

Bericht Nr. 88, Juni 1997

Arbeitspapiere des
Sonderforschungsbereichs 340

ISSN 0947-6954/97

SPRACHTHEORETISCHE GRUNDLAGEN FÜR DIE COMPUTER LINGUISTIK

Arthur Merin & Christine Bartels :
Decision-Theoretic Semantics for Intonation



Universität Stuttgart

IBM Deutschland Informationssysteme GmbH

Universität Tübingen



DECISION-THEORETIC SEMANTICS FOR INTONATION

ARTHUR MERIN AND CHRISTINE BARTELS

ABSTRACT. We propose a decision-theoretically grounded, transcontextual semantics for intonation contours of natural languages, in particular, of English. We critically examine a recent set of proposals for a tonal semantics (Pierrehumbert & Hirschberg 1990; Hobbs 1990) which identifies smallest independently meaningful units with level tone phonemes and, like other current approaches (e.g. Gussenhoven 1984), finds denotata in epistemic and discourse relations. Retaining the now standard autosegmental description for intonation phonology (Pierrehumbert 1980) that underlies these discourse-epistemic accounts of tonal meaning, we propose instead a domain of possible denotata in a space of elementary social relations characterizing negotiation (Nash 1953) of joint deontic-boulomaic or epistemic commitments between cooperating autonomous agents (Merin 1994). Intonational morphemes are tone-level transitions, coding ostensible allocation of power of choice of negotiable situational parameters. The basic symbolism is iconic of ontogenetically salient associations of pitch and social power; standard default associations of preference and initiative parameters engage 'emotive', 'discourse-anaphoric' and, via Peirce/Hintikka games, 'logical' concomitants. We exemplify with a variety of classic examples (questions, negation scope) including the so-called Contradiction Contour (Lieberman & Sag 1974) familiar to a wider audience from Johnson-Laird (1988).

1. STATE OF THE ART

Description of intonational units, as Pierrehumbert & Hirschberg (1990) [PH] observe, cannot get far without a semantics for them. Autosegmental phonology for intonation (Pierrehumbert 1980) has been well motivated on intraphonological grounds. Phonological units instantiate the finitary schema $X@$ where X instantiates to one of two 'static, level' tones in the set $\{H,L\}$ – to be realized phonetically by vocal pitch, i.e. fundamental f_0 frequency – and $@$ is a stress- or position-related feature with values in $\{*,-, \%$.

ii. Did the suspect speak French or Flemish?

L* H* L-L%

iii. Did the suspect speak French(L*) or Flemish(H* L-L%)?

‘French’ is ‘given’ or ‘not-new’ and ‘Flemish’ is ‘new’ to the discourse context. This is not at all borne out by intuitions of given/new-ness, which are not, and could not be, claimed to be a matter of intrasentential syntactic precedence among disjuncts. Permuting ‘French’ and ‘Flemish’ in (2) changes nothing intuitable regarding ‘given/newness’ but would force a switch in accent tones. This is unlike the case in (1), where a change in clause sequence would not affect tone assignment to accented elements.

In other cases the PH semantics fails to make requisite predictions. Examples include variations in phrasal intonation of questions. The term ‘question intonation’ standardly refers to contours characterized by a final rise. Yet AQs such as (1) above must obligatorily fall on the last disjunct. Yes-no questions (YNQs) such as those in (3a,b)

(3)a. I didn’t know John took a job all the way over
in Redwood City.

now?

CAR

Does he have a

[Does he have a car(H*) now(H-H%)?]

b. John did good work for us last year. But I doubt that
we could still have him going round on his bicycle.

CAR

Does he have a

now?

[Does he have a car(H*) now(L-L%)?]

may either rise or fall. (By corpus statistics, 2/3rds fall. Note: each of (3a,b) has a single speaker.) The same goes for wh-questions (WHQs) such as (4a,b),

(4) A: I still have that mysterious backache. It simply
won’t go away. I even went to see an orthopedic
specialist yesterday.

say?

HE have to

a. E: And what did

[And what did he(H*) have to say(H-H%)?]

HE

- b. E: And what did have to
 say?
 [And what did he(H*) have to say(L-L%)?]

although these tend more strongly towards a falling pattern. Only in their use as ‘echo questions’ (5a,b) do the latter two question types consistently end with a rise:

- (5)a. A: Did Amy get the summer job at the embassy?

sy?

EMbas

- E: Did she get the job at the
 [Did she get the job at the embassy(H* H-H%)]
 (...Was that your question?)

- b. A: Amy started her job at the embassy last week.

week?

WHERE last

- E: She started her job
 [She started her job where(H*) last week(H-H%)?]

Further, a purely epistemic or discourse-syntagmatic semantics does not directly engage, and can rarely be construed to implicate, psychosociological aspects of intonational meaning that are traditionally treated in the sociological category of ‘register’ or a psychological one of ‘emotive’ meaning. Examples are intuitively verifiable glosses such as ‘less-than-best’, ‘defensive’ or ‘submissive’ attaching to the vocal response of, say, one Evita at a job interview with All-Boss Inc. Our example employs an instance of what is traditionally called a Fall-Rise (FR) to render musical Hirschberg’s (1985) – originally non-intonational – dialogue

- (6) A: Do you speak Portuguese?

HUS

does.

- E: My band
 H* L-H%
 [My husband(H*) does(L-H%).]

2. DECISION-THEORETIC SEMANTICS FOR INTONATION

The present, alternative proposal leans towards more traditional approaches in taking kinetic tones, i.e. transitions or movements of perceived pitch, to be the basic meaning-bearing units, i.e. intonational morphemes. The motivation, by definition, will thus be to a large extent semantic. In the present case, where the PH phonology is taken for granted, it is exclusively so: no intra-phonological or suprasegmentally morphophonemic points of critique are addressed.

2.1 Construction of a Domain of Denotata

We choose a ‘top-down’ approach. A denotation space for intonation must satisfy constraints of

1. requisite variety,
2. ontogenetic significance,
3. phylogenetic plausibility.

The first rules out accounts based on physiological excitation (Bolinger 1982) as much too non-specific. (There is very little one cannot get excited about, i.e. arousal is cognitively and emotively non-specific.) It also rules out given/new-style discourse relations, as does the second and, by speculative extension thereof, the third. In return, the die-hard favourites of speculative phylogeny lack any structure of accounts reckoned: alarm cries and mating calls do not announce the road to reason.

If intra-individual physiology is too broad, and micro-epistemic discourse relations seem to narrow, to provide semantic structure of just the right level of boulomaic-deontic-epistemic generality, what remains? We propose fundamental socio-political relations governing the establishment, maintenance and negotiation of cooperation among potentially autonomous actors as the natural target domain for a semantics of intonation (cf. Merin 1983, 1985). Humanity had to negotiate as it came into existence, and individual people have to start negotiating soon enough after taking to the world.

Cooperating actors in a minimal, i.e. bi-person social situation – call them [E]go and [A]lter – have to establish a common ground [CG] of joint deontic-boulomaic or epistemic commitments. (For linguistic application our default identification is for Ego with present Speaker, and for Alter with Addressee.) To the extent of being autonomous, E and A are in need of persuasion. The paradigmatic question is always:

‘Why (<expletive>) should I (do/believe that)?’.

The need for persuasion, the original *principium reddendae rationis*, implies that E's and A's preferences are formally inverse regarding points at issue. Indifference or consonance means, by definition, there is no issue.

In the most basic model, predicated on the sociopolitical equivalent of the physicist's extensionless mass-point, the *homo oeconomicus*, negotiations are in essence bargaining games (Nash 1953), i.e. social situations in which interests are neither wholly opposed nor wholly consonant. (Example: How do we divide a cake on a shelf it takes both of us to reach?)

Negotiations on what becomes CG proceed by Elementary Social Acts (ESAs) of Claim, Concession, Denial, Retraction (of a Claim).² These are transitions to (and from) negotiation states. For a given proposition, the state set is that of a finite state automaton (transition system) modelling conventional sequences of moves (e.g. Ego's Concession or Denial is preceded by Alter's Claim). Act-descriptions thus map to constant partial functions on the state set. Negotiation states c are characterized by n -tuples (vectors) consisting of a proposition (Θ , or $\neg\Theta$) and four binary decision-theoretic parameters of negotiation states (Merin 1994).

Parameters are functions of states c allocating ostensible actor-role [S]; preference [P] (deemed inverse w.r.t. propositions under negotiation); dominance [D] with respect to balance of incentives/warrant; and initiator-role [I] among E and A. Thus a state vector is of general form $\langle S(c), \Theta, P(c), D(c), I(c) \rangle$. Example: Ego's Claim for Θ to become mutually binding constraint is explicated as $\langle E, \Theta, E, E, E \rangle$. Here $S(c) = P(c) = D(c) = I(c) = E$. Compare Ego's Concession := $\langle E, \Theta, A, A, A \rangle$ of Θ and check for naturalness; similarly for Denial and Retraction.

A fixed equational parameter constraint, $P(c)=I(c)$, determines the above act-typology: conventionally short-sighted *homini oeconomici* do not undertake counterpreferential initiatives. However, other act-typologies (including Entreaty, Offer,...) are obtainable by relaxing or even inverting the constraint.

The relation to CG is: once a claim is (vocally or tacitly) conceded, the propositional object becomes, as legal parlance originally had it, 'common ground'. Various conventions for retraction (upon denial) are conceivable (default against the claimant, else mutual non-commitment.) In Merin (1994, 1996) the explanatory advantages of hypothesizing such an underlying 'Discourse Institution' (DI) of ESAs are explored for non-intonational semantics and pragmatics.

²Note the deontic/epistemic generality of the latinate English terms.

In bargaining situations proper, of which basic DIs are highly simplified representations, extensions of the simple D and P parameters are determined in terms of cardinal (dis-)utilities. This seems plausible for typical imperatives, e.g.

‘Give me your wallet! (Else be a casualty)’,

but specious for indicatives. Beliefs are, if anything, dispositions to action, not actions.

In the present model, therefore, a measure of epistemic state change potential (Merin 1996), viz. stochastic evidential *relevance*, explicated by the Peirce/Turing/Good log-likelihood-ratio or the Johnson/Keynes coefficient, instantiates the role of a utility. The proposition is a more or less useful *argument* for or against an ulterior constraint on belief or action, i.e. a proposition at issue.

2.2 Intonational Morphemes and their Semantics

Intonational (i.e. suprasegmental) morphemes are kinetic tones: Falls, Rises and some of their Compounds. They denote, in the first and core instance, (re-)allocations of the [D]-parameter value – i.e. of the power of choice – either regarding propositions expressed by a whole (segmental) sentence or clause or regarding focus-identified subsentential items (usually syntactic constituents) that co-determine propositions.

The cultural association of High mean pitch with low relative social power (of children vs. adults, or – induced, as evidenced by widespread stereotypic socialization accentuating differences of laryngeal physiology – between women and men) provides a ready substrate for association of tone levels: High codes dominance by Alter, Low by Ego. Supposing isomorphism, transitions to levels should thus code (re-)allocations of power to choose instantiation of a variable under negotiation. Variably defeasible or ‘bleachable’ default associations for acts in DIs introduce preference (‘scale’) and initiative (‘anaphoricity’)-related aspects. (Recall, e.g., that Claim := $\langle E, (\cdot), E, E, E \rangle$ is Ego’s least marked act type.)

2.3 Specifics of Intonation Semantics

Recovering familiar ‘register’ universals, Rise R = [L*H@] alienates choice to Alter, Fall F = [H*L@] appropriates it. Final X- codes target level (Ego-centric mnemonics : H- ‘exo’, L- ‘endo’) and thus dominates in the mapping from pairs of static phonemes to kinetic morphemes (X*Y-). Rightmost dominance (Sweet 1878) also holds for X% and for compound morphemes. Compounding may occur whenever a tonal movement not associated with an X* follows one that is, with no prosodic phrase boundary intervening. The most distinctive compound morpheme is Fall-Rise

(FR) = [H* L-H%] We interpret this as an embedding of allocation, analogous to ‘free indirect speech’: a choice (F) is *imputed* by Ego (via R) to Alter (Merin 1985).

3. QUESTIONS

The decision-theoretic model offers a plausible account of the data mentioned above for which the discourse-epistemic one fails to predict.

Variability of final pitch movement in YNQs and WHQs reflects the fact that Ego may foreground either of two choice-related aspects inherent to questions:

- in asking, Ego is alienating choice among alternatives (sets of possible worlds) to Alter, i.e. making a Concession;
- in requiring an answer from Alter, Ego is forcing Alter to commit himself to one mutually binding alternative (thus banning others from inclusion in the CG) and is thereby making a Claim.

Ostensibly concessive allocation of choice to Alter and ostensible demand for commitment are conveyed through final rise (e.g. (3a,4a)) vs. final fall (e.g. (3b,4b)), respectively.

The former goes with Ego preferring the inverse of the posed proposition, the latter with Ego being neutral. The greater excursion of a low rise relative to a high rise³ reflects a greater perceived discrepancy between Ego’s epistemic state or preference and Alter’s choice as anticipated by Ego:

(7) I can’t believe you went ahead and got
yourself a job at the embassy in Lisbon.

guese?
Do you speak u
 PORT
[Do you speak Portuguese(L* H-H%)?]

WHQs tend relatively more strongly than YNQs towards showing a final fall, since their existential presuppositions constrain the addressee’s options even as choice of instantiation for the variable represented by the

³Rises are measured beginning with the target level of X* (pre-accent movements do not count), in keeping with the general assumption (cp. Gussenhoven 1984) that the left edge of unary as well as compound tonal morphemes is always associated with an X*.

wh-expression is allocated to him (Bolinger 1982; Merin 1983). Evidence for our account are so-called ‘reference questions’, e.g.

(8) A: I just talked to him last night.

WHO

B: You talked to last
 night?
[You talked to who(H*) last night(L-L%)]

which have the express purpose of producing previously evaded commitment from Alter to a specific extension of a designating expression (Bartels 1997a,b). These are obligatorily falling, as predicted.

Contrast now echo questions. Whatever their syntax, they typically end on either plain R (5a,b above) or FR (9a,b):

(9)a. A: Did Amy get the summer job at the embassy?

EM

sy?

B: Did she get the job at the bas
[Did she get the job at the embassy(H* L-H%)?]

b. A: Amy started her job at the embassy last week.

WHERE

week?

B: She started her job last
[She started her job where(H*) last week(L-H%)?]

Explanation: Ego cannot felicitously demand (claim) of Alter commitment to a proposition to which Alter has already made a commitment by his original utterance. Rather, Ego ostensibly offers Alter another choice to accept the posed sentence as an accurate echo of his original utterance or not (cf. Bartels 1997a,b). If FR is used rather than a plain R, Ego ostends (read as short for: ‘ostensibly intends to show’) that it is Alter (or at least a ‘virtual’ Alter) who has already specified the extension of the accented (H*) expression; the result is an added connotation of anaphoricity (Merin 1985).

In AQs such as (2), (high or low) rises on nonfinal disjuncts ostensibly concede to the addressee (Alter) the choice of whether the respective

proposition is to become a mutually binding constraint, i.e. part of the CG. However, the last disjunct ('Flemish') represents a proposition which *must* be added to the CG if none of the preceding ones have been. It is a demand (Claim) for the addressee to commit himself; hence the obligatory final fall.

As before, if a nonfinal disjunct shows an FR rather than a plain R, evoking anaphoricity, the speaker ostends that it is ('virtual') Alter who has specified the extension for the expression bearing the fall. On the other hand, if a disjunctive question shows a plain F on nonfinal accented constituents,

(10)	Did the subject speak	FRE		FLE	
			or		
		NCH		mish?	

it is best seen as a sequence of independent demands for commitment. And indeed, intonation does not now suggest mutual exclusivity of alternatives.

4. FALL-RISE AND INSINUATION ('CONTRADICTION')

These are some of the most widely discussed 'tunes' of English and one claim to be defended is that both are FRs.

4.1 Fall-Rise and 'Second-Best'

In Ex. (6) the answer

'My husband(H*) does(L-H%)'

in response to 'Do you speak Portuguese?' will generate an implicature

'Speaker (Evita) does not speak Portuguese'

(Hirschberg 1985). We believe the correct explanation (cf. Merin 1996) is that Evita's speaking Portuguese would be positively relevant to her being hired; hence, her not speaking it, negatively. She avoids, as it were, incriminating herself, by offering instead a plainly less positively relevant but still positive alternative that she *can* muster. Husbands, relatives, friends can be called upon if necessary.⁴ The implicature is then what is imputed to All-Boss, the sceptical Alter, as a counterclaim which Evita (Ego) concedes by default. This much can be had without reference to intonation.

What Evita's FR intonation ostends is Evita imputing All-Boss with prior choice of 'husband', of whom Evita can smartly assert facility in

⁴Unlike Kim Basinger or the President of Brazil: hence 'Kim (Bloggs) does' is acceptable, and 'Kim Basinger does' is not.

- (14)a. ALL <'all-not'>
 the men didn't
 go.
 [All(H*) the men didn't go(L-L%).]

and Fall-Rise ('B accent')

- (14)b. ALL <'not-all'>
 go.
 the men didn't
 [All(H*) the men didn't go(L-H%).]

The decision-theoretic account engages several *cooperating* features of FR which lend the example type its semantically distinctive features. First, FR induces 'scalar' intuitions of meaning (Ward & Hirschberg 1985). Why should this be so? A priori a plain F is no less able to bear on an item selected from a set ordered by entailment, situational relevance, utility, or the like. However, a non-extreme item (recall: 'not the best'⁵) draws, like any privation, attention to the alternative. (The cognitive scientist should not ignore commonsense or folk psychology as a reservoir of *topoi*.)

On the 'Aristotelian' square of oppositions, $\neg\forall$ yields the contradictory, and $\forall\neg$ the contrary of \forall . In terms of relevance to an ulterior issue $\{H, \neg H\}$, $\forall\neg$ will in general be more strongly positive than $\neg\forall$ if $\neg\forall$ is positive. If \forall is positive, $\neg\forall$ must be negative; $\forall\neg$ need not always be negative, but usually *is* more strongly negative than $\neg\forall$. Typically, then $\neg\forall$ is a less-than-maximally-strong counter. It is a denial of a positive rather than an assertion (claim) of the negative. (The import of this statement about quantified sentences will be explored elsewhere.) Denial, we have seen, is an ESA *anaphoric* to a prior claim. Focus of FR in (14b) is 'all', picked out as the choice of (formally agonistic) Alter.

Now the second, Rise component of FR comes into play once more. (Over and above marking the item 'all' in its focus as chosen by Alter.) Its basic meaning, we noted, is always that of conceding – something or other of situational importance. In the present case, as in the ones above, it is the scalar implicature that represents the concession? In (14b) the implicature might be (with potential vocalization added) that

Some men did(H*) come(L-L%).

⁵Here is also an explanation why 'any' meaning 'just any old' goes with, even needs, an FR: the choice is not a friendly soul's.

So everything militates for negation to be attracted, in logical reconstruction of what goes on, to ‘all’, and having wide scope. And it can give in to attraction as long as no separate focus (by hypothesis to be explored elsewhere: no non-!H*-accent) intervenes between the negation morpheme and the FR-accented constituent.

4.3 Generic vs. Particular Readings of Indefinites

Phonemic L* (hence an L*H-, a [R]ise) on indefinite sentence topics as in

(15) a. STRA
A ger is allergic to
TI W.
 [A tiger(L* H-) is allergic to straw(H* L-L%).]

does not induce the interactive (‘not new’) interpretation postulated by PH, but rather forces a universal/generic reading of ‘A tiger is allergic to straw’. (Similarly for FR.) By contrast, an H* (hence an H*L-, a [F]all)

(15) b. TI STRA
A
ger is allergic to W.
 [A tiger(H* L-) is allergic to straw(H* L-L%).]

favours an existential/particular one, though a generic reading *is* available too.

For PH one might now argue that L* does spell genericity *qua* universal idea everyone has a grasp of, while H* indicates a specific beast not everyone knows about. But this would be specious: even if there is a well-established tiger in common ‘referential’ ground, the preferred interpretation of R is still the universal one.

The decision-theoretic explanation is in line with an interactive semantics for quantifiers proposed by Peirce and developed by Beth, Lorenzen and, best known though a-socially, Hintikka (1968). If Ego asserts a universal sentence ($\forall x[Qx]$), agonistic Alter has free choice of any singular instance (i.e. a named individual). Suppose Alter picks a . Then Ego has to prove Qa . Failure amounts to a win for Alter: $\forall x[Qx]$ is falsified. Dually, Ego’s assertion of an existential $\exists x[Qx]$ allows Ego free choice of instantiating instance. Say he chooses e . Then $\exists x[Qx]$ is verified if (though not only if) Qe is.

Generic readings of indefinite noun phrases are, for most inferential purposes, tantamount to universal ones. Similarly for singular, specific

context even from a kazoo performance (cf. Johnson-Laird 1988). However, it is no less (and no more) compositional as a tune than any other instance of FR.

5. REFERENCES

- Bartels, C. (1997a). *Towards a compositional interpretation of English statement and question intonation*. Doctoral dissertation, University of Massachusetts, Amherst.
- Bartels, C. (1997b). *The pragmatics of wh-question intonation in English*. Paper presented at the 21st Penn Linguistics Colloquium, Philadelphia, February 1997.
- Beckman, M., and G. Ayers (1994). *Guidelines for ToBI Labeling*. Ms. Ohio State University, Columbus.
- Beckman, M., and J. Pierrehumbert (1986). Intonational structure in Japanese and English. *Phonology Yearbook* 3, 15-70.
- Bolinger, D. (1982). Intonation and its parts. *Language* 58(3), 505-533.
- Cutler, A. (1977). The context-dependence of 'intonational meanings'. In *CLS* 13, 104-115.
- Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
- Hintikka, K.J.J. (1968). Language games for quantifiers. In N. Rescher (ed.) *Studies in logical theory*. Oxford: Blackwell.
- Hirschberg, J. (1985). *A theory of scalar implicature*. Doctoral dissertation, University of Pennsylvania. Publ. New York: Garland 1991.
- Hobbs, J. (1990). The Pierrehumbert-Hirschberg theory of intonational meaning made simple: Comments on Pierrehumbert and Hirschberg. In P.R. Cohen et al. (eds.), *Intentions in communication*. Cambridge MA: MIT Press.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge MA: MIT Press.
- Johnson-Laird, P.N. (1988). *The computer and the mind*. Cambridge MA: Harvard University Press.
- Ladd, D.R. (1980). *The Structure of intonational meaning*. Indiana University Press, Bloomington.
- Liberman, M. & Sag, I. (1974). Prosodic form and discourse function. *CLS* 10, 416-427.
- Merin, A. (1983). Where it's at (is what English intonation is about), *CLS* 19, 287-298.

- Merin, A. (1985). *Elementary social relations, duality principles, and modal paradigmata: sociomorph structures of English grammar*. Doctoral Dissertation, University of Sussex.
- Merin, A. (1992). Permission sentences stand in the way of boolean and weaker lattice-theoretic semantics. *Journal of Semantics*, 95-162.
- Merin, A. (1994). Algebra of elementary social acts. In S.L. Tsohatzidis (ed.), *Foundations of speech act theory*. London: Routledge.
- Merin, A. (1996). Information, relevance and social decisionmaking. In L. Moss (ed.), *Proc. ITALLC-96*, London. [expd. to appear CSLI]
- Nash, J.F. (1953). Two-person cooperative games. *Econometrica* 21, 140-152.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Doctoral Dissertation, MIT.
- Pierrehumbert, J., and J. Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen et al. (eds.), *Intentions in Communication* Cambridge MA: MIT Press.
- Sweet, H. (1878). *Handbook of phonetics*. Oxford: Clarendon Press.
- Ward, G. & Hirschberg, J. (1985). Implicating uncertainty: the pragmatics of fall-rise intonation. *Language* 61(4), 747-776.

[This is a text-identical reformatted edition of a paper that has been circulating in double-column format as of February 1979.]

ARTHUR MERIN, INSTITUT FÜR MASCHINELLE SPRACHVERARBEITUNG (IMS), UNIVERSITÄT STUTTGART, AZENBERGSTR. 12, 70174 STUTTGART, GERMANY
E-mail address: `arthur@ims.uni-stuttgart.de`

CHRISTINE BARTELS, DEPARTMENT OF LINGUISTICS, UNIVERSITY OF OREGON, EUGENE, OR 97403, USA
E-mail address: `bartels@darkwing.uoregon.edu`