silverman. *The Structure and Processing of Fundamental Frequency Contours*. PhD thesis, Cambridge University, Cambridge UK, 1987.
- [18] Marc Swerts and Mari Ostendorf. Discourse prosody in human-machine interactions. In *Proceedings ESCA Workshop on Spoken Dialogue Systems: Theories and Applications*, pages 205–208, Visgo, Denmark, May/June 1995. ESCA.
- [19] E. Uldall. Dimension of meaning in intonation. In D. Abercrombie, D. B. Fry, P. A. C. MacCarthy, N. C. Scott, and J. L. Trim, editors, *In Honour of Daniel Jones: Papers Contributed on the Occasion of his Eightieth Birthday*, pages 271–279. Longman, London, 1964.
- [20] A. Wichmann, J. House, and T. Rietveld. Peak displacement and topic structure. In *Proceedings of the ESCA Workshop on Intonation*, pages 329–332, Athens, 1997.