Interpreting English Pitch Contours in Context

Interpreting English Pitch Contours in Context

Julian J. Schlöder
Institute for Logic, Language & Computation
University of Amsterdam
julian.schloeder@gmail.com

Alex Lascarides
School of Informatics
University of Edinburgh
alex@inf.ed.ac.uk

Abstract

This paper presents a model of how pitch contours influence the illocutionary and perlocutionary effects of utterances in conversation. Our account is grounded in several insights from the prior literature. Our distinctive contribution is to replace earlier informal claims about the implicatures arising from intonation with logical derivations: we validate inferences in the SDRT framework that resolve the partial meaning we associate with a pitch contour to different specific interpretations in different contexts.

1 Introduction

In this paper, we give a formal semantics of pitch contour in spoken dialogue, implemented in SDRT (Asher and Lascarides, 2003). Our main claim is that the pitch contour of an utterance conveys cognitive attitudes, and that taking these attitudes into account perturbs calculable implicatures. The following examples, adapted from Steedman (2000), are cases in point:[1]

(1) A: You're a millionaire.
  a. B: I'm a MILLIONAIRE. (H* LL%)
  b. B: I'm a MILLIONAIRE? (H* LH%)
  c. B: I'm a MILLIONAIRE? (L* LH%)

(2) A: Are you rich?
  a. B: I'm a MILLIONAIRE? (H* LH%)

The utterance in (1a) is an assertion with the 'high focus, final fall' contour (H* LL%). Conventionally, this commits B to the proposition *'B is a millionaire'* and thereby establishes agreement between A and B. The same utterance with a final rise (LH%) in (1b) is a question (*'Am I?'*) that does not make any such commitment. The low pitch accent (L*) in (1c) additionally reveals that B is somewhat surprised or doubtful about A's assertion (*'Am I? Really?'*). While (2a) has exactly the same form as (1b), in the context of (2) B *does* make a commitment to being a millionaire, but displays uncertainty on whether this answers A's question. These examples show that the intonation of an utterance can influence both illocutionary and perlocutionary inferences: in (1b) an indicative mood utterance is a clarification request, and (1c) expresses the failure of belief transfer after an assertion. They also show that such inferences are highly context-sensitive.

Our formal account makes specific, computable predictions on what attitudes are displayed by pitch, and when a particular pitch contour is licensed. It achieves this by leaving the compositional semantics of pitch deliberately underspecified, with contextual information and inference supporting a specific and complete interpretation in context. We believe that our model is novel in its formal precision, with previous work resorting to semi-formal paraphrases of how intonation gives rise to implicatures.

Steedman (2014) formalises pitch contours in terms of their effect on common ground, and claims that the effects outlined above are derivable from general principles of truth maintenance. However, he does not give a formal account of these derivations. Our model proposes a compositional semantics for individual pitch accents in terms of public commitment; our semantic postulates are inspired by Steedman's, but we *formally derive* their specific contribution in context. However, we abstract away from grammatical parsing and assume that a grammar is in place which connects to our semantics of pitch. This means that we do not take the lexical placing of the focus accent into account and assume that the foreground proposition of an utterance is computed elsewhere.

In the next section, we expand our informal discussion to further examples. We give a brief introduction to the formal framework of SDRT in section 3, including some amendments to SDRT's cognitive logic. We present our formal theory in

[1] To describe our examples, we use the ToBI annotation scheme throughout (Silverman et al., 1992).

section 4 and show that it corresponds to the analyses surveyed in section 2. In section 5 we conclude and give pointers towards further work.

2 Informal Discussion

The meaning of pitch contours in English has received substantial attention in the literature. We briefly review some prevalent discussions and use them to motivate our formal model. The data we present is not new: it is derived from a number of earlier analyses that comprehensively survey the phenomena we are interested in. Where we constructed an example for the sake of exposition, we verified that our reading is in accordance with earlier accounts.

2.1 Final Rise

We follow the discussion in Schlöder (2015). A typical interpretation of the final rise in English is that it signals 'insufficiency' in various senses. Hobbs (1990) and Bolinger (1982) characterised the meaning of a final rise as signalling 'incompleteness.' Šafářová (2005) has characterised this incompleteness as displaying, *i.a.*, an 'uncertain attitude.' Westera (2013) has further specified such uncertainty by relating it to the Gricean maxims: the speaker displays an uncertainty regarding the truthfulness (Quality), specificity (Quantity) or appropriateness (Relation) of their utterance. Given these observations, we characterise the final rise as marking an utterance as incomplete, but consider incompleteness itself to be an underspecified notion, *i.e.*, incompleteness can be resolved in different ways. The following two possible continuations, adapted from Hirschberg and Ward (1995), exemplify this variation in resolution:

(3) a. A: Where are you from?
b. B: I'm from SKOKIE. (H* LH%)
c. B: That's in Illinois.
c.' A: Okay, good.

The final rise in (3b) is interpreted to signal incompleteness. Specifically, B is uncertain if the answer in (3b) pragmatically resolves A's question. Both follow-ups in (3c) and (3c') resolve this incompleteness in different ways: in (3bc), B is supplying additional information himself, signalling that *'Skokie'* is not the full answer, but that *'Skokie, Illinois'* is. In (3bc'), B is waiting for A to comment on whether she considers (3b) to be sufficient. Thus in (3bc'), B's utterance is taken to have question force (in some sense), but cannot be glossed as *'am I? [from Skokie];'* rather, it is *'I'm from Skokie—does that answer your question?'* So here, the incompleteness is resolved in the dialogue as posing a question that needs be answered by A. Similarly, the final rise in (1b) is interpreted as a clarification request. However, its illocutionary force *can* be paraphrased as *'Am I? [a millionaire].'* So a proposition p presented in indicative mood with final rise can have the illocutionary effect of the polar question $?p$ (*'Is it the case that p?'*) in some but not all contexts.

Hence, our basic idea to describe incompleteness goes as follows: the final rise always *projects* a follow-up, *i.e.*, it demands some kind of resolution, but we leave open what kind of speech act is being projected and what precisely its contribution to the discourse is (*cf.* Malamud and Stephenson (2011)). We then infer the specific illocutionary force of the final rise utterance and of its follow-up via a logic of defeasible inference that draws on the compositional and lexical semantics of the utterances involved. Where appropriate, we default to clarification questions, *i.e.*, our logical axioms validate a defeasible inference that the preferred follow-up to (1b) is a confirmation. But we leave room to resolve the projection in another way in different contexts, *e.g.*, by elaboration in (3bc) or acceptance in (3bc').

2.2 Pitch Accents

We are primarily interested in the meaning of a low pitch accent (L*). A typical interpretation of the low pitch in English is that it indicates that belief transfer has failed in some way: Hobbs (1990) takes L* to indicate that the foreground proposition is either known or false, and Steedman (2000; 2014) takes it to mean that grounding fails. We are also interested in the low pitch with high border accents (H+L* and L*+H). Consider the following possible responses in (4):

(4) A: France has a king!
a. B: France is a MONARCHY. (L* LL%)
'It is not, this is obvious.'
b. B: France is a MONARCHY. (H+L* LL%)
'It is not, this is obvious.'
c. B: # France is a MONARCHY. (L*+H LH%)

The utterances in (4a) and (4b) are intonated to openly display sarcasm, and are hence disagreeing with A's assertion. However, using the so-

called contradiction contour (*cf.* Liberman and Sag (1974)) as in (4c) is incoherent. In dialogues where—abstracting away from pitch—the semantics of B's response contradicts A's, the felicity and meaning change:[2]

(5) A: France has a king!
  a. B: France is a REPUBLIC. (L* LL%)
    'It is obviously a republic.'
  b. B: # France is a REPUBLIC. (H+L* LH%)
  c. B: France is a REPUBLIC. (L*+H LH%)
    'You are wrong, it is a republic.'
  d. B: # France is a REPUBLIC. (H* LH%)
  e. B: # France is a REPUBLIC. (H* LL%)

Unsurprisingly, the contradiction contour in (5c) is licensed now. The L* pitch that earlier gave rise to a sarcastic reversal of meaning now yields a rather condescending correction in (5a), but (5b) is incoherent. Perhaps unexpectedly, a high pitch is also incoherent here. Clearly, the question in (5d) cannot be a clarification request, but the case for the incoherence of (5e) is more opaque: speaker A asserts something, and B asserts the opposite. *Prima facie* this is alright, but with these pitch contours it seems as if A and B are talking past each other.

We conclude that while the low pitch generally indicates a problem with belief transfer, its meaning depends on its surrounding border accents: the simple low pitch (L*) signals disagreement irrespective of the utterance's actual propositional meaning, whereas the complex pitches (H+L* and L*+H) change meaning (and felicity).

Hobbs (1990) defines the effect of H* as marking the foreground proposition as 'new.' Since 'not new' to him means grounded or false, 'new' should denote ungrounded *and* true. Since, in dialogue, truth *as far as the participants are concerned* takes precedence, we paraphrase Hobbs' 'new' instead as 'uncontroversial.' Steedman's (2014) gloss of the H* LL% contour is *'I succeed in making common ground that p.'* We are reluctant to accept that a speaker can make a public commitment for *another* speaker (*pace* Gunlogson (2003)). Hence, one cannot *individually* succeed in making something common ground, and we can only consider Steedman's gloss to mean that the speaker *assumes* that belief transfer will succeed, *i.e.*, that he has no reason to believe otherwise. We again can paraphrase this as 'p is uncontroversial.' This link from H* to 'uncontroversial' explains the incoherence of (5d) and (5e): after A commits to France having a king, the proposition *'France is a republic'* cannot be considered uncontroversial.

Example (6), taken from Ladd (1980), shows that intonation has more subtle effects when we are interested in more nuanced differences than assent *vs.* contradiction.

(6) A: Harry's the biggest liar in town.
  a. B: The biggest FOOL maybe. (H* LL%)
  b. B: The biggest FOOL maybe. (L*+H LH%)

The contour (6a) has B agreeing with A's assertion by *elaborating* it, *i.e.*, B is publicly committed to Harry being both the biggest liar *and* maybe the biggest fool. The contour in (6b) does not commit B to Harry being the biggest liar, but it is not an outright denial either. Instead, B is committed to Harry being the biggest fool, and that this makes *'Harry is the biggest liar'* less believable. Hence, the utterance in (6b) has a different illocutionary force than (6a), and attaches to the discourse as *counterevidence*.

While the H* LH% contour in (7a) again displays uncertainty, the examples in (7b,c) reveal something new about the low pitch: B conveys something about how his beliefs have changed *in the aftermath* of A making her previous utterance:

(7) A: Did you read the first chapter?
  a. B: I read the THIRD chapter? (H* LH%)
    'Does that suffice?'
  b. B: I read the THIRD chapter? (L*+H LH%)
    'Wasn't I supposed to read the third?'
  c. B: I read the whole DISSERTATION. (L*+H LL%)
    'And you should know that I did.'

In (7b), B believed it to be common ground that he is to read the third chapter, and in (7c) he believed that A knew that he had read the whole dissertation. To make sense of these perlocutions, a listener needs to draw inferences of the form *'before A said u, B must have believed p.'* Such reasoning is also useful to describe the meaning of contours signalling surprise. In the next section, we describe how we formalise such *hindsights* in SDRT's cognitive modelling logic.

[2]The glosses/interpretations of the utterances in (4) and (5) are our own. We additionally verified that they are consistent with Steedman's (2014) explanations of the grounding-related effects of pitch contours.

3 Framework

Our theory of pitch contours is implemented in Segmented Discourse Representation Theory (Asher and Lascarides, 2003; Lascarides and Asher, 2009). Our rationale goes as follows. SDRT models back-and-forth information flow between three interconnecting languages and associated logics: the *language of information content*, the *glue logic*, and the *cognitive modelling logic*. By manipulating this information flow, we gain fine-grained control over different aspects of an utterance's interpretation, allowing us to model perturbations of the standard interpretations. Each of the logics in SDRT is designed for a specific task, and we briefly describe each of them in turn.

The *language of information content* is used to express the logical form of a discourse, capturing its pragmatically resolved, specific interpretation. The dynamic semantics of this language models the truth conditions of the public commitments that speakers make through their utterances. The language of information content includes *rhetorical relations* (*e.g.*, $Explanation$ or $Elaboration$) that connect the representations of individual discourse units.

A logical form in SDRT is an SDRS: a set of labels Π, where each label stands for a discourse segment, and a mapping $\mathcal{F}$ from each label in Π to a formula representing that segment's content (we will sometimes write $\mathcal{F}(\pi)$ as K_π). Since these formulae include rhetorical relations among labels, $\mathcal{F}$ imposes an ordering on Π: π_1 *immediately outscopes* π_2 if $\mathcal{F}(\pi_2)$ features $R(\pi_1, \pi)$ or $R(\pi, \pi_2)$ as one of its conjuncts. We write $\pi \succeq \pi'$ for the transitive closure of this relation. A well-formed SDRS imposes the constraint that this partial order is rooted, *i.e.*, there is a single discourse segment consisting of rhetorically connected subsegments.

In dialogue, participants make *public commitments* to SDRSs. Specifically, the logical form of a dialogue turn is a set of SDRSs, one for each dialogue participant. When a speaker utters a unit π, he commits to the *rhetorical relation* that connects π to the prior context. In effect, this makes speakers publicly committed to the illocutionary contribution of their moves. The logical form of a dialogue is the logical forms of its turns. For example, the logical form of (8) is as follows:

(8) A: Max fell.
B: John pushed him.

Turn	A's SDRS	B's SDRS
1	$\pi_1 : \mathit{fall}(e, m)$	$\emptyset$
2	$\pi_1 : \mathit{fall}(e, m)$	$\pi : \mathit{Explanation}(\pi_1, \pi_2)$ $\pi_2 : \mathit{push}(e', j, m)$

The dynamic semantics of $Explanation(\pi_1, \pi_2)$ entails the contents of π_1 and π_2 in dynamic conjunction, and that the latter answers the question *'why is π_1 true?'* Being publicly committed to $Explanation(\pi_1, \pi_2)$ thus makes B publicly committed to the content of π_1. This means that A and B agree that Max fell—they share a public commitment to it—even though this is an implicature of B's contribution and not linguistically explicit.

There are several parts of the logical form of (8) that go beyond the compositional and lexical semantics of its individual units: the pronoun *'him'* is resolved to m, and the illocutionary contribution of B's utterance is to provide an *Explanation* to A's. These inferences are about the *construction* of logical form (as opposed to their *truth*). As input, they take *underspecified logical forms* (ULFs), which are in turn computed from an utterance's linguistic surface form. This construction is modelled in the *glue logic* which we discuss next.

The glue logic validates defeasible inferences from partial descriptions of logical forms (*i.e.*, ULFs) to fully specified discourses (*i.e.*, SDRSs like those in 8). SDRSs capture the pragmatically preferred and complete interpretation of the discourse. These inferences are facilitated by axioms of the following form:

$$(\lambda :?(\alpha, \beta) \wedge \mathit{Info}(\alpha, \beta)) > \lambda : R(\alpha, \beta).$$

The $>$ denotes a default conditional and we use Greek letters to label discourse segments. So, informally, the above formula says: 'if α and β are rhetorically connected to form a part of the extended discourse segment λ, and their ULFs satisfy *Info*, then normally, their rhetorical connection is R.' Such default axioms are justified by word meaning, world knowledge and cognitive states.

The default conditional $>$ yields a nonmonotonic proof theory $\mid\!\sim_G$. To ensure that the glue logic remains decidable, it reasons about ULFs (*i.e.*, the (partial) *form* of a logical form), but has only limited access to what those logical forms *mean* in the logic of information content. Keeping the glue logic decidable accounts for how people by and large agree on *what* was said, if not on whether it is true.

The glue logic also has access to information in the *cognitive modelling logic*. This logic in-

cludes a number of modal operators: KD45 modal operators for beliefs (B_S for a speaker S); K45 operators for public commitment (P_S); and special modal operators for intentions (I_S).[3] Also, for each action term δ, there are two modal operators $[\delta]$ ('after δ') and $[\delta]^{-1}$ ('before δ'); see Asher and Lascarides (2008) for a discussion of this logic. The only action term we will be concerned with is the act of uttering something, $\delta = s_S(\pi)$ for an utterance label π and its speaker S. For our purposes, these operators are sufficiently specified by postulating the following axioms:

Glue to Cognitive Logic (GL to CL).
Let $\pi_1 \dots \pi_n$ be elementary discourse units spoken by $S_1 \dots S_n$, and Γ_n be the context after π_n (*i.e.*, their ULFs plus facts and axioms). Let $\vdash_G$, $\mid\!\sim_G$ be the monotonic and nonmonotonic proof theories of the glue logic. Let $\vdash_C$ and $\mid\!\sim_C$ be the ones for the cognitive modelling logic.
If $\Gamma_n \vdash_G \varphi$, then
$$\Gamma_n \vdash_C [s_{S_1}(\pi_1)] \dots [s_{S_n}(\pi_n)] P_{S_n} \varphi.$$
If $\Gamma_n \mid\!\sim_G \varphi$, then
$$\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \dots [s_{S_n}(\pi_n)] P_{S_n} \varphi.$$

Persistence.
If $\Gamma \mid\!\sim_C P_A \varphi$ and $A \neq S$, then $\Gamma \mid\!\sim_C [s_S(\pi)] P_A \varphi$.
A person's public commitments are unaffected by another speaker's utterance.

Hindsight.
If $\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \dots [s_{S_n}(\pi_n)][s_{S_i}(\pi_i)]^{-1} B_S \varphi$, then $\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \dots [s_{S_{i-1}}(\pi_{i-1})] B_S \varphi$.
'Before'-operators cancel up to a corresponding 'after'-operator.

Conservativity.
$([s_S(\pi)] B_{S'} \varphi) \rightarrow (B_{S'} \varphi \vee B_{S'}((P_S K_\pi) > \varphi))$.
Beliefs after an utterance are either carried over from before, or are inferred from that utterance.

Reduction.
$(B_{S'}[s_S(\pi)]\varphi) > ([s_S(\pi)] B_{S'} \varphi)$, and
$(B_{S'}[s_S(\pi)]^{-1}\varphi) > ([s_S(\pi)]^{-1} B_{S'} \varphi)$.
Beliefs usually transfer to hindsight and foresight judgements, *i.e.*, if a speaker believes that after/before the act π, the proposition φ holds, they have that belief in foresight/hindsight.

The axioms *GL to CL* and *Persistence* together ensure that glue logic inferences about the illocutionary act that a speaker performs matches their (current) public commitments in the cognitive logic; so if A has asserted that p then in the cognitive logic A is publicly committed to p. Conversely, defeasible inferences made in the glue logic can also be *blocked* by *facts* from the cognitive modelling logic, *e.g.*, if $\Gamma \vdash_C P_S \neg\varphi$, then the glue logic cannot defeasibly infer a discourse relation in S's SDRS that would entail φ.

Note that the context Γ_n in *Hindsight* does not change. The axiom models inferences that interlocutors can make about previous cognitive states *from their current knowledge* Γ_n, which *extends* their prior knowledge Γ_{i-1}. In particular, it is possible that the axiom applies in Γ_n, but that $\Gamma_{i-1} \not\mid\!\sim_C [s_{S_1}(\pi)] \dots [s_{S_{i-1}}(\pi_{i-1})] B_S \varphi$. Also note that the hindsight-inferences formalised by the *Hindsight* and *Reduction* axioms are scoped by a belief modality. Since defaults support belief revision (*i.e.*, it is possible that $\Gamma \mid\!\sim B_S \varphi$ while $\Gamma \wedge \psi \mid\!\sim B_S \neg\varphi$), the above axioms support revision *in hindsight*. We go more in-depth on these phenomena in the next section.

4 Formal Model of Pitch Contours

We now give a precise, formal account of the effects we discussed in section 2. We first give a brief account of cooperative principles in SDRT and how they are used to compute the perlocutionary effects of utterances. This initial presentation will discuss the standard (unperturbed) inferences. We then present our semantics for pitch contours, and afterwards show how we derive their pragmatic effects.

4.1 The Standard Reasoning

Our main concern are the *perlocutionary* contributions of pitch contours, which we model in SDRT's cognitive modelling logic. In SDRT, such effects (like belief transfer) are specified by stipulating axioms affecting the cognitive models of the speakers (Asher and Lascarides, 2003; Asher and Lascarides, 2013). The following axioms give a Gricean account of cooperativity:

Sincerity (a). $P_S \varphi > B_S \varphi$.
Sincerity (b). $B_S \neg\varphi > \neg I_S P_S \varphi$.
Cooperativity. $P_S I_S \varphi > I_H \varphi$.
Intention Transfer. $P_S \varphi > P_S I_S P_H \varphi$.

Sincerity states that public commitments are usually truthful regarding the interlocutor's beliefs, *Cooperativity* that publicly announced intentions are usually adopted by their addressee, and *Intention Transfer* that a public commitment is usually intended to be grounded, *i.e.*, to become a shared public commitment. In SDRT, both interlocutors

[3] Glossing over the details, we write $I_A B_B \varphi$ if A wants B to believe that φ, and $I_A P_B \varphi$ if A wants B to commit to φ.

maintain their own private model of the cognitive modelling logic, *i.e.*, their individual representation of the public commitments, beliefs and intentions of everybody involved in the conversation. We assume that everyone agrees on the above axioms, and that this fact is mutually known.

As an example, suppose that a speaker S asserts p to a hearer H. By *GL to CL*, S and H infer $P_S p$ in the cognitive model. Then, H can infer that S actually believes that p by *Sincerity*. Further, both can infer by *Intention Transfer* that S wants H to make the same commitment, *i.e.*, $P_S I_S P_H p$. By *Cooperativity* the speaker S can infer that $I_H P_H p$ and so expects an agreement move (establishing H's commitment to p) next.

4.2 Final Rise

Based on our discussion in section 2, we take the final rise to have an influence on: (i) the *structure* of the dialogue by demanding a follow-up (incompleteness); (ii) the illocutionary force of an utterance (*e.g.*, an inferred question force); and (iii) the inferred attitudes of the speaker (uncertainty). We refine the model of Schlöder (2015). The following mapping formalises incompleteness:[4]

Semantics of the Final Rise.

$$\pi(LH\%) \mapsto \pi' = ? \wedge \pi'' = ? \wedge R = ? \\ \wedge R(\pi', \pi'') \wedge \pi' \succeq \pi.$$

That is, the final rise semantics enforces that there is a yet unknown follow-up response standing in some relation to the final rise unit π. We leave open what discourse relation is projected, and we allow it to attach to a wider discourse segment as long as it includes π as a part. This is required to model cases where it is the *discourse relation itself* that is uncertain. For example, in (3bc'), where A accepts the *Question-Answer-Pair* (QAP) relation itself (*i.e.*, that 3b answers her question 3a), and not just the contents of (3b). That is, A's move is *Accept*(π, c), where π : *QAP*(a, b). In (3bc), however, the projected relation is $Elaboration(b, c)$, directly attaching to the final rise utterance (3b).

In addition, we stipulate a glue logic axiom that, where truth-conditionally appropriate, defeasibly infers from a final rise that a question is being asked. The following rule serves to interpret an indicative mood utterance with content p as the polar question $?p$ (as in example 1a):[5]

Clarification from Final Rise.

$$\big(\beta : LH\% \wedge \lambda : ?(\alpha, \beta) \wedge \Box(K_\alpha \rightarrow \mathit{prop}(K_\beta))\big) \\ > \lambda : CR(\alpha, \beta).$$[6]

In this axiom, $\pi : LH\%$ means that the label π includes the final rise semantics. So this axiom stipulates that if an utterance has a final rise, and its core propositional content is entailed by that of its attachment point, then normally, it is a clarifying polar question.[7] The entailment $K_\alpha \rightarrow prop(K_\beta)$ is required to explain the incoherence of (9b):

(9) A: You are rich.
  a. B: I'm rich? *'Am I?'*
  b. B: # I'm a millionaire?

(10) A: You are a millionaire.
  a. B: I'm rich? *'Am I?'*
  b. B: I'm a millionaire? *'Am I?'*

Both answers in (10) are licensed because, conventionally, *'millionaire'* implies *'rich,'* hence the question in (10a) is reasonable. Conversely, *'rich'* does not necessarily imply *'millionaire'*, so B's utterance in (9b) does not support an interpretation as a clarification request. Lastly, we model 'uncertainty' in SDRT's cognitive modelling logic. Here, the functions $S(\pi)$ and $H(\pi)$ map a label to its speaker and hearer, respectively.

Cognitive Contribution of the Final Rise.

$$\pi : LH\% \wedge \lambda : R(\alpha, \pi) \wedge \neg\pi : ?\mathit{prop}(K_\pi) \\ > P_{S(\pi)} \neg B_{S(\pi)} I_{H(\pi)} P_{H(\pi)} R(\alpha, \pi).$$

This stipulates that if the utterance with the final rise directly attaches to an antecedent, but is not a question,[8] then the speaker publicly displays uncertainty about whether the hearer is willing to commit to that relation. This is in particular true if the relation is *Correction* (as in, *e.g.*, 5c), but also applies to uncertain answers as in (3b).

As discussed in section 4.1, the combined application of *Cooperativity* and *Intention Transfer* would normally yield $I_{H(\pi)} P_{H(\pi)} R(\alpha, \pi)$, *i.e.*, the hearer will establish a shared commitment on the discourse relation R in the next turn. The cognitive contribution of the final rise conveys that the speaker S was unable to make that inference—for whatever reason. We take this to be the underspecified uncertainty that a final rise communicates.

[4] *cf.* Pierrehumbert and Hirschberg (1990): 'to interpret an utterance with particular attention to subsequent utterances.'

[5] Note that if an utterance is in interrogative mood, then the axioms of Asher and Lascarides (2003) will already support an inference that the utterance has the force of a question.

[6] CR $\simeq$ Clarification Request. CRs have question semantics, *i.e.*, $\pi : ?K_\pi$, and are *sincere* (not rhetorical): $\neg P_S K_\alpha \wedge \neg P_S \neg K_\alpha$. *CR* has the dynamic semantics of *elaborating questions* (Asher and Lascarides, 2003, p. 468).

[7] It is necessary to map K_β to its propositional content, as once question force is inferred, K_β is a question.

[8] On questions, the final rise is part of the default contour and cannot be taken to convey uncertainty.

4.3 Pitch Accents

We only discuss the cognitive functions of nuclear pitch accents, abstracting away from pre-nuclear pitches and lexical position. We are furthermore only concerned with pitch accents on indicative utterances (including those interpreted as questions), but not with interrogatives. We stipulate the following cognitive contributions (we simplify notation by setting $S = S(\pi)$, $H = H(\pi)$):

Cognitive Contributions of Nuclear Accents.

- $\pi(H^*) \mapsto P_S(\neg B_S B_H \neg K_\pi)$.
  'I don't think what I'm saying is controversial.'
- $\pi(H{+}L^*) \mapsto P_S(\neg I_S P_S K_\pi)$.
  'I'm not committing to what I just said.'
- $\pi(L^*) \mapsto \lambda : ?(\alpha, \pi) \to \Big(P_S\big(\neg I_S P_S K_\alpha\big) \wedge P_S\big([s_H(\alpha)]^{-1} B_S \neg I_H P_H K_\alpha\big)\Big)$.
  'I didn't think you'd want to commit to what you just said, and I'm unwilling to.'
- $\pi(L^*{+}H) \mapsto \lambda : ?(\alpha, \pi) \to \Big(P_S\big(B_S B_H \neg K_\pi\big) \wedge P_S\big([s_H(\alpha)]^{-1} \neg B_S B_H \neg K_\pi\big)\Big)$.
  'I didn't think what I'm saying is controversial, but now I do.'

Note that the postulate for H* states that the speaker S assumes that belief transfer, as formalised by the successive application of *Intention Transfer* and *Cooperativity*, will succeed. To be precise, if S's cognitive model would include $B_S B_H \neg K_\pi$, then S would infer by *Sincerity (b)* that $B_S \neg I_H P_H \varphi$, and would hence believe that *Cooperativity* would not apply, *i.e.*, S would not expect an agreement move next. Intonating H* is therefore the default pitch insofar that S explicitly communicates that she sees no reason why the standard grounding process should *not* obtain, yielding the implicature *'I expect you to agree.'* Such an expectation is unwarranted if the utterance is a correction move; in section 4.4 we show how this explains the incoherence of (5e).

Conversely, the first conjunct of the L*+H contribution has the speaker conveying the opposite: she assumes that her utterance's content *is* controversial. Accordingly, the L*+H contour features prominently in utterances that put two propositions in contrast, *e.g.*, in denials. We give a formalisation of example (6b) in section 4.4.

In the first clause of the L* contribution, however, a speaker is explicitly announcing that the *Cooperativity* axiom has failed on *their* side of the model, and belief transfer on H's earlier statement (labelled α) has failed. The cognitive contribution of the H+L* pitch has the same form, but relates to the current utterance (labelled π): the speaker is indicating that she does not intend that her own utterance's contents be grounded. [9] If the propositional content of π is the same as that of α, the result is a sarcastic rejection (as in 4b). Usually, such a rejection is taken to mean that the speaker actually believes the opposite. Hence we include the following negation-strengthening axiom:

Sarcasm. $P_S \neg I_S P_S \varphi > P_S \neg \varphi$.

This reads as follows: if someone makes the explicit public commitment to *not* make a particular commitment, they are usually taken to commit to the opposite. This accounts for the actual reversal of meaning in a sarcastic utterance, instead of a mere refusal to ground. In the next section, we show how this axiom separates the sarcastic rejection (4a) from the loaded question (1c).

What is left to discuss are the second clauses of the L* and L*+H contributions, respectively. These clauses convey something about earlier beliefs, allowing for hindsight inferences. By uttering something, a speaker incurs a public commitment and the second clause of the L* contribution conveys that the next speaker did not expect this commitment. The second clause of the L*+H contribution relates to an utterance's content being thought uncontroversial, but that belief has now changed—thereby allowing inferences on the speaker's beliefs *before* the utterance in hindsight.

4.4 Applications

We now verify derivations of the effects of pitch contours for four of our earlier examples.

Ex. (5e) A: France has a king!
B: # France is a REPUBLIC.
H* LL%

From A's utterance we can infer:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)]\Box(K_\alpha \to \neg K_\pi)$ (fact).

$\Gamma \vdash [s_A(\alpha)] P_A K_\alpha$ (GL to CL).

$\Gamma \mathrel{|\!\sim} [s_A(\alpha)][s_B(\pi)] P_A K_\alpha$ (Persistence), hence

$\Gamma \mathrel{|\!\sim} [s_A(\alpha)][s_B(\pi)] P_A \neg K_\pi$.

$\Gamma \mathrel{|\!\sim} [s_A(\alpha)][s_B(\pi)] B_A \neg K_\pi$ (Sincerity)

$\Gamma \mathrel{|\!\sim} B_B [s_A(\alpha)][s_B(\pi)] B_A \neg K_\pi$
(axioms are mutually believed).

[9] *cf.* Steedman (2014) *'I fail to make it common ground.'*

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] B_B B_A \neg K_\pi$ (Reduction).

From B's utterance, including its H*, we get:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B \neg B_B B_A \neg K_\pi$

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] B_B \neg B_B B_A \neg K_\pi$ (Sincerity).

$\Gamma \mid\!\sim [s_H(\alpha)][s_B(\pi)] \neg B_B B_A \neg K_\pi$ (B is KD45[10]).

Hence we infer that one of A or B are insincere. Since it is indeterminable to an overhearer *who* is insincere, *i.e.*, which application of Sincerity is blocked, the dialogue appears incoherent.

Ex. (6b) A: Harry's the biggest liar in town.
B: The biggest FOOL maybe. (L*+H LH%)

The intended reading is that B is putting his utterance in contrast to A's utterance. We start with the second conjunct of the L*+H semantics:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B\big([s_A\alpha]^{-1} \neg B_B B_A \neg K_\pi\big)$.

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] B_B\big([s_A\alpha]^{-1} \neg B_B B_A \neg K_\pi\big)$ (Sincerity).

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)][s_A\alpha]^{-1} B_B \neg B_B B_A \neg K_\pi$ (Reduction).

$\Gamma \mid\!\sim B_B \neg B_B B_A \neg K_\pi$ (Hindsight).

$\Gamma \mid\!\sim \neg B_B B_A \neg K_\pi$ (B is KD45)[11] $(*)$.

Now, the first conjunct of the model for L*+H:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B\big(B_B B_A \neg K_\pi\big)$.

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] B_B\big(B_B B_A \neg K_\pi\big)$ (Sincerity).

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] B_B\big(B_A \neg K_\pi\big)$ (B is KD45).

$\Gamma \mid\!\sim [s_A(\alpha)]\big(B_B B_A \neg K_\pi \vee B_B(P_B K_\pi > B_A \neg K_\pi)\big)$ (Conservativity).

$\Gamma \mid\!\sim [s_A(\alpha)] B_B B_A \neg K_\pi$ ($\vee$-elimination).[12]

$\Gamma \mid\!\sim B_B\big(P_A K_\alpha > B_A \neg K_\pi\big)$ (Conservativity + $*$).

$\rightsquigarrow$ *'That you told me he is a liar tells me that you don't think he is a fool.'*

Ex. (4a) A: France has a king!
B: France is a MONARCHY. (L* LL%)

By the first conjunct of the model for L*:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B \neg I_B P_B K_\alpha$.

$\Gamma \mid\!\sim [s_A(\alpha)][s_B(\pi)] P_B \neg K_\alpha$ (Sarcasm).

$\rightsquigarrow$ *'You are wrong.'*

This inference channels back into the glue logic, which now validates the discourse relation $Correction(\alpha, \pi)$, entailing $\neg K_\alpha$, in B's SDRS.

Also, similar to the derivation of $(*)$, applying Sincerity, Hindsight and Reduction to the second conjunct yields:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B [s_A(\alpha)]^{-1} B_B \neg I_A P_A K_\alpha$.

$\Gamma \mid\!\sim B_B \neg I_A P_A K_\alpha$.

$\rightsquigarrow$ *'I thought you wouldn't say that.'*

In sum, B is communicating to A in (4a) that he is correcting A and that he did not expect that he would have to do so.

Ex. (1c) A: You're a millionaire.
B: I'm a MILLIONAIRE? (L* LH%)

Here, B's utterance has question force, but it is read with a bias towards the negative answer. First, the axiom *Clarification from Final Rise* renders B's utterance as a CR. Hence, the cognitive contribution of the final rise does not apply. Now, consider the second conjunct of the L* contribution and apply, as before, Reduction and Hindsight:

$\Gamma \mid\!\sim B_B \neg I_A P_A K_\alpha$

$\rightsquigarrow$ *'I thought you wouldn't say that'* ($\approx$ surprise).

Then, by the first conjunct of the model for L*:

$\Gamma \vdash [s_A(\alpha)][s_B(\pi)] P_B \neg I_B P_B K_\alpha$.

$\rightsquigarrow$ *'I'm unwilling to agree with what you just said.'*

In contrast to (4a), Sarcasm cannot be applied here, because it is blocked by the dynamic semantics for clarification requests (CRs must be sincere questions). Hence, A is not communicating that B is *wrong*, but just that B is not ready to agree.

5 Conclusion

We have presented a unified, formal account of the perlocutionary effects of pitch contours in colloquial English as discussed in the literature. The novel contribution of our model is the formal derivability of these effects. Our stipulations of cognitive effects are independently motivated and in line with previous analyses of these effects. By connecting them with the logics of SDRT, we obtain concrete derivations of implicatures communicated by pitch. In future work, we plan to extend this analysis to interrogatives and imperatives. Further, we have ignored the focus effects of the lexical placement of pitch accents here. To integrate these effects into our account, we plan to extend SDRT's glue and cognitive logics to reasoning with the contents of sub-clausal units.

[10] KD45 models introspection, *i.e.*, $B_B\varphi \rightarrow B_B B_B \varphi$. Hence, if $B_B \neg B_B \varphi$, then $B_B \varphi$ must fail.

[11] The analoguous derivation for examples (7b) and (7,c) accounts for the *'I thought you knew'* implicature.

[12] By Int. Transfer+Cooperativity, $P_B K_\pi \mid\!\sim I_A P_A K_\pi$, and by Sincerity (b), $B_A \neg K_\pi \mid\!\sim \neg I_A P_A K_\pi$, hence the second disjunct normally does not apply.

References

Nicholas Asher and Alex Lascarides. 2003. *Logics of conversation*. Cambridge University Press.

Nicholas Asher and Alex Lascarides. 2008. Commitments, beliefs and intentions in dialogue. In *12th Workshop on the Semantics and Pragmatics of Dialogue (LONDIAL)*.

Nicholas Asher and Alex Lascarides. 2013. Strategic conversation. *Semantics and Pragmatics*, 6(2):1–62.

Dwight Bolinger. 1982. Intonation and its parts. *Language*, pages 505–533.

Christine Gunlogson. 2003. *True to form: Rising and falling declaratives as questions in English*. Routledge.

Julia Hirschberg and Gregory Ward. 1995. The interpretation of the high-rise question contour in english. *Journal of Pragmatics*, 24(4):407–412.

Jerry Hobbs. 1990. The pierrehumbert-hirschberg theory of intonational meaning made simple: Comments on pierrehumbert and hirschberg. In J. Morgan P. R. Cohen and M. E. Pollack, editors, *Intentions in Communication*, pages 313–323. MIT Press.

D Robert Ladd. 1980. *The structure of intonational meaning: Evidence from English*. Indiana University Press.

Alex Lascarides and Nicholas Asher. 2009. Agreement, disputes and commitments in dialogue. *Journal of Semantics*, 26(2).

Mark Liberman and Ivan Sag. 1974. Prosodic form and discourse function. In *10th Regional Meeting of the Chicago Linguistics Society*, pages 416–427.

S. Malamud and T. Stephenson. 2011. Three ways to avoid commitments: declarative force modifiers in the conversational scoreboard. In *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)*, pages 74–83. Los Angeles.

Janet Pierrehumbert and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In J. Morgan P. R. Cohen and M. E. Pollack, editors, *Intentions in communication*, pages 271–311. MIT Press.

Julian J Schlöder. 2015. A formal semantics of the final rise. In *Student Session of the European Summer School in Logic, Language and Information 2015*.

Kim EA Silverman, Mary E Beckman, John F Pitrelli, Mari Ostendorf, Colin W Wightman, Patti Price, Janet B Pierrehumbert, and Julia Hirschberg. 1992. ToBI: a standard for labeling english prosody. In *2nd International Conference on Spoken Language Processing*.

Mark Steedman. 2000. *The syntactic process*. MIT Press.

Mark Steedman. 2014. The surface-compositional semantics of english intonation. *Language*, pages 2–57.

M. Šafářová. 2005. The semantics of rising intonation and interrogatives. In E. Maier, C. Bary, and J. Huitink, editors, *Proceedings of SuB9*, pages 355–369.

Matthijs Westera. 2013. Attention, im violating a maxim! A unifying account of the final rise. In *17th Workshop on the Semantics and Pragmatics of Dialogue (DialDam)*.