Understanding Focus: Tune, Placement, and Coherence

Julian J. Schlöder
ILLC
University of Amsterdam
julian.schloeder@gmail.com

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

Abstract We present a semantic analysis of focal stress and prosodic tunes in English dialogue. We use data involving two distinct tunes to argue for modelling focus and tune jointly. This allows us to combine some independently motivated principles to tackle some previously confounding problems. We model focus as triggering a presupposition (Geurts & van der Sandt 2004a), which in turn is interpreted via its rhetorical connections to the discourse context and the proffered content of the utterance that introduced it; and we assign tunes perlocutionary effects that also interact with coherence (Schlöder & Lascarides 2015). We regiment this in Segmented Discourse Representation Theory (SDRT, (Asher & Lascarides 2003)), and use the formalism to predict intuitively compelling anomalies and implicatures of various tunes in various contexts.

1 Introduction

The intonation of an utterance contributes to its meaning, but in highly complex ways. Some researchers aim for compositional meanings of discretised accents (e.g., via the ToBI annotation scheme; Silverman et al. 1992). But others provide theoretical and empirical evidence for a non-discrete *spectrum* of possible tunes with varying meaning, which cannot be treated compositionally (Ladd 1980, Bolinger 1982, Calhoun 2007). Nevertheless, there are strong intuitions regarding the felicity and meaning of some tunes in some contexts: (1) demonstrates the basic intuition that in the context of a *wh*-question, focal placement in an answer follows the structure of the question; in (2), the intonation leads to an *as-opposed-to* implicature (Pierrehumbert & Hirschberg 1990).[1]

(1) a. Harvey: Who likes Michael?
b. Jessica: $\underline{\text{Rachel}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
#b.′ Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H}\cdot\text{LL\%}}$

1 In all examples, we underline the word on which the nuclear accent is placed, adding its pitch type as a subscript. This discretised notation obscures intensity: the examples should be intonated with strong and prominent stress; for fall–rise, the most intense part can be on the low or high pitch.

(2) a. Louis: Is Harvey going to fire me?
b. Donna: Harvey$_{\text{L+H}}$ is not going to fire you.$_{\text{LH\%}}$
$\rightsquigarrow$ *but someone else is*

The judgements on these and similar cases appear to be robust enough to be amenable to systematisation and formal treatment, at least in principle.

Following Steedman (2014) and others, our goal is to associate each intonational form with a *single* logical form, with its distinct implicatures in distinct contexts being derivable via independently motivated principles of pragmatics. Our account is distinct from prior work, however, in achieving a *combination* of two things: (i) making the logical form derived from form sensitive to *both* the tune and accent placement; and (ii) providing *formally precise logical derivations* from these logical forms to pragmatic implicatures, via existing axioms of pragmatic inference that have already been used to model other pragmatic phenomena (e.g., presupposition and anaphora). Existing work that formally derives implicatures ignores either the effects of accent placement (e.g., Schlöder & Lascarides (2015)) or of tune (e.g., Roberts (2012)), or applies to only one tune (Reese 2007). On the other hand, existing work where logical forms are sensitive to both accent placement and tune offer no formal derivation from those logical forms to implicatures (e.g., Steedman 2014).

Our goal is to expand semantic research on *focus* by linking it to the semantics of *tunes*. We cannot capture the full spectrum of tunes in the present work, so for the sake of clarity we avoid overlapping or vague tunes, and consider only two exaggerated and diametrically opposed tunes:

- a *falling* tune with a single clearly discernible high pitched nucleus, which we annotate as H LL%; and
- a *fall–rise* tune with a single clearly discernible low–high pitched nucleus, which we annotate as L+H LH%.[2]

This means in particular that we ignore several relevant phenomena, including pre-nuclear pitches, boundary tones within clauses, and multiple-focus constructions. All our basic data will consist of a single clause with only one (nuclear) accent, where we vary: (i) the placement of that accent; and (ii) whether the overall tune is falling or fall–rise.

We will proceed as follows. Section 2 argues for a main tenet of this work—that focus and tune need to be modelled jointly. In Sections 3 and 4, we informally motivate and describe a semantics for the above two tunes that draws on familiar

2 Gunlogson (2003) and (Steedman 2014) assign the final rise a distinct semantics from the rest of the tune. Following Calhoun (2007) we prefer a holistic approach, but our framework is flexible enough to *add* a distinct final rise semantics, if data should surface that requires this (Schlöder 2015).

pragmatic concepts, such as discourse coherence and presupposition. We formalise our model in Sections 5 and 6 within Segmented Discourse Representation Theory (SDRT, Asher & Lascarides (2003)), by exploiting its existing model of interaction between discourse coherence, the interpretation of presuppositions (Asher & Lascarides 1998) and perlocutionary effects (Asher & Lascarides 2013, Schlöder & Lascarides 2015). In Section 7 we demonstrate the predictive power of our model by analysing some of the examples that motivated it.

2 Tunes and Accent Placement

We now argue for assigning meaning along two dimensions: the *accent placement* within a clause and the *meaning of the tune itself* (in the sense of Ladd's intonational lexicon (1980: ch.7)).

2.1 Congruence, Focus and Tune

Example (1) motivates the principle of *question-answer congruence* (Halliday 1967, Büring 2007, Roberts 2012): i.e., that focus indicates the *wh*-question an assertion answers. The *congruent* question is the one obtained by placing the *wh*-element on the focal constituent. So (1b) succeeds in answering (1a) but (1b′) does not. However, contrary to question-answer congruence, there are felicitous answers to *wh*-questions where accent placement does *not* match the *wh*-element of the question: (3ab) is infelicitous, but (3ab′) and (3ab″) are acceptable and indeed natural in a context where Jessica thinks anyone liking Michael is absurd.[3]

(3) a. Harvey: Who likes Michael?
#b. Jessica: Nobody likes $\underline{\text{Michael}}_{\text{H}\cdot\text{LL\%}}$
b.′ Jessica: Nobody likes $\underline{\text{Michael}}_{\text{L+H}\cdot\text{LH\%}}$
b.″ Jessica: Nobody $\underline{\text{likes}}_{\text{L+H}}$ Michael.$_{\text{LH\%}}$

Roberts (2012: p34) claims that placing stress on "likes" or on "Michael" is infelicitous; but she only marks focus—not tune—in her annotation. Therein lies a methodological problem that we want to address. Considering focus without its *tune* leads to confusion. This point is acknowledged (Beaver & Clark 2009, Roberts 2012), but we want to give it prominence.

One might claim that annotating focus without its tune categorically denotes a falling tune. But this does not square with how focus is discussed *across* the literature. To see this, we'll argue that Kratzer's (1989) claims about (4) apply only when they're uttered with a fall–rise tune:

3 Roberts' (2012) account is more sophisticated, allowing the accommodation of congruent questions; we discuss this in Section 3.1.

(4) a. Paula does not live in Paris.
b. Paula does not live in Paris.

She claims that (4a) presupposes that *someone who is not Paula lives in Paris* (and this contrasts with the proffered content) whereas (4b) presupposes that *Paula lives somewhere that is not Paris* (which again contrasts the proffered content). But in fact, the felicity and presupposition vary when placing (4a) in different contexts with different tunes:

(5) a. William: Does Paula live in Paris?
b. Edith: Paula$_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *someone (else) does live in Paris*
#b.′ Edith: Paula$_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$
b.″ Edith: Paula does not$_{\text{H}}$ live in Paris.$_{\text{LL\%}}$
$\not\rightsquigarrow$ *someone (else) does live in Paris*

(6) a. William: Who does not live in Paris?
b. Edith: Paula$_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *But is this what you wanted to know?*
b.′ Edith: Paula$_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$
$\not\rightsquigarrow$ *someone (else) does live in Paris*

Focus placement and tune work jointly here: since (5b″) is felicitous, it is not the falling tune in itself that is bad in (5b′), but the falling tune with accent on *Paula*.

Now, on Kratzer's reading, (4a) has the truth conditions of (7a) (see also (5b)). But (6b′) shows that (4a) can be interpreted in a way similar to (7b) too.

(7) a. It is not Paula who lives in Paris.
$\rightsquigarrow$ *someone (else) does live in Paris*
b. It is Paula who does not live in Paris.
$\not\rightsquigarrow$ *someone (else) does live in Paris*

Furthermore, the fact that (7a) is an anomalous response to the question (6a) and an acceptable response to (5a), while it is the other way round for (7b), suggests that (7a) broadly corresponds to the fall–rise tune and (7b) to the falling tune (Section 3.2 addresses how (6ab) defies this correspondence). Since (5a) is a *prima facie* more natural context than (6a), it seems reasonable to assume that one tends to read (4a) in its null context with the fall–rise tune given in (5b); this would explain Kratzer's intuitions. We conjecture that in general, the intuitions linked to examples like (4) where stress is annotated but no tune and no discourse context are given are actually the intuitions associated with a tune that is felicitous in the most natural discourse context. Overall, discussion is clouded by the inherent ambiguity in considering focus without also considering tune. We aim to develop a model that considers focus and tune *jointly*.

2.2 Interest

Our model will follow prior work that links intonation to *cognitive attitudes* (e.g., Pierrehumbert & Hirschberg (1990), Hobbs (1990), Gunlogson (2003), Steedman (2014)). Notably, Bolinger (1972, 1985) claims that accent placement indicates something *interesting* to the speaker. This offers an alternative explanation to question-answer congruence for predicting (1). It's quite intuitive to assume that Rachel is significant (*interesting*) while Michael is not, because dialogue participants share an expectation that questions will be answered, and basic principles of information theory then predict the significance of 'Rachel' and the insignificance of 'Michael' (Weaver & Shannon 1963). Bolinger's notion of interest can potentially explain (3) too. Assuming for the time being that the fall–rise tune implicates that Jessica finds anybody liking Michael absurd, the accent placement then indicates Jessica's interest: in (3b′) she thinks it absurd that of all the people one could like, one would choose Michael; in (3ab″), that of all the things one could do to Michael, one would choose liking.

Explaining focus in terms of *interest* may be compelling, but is far too vague to systematically predict infelicity judgements or compute implicatures. Clearly, we cannot *define* what it means for a speaker to find something interesting; Bolinger's (1972) point is that interest (and thus focus) is not predictable. However, we can identify and formalise conditions that are *necessary* for something to be marked as interesting. Bolinger (1972) provides us with one such condition that *is* predictable: what is obvious cannot be interesting. Our account will formalise this necessary condition on interest, and interface it with an appropriate tune semantics.

Accordingly, we take the non-focal parts of an utterance to be *not marked as interesting* and require them to function as a coherent presupposition (Jackendoff 1972, Geurts & van der Sandt 2004a). These ideas can be made predictive via independently motivated theories of presupposition (van der Sandt 1992, Asher & Lascarides 1998). Thus focus does not cleanly separate the new from the given content. Just because something is not *marked* as interesting does not mean that it *is not* interesting; in particular, it can be new. So our account allows non-focal information to be discourse-new (contrary to what Büring (2006) describes as the "standard view" on focus projection). In such cases, the presupposition triggered by the non-focal parts needs to be accommodated. Our semantics for intonation as derived from its form is sufficiently underspecified to model readings that might be treated via focus projection on other accounts (see Section 4.2 for details).

The following examples motivate interest interacting with *coherence*:

(8) a. Harvey: Does Rachel like Michael?
b. Jessica: Rachel $\underline{\text{does}}_{\text{H}}$ like Michael.$_{\text{LL\%}}$
b.′ Jessica: Rachel does $\underline{\text{not}}_{\text{H}}$ like Michael.$_{\text{LL\%}}$

(9) a. Harvey: Does Rachel like Michael?
b. Jessica: $\underline{\text{Rachel}}_{\text{L+H}}$ does not like Michael.$_{\text{LH\%}}$
$\rightsquigarrow$ *but someone else does*.
b.′ Jessica: Rachel does not $\underline{\text{like}}_{\text{L+H}}$ Michael.$_{\text{LH\%}}$
$\rightsquigarrow$ *but she does something else regarding Michael*.
b.″ Jessica: Rachel does not like $\underline{\text{Michael}}_{\text{L+H}}$.$_{\text{LH\%}}$
$\rightsquigarrow$ *but someone else*.

(10) a. Harvey: Does Rachel like Michael?
b. Jessica: $\underline{\text{Katrina}}_{\text{L+H}}$ likes Michael.$_{\text{LH\%}}$
$\rightsquigarrow$ *Rachel is less interesting here than Katrina*.
or *Jessica does not know whether Rachel does*.
#b.′ Jessica: Katrina $\underline{\text{likes}}_{\text{L+H}}$ Michael.$_{\text{LH\&}}$
#b.″ Jessica: Katrina $\underline{\text{likes}}_{\text{H}}$ Michael.$_{\text{LL\%}}$

The answers in (8) are congruent; in Bolinger's terms, Jessica marks as interesting the polar issue in question (8a). But the answers in (9) indicate interest in other elements, leading to certain implicatures. In (9b), Jessica seems to indicate that whether Rachel likes Michael is less interesting than *who* likes him; in (9b′) *what* Rachel thinks of Michael; and in (9b″) *who* Rachel likes. Thus Jessica answers the question, but what the intonation marks as interesting is *coherently related* rather than *equivalent* to the polar issue in question (9a). That Jessica's marked interest must be *coherent* predicts why focus placement in her responses (10) are more constrained. (10b) is just like (9b); i.e., Jessica places interest on *who* likes Michael. But in (10b′) and (10b″), she places interest on what Katrina thinks of Michael, which bears no discernible relation to (10a); it sounds *off-topic*.

3 Our Semantics for Intonation

We argue for defining intonation in terms of coherence (see also Hobbs (1990), Reese (2007)), and that the *background* and *foreground* derived from intonational form are dependent on *both* the accent placement and the tune (Steedman 2014).

3.1 Background and Foreground

Most accounts of focus separate a *foregrounded* (focal, rhematic) constituent from a *backgrounded* (given, thematic) one (e.g., Gussenhoven (1983), Krifka (1992) and others).[4] The accounts differ on how these two parts interact with the context and with each other. Following Geurts & van der Sandt (2004a), we make the background

4 Vallduví & Engdahl (1996) also partition the background into two parts, but we gloss over this here.

trigger a presupposition, but our account differs on what is presupposed, particularly for the fall–rise tune. We propose the following focus semantics for the falling tune.

(I) *Focus Semantics* (falling tune)
Focal placement separates an utterance into a *foreground* f and a *background* φ, where a variable x of the same type as f occurs freely in φ. Updating a discourse with an utterance that has a falling tune with nuclear accent on f proceeds as follows:

- Update with the *presupposition* φ; that is, its free variable x must be resolved anaphorically (it's either bound or accommodated as $\exists x.\varphi$).
- Update the result with the *proffered* content $(\lambda x.\varphi)(f)$ (and all its presuppositions), such that the proffered content and φ are coherently connected to form a common topic (this means that the proffered content must *elaborate* the presupposition or form a *continuation* with it (Asher & Lascarides 2003)).

Note that we treat presuppositions as anaphora—i.e., a presupposition must be bound to an available unit in the discourse context or accommodated by coherently relating it to such a unit (Asher & Lascarides 1998). It's this treatment of presuppositions that leads to treating the free variable x in the background φ as an anaphor, which needs to be bound or existentially closed (van der Sandt 1992).

In "$\underline{\text{Rachel}}_{\text{H}}$ likes Michael", $f = \textit{Rachel}$ and $\varphi = x \textit{ likes Michael}$ (the foreground triggers a presupposition via the proper name, but by Rule (I) this updates the context *after* φ). After φ updates the context, the proffered content $\varphi(f)$ must attach to it with *Elaboration* (making their common topic φ) or *Continuation* (their common topic is a generalisation of their distinct but related contents). In (1), the question presupposes *someone likes Michael*, and so by the dynamic semantic approach to interpreting presuppositions, the background x *likes Michael* binds to this, with x bound to the existential quantifier (van der Sandt 1992). The proffered content then attaches to this background with *Elaboration*—colloquially, *someone likes Michael—specifically, Rachel does*.

Other accounts also predict that the background is presupposed. To wit, $?\lambda x.\varphi(x)$, where φ is the background defined in Rule (I), is the *wh*-question that's required by a Question Under Discussion (QUD) account (Roberts 2012), and $\{x \mid \varphi(x)\}$ is the set of alternatives in Alternative Semantics (Rooth 1992). Under the reasonable assumptions that *wh*-questions presuppose at least one true answer and that alternatives sets are non-empty, both accounts generate the presupposition that there is an x such that $\varphi(x)$. In this sense, presupposing the background is a minimal assumption and one might think that drawing on coherence relations is redundant. But we intend

to argue that coherence is needed: it accurately constrains how presuppositions get resolved in context (Asher & Lascarides 1998), and it also characterises the semantic dependencies between background and focus in ways that are sensitive to the tune.

Geurts & van der Sandt (2004a) argue convincingly that standard models for how presuppositions get bound or accommodated make the right predictions for focus (though we will argue against their treatment of negation in Section 3.4). Our informal analysis of (1) was an example of binding; (11) (adapted from Geurts & van der Sandt (2004b: ex.3)) is an example of accommodation.

(11) a. Harvey: Does anybody like Michael?
b. Jessica: $\underline{\text{Rachel}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$

Unlike (1a), the question (11a) doesn't generate an existential presupposition, so *x likes Michael* is accommodated: in standard dynamic semantics, this is equivalent to adding an existential quantifier $\exists x$. This accommodation is similar to Roberts' (2012) idea that focus triggers the presupposition of a congruent question and that this question can be *accommodated* as a QUD. There must be constraints on accommodating questions; otherwise, any accent placement is acceptable. Roberts (2012: pp14–15) allows a question to be accommodated if any complete answer to it partially answers the previous QUD. In (11), Jessica's utterance requires *Who likes Michael?* to be accommodated, and this succeeds because each complete answer to it also answers (11a).

By and large, Rule (I) makes the same predictions as the QUD model. But the following examples motivate replacing congruence with something else, as we have done in Rule (I).

(12) a. Harvey: Who likes Michael?
#b. Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H}}$.$_{\text{LL\%}}$

(13) a. Harvey: Does Rachel like Michael?
b. Jessica: $\underline{\text{Rachel}}_{\text{L+H}}$ does not like Michael.$_{\text{LH\%}}$
#b.′ Jessica: $\underline{\text{Rachel}}_{\text{H}}$ does not like Michael.$_{\text{LL\%}}$

Roberts (2012: p2) claims her model predicts the infelicity in (12). This would be the case if question-answer congruence were the *only* principle of discourse interpretation. But Roberts (2012: p8) acknowledges the need for implicit questions to be accommodated (see also (11)). So to predict that (12b) is infelicitous, accommodating *Who does Rachel like?* must be ruled out. But the constraint on accommodating questions that we described earlier is satisfied: each complete answer to *Who does Rachel like?* entails either that Rachel likes Michael, or that she does not, and so partially answers (12a). Further, *both* (13b) and (13b′) would involve accommodating *Who does not like Michael?*; so merely computing how this question relates to

question (13a) cannot tease apart their differences in felicity. So it's not clear how her model makes the right predictions in (12) and (13). We will account for them by assigning a different semantics to fall–rise tunes in Section 3.2, and by defining a weakened notion of question–answer congruence in Section 4.

3.2 Negation and Contrast

In Section 2.1 we argued (5b) has the truth-conditions of (7a) and (6b′) of (7b).

(5) a. William: Does Paula live in Paris?
b. Edith: $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *someone (else) does live in Paris*
#b.′ Edith: $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$

(6) a. William: Who does not live in Paris?
b. Edith: $\underline{\text{Paula}}_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *But is this what you wanted to know?*
b.′ Edith: $\underline{\text{Paula}}_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$
$\not\rightsquigarrow$ *someone (else) does live in Paris*

(7) a. It is not Paula who lives in Paris.
$\rightsquigarrow$ *someone (else) does live in Paris*
b. It is Paula who does not live in Paris.
$\not\rightsquigarrow$ *someone (else) does live in Paris*

But in (6) the fall–rise tune is acceptable and its meaning is *not* that of the *it*-cleft (7a). Thus the fall–rise tune doesn't *mandatorily* result in a presupposition corresponding to that of (7a); we must account for this. Further, the as-opposed-to reading of the fall–rise tune in (5b) can arise in the absence of any overt negation:

(14) a. William: Does Paula live in Paris?
b. Edith: $\underline{\text{Paula}}_{\text{L+H}}$ lives in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *but someone (else) does not*

There is no negation in (14b), so we cannot attribute its implicature to determining the relative scope of a presupposition to a linguistically-introduced negation. Rather, this as-opposed-to reading derives from *adding* negation to the background content, and determining its relative scope. We regiment this as follows: if the tune is fall–rise, we leave the *polarity* of the background underspecified; further, to obtain the intuitive readings in (6b) and (14b), we specify that the foreground is in *contrast* to the background. Thus the way the polarity in the presupposition gets resolved will depend on how it coherently relates to its discourse context, and how it can support a contrast with the proffered content. This semantics is expressed as follows.

(II) *Focus Semantics* (fall–rise tune, first attempt)
The focus placement separates an utterance into a *foreground* f and a *background* φ with free variable x. Updating a discourse with an utterance that has a fall–rise tune with nuclear accent on f proceeds as follows:

- Update with the *presupposition* $y(\varphi)$ where y is an underspecified variable of type *polarity*; i.e. $y \in \{\top, \neg\}$.
- Update with the *proffered* content $(\lambda x.\varphi)(f)$ such that the proffered content *contrasts with* the presupposition.

Rule (II) sets our account apart from competing models in a significant way. Recall example (5ab). On the face of it, (5b) corresponds to the alternatives set *people who do not live in Paris* or the QUD *Who does not live in Paris?*, respectively. Both accounts can, in principle, predict the implicature (as Rule (II) does) by accommodating the alternatives set *people who live in Paris* or the QUD *Who lives in Paris?* respectively. But whatever the grounds for accommodating this question, it can't be congruence: (5b) is not congruent to this question.

On our account, where no congruence is required, we predict the implicature as follows. The utterance (5b) presupposes y(*x does not live in Paris*). The only available referent to bind x to is *Paula*, but this is blocked (regardless of how y is resolved) by the requirement that it contrast the proffered content. Thus, the presupposition must be accommodated as $\exists x.y$(*x does not live Paris*). Both $y = \top$ and $y = \neg$ would be permissible here, but $y = \neg$ establishes a stronger contrast to the proffered content *Paula does not live in Paris* and is therefore preferred on the independently motivated principle that people interpret discourse in a way that *maximises coherence* (Asher & Lascarides 2003). By double negation elimination this results in the presupposition that *someone else lives in Paris*.

To see the maximisation of contrast, compare these approximate representations:

(15) a. *There is someone who does not live in Paris, but that someone is not Paula.*
a.′ *There is someone who lives in Paris, but Paula does not.*

While (15a) *can* contrast 'someone' with 'Paula', the contrast in (15a′) is better. The dynamic semantics of *Contrast* and the principle Maximise Discourse Coherence in SDRT formally express a preference for resolving underspecified elements in the discourse units to specific values that maximise coherence, and so predicts that the pragmatic interpretation of (5b) can be paraphrased as (15a′) rather than (15a). Similar reasoning about maximising contrast captures the implicature in (14) as well. We will need additional machinery from Section 4 to explain (5b′).

Now consider (6ab). The *wh*-question (6a) presupposes *some e does not live in Paris*. By Rule (II) (6b) presupposes y(*x does not live in Paris*). This presupposition

can be bound to the presupposition of (6a) by making $y = \top$ and $x = e$ (since binding a presupposition is preferred, $y \neq \neg$). However, the proffered content must *contrast* with this resolution, and so binding $e = x =$*Paula* is blocked: this would result in the presupposition *Paula does not live in Paris* and the identical, therefore non-contrasting, proffered content *Paula does not live in Paris*. Thus, $e \neq$*Paula*, resulting in the reading (15a). In other words, Edith implicates that her answer *Paula* doesn't resolve William's question—the desired implicature of (6b).

Typically $y = \neg$ results in a better contrast with the foreground and is thus often a part of the pragmatic interpretation. In particular, in the null context $y = \neg$ is preferred, yielding the as-opposed-to reading of fall–rise. Only in highly particular contexts, such as (6a), does binding x to an available antecedent and resolving y to $\top$ yield a more coherent discourse. The principle of Maximise Discourse Coherence is the common thread: it captures how the underspecified background content that we stipulate gets resolved to capture the intuitive implicatures of the fall–rise tune in its various contexts.

Finally, the negation that y resolves to can be *metalinguistic*.

(16) a.William: We bought po-tah-toes.
b.Edith: We bought po-tay-toes$_{\text{L+H·LH\%}}$
$\rightsquigarrow$ not "po-tah-toes"

Clearly, in (16b) Edith is not denying the propositional content of (16a) (Carston 1996). We can account for these cases by allowing the $\neg$ in Rule (II) to be metalinguistic and the x in Rule (II) to resolve to prior *use* or *mention*. This option also accounts for fall–rise signalling a speaker taking issue with the *presentation* of a proposition: While Edith logically gives a *positive* answer to (17a), her answer implicates that "in the US", while true, mischaracterises the circumstances.

(17) a.William: Do you live in the US?
b.Edith: I live in New York City$_{\text{L+H·LH\%}}$
$\not\rightsquigarrow$ not in the US.
$\rightsquigarrow$ not "the US".

3.3 Uncertainty Readings

Rule (II), as it stands, fails to model *uncertainty readings*.

(18) a. William: Did Paula eat all the cookies?
b. Edith: Paula ate some$_{\text{L+H}}$ of the cookies.$_{\text{LH\%}}$
$\rightsquigarrow$ *but not all of them*;
or *but Edith isn't sure whether it was all the cookies*.

(19) a. William: Is Michael coming to the party?
b. Edith: He is $\underline{\text{invited}}_{\text{L+H·LH\%}}$
$\rightsquigarrow$ *but he is not coming*;
or *but Edith doesn't know whether Michael is coming*.

Edith's utterances in (18b) and (19b) are ambiguous: they can be interpreted as indirect *negative answers* or as indicating that Edith *is uncertain about* the answer to William's question; in the latter case, Edith is giving information that she has marked (by way of intonation) as perhaps insufficient to resolve the question.

Due to cases like these, the fall–rise tune has been associated with 'uncertainty': Ward & Hirschberg (1985) associate it with the whole tune; Pierrehumbert & Hirschberg (1990) with the L+H accent; and Šafářová (2005) with the final rise. In some accounts this uncertainty is associated exclusively with scalar items (Pierrehumbert & Hirschberg 1990), but a scalar relationship isn't always necessary:

(20) a. Julian: Is Nicholas coming to my talk?
b. Alex: $\underline{\text{Ernie}}_{\text{L+H}}$ is coming to your talk.$_{\text{LH\%}}$

Alex' utterance is ambiguous, like (19). It can express that Nicholas is not coming, or that Nicholas *might* be coming (but Alex does not want to commit to the proposition that he definitely is).[5]

Which of the two readings is preferred seems to vary with the intensity of the intonation and the steepness of the rises (Ward & Hirschberg 1988), as well as contextual knowledge of the speaker's knowledge or intentions. Our notation for fall–rise tunes underspecifies such features. Thus we need to make sure that our semantics makes both readings *available*. The first version of the Focus Semantics for fall–rise (Rule II) only predicts the indirect answer reading; i.e., that Edith is saying that *Paula did not eat all the cookies (only some)* in (18) and that *Michael is not coming (despite being invited)* in (19).

To account for such uncertainty readings, we amend the Focus Semantics by adding the option that the underspecified polarity y might resolve to the modal $\Diamond$ introduced by Šafářová's (2005) semantics of final rise, as well as to $\top$ or to $\neg$.[6] Since our model must include underspecification of auxiliaries anyway (see Section 3.4), this incurs no additional formal overhead.

5 Other readings are also available. For instance, if everyone knows that Nicholas goes wherever Ernie goes, (20b) implicates a positive answer (i.e., Nicholas is coming) and that Julian should have known this. We'll predict this reading in Section 6. A further reading is that Alex is giving *better* information with regards to Julian's intentions; i.e., that it doesn't matter whether Nicholas is coming because Ernie is. Since this requires very specific knowledge of Julian's intentions, we leave it aside.

6 The modal operators allow for counterfactuals. Thus, if something is P, it is *definitely* P ($P(x) \models \Box P(x)$), but even if something is P, it is (counterfactually) possible that it is not. That is, $P(x) \wedge \Diamond\neg P(x) \not\models \bot$; note that the conjunction is interpreted dynamically. Also see Groenendijk et al. (1996).

(II) *Focus Semantics* (fall–rise tune, final version)
The focus placement separates an utterance into a *foreground* f and a *background* φ, which contains a free variable x of the same type as f. Updating a discourse with an utterance that has a fall–rise tune with nuclear accent on f proceeds as follows:

- Update with the *presupposition* $y(\varphi)$ where y is an underspecified variable of type (alethic) *modality*; i.e. $y \in \{\top, \Diamond, \neg\}$.
- Update with the *proffered* content $(\lambda x.\varphi)(f)$ such that the proffered content *contrasts with* the presupposition.

Note that, as before, *maximising contrast* typically favours the $\neg$ reading: "something isn't P but C is P" is typically a better contrast than both "something is P but C isn't that something" and "something is possibly P but C is (definitely) P". So the interpretations we outlined in Section 3.2 are replicated by this final version of the Focus Semantics. Thus our semantics favours interpreting (18b), (19b) and (20b) as indirect negative answers. But the reading where $y = \Diamond$ is available to interpret (19b) (for instance) as *possibly Michael is coming, but he (definitely) is invited*, and this reading arises if the indirect answer reading is pragmatically *blocked*, for instance by the knowledge that Edith cannot know for sure that Michael is coming.

More generally, real world knowledge can substantially affect how y resolves in context. Example (21) is a case where the uncertainty reading is preferred:

(21) a. Amy: Does Paula like opera?
b. Bob: She likes $\underline{\text{Wagner}}_{\text{L+H·LH\%}}$
$\not\rightsquigarrow$ *Paula does not like opera.*
$\rightsquigarrow$ *Paula possibly likes opera.*
b.′ Bob: She likes $\underline{\text{Wagner}}_{\text{H·LL\%}}$
$\rightsquigarrow$ *Paula likes opera.*

Axioms of rationality and cooperativity predict that responses to polar questions provide evidence for a positive answer or for a negative answer (Asher & Lascarides 2003: p403–405); when the evidence proffered is conclusive, a particular answer is implied. Combining this expectation with the real world knowledge that *liking Wagner* is strong evidence for *liking opera* makes the reading *Paula does not like Opera, but she likes Wagner* dispreferred. So, intuitively, this real world knowledge predicts that Bob has offered evidence for a *positive* answer, and so in this context the fall–rise intonation conveys that he doesn't quite commit to a positive answer (*Paula possibly likes opera, but definitely likes Wagner*).

This contrasts with (21b′) uttered with falling intonation, where the (same) evidence for a positive answer, provided by real world knowledge about Wagner

and opera, commits Bob to a positive answer (*Paula likes opera, specifically, she likes Wagner*). These differences are predicted by our tune semantics: the fall–rise tune demands a contrast between given and proffered content, while a falling tune demands Elaboration or Continuation. However, such readings are a matter of degree: (22b) (from Steedman (2014)) is arguably ambiguous as to which answer it implicates because we cannot decide with sufficient confidence whether Bob's assumptions about the commonsense relations between musicals and opera lead him to believe that *liking musicals* is positive evidence for *liking opera*, or if he intends it to be negative evidence.

(22) a. Amy: Does Paula like opera?
b. Bob: She likes $\underline{\text{musicals}}_{\text{L+H·LH\%}}$

3.4 Some Challenges to a Presupposition Approach

Explaining focus by presupposing the background has been criticised (Dryer 1996, Rooth 1999). Even Geurts & van der Sandt (2004a: pp28–30) criticise it for focussed quantifiers: a naïve reading of our semantics can yield the faulty readings in (23) (x's type appears as a subscript).

(23) a. $\underline{\text{Nobody}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
?? background: $x_{\textbf{entity}}$ *likes Michael*.
b. $\underline{\text{Somebody}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
?? background: $x_{\textbf{entity}}$ *likes Michael*.

The background in (23a) contradicts the proffered content: $x_{\textbf{entity}}$ denotes an individual in the model and so cannot be nobody. The background in (23b) makes the proffered content not new (or interesting). In both cases presupposing the background is absurd, and so Geurts & van der Sandt (2004a: pp28–30) argue instead for a *polarity* focus, thereby yielding for both (23a) and (23b) the tautological presupposition *either nobody likes Michael or somebody does*. They justify this as follows:

> The non-logical part of the semantic content of words like 'somebody' and 'nobody' is so general that it is unlikely to attract the focus of a statement; 'somebody' cannot be used to mean 'some person, as opposed to some vehicle' (say). What remains to be focused is the negative part of 'nobody' and the corresponding positive component of 'somebody', which is what determines the polarity of the sentence.

However, contrary to this: (i) there are cases where *somebody* gets an 'as opposed to' reading; and (ii) there are cases where focus on an existential quantifier is not polar (i.e. not contrasting with 'nobody' or 'nothing'). To see (i) consider (24) from Walker (1996) (we've added an appropriate tune).

(24) a. William: There is a man in the garage.
b. Edith: There is <u>something</u>$_{\text{L+H}}$ in the garage.$_{\text{LH\%}}$

(25) a. William: There is something in the garage.
b. Edith: There is some<u>body</u>$_{\text{L+H}}$ in the garage.$_{\text{LH\%}}$

Intuitively, (24b) means *it need not be a man* and so corrects (24a). Walker calls this *implicature rejection*: since *something* is less informative than *a man*, one infers the desired reading via a Quantity implicature. But one cannot analyse (25) this way—*somebody* is not less specific than *something*—and yet its meaning is exactly what Geurts & van der Sandt (2004a) deny. Dialogue (26) from Schlöder & Fernández (2015) is an example of case (ii): [7] Danny's denial move with focus on 'some' cannot be about polarity: the issue is not between *some* and *none*, but between *some* and *all*.

(26) a. James: (. . .) we're all mad, aren't we?
b. Danny: Well, <u>some</u>$_{\text{L+H}}$ of us.$_{\text{LH\%}}$

We propose that in (24–26), the second speaker takes issue with the first speaker's choice of *specific quantifier*. So we allow quantifiers as foregrounds, and hence as free variables in the presupposition that's triggered by our Focus Semantics:[8]

(27) a. <u>Nobody</u>$_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
background: $x_{\textbf{quantifier}}z.(z\ \textit{likes Michael})$.
b. <u>Somebody</u>$_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
background: $x_{\textbf{quantifier}}z.(z\ \textit{likes Michael})$.

The presupposition in (27) can be accommodated to form a tautology, since 'there is a quantifier x such that $x(p)$' is true of any proposition p. Thus we obtain the reading of Geurts & van der Sandt. However, this semantics is also compatible with the more specific readings of the backgrounds in (24–26). Note moreover that a similar phenomenon applies to modal operators.

(28) a. William: Do I have to attend class?
b. Edith: You <u>may</u>$_{\text{H}}$ attend class;$_{\text{LL\%}}$ we do not take attendance.
background: *William* $x_{\textbf{aux}}$ *attend class*.

To the best of our understanding, an Alternative Semantics would raise a set of alternative modalities and a QUD-account would take (28b) to answer 'What are

7 This is from the British National Corpus, file HUV, lines 1468–1469. The tune is constructed by us; full audio is not available.

8 We assume standard compositional semantics for (generalised) quantifiers. To wit $[\![x_{GQ}z.\varphi]\!] = 1$ iff $\{z \mid \varphi\} \in [\![x_{GQ}]\!]$. For instance $[\![\forall]\!]$ is the singleton universal set and $[\![\exists]\!]$ is the set of all nonempty sets.

Edith's obligations regarding attending class?'. If modal operators can be foregrounds f, then there is little harm in also having foregrounded quantifiers.

Geurts & van der Sandt (2004a) avoid further tricky cases by making the background presupposed *by default*—the presupposition gets what they call *suspended* in certain cases. But making the representation of the presupposition sensitive to the tune explains the cases of apparent suspension while keeping the Focus Semantics *universal*: i.e., the background is *always* presupposed. Geurts & van der Sandt (2004a: p12) motivate suspended presuppositions by claiming that (29a) presupposes *somebody stole the tarts* but (29b) does not:

(29) a. Fred's wife didn't steal the tarts.
b. I'm still not convinced that the tarts were stolen, but surely Fred's wife didn't steal them.

But we think this difference arises from reading (29a) with a fall–rise tune.

(30) ?a.′ Fred's wife$_{\mathrm{H}}$ didn't steal the tarts.$_{\mathrm{LL\%}}$
a.″ Fred's wife$_{\mathrm{L+H}}$ didn't steal the tarts.$_{\mathrm{LH\%}}$
b.′ I'm still not convinced that the tarts were stolen, but surely Fred's wife$_{\mathrm{H}}$ didn't steal them.$_{\mathrm{LL\%}}$
b.″ I'm still not convinced that the tarts were stolen, but surely Fred's wife$_{\mathrm{L+H}}$ didn't steal them.$_{\mathrm{LH\%}}$

It sounds odd to utter (30a′) out of the blue: one is inclined to read (29a) as belonging to a context in which Fred's wife is suspected of tart stealing, and a fall–rise tune as in (30a″) is the intuitive tune for denying this suspicion (and implicate that someone else stole them).[9] According to our Focus Semantics (Rule II), (30a″) generates the presupposition y(*x didn't steal the tarts*) and its preferred coherent interpretation yields $y = \neg$: *there's someone who did steal the tarts* forms the best available contrast with the proffered content. Compare this with the alternatives: $y = \top$ (??*someone didn't steal the tarts but that someone is not Fred's wife and Fred's wife didn't*) and $y = \Diamond$ (?*someone possibly didn't steal the tarts, but Fred's wife (definitely) didn't steal the tarts*) support only a weak contrast with the proffered content.

Now, (30b′) generates the presupposition *surely x didn't steal the tarts*; this is consistent and coherent in its context (that you're not convinced the tarts were stolen). (30b″) generates the presupposition y(*surely x didn't steal the tarts*). Setting $y = \neg$ and accommodating x to $\exists x$ yields the most coherent interpretation (we interpret *not surely not* as *possibly*): *I'm still not convinced that the tarts were stolen, but someone*

9 One can construct a context where the issue *who did not steal the tarts* is salient; for instance when everyone knows many people stole tarts and one wonders whether anyone didn't. That is the context needed for (30a′).

possibly stole them though it surely wasn't Fred's wife. Thus, in no case do we need to suspend a presupposition.

4 Pragmatics of Focus and Tune

We now specify which parts of an utterance can be considered *interesting* and discuss the *perlocutionary effects* of our two tunes.

4.1 Given and Interest

Any adequate semantics must explain the infelicity of (1b′) and (1b″).

(1) a. Harvey: Who likes Michael?
b. Jessica: $\underline{\text{Rachel}}_{\text{H}}$ likes Michael.$_{\text{LL\%}}$
#b.′ Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H·LL\%}}$
#b.″ Jessica: Rachel likes $\underline{\text{Michael}}_{\text{L+H·LH\%}}$

We proposed earlier to explain it by formalising a necessary condition on *interesting*: intuitively, one cannot place focus on 'Michael' because in the context of (1a), 'Michael' is *obvious* and hence not interesting.

Thus we propose that to be interesting it is *necessary* (but not sufficient) to be *not given*. To make such a constraint predictive, we need to precisely define 'given'. Because focus marks the *speaker's* interest, we take the backgrounded content of the current utterance to contribute to what is considered given.

(III) *Givenness Rule*
The given information is the content that results from (coherently) updating the content of the prior discourse with the (presupposed) background of the current utterance.

Thus, if the most coherent way to update the discourse context with the background content results in a meaning that entails the proffered content, then what is proffered is *given*. Rule (IV) now makes this anomalous:

(IV) *Necessary Condition for Interest*
A foreground–background pair $\langle \varphi, f \rangle$ is *not interesting* if the proffered content $\varphi(f)$ is given. This means:

- for the falling intonation, which presupposes φ with free variable x, the utterance is *not interesting* if it is given that $x = f$.
- for the fall–rise intonation, which presupposes $y(\varphi)$ with free variable x and modality variable y, the utterance is *not interesting* if $y = \top$ and $x = f$ is given.

This seems to stipulate the old wisdom that focal information cannot be given (Büring 2006, Beaver & Clark 2009), but our definition of Givenness (III) differs from prior accounts. We do not say that what can be background is *constrained* by what is given, but instead *add* the background to *compute* what is given (compare in particular with Schwarzschild (1999)). Thus, the background may contain *new* information (i.e. that is not in the prior context) which we then *accommodate* as given. Thus, what Rule (IV) essentially says is that you can't *mark* as interesting content that you (also) present as given. As said, we don't *define* interest. Instead, we use notions already available—presupposition, coherence, *etc*—to *constrain* it enough to explain our data.

Rules (III) and (IV) together are a weakening of question–answer congruence. First, they predict that (1b′,b″) sound odd. In the context of (1a), the most coherent way to instantiate x in the presupposition *Rachel likes x* that's triggered by (1b′) is to bind it to *Michael*; for the fall–rise case (1b″), one also resolves y to $\top$. Thus the proffered content with both tunes is given, violating Rule (IV). Note that in computing what is given, the proffered content is not considered; so it doesn't matter that the resolution *Rachel likes Michael* cannot contrast the proffered content in (1b″).

However, these rules arguably amount to a *weaker* constraint than congruence, because they also account for why (31c) succeeds in answering (31a), even though (31ac) *without* (31b) and (31d) is the same as (1ab′).

(31) a. Harvey: Who likes Michael?
b. Harvey: And who likes Louis?
c. Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H·LL\%}}$
d. Donna: Katrina likes $\underline{\text{Louis}}_{\text{H·LL\%}}$

On the congruence view, the focus determines the *form* of the question that is being answered. In (31) this would be *Who does Rachel like?*, which doesn't seem plausible here.[10] Rather, it's more natural to assume that the discourse function of focus placement in (31c) is to identify which question *available in the context* is being addressed. In other words, focus can *select* the contextual element it relates to, rather than just determine the *form* of that contextual element.

Our proposed rules support this. (31c) presupposes *Rachel likes x*. In the context of (31ab), x could be bound to *Michael* or to *Louis*. However, it is not determinable from the given information *alone* which of these instantiations of x is preferred: i.e., coherent interpretations of the given content can be computed, but it remains ambiguous. Thus, the foreground content (*Paula likes Louis*) is not entailed by the given content, and so by Rule (IV) it is felicitous.

10 One could alternatively propose that (31a) and (31b) are *jointly* a question that allows accommodation of the congruent questions of (31c) and (31d). We would not know how to execute this.

Computing given information via coherence plays a further useful role in predicting felicity judgements:

(10) a. Harvey: Does Rachel like Michael?
b. Jessica: Katrina$_{\text{L+H}}$ likes Michael.$_{\text{LH\%}}$
#b.′ Jessica: Katrina likes$_{\text{L+H}}$ Michael.$_{\text{LH\%}}$
#b.″ Jessica: Katrina likes$_{\text{H}}$ Michael.$_{\text{LL\%}}$

Unlike (31), where two coherent interpretations of the given information can be computed (though there's ambiguity as to which is intended), *no* coherent interpretation can be computed of the presuppositions triggered by (10b′,b″) in the context (10a), as there is no x that would resolve *Katrina x'es Michael* as a coherent response to (10a). Thus the independently motivated principle that accommodating a presupposition must coherently connect it to its context (Asher & Lascarides 1998) explains (10b′,b″). Moreover, just like presuppositions generally, the coherent interpretation of given information needn't be *unique* nor *final*, but there must be *at least one* way of establishing its coherence.

Dialogue (5) is a case where a highly salient coherent interpretation of given information gets overridden by proffered content:

(5) a. William: Does Paula live in Paris?
b. Edith: Paula$_{\text{L+H}}$ does not live in Paris.$_{\text{LH\%}}$
$\rightsquigarrow$ *but someone (else) does*.

The presupposed content of (5b) is y(*x does not live in Paris*). The most coherent update of (5a) with the (underspecified) presupposition *on its own* yields $x =$ *Paula* but doesn't resolve y uniquely: $y = \top$ and $y = \neg$ both supply (full) answers to the question, while $y = \Diamond$ provides a partial answer (and so arguably is less coherent and hence dispreferred). Either way, the presupposition is coherent (though ambiguous), with $x =$ *Paula*. However, when updating this with the proffered information, the Focus Semantics for fall–rise demands a *Contrast*, and so the defeasible inference that $x =$ *Paula* is overridden: x gets existentially bound and y resolves to $\neg$ (resulting in *someone else lives in Paris*). This (still) coherently attaches to the question (as a commentary rather than an answer) but it also contrasts the proffered content, as demanded by the focus semantics.

Since the tunes trigger different presuppositions, the focal placement alone doesn't determine whether the foreground can be interesting. Rather, interest is governed by the focal placement and tune in combination:

(5) a. William: Does Paula live in Paris?
#b.′ Edith: Paula$_{\text{H}}$ does not live in Paris.$_{\text{LL\%}}$
b.″ Edith: Paula does not$_{\text{H}}$ live in Paris.$_{\text{LL\%}}$

The presupposition of (5b′) is *x does not live in Paris*; in the context of (5a), the most coherent resolution is *x*=*Paula*. So by Rule (IV), (5b′) is not interesting and thus incoherent. In contrast, (5b) generates the presupposition *y*(*x does not live in Paris*). As with (14), resolving *y* to $\top$ or $\neg$ are equally coherent ($y = \neg$ forms a better contrast with the proffered content, but Rule (III) computes what's given *independently* of what the proffered content might be). Thus (5b) can be interesting, according to Rule (IV). *Mutatis mutandis* this also explains why (5b″) is coherent.

Bolinger (1972) collected some peculiar cases similar to (32) to challenge syntactic approaches to focus. We can explain some of them.

(32) a. William: What did John do?
#b. Edith: John killed $\underline{\text{someone}}_{\text{H*}\cdot\text{LL\%}}$
b.′ Edith: John $\underline{\text{killed}}_{\text{H*}}$ someone.$_{\text{LL\%}}$
b.″ Edith: John killed a $\underline{\text{policeman}}_{\text{H*}\cdot\text{LL\%}}$
#b.‴ Edith: John $\underline{\text{killed}}_{\text{H*}}$ a policeman.$_{\text{LL\%}}$

Even though *killed someone* and *killed a policeman* are both discourse-new, the former requires focus on *killed* and the latter on *policeman*. In (32b) the presupposed background is *John killed* $x_{\textbf{quantifier}}$. Since *John killed nobody* is a dispreferred response to (32a), it is given that *x* resolves to a nonempty quantifier. Thus *John killed someone* is given and so by Rule (IV) (32b) is anomalous and (32b′) is the correct way to mark interest. However, *John killed a policeman* is not given; thus explaining why (32b‴) is infelicitous requires principles that go beyond interest.

4.2 Beyond Interest

Our necessary condition on interest is at best a small component of a model of interest. At its heart, interest is paralinguistic and subject to individual variation: Bolinger (1985) suggests the Boston Strangler might utter (33a) while a sane individual might describe the situation as (33b):

(33) a. I'm looking for a $\underline{\text{girl}}_{\text{H*}}$ to strangle.$_{\text{LL\%}}$
a. He's looking for a girl to $\underline{\text{strangle}}_{\text{H*}\cdot\text{LL\%}}$

For the Boston Strangler, it's a matter of course he will strangle someone, and *who* exactly he will strangle is his matter of interest. We cannot account for such variations. Nonetheless, there *are* linguistic constraints on what *can* be marked as interesting. Our necessary condition (IV) is one of them. But there are other linguistic constraints on focus that Bolinger was not concerned with. Notably, if there are multiple parts of a constituent that are not given (and so potentially interesting) there is an observed tendency to place focus on the right-most part.

(34) a. Stephen: What did you see?
b. Lucy: I saw a white $\underline{\text{cat}}_{\text{H*}\cdot\text{LL\%}}$
#b.′ Lucy: I saw a $\underline{\text{white}}_{\text{H*}}$ cat.$_{\text{LL\%}}$

Neither 'white' nor 'cat' are given in (34), so at least as far as interest is concerned, it would be legitimate to put focus on either of them. However, there is a clear preference for (34b) over (34b′). This preference is preserved even in the romance languages, where adjectives are to the right of the noun they modify. (35) is an example from Spanish.

(35) a. Esteban: ¿Que viste? *What did you see?*
b. Lucia: Vi un gato $\underline{\text{blanco}}_{\text{H*}\cdot\text{LL\%}}$ *I saw a <u>white</u> cat.*
#b.′Lucia: Vi un $\underline{\text{gato}}_{\text{H*}}$ blanco.$_{\text{LL\%}}$ *I saw a white <u>cat</u>.*

This does not mean that English speakers find nouns interesting while Spanish speakers are fascinated by adjectives. Rather, the two languages' grammars affect the rules for where to place focus *among* the potentially interesting information. Note that both languages forbid the focus marking of given information, in accordance with Rule (IV):

(36) a. Stephen: What kind of cat did you see?
#b. Lucy: I saw a white $\underline{\text{cat}}_{\text{H*}\cdot\text{LL\%}}$
b.′ Lucy: I saw a $\underline{\text{white}}_{\text{H*}}$ cat.$_{\text{LL\%}}$

(37) a. Esteban: Viste algo blanco? *Did you see something white?*
#b. Lucia: Vi un gato $\underline{\text{blanco}}_{\text{H*}\cdot\text{LL\%}}$ *I saw a <u>white</u> cat.*
b.′ Lucia: Vi un $\underline{\text{gato}}_{\text{H*}}$ blanco.$_{\text{LL\%}}$ *I saw a white <u>cat</u>.*

Comparing (34) and (35) can be used to support focus projection: by projecting focus to the head of a noun phrase, both (34b) and (35b) focus *white cat*. However, we do not think that this is necessary. The logical forms that our Focus Semantics assigns to these examples can be paraphrased as follows:

(38) (34b): *L saw something white—it was a cat.*
(35b): *L saw a cat—it was white.*

These logical forms are, to all intents and purposes, equivalent; both semantically and pragmatically (e.g. where accessibility to anaphora are concerned). Thus we claim that our semantics is sufficiently flexible to avoid the need for focus projection. However, this departs from Bolinger's account: we do not, and cannot, say that 'cat' in (34b) or 'blanco' in (35b) is *the* or *the most* interesting constituent. Rather, it is *one* of the constituents that can be marked as interesting, and other principles (that possibly don't relate to interest at all) govern the placement among these.

4.3 Perlocutionary Tune Semantics

Responses (39b) and (39b′), from Ward & Hirschberg (1985), have the same illocutionary force: they are both positive, indirect answers to (39a).

(39) a. William: Did you read the first chapter?
b. Edith: I read the entire dissertation$_{\text{H·LL\%}}$
b.′ Edith: I read the entire dissertation$_{\text{L+H·LH\%}}$
$\rightsquigarrow$ *I thought you knew this already.*

However, intuitively, they carry distinct *perlocutionary effects*: (39b′) carries the implicature we've paraphrased above, while (39b) does not. Similarly, (40b) and (40b′) both deny (40a), but with similar distinct perlocutionary effects:

(40) a. Amy: You can't afford that.
b. Bob: I'm a millionaire$_{\text{H·LL\%}}$
$\not\rightsquigarrow$ *I thought you knew this already*
b.′ Bob: I'm a millionaire$_{\text{L+H·LH\%}}$
$\rightsquigarrow$ *I thought you knew this already*

More generally, tunes give voice to certain *cognitive attitudes* (Liberman & Sag (1974), Ladd (1980) and many others). Here, where we address only falling *vs.* fall–rise tunes, we attribute the following perlocutionary consequences to them (largely in line with our own earlier work (Schlöder & Lascarides 2015) and that of others):

(V) *Tune Semantics*

The falling tune marks a proposition as *informative*.

The fall–rise tune makes a contribution that can be glossed as: what you just said leads me to believe that you don't know what I'm saying now, but I thought you did know what I'm saying now.[11]

In (39) and (40), the distinct perlocutionary consequences of the two tunes don't yield different inferences about the speaker's *illocutionary* act. This is because semantic relationships between the linguistic contents—entailment between *reading the entire dissertation* and *reading its first chapter*, and divergence between *being unable to afford something* and *being a millionaire*—suffice to infer the same specific coherence relations between William's move and Edith's response. In (41ab) and (41ab′), however, where world knowledge supplies no (prior) logical relationship between being a liar and being a fool, both the illocutionary and the perlocutionary effects are different, as discussed by Ladd (1980):

11 If a fall–rise is interpreted discourse-initially, then it can either attach to nonlinguistic antecedents (such as actions indicating particular beliefs) or simply fail to apply for lack of an antecedent.

(41) a. Amy: Harry is the biggest liar in town.
b. Bob: The biggest $\underline{\text{fool}}_{\text{H}}$, maybe.$_{\text{LL\%}}$
b.′ Bob: The biggest $\underline{\text{fool}}_{\text{L+H}}$, maybe.$_{\text{LH\%}}$

In (41ab), Bob is not denying that *Harry is the biggest liar*; he *adds* that *Harry may be the biggest fool* (as well). But in (41ab′), Bob doesn't agree with (41a): he offers the proposition that Harry is the biggest fool *instead*, and moreover implicates something like *before you said what you did, I thought you knew this.* We'll show in Section 7 that the denial move in (41ab′) is derivable from the formalisation of our rules, with the Tune Semantics (Rule (V)) *voicing* what the denial implicates: that Bob believes that believing Harry is a liar normally means you don't believe he's a fool. *Ceteris paribus*, we can give the same analysis of (42),[12] the only difference being that there is already a commonsense relationship between *being good at badminton* and *not being a klutz* (whereas *fool* and *liar* are largely independent).

(42) a. William: Alan's such a klutz.
b. Edith: He's a good at $\underline{\text{badminton}}_{\text{L+H·LH\%}}$

4.4 Summary: Intonated Discourse Update

In sum, our proposed analysis of intonation is as follows:

i. The grammar produces a foreground–background pair $\langle f, \varphi \rangle$, where φ features a free variable x of the same semantic type as f.

ii. From f and φ, the *proffered* and (underspecified) *presupposed* content is computed, according to the Focus Semantics (I, II).

iii. To interpret the utterance, one must first compute the *given* information, defined by the Givenness Rule (III), via the general and independently motivated principles for computing a coherent interpretation of the (prior) discourse context with the (underspecified) presupposed background. The result must be coherent and make the proffered content not given (Rule (IV)).

iv. If all is well, one updates the discourse with both the presupposed and proffered content (again via reasoning about discourse coherence), ensuring that the result is consistent with the coherence relations entailed by the focus semantics (Rules I and II) and the (cognitive) meaning postulates for the tunes (V).

12 (42b) is usually given in the form "He's a good $\underline{\text{badminton}}_{\text{L+H}}$ $\underline{\text{player}}_{\text{L+H·LH\%}}$" To fit our simplifying assumptions, we have transformed it to a single–focus construction.

Steps (i–ii) serve to define the logical form of the utterance given its intonational form (though we forego *deriving* these within the grammar): φ and f are simply computed by λ-abstracting the constituent with the nuclear accent. Step (iii) is a *check* on the felicity conditions of those proposed (underspecified) logical forms. This check makes use of notions related to coherent discourse update, but does not amount to an *actual* update to the current context; this is executed in Step (iv) only if the felicity conditions in (iii) are satisfied.

5 Formal Preliminaries

We now formally regiment the above analysis within Segmented Discourse Representation Theory (SDRT, Asher & Lascarides (2003)). We use SDRT because it has already been used extensively to model the interaction between discourse coherence, presuppositions and cognitive states. Our informal analysis predicts the pragmatic interpretation of an intonated utterance by finding a *maximally coherent* specific interpretation of an underspecified semantics, as derived from linguistic and intonational form. We also use incoherence to predict when intonation is anomalous. We therefore start by giving a brief description of SDRT, and then we'll use it to support the interpretation of intonational meaning as set out by Rules (I–V).

5.1 Discourse Structure

SDRT models discourse structure by connecting the contents of utterances with *rhetorical relations*; e.g., *Narration*, *Elaboration* and *Correction*. Logical forms in SDRT consist of a set of *labels* $\pi_1, \pi_2, \ldots$ that each represent a unit of discourse, and an *assignment function* $\mathcal{F}$ that associates each label π with a formula φ, representing the unit's interpretation. We may write $\mathcal{F}(\pi) = \varphi$ or $\pi : \varphi$ to express this mapping. The content φ assigned to a label can consist of rhetorical relations among labels, and so the labels form a *partial order*: π_1 immediately outscopes π_2 if $R(\pi_1, \pi_3)$ or $R(\pi_3, \pi_1)$ is a part of the formula φ where $\mathcal{F}(\pi_1) = \varphi$. A *coherent* logical form—known as a Segmented Discourse Representation Structure or SDRS—has a unique root under this partial order: in other words a coherent discourse consists of a single segment of rhetorically connected subsegments.

Cue phrases (e.g. *then*, *therefore*, *but*) can entail coherence relations, but frequently they're inferred via commonsense reasoning with linguistic and non-linguistic information. Even so, ambiguity can persist: simplifying somewhat (we ignore presuppositions, tense and so on), (43′) disambiguates (43) to an interpretation that is equivalent to having *therefore* connecting the clauses; (43″) is equivalent to *because*.

(43) The meeting is cancelled. Nicholas stayed at home.

(43′) $\pi_0 : Result(\pi_1, \pi_2)$
$\pi_1 : (\iota x)(meeting(x) \wedge cancel(e_1, x))$
$\pi_2 : (\iota y)(stay(e_2, n) \wedge home(y) \wedge at(e_2, y))$

(43″) $\pi_0 : Explanation(\pi_1, \pi_2)$
$\pi_1 : (\iota x)(meeting(x) \wedge cancel(e_1, x))$
$\pi_2 : (\iota y)(stay(e_2, n) \wedge home(y) \wedge at(e_2, y))$

SDRSs are assigned a dynamic semantics, where one starts to unpack its content from its (unique) root label. The dynamic semantics of a rhetorical connection $R(\pi_1, \pi_2)$ is defined in terms of its arguments' contents (i.e., $\mathcal{F}(\pi_1)$ and $\mathcal{F}(\pi_2)$). For instance, the general rubric for *veridical* relations like *Explanation* and *Result* is given in (44), where $\mathcal{C}$ and $\mathcal{C}'$ are the contexts of interpretation (typically, sets of world-assignment pairs), $\wedge$ corresponds to *dynamic conjunction* (and so $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \circ [\![\psi]\!]$), and $\varphi_{R(\pi_1,\pi_2)}$ is content that is specific to the rhetorical relation R and is specified in terms of $\mathcal{F}(\pi_1)$ and $\mathcal{F}(\pi_2)$:

(44) $\mathcal{C}[\![R(\pi_1, \pi_2)]\!]\mathcal{C}'$ iff $\mathcal{C}[\![\mathcal{F}(\pi_1) \wedge \mathcal{F}(\pi_2) \wedge \varphi_{R(\pi_1,\pi_2)}]\!]\mathcal{C}'$.

For example, *Background* is a veridical relation, where $\varphi_{Background(\pi_1,\pi_2)}$ is equivalent to the condition that the event e_1 described by π_1 spatially and temporally overlaps the event e_2 described by π_2. Note that $R(\pi_1, \pi_2)$ will trigger changes to the input context whenever the contents of its arguments do (e.g., when $\mathcal{F}(\pi_1) = \exists x \varphi$).

5.2 Construction of Logical Form and Maximising Discourse Coherence

Logical form and its dynamic semantics capture how to evaluate a *representation* of the discourse against the real world or a model. But *constructing* which logical form is the intended interpretation is *defeasible*, and this task is carried out in a separate but related *glue logic* (Asher & Lascarides (2003) provide detailed motivation for this separation).

The glue logic consists of *default axioms* that model how commonsense reasoning with both linguistic and non-linguistic information yield (defeasible) inferences about which available unit(s) in the context the current unit connects to, which coherence relations connect them, and how other semantic elements that are left underspecified by linguistic form get resolved to specific values (e.g., the relative semantic scope of presuppositions and pronominal reference). The glue logic thus reasons over *underspecified logical forms* (ULFs), which in turn express *partial descriptions* of fully specific logical forms—i.e., SDRSs. The default axioms thus support defeasible inferences from ULFs to more specific and pragmatically preferred ULFs.

Since the glue logic reasons with ULFs, its language can underspecify semantic scope in the usual way: i.e., each predication is assigned a label ($l_1, l_2 \ldots$), which denotes the scopal position of that predication in the (fully specific) SDRS, and (partial) constraints on their relative scope are expressed with the formula $l_1 \succeq l_2$. Pronouns introduce a condition $x =?$, which means that x must be co-referent with an available antecedent, but *exactly which* antecedent isn't known. More generally, a ULF uses a variable ? of an appropriate sort whenever the specific value of the constructor in the (fully specific) SDRS that the ULF describes isn't known. For instance, $\lambda :?(\alpha, \beta)$ means that β is connected to α, forming part of the segment labelled by λ, but the value of their coherence relation isn't known.

The glue logic's *default axioms* are expressed with a defeasible conditional $>$ ($\varphi > \psi$ means *If* φ *then normally* ψ). For example, `IQAP` is a glue-logic default axiom which stipulates that normally, a response to a question is an indirect answer (IQAP stands for *Indirect Question Answer Pair*):

$$(\texttt{IQAP})\ (\lambda :?(\alpha, \beta) \wedge \mathsf{interrogative}(\alpha) \wedge \mathsf{spk}(\alpha) \neq \mathsf{spk}(\beta)) > \lambda : \mathit{IQAP}(\alpha, \beta)$$

In words, if β is rhetorically connected to α but we don't (yet) know with what coherence relation, α is an interrogative and α and β are said by different people, then normally, the relation between them is *IQAP*. Another general default axiom stipulates that when the contents associated with α and β satisfy the necessary semantic consequences of $R(\alpha, \beta)$, then normally they connect with R (Asher & Lascarides 2003: p403).

The conditional $>$ defines a nonmonotonic proof theory $\mid\!\sim_g$ (Asher & Lascarides 2003: ch5), which validates a number of intuitively compelling patterns of defeasible inference, such as Defeasible Modus Ponens ($\varphi, \varphi > \psi \mid\!\sim_g \psi$) and Specificity (If $\varphi \vdash_g \varphi'$ then $\varphi, \varphi > \neg\psi, \varphi' > \psi \mid\!\sim_g \neg\psi$).

In addition to the default axioms, one of SDRT's most important principles for LF construction is in fact *monotonic*: one (always) interprets the discourse in a way that *Maximises Discourse Coherence* (MDC). For instance, Asher & Lascarides (1998) use MDC to predict when a presupposition gets locally accommodated, overriding the default that it projects from its syntactic embedding. As we have implied throughout this paper, discourse coherence is not a yes/no matter; it can vary in quality. SDRT's principle MDC defines factors that affect that quality. Roughly put, they are as follows (formal details are in Asher & Lascarides (2003: p233)):

Maximise Discourse Coherence (MDC).

> Suppose that one is updating the logical form of a discourse context c with new information σ. Then the SDRS K that results from this update operation must be *maximally coherent*. An SDRS K is at least as coherent as an SDRS K', $K' \leq^c K$, if and only if all of the following hold:

i. If K' is consistent, then so is K.

ii. *Prefer rich structure:* K has at least as many coherence relations as K'.

iii. *Prefer flat structure:* K has at most as many labels as K' unless K' has a *semantic clash* and K does not. The SDRS (45c) is an example of a semantic clash, because the content of π_2 is made veridical by *Parallel* but non-veridical by the *If* (-then) relation; (45b) is thus more coherent (despite having more labels).

(45) a. π_1: If a shepherd goes to the mountains,
π_2: he normally brings his dog.
π_3: He brings a good walking stick too.
b. π_0: $If(\pi_1, \pi)$
π : $Parallel(\pi_2, \pi_3)$
c. π_0: $If(\pi_1, \pi_2) \wedge Parallel(\pi_2, \pi_3)$

iv. *Prefer better relations:* Each rhetorical connection in K is at least as coherent as those in K'. Note that we have talked of a *Contrast* connection varying in quality: some contrasts are better than others. Formally, the more the contrasting contents are structurally isomorphic and the more the parts in the isomorphic mapping are semantically distinct, the better the contrast. Similarly, a *Continuation* is better the more specific the common topic (or generalisation) of the contents it relates are.

v. *Prefer resolution:* K resolves (as computed by dynamic update through the coherence relations) at least as many underspecifications as K' does.

While MDC constrains interpretations to be maximally coherent, it does not entail that there is a *unique* maximally coherent interpretation, even in context. This allows for misunderstandings to surface in the dialogue. For instance, MDC does not distinguish between the two alternative SDRSs (43′) and (43″) of (43). But a subsequent utterance may serve to resolve the ambiguity: *He had no other reason to come in* favours (43′) and *It makes no sense to meet without him* favours (43″).

The glue logic proof theory $\mid\sim_g$ defines Discourse Update:

Definition 1 (Update)**.** Let Γ be a ULF for the discourse context and $\pi : \mathcal{K}$ a ULF representing new information. Then $update(\Gamma, \pi : \mathcal{K})$ is the set of all (and only) those SDRSs that satisfy the glue logic consequences of attaching $\mathcal{K}$ to some available segment α in Γ. More formally: $K \in update(\Gamma, \pi : \mathcal{K})$ iff K is an SDRS and there is an available segment α in Γ where for all formulae φ

$$\text{If } \Gamma, \pi : \mathcal{K}, \lambda :?(\alpha, \pi) \mid\sim_g \varphi, \text{ then } K \models_g \varphi.$$

The static glue logic from Asher & Lascarides (2003) doesn't axiomatise MDC, and so the fully specific SDRSs in $update(\Gamma, \mathcal{K})$ are ordered via the logically extraneous MDC. In its dynamic version (Asher & Lascarides 2011), MDC is axiomatised within the glue logic itself. Either way, we will assume from now on that all SDRSs in a discourse update are maximally coherent, as defined by MDC.

5.3 Presuppositions

Following van der Sandt (1992), SDRT assumes that the linguistic grammar derives a ULF in which proffered content is separated from presupposed content and their relative semantic scope is underspecified. For instance, in both theories the presupposition trigger *regret* yields the logical form for (46a) given in (46b):

(46) a. A man didn't regret smoking.
b. proffered: $\pi_1 : \exists x(man(x) \wedge \neg regret(e, x, {}^{\wedge}smoke(e', x))$
presupposed: $\pi_2 : smoke(e', x)$

The glue logic axioms must then validate how the presupposed and proffered contents coherently relate to their context (and each other). These axioms ensure that presupposed vs. proffered components are treated differently. First, the utterance's presuppositions update the discourse before its proffered content. Secondly, the glue logic axioms make presuppositions (defeasibly) *bind* to a prior unit over coherently relating to one; and in the case where binding isn't consistent, the presupposition is (defeasibly) coherently related to an available unit in the context that outscopes other units (proviso the result satisfies the monotonic axiom MDC). These defaults don't apply to proffered content.

For example, simplifying somewhat (e.g., only the ULF labels that contribute to anaphoric or semantic scope ambiguities are shown), the grammar generates for the two sentences in (47a) the ULFs (47b) and (47c) (where presupposed content is marked with ∂, following Beaver (1997)).

(47) a. A man had a health scare. But he didn't regret smoking.
b. $\pi_1 : \exists x \exists y(man(x) \wedge health\text{-}scare(y) \wedge have(e_1, y, x))$
c. $\pi_2 : Contrast(\pi, \pi_3), \pi = ?$
$\pi_3 : \neg regret(e_2, z, {}^{\wedge}smoke(e_3, z))$
$\pi_4 : \partial smoke(e_3, z), z = ?$
d. $\pi_2 : Background(\pi_1, \pi_4) \wedge Contrast(\pi_1, \pi_3)$

Sentence-initial *but* introduces a *Contrast* relation whose first argument is anaphoric ($\pi = ?$). Given number and gender constraints (which we have omitted here), the only candidate for resolving $z = ?$ is $z = x$. So by MDC, x must be made available, which means that the presupposition π_4 must connect to π_1. Both these discourse

units describe *states*, and so the glue logic axioms validate a (defeasible) inference that they connect as $Background(\pi_1, \pi_4)$. MDC then predicts that $\pi =?$ resolves to $\pi = \pi_1$: this forms a better quality *Contrast* (and a flatter structure) than the alternative (i.e., $Contrast(\pi, \pi_3)$ where $\pi : Background(\pi_1, \pi_4)$). So the final SDRS is as shown in (47d): this entails that smoking occurred, even though *smoking* was syntactically outscoped by *not*.

5.4 Cognitive Modelling Logic

SDRT's glue logic inferences are additionally aided and constrained by information about the participants' *cognitive states*. The *cognitive logic* features a number of modal operators: KD45 operators for belief (B_S for a speaker S); K45 operators for public commitment (P_S); and higher-order operators for intentions (I_S, Schlöder et al. (2017)). Furthermore, there are action operators $[s_S(\pi)]$ ('after S uttered the discourse unit π') and $[s_S(\pi)]^{-1}$ ('before S uttered π'). These operators are not *updates* in the cognitive logic; rather they are *accesses* to different states (at different times) in the cognitive model. So we remain with a static modal model theory.

The cognitive logic contributes to modelling the interaction between locutionary, illocutionary and perlocutionary effects. Following Hamblin (1970), SDRT makes a speaker S who conveys a message φ publicly committed to φ (Lascarides & Asher 2009). Using the above operators, we can formalise some basic pragmatic principles:

Sincerity. (a) $P_S\varphi > B_S\varphi$.
(b) $B_S\neg\varphi > \neg I_S P_S\varphi$.
Normally, you believe what you commit to; and you don't intend to commit to what you don't believe.

Intention Transfer. $P_S\varphi > P_S I_S P_H\varphi$.
Normally, you intend to make your commitments shared.

Cooperativity. $P_S I_S\varphi > I_H\varphi$.
Normally, intentions are kept aligned.

Sincere Questions. (a) $\mathsf{interrogative}(\varphi) \rightarrow (P_S\varphi > \neg B_S\mathsf{resolved}(\varphi))$.
(b) $\mathsf{interrogative}(\varphi) \rightarrow (B_S\mathsf{resolved}(\varphi) > \neg I_S P_S\varphi)$.
Normally, questions sincerely ask for unknown information.

To make these axioms have traction when interpreting discourse, glue-logic inferences about the semantic representation of the discourse get transferred into the cognitive logic as public commitments:

Commitment. Let $\pi_1 \ldots \pi_n$ be elementary discourse units spoken by $S_1 \ldots S_n$, 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 $\Gamma_n \vdash_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] P_{S_n} \varphi$.
– If $\Gamma_n \mid\!\sim_G \varphi$, then $\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] P_{S_n} \varphi$.

A simple application is the situation in which a speaker S makes an assertion to an addressee H. Suppose the glue logic can resolve the speaker's signal to the (maximally coherent) logical form $\pi : \mathcal{K}_\pi$. Then *Commitment* transfers this to the cognitive modelling logic as $[s_S(\pi)] P_S \mathcal{K}_\pi$. By *Sincerity (a)*, H can now (defeasibly) infer that *S* believes that $\mathcal{K}_\pi$. Furthermore, $[s_S(\pi)] P_S I_S P_H \mathcal{K}_\pi$ follows by *Intention Transfer* (i.e., S wants H to agree with $\mathcal{K}_\pi$). Finally, by *Cooperativity*, the speaker S can infer that H will (cooperatively) do so, $[s_S(\pi)] I_H P_H \mathcal{K}_\pi$, *unless* this inference is blocked by other information. For instance, if it is known that H believes that $\mathcal{K}_\pi$ is false then *Sincerity (b)* would block the *Cooperativity* inference (when the antecedents of two default axioms are satisfied but these antecedents aren't logically related and their consequences are contradictory, then $\mid\!\sim_g$ validates *neither* consequence).

The action operators can formally express *surprise*: $[s_H(\pi)]^{-1} B_S \neg I_H P_H \mathcal{K}_\pi$ expresses that *before* H's speech act π, S thought that H didn't intend to commit to $\mathcal{K}_\pi$. That is, H's action defied S's prior expectations. The meaning of these operators is further specified by the following axioms (Schlöder & Lascarides 2015).

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 $\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_n}(\pi_n)] [s_{S_i}(\pi_i)]^{-1} B_S \varphi$,
then $\Gamma_n \mid\!\sim_C [s_{S_1}(\pi_1)] \ldots [s_{S_{i-1}}(\pi_{i-1})] B_S \varphi$
'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 normally have that belief in foresight/hindsight.

These axioms ensure that inferences about previous dialogue states work correctly. Note that the context Γ_n in *Hindsight* does not change. It models hindsight inferences that interlocutors can make *about* previous cognitive states *from* their current knowledge Γ_n, which *extends* their prior knowledge Γ_{i-1}. That is, the axiom may apply

in Γ_n, but $\Gamma_{i-1} \not\sim_C [s_{S_1}(\pi)] \ldots [s_{S_{i-1}}(\pi_{i-1})] B_S \varphi$. Thus a speaker A can derive that their interlocutor B must have held a particular belief earlier on *without* being aware of that belief at the time. For instance, Bob's contribution in (48), if it expresses surprise, triggers the inferences from that surprise given below:

(48) α: Amy: I want you to come to the ball with me.
π: Bob: Really?!

$[s_A(\alpha)][s_B(\pi)] P_B [s_A(\alpha)]^{-1} B_B \neg I_A P_A \mathcal{K}_\alpha$.

$\sim$ $[s_A(\alpha)][s_B(\pi)] B_B [s_A(\alpha)]^{-1} B_B \neg I_A P_A \mathcal{K}_\alpha$ (Sincerity a).

$\sim$ $[s_A(\alpha)][s_B(\pi)][s_A(\alpha)]^{-1} B_B B_B \neg I_A P_A \mathcal{K}_\alpha$ (Reduction).

$\sim$ $B_B B_B \neg I_A P_A \mathcal{K}_\alpha$ (Hindsight).

$\sim$ $B_B \neg I_A P_A \mathcal{K}_\alpha$ (Belief modal).

$\rightsquigarrow$ "At the beginning of the dialogue, Bob thought Amy wouldn't say that!"

6 Intonation in Discourse

With the formal preliminaries in place, we now formalise our model from Section 3. Definition 2 formalises the Focus Semantics (I, II).

Definition 2 (Focus Semantics)**.** Let $\langle \varphi, f \rangle$ be the foreground–background pair of the current utterance. The discourse update associated with $\langle \varphi, f \rangle$ is an update with $\pi^b : \partial \mathcal{K}_{\pi^b}$, $\pi^f : \mathcal{K}_{\pi^f}$ and $\pi : R^{tune}(\pi^b, \pi^f)$, where:

- $\mathcal{K}_{\pi^f}$ is the ULF corresponding to $\varphi(f)$.
- If the tune is falling, then $R^{tune} = \mathit{Continuation} \vee \mathit{Elaboration}$ and $\mathcal{K}_{\pi^b}$ is the ULF corresponding to $\varphi(x)$, where x is free in φ and of the same type as f, and the semantic index (an eventuality term) of φ is syntactically distinct in $\mathcal{K}_{\pi^b}$ and $\mathcal{K}_{\pi^f}$ (although they can denote the same eventuality).
- If the tune is fall–rise, then $R^{tune} = \mathit{Contrast}$ and $\mathcal{K}_{\pi^b}$ is the ULF corresponding to $?_{mod}(\varphi(x))$, where x is free in φ and of the same type as f, the semantic index (an eventuality term) of φ is syntactically (but not necessarily semantically) distinct in $\mathcal{K}_{\pi^b}$ and $\mathcal{K}_{\pi^f}$, and $?_{mod}$ underspecifies *modality*—that is, it can resolve to $\top$, $\Diamond$ or $\neg$.

Givenness (Rule III) corresponds to updating the discourse context with the (presupposed) background content, which in SDRT is the maximally coherent interpretation of their combination. Accordingly, Interest (Rule IV) is simply that the foreground isn't entailed by this.

Definition 3 (Givenness and Interest)**.** The *given information* is what is inferable from updating the discourse context Γ with the (presupposed) ULF $\pi^b : \partial\mathcal{K}_{\pi^b}$ representing the background content (in a maximally coherent way). That is, ϕ is given iff for any (fully specific) SDRS $K \in update(\Gamma, \pi^b : \mathcal{K}_{\pi^b})$, $K \models_g \phi$.

The proffered content $\mathcal{K}_{\pi^f}$ (which is $\varphi(f)$) is *interesting* only if it is not given. I.e., there is a $K \in update(\Gamma, \pi^b : \mathcal{K}_{\pi^b})$ such that $K \not\models_g \mathcal{K}_{\pi^f}$.

The cognitive effects of the tunes (Rule V) is expressed in SDRT's cognitive logic. We formalise the constraint that the foreground of a falling tune is *informative* as the speaker's belief that this content is not yet mutually accepted. Following Schlöder & Lascarides (2015), we formalise the cognitive contribution of fall–rise as follows: the attachment point α of either the foreground or the entire *Contrast* segment is taken to be the utterance that the current speaker S is marking as triggering a change in S's beliefs about H. The hindsight formula expresses that before α, S did not believe that H wouldn't know $\mathcal{K}_{\pi^f}$, but that α changed that belief. The antecedent of the term for fall–rise intonation finds the right attachment point α, and the consequent then defines the belief-change triggered by α.

Definition 4 (Tune Semantics)**.** Where S utters an intonational phrase with proffered content π^f, the different tunes impose the following restrictions in the cognitive logic, where cg stands for *it is common ground that* (Asher & Lascarides 2008).

fall: $C^{tune} = B_S \neg \mathrm{cg}(K_{\pi^f})$.

fall–rise: $C^{tune} = \big((\lambda : ?(\alpha, \pi^f) \vee (\lambda : ?(\alpha, \pi) \wedge \pi : Contrast(\pi^b, \pi^f))) \wedge S(\alpha) = H\big) \rightarrow P_S\big(([s_H(\alpha)]^{-1} \neg B_S \neg B_H \mathcal{K}_{\pi^f}) \wedge (B_S \neg B_H \mathcal{K}_{\pi^f})\big)$.

Definition 5 combines these as specified in Section 4.4:

Definition 5 (Intonated Discourse Update)**.** Let Γ be the prior context and $\langle \varphi, f \rangle$ be the background–foreground pair of the current utterance.

i. Compute the ULFs $\mathcal{K}_{\pi^f}$ and $\mathcal{K}_{\pi^b}$ as in Definition 2.

ii. If the glue logic supports no coherence relation between $\mathcal{K}_{\pi^b}$ and an available unit in Γ, break.

iii. If $\mathcal{K}_{\pi^f}$ isn't *interesting*, (i.e., not given, Definition 3), break.

iv. Add C^{tune} (Definition 4) to the cognitive logic.

v. Do discourse update on Γ with $\pi^b : \partial K_{\pi^b}$, $\pi^f : K_{\pi^f}$ and R^{tune}.

The constraints in (ii–iv) can all be expressed as monotonic axioms in the glue logic: they are then effectively a part of the *lexical semantics of tunes*. We can then understand these semantics as affecting the standard SDRT-update (Definition 1). For the sake of simplicity, however, we use Definition 5 to compute discourse update (without explicitly reducing it to Definition 1).

7 Some Formal Analyses

Let's start with a formal analysis of the distinct illocutionary and perlocutionary effects of (41ab) vs. (41ab′).

(41) a. Amy: Harry is the biggest liar in town.
b. Bob: The biggest $\underline{\text{fool}}_{\text{H}}$, maybe.$_{\text{LL\%}}$
b.′ Bob: The biggest $\underline{\text{fool}}_{\text{L+H}}$, maybe.$_{\text{LH\%}}$

Simplifying somewhat (in that we treat "maybe" as equivalent to $\Diamond$, omit tense information, treat subject pro-drop like a pronoun and considerably over-simplify the semantics of "biggest X in town") the ULF of (41a) is α and the Focus Semantics (Definition 2) yields the ULFs (49b) and (49b′) for (41b) and (41b′):

(49) a. $\alpha : liar(e_\alpha, h) \wedge biggest(e_\alpha)$ *Harry is the biggest liar.*
b. $\pi : ?_{fall}(\pi^b, \pi^f)$.
$\pi^b : \partial\Diamond(P(e'_\beta, x) \wedge biggest(e'_\beta)) \wedge P = ?$ *he is maybe the biggest P.*
$\pi^f : \Diamond(fool(e_\beta, x) \wedge biggest(e_\beta))$ *he is maybe the biggest fool.*
b′. $\pi' : Contrast(\pi'^b, \pi'^f)$.
$\pi'^b : \partial ?_{mod}\Diamond(P(e'_\beta, x) \wedge biggest(e'_\beta))$ *y(he is maybe the biggest P).*
$\wedge P = ?$
$\pi'^f : \Diamond(fool(e_\beta, x) \wedge biggest(e_\beta))$ *he is maybe the biggest fool.*

In (49ab), one must first ensure that π^b can coherently update the context α to yield given information that doesn't entail π^f. MDC prefers binding P to the available antecedent *liar* (and x to the available antecedent Harry) rather than resolving P via existential quantification. Thus the proffered content isn't given and discourse update can proceed. First, resolving P to *liar* validates the relation *Accept* between α and π^b. Then, the underspecified relation $?_{fall}$ that's introduced by the tune semantics resolves to *Continuation*: the glue logic axioms don't validate inferring *Elaboration* ('liar' and 'fool' are conventionally not in an entailment relation) but do validate *Continuation* (because π^b and π^f share a common topic). Since flat structures are preferred by MDC, the final discourse structure is: $\pi_0 : Accept(\alpha, \pi^b) \wedge Continuation(\pi^b, \pi^f)$. Its dynamic semantics entail that Bob is committed to Harry being the biggest liar and (also) maybe the biggest fool. In addition, given the tune, the cognitive logic derives

$B_B\neg cg(\mathcal{K}_{\pi^f})$: Bob believes that *Harry is maybe the biggest fool* is not (yet) common ground.

Now consider (41ab′), where the ULFs are α and (49b′). We again start by computing how (just) the presupposed background content π'^b would update α. As with (49ab), binding P to the available antecedent *liar* and binding x to Harry is preferred (via MDC). The underspecified modality $?_{mod}$ can resolve to $\top$, $\Diamond$ or $\neg$. Any of these resolutions produce a coherent update, and render the proffered content π'^f not given. So the example passes the Interest tests. However, when updating the context with both π'^b and π'^f a clearly preferred resolution of the modality emerges via MDC.

For simplicity, we'll assume that $\Diamond\Diamond\phi$ is equivalent to $\Diamond\phi$ and $\neg\Diamond$ is $\Box\neg$; so resolving $?_{mod}$ to $\top$ or to $\Diamond$ are equivalent, and resolving it to $\neg$ leads to a 'definitely not' reading. Then we can see that $?_{mod} = \neg$ maximises contrast, for compare:

(50) a. *Harry is definitely not the biggest liar, but he is maybe the biggest fool*
b. *Harry is maybe the biggest liar, but (also) maybe the biggest fool.*

Saying what Harry is *vs*. what he is *not* yields a better contrast then contrasting two things that he is. So by MDC, $?_{mod}$ resolves to $\neg$. As a consequence, the contents of α and π'^b satisfy the necessary consequences of *Correction* (for the latter entails the negation of the former). So SDRT's glue logic yields the logical form (51).

(51) $\pi_0 : Correction(\alpha, \pi'^b) \wedge Contrast(\pi'^b, \pi'^f)$.
$\alpha : liar(e_\alpha, h) \wedge biggest(e_\alpha)$ *Harry is the biggest liar.*
$\pi'^b : \neg\Diamond(liar(e_\alpha, h) \wedge biggest(e_\alpha))$ *Harry is not the biggest liar.*
$\pi'^f : fool(e_\beta, h) \wedge biggest(e_\beta)$ *Harry is the biggest fool.*

There is no commonsense contrast between liar and fool. However, the perlocutionary semantics entails that Bob sees exactly such a contrast. This derivation in the cognitive logic is as follows (see also Schlöder & Lascarides (2015)). First, there is a (hindsight) derivation from Bob's utterance to Bob's beliefs about Amy before Amy said (41a):

$\vdash [s_A(\alpha)][s_B(\pi)]P_B[s_A(\alpha)]^{-1}\neg B_B\neg B_A\,\mathcal{K}_\pi$. ($C^{fallrise}$).

$\mid\!\sim [s_A(\alpha)][s_B(\pi)]B_B[s_A(\alpha)]^{-1}\neg B_B\neg B_A\,\mathcal{K}_\pi$. (Sincerity a).

$\mid\!\sim [s_A(\alpha)][s_B(\pi)][s_A(\alpha)]^{-1}B_B\neg B_B\neg B_A\,\mathcal{K}_\pi$. (Reduction).

$\mid\!\sim B_B\neg B_B\neg B_A\,\mathcal{K}_\pi$. (Hindsight).

$\mid\!\sim \neg B_B\neg B_A\,\mathcal{K}_\pi$. ($B_B$ is KD45).

Second, there is a derivation about how that belief changes, given Amy's move:

$\vdash\ [s_A(\alpha)][s_B(\pi)](B_B\neg B_A\,\mathcal{K}_\pi)\ (C^{fallrise})$.

$\mid\!\sim\ [s_A(\alpha)]\big((B_B\neg B_A\,\mathcal{K}_\pi \vee B_B(P_B\,\mathcal{K}_\pi > \neg B_A\,\mathcal{K}_\pi)\big)$. (Conservativity).

$\mid\!\sim\ [s_A(\alpha)](B_B\neg B_A\,\mathcal{K}_\pi)$. ($\vee$-Elimination).
The second disjunct is (normally) false, as $P_B\,\mathcal{K}_\pi \mid\!\sim I_A P_A\,\mathcal{K}_\pi$ by Cooperativity and Intention Transfer and so $P_B\,\mathcal{K}_\pi > \neg B_A\,\mathcal{K}_\pi$ clashes with Sincerity b.

$\mid\!\sim\ (B_B\neg B_A\,\mathcal{K}_\pi \vee B_B(P_A\,\mathcal{K}_\alpha > \neg B_A\,\mathcal{K}_\pi)\big)$. (Conservativity).

$\mid\!\sim\ B_B(P_A\,\mathcal{K}_\alpha > \neg B_A\,\mathcal{K}_\pi)$ ($\vee$-Elimination, given the above derivation).

$\approx$ "That you told me α tells me that you don't know π."

Thus, by way of intonation, Bob conveys at the cognitive level a strong contrast between *liar* and *fool*, which makes the *Contrast* relation in (51) even more coherent (for it is supported by the tune's implicature, that Bob believes a commitment to Harry being a liar normally entails you don't believe he's a fool). This contrast is *absent* in (41ab)—the focus semantics doesn't entail *Contrast*, nor does the tune carry the 'surprise'-type perlocutionary effects that are associated with fall–rise intonation.

In (40), the tune changes the perlocutionary effects but not the illocutionary ones.

(40) a. Amy: You can't afford that.
b. Bob: I'm a $\underline{\text{millionaire}}_{\text{H·LL\%}}$
$\not\rightsquigarrow$ *I thought you knew this already*
b.′ Bob: I'm a $\underline{\text{millionaire}}_{\text{L+H·LH\%}}$
$\rightsquigarrow$ *I thought you knew this already*

The glue logic (defeasibly) yields *Correction* in both (40ab) and (40ab′), on the grounds that the contents of the sentences satisfy *Correction*'s necessary consequences. The perlocutionary derivation is slightly different from the above, however. We can make the same Hindsight derivations from the fall–rise tune:

$\mid\!\sim\ (\neg B_B\neg B_A\,\mathcal{K}_\pi) \wedge (B_B(P_A\,\mathcal{K}_\alpha > \neg B_A\,\mathcal{K}_\pi))$.

However in contrast to (41), in (40) the second conjunct is already given, since *cannot afford* and *millionaire* are, conventionally, contrary. Thus a commitment to *cannot afford* normally means that one does not believe *millionaire*: i.e., $P_A\,\mathcal{K}_\alpha > \neg B_A\,\mathcal{K}_\pi$ is derivable independently of the tune. However, we can reason further with the first conjunct that's derivable from the tune, by including the fact that millionaires can afford things:

$\vdash\!\sim$ $\neg B_B \neg B_A \mathcal{K}_\pi$ ($\wedge$-elimination from the above derivation).

$\vdash\!\sim$ $\neg B_B \neg B_A \neg \mathcal{K}_\alpha$. (by $\Box(\mathcal{K}_\pi \rightarrow \neg \mathcal{K}_\alpha)$ and belief is KD45).

$\vdash\!\sim$ $\neg B_B \neg\neg I_A P_A \mathcal{K}_\alpha$ (Sincerity b and belief is KD45).

$\vdash\!\sim$ $\neg B_B I_A P_A \mathcal{K}_\alpha$ (double negation).

$\approx$ "I didn't believe you would commit to α."

Thus, we obtain the desired perlocutionary effect of the fall–rise intonation in (40ab′). Similar computations can be made for (39ab′) using *Sincere Questions (b)* instead of *Sincerity (b)*.

(39) a. Amy: Did you read the first chapter?
b.′ Bob: I read the entire dissertation$_{\text{L+H·LH\%}}$
$\rightsquigarrow$ *I thought you knew this already.*

Finally, we show how to predict felicity judgements and implicatures in the context of *wh*-questions.

(3) a. Harvey: Who likes Michael?
b.′ Jessica: Nobody likes Michael$_{\text{L+H·LH\%}}$

Definition 2 yields the ULF of (3b′) given by π, π^b and π^f below, and these must update Harvey's move α (where we have already resolved the presupposition triggered by the *wh*-question as shown):

(52)
α : $Background(\alpha^b, \alpha^f)$.
α^f : $?\lambda x.like(x,m)$ — *Who likes Michael?*
α^b : $\exists y.like(y,m)$ — *someone likes Michael*
π : $Contrast(\pi^b, \pi^f)$.
π^b : $\partial ?_{mod}(\neg\exists z.like(z,x)) \wedge x = ?$ — *y(nobody likes x)*
π^f : $\neg\exists z.like(z,m)$ — *nobody likes Michael*

We must first check that proffered content isn't given (Definition 3). The glue logic predicts that π^b *binds* to α^b (on the grounds that binding presuppositions is maximally coherent), thereby setting $?_{mod} = \neg$ and $x = m$. So interest is satisfied—$\mathcal{K}_{\pi^f}$ isn't entailed by this result. But $x = m$ *cannot* be a part of the *final* update with π^f, whatever the resolution of $?_{mod}$: (53c) is inconsistent with π^f, and the resolutions in (53a) and (53b) are entailed by π^f and therefore fail to establish a contrast.

(53) a.$?_{mod} = \top \rightsquigarrow \pi^b$: *nobody likes Michael*.
b.$?_{mod} = \Diamond \rightsquigarrow \pi^b$: *possibly, nobody likes Michael*.
c.$?_{mod} = \neg \rightsquigarrow \pi^b$: *somebody likes Michael*.

So when executing the update of the context with both the background and proffered contents, the final logical form must *accommodate x* via an existential quantifier (rather than binding it to *m*). In the usual ways, we can see that resolving $?_{mod}$ to $\neg$ maximises the contrast with π^f (*someone likes someone (other than Michael), but nobody likes Michael*), to yield (54):

(54) $\pi_0 : \mathit{Background}(\alpha^f, \pi^b) \wedge \mathit{Contrast}(\pi^b, \pi^f) \wedge$
$\mathit{Correction}(\alpha^b, \pi^f) \wedge \mathit{QAP}(\alpha^f, \pi^f)$

$\alpha^f : ?\lambda z.\mathit{like}(z,m)$	*Who likes Michael?*
$\alpha^b : \exists z.\mathit{like}(x,m)$	*someone likes Michael.*
$\pi^b : \exists x.\exists z.\mathit{like}(z,x)$	*there is someone that somebody likes.*
$\pi^f : \neg\exists z.\mathit{like}(z,m)$	*nobody likes Michael.*

The discourse *structure* is one where Jessica is answering Harvey's question but correcting its presupposition. Note that due to the resolution of π^b Jessica tacitly acknowledges that there was a liking, but she denies that anyone likes Michael. If (3b′) were uttered with a falling tune, then the background ULF is $\partial\neg\exists y\mathit{like}(y,x)$, and the most coherent way to interpret this sets $x = m$, so that it corrects α^b. But this entails the proffered content, and so by Definition 3 it is anomalous. In other words, with stress on "Michael", one tune successfully voices the denial of the question's presupposition and the other tune doesn't.

Finally, we analyse the infelicity judgement in (1).

(1) a. Harvey: Who likes Michael?
#b.′ Jessica: Rachel likes $\underline{\text{Michael}}_{\text{H·LL\%}}$

As before, the question (1a) yields the SDRS rooted at α in (55); (1b′) yields the ULFs involving π, π^b and π^f:

(55) α : $\mathit{Background}(\alpha^b, \alpha^f)$
α^b : $\exists x.\mathit{like}(e_{\pi_1}, x, m)$
α^f : $?\lambda y.\mathit{like}(e_{\pi_1}, y, m)$
π : $?_{\mathit{fall}}(\pi^b, \pi^f)$
π^b : $\partial\mathit{like}(e'_\beta, r, z) \wedge z = ?$
π^f : $\mathit{like}(e_\beta, r, m)$

As before, the Interest test amounts to examining the most coherent way of interpreting π^b (alone) in the context of α, and checking that this interpretation is coherent and does not entail π^f. Here, the most coherent update resolves z to m, as this can attach as an answer to α^f. So the Interest test fails: the most coherent interpretation of the background entails the foreground. Thus, the intonation is anomalous.

If Jessica's response has a fall–rise tune, then the ULFs are as above except that π^b features the underspecified modality $?_{mod}$ and *Contrast* replaces $?_{\mathit{fall}}$. The most

coherent update of α with π^b still resolves z to m and $?_{mod}$ to $\top$ because this results in an answer to (1a), whereas $?_{mod} = \neg$ or $?_{mod} = \Diamond$ provide only *partial* answers (SDRT makes complete answers more coherent). Thus, by failing the Interest test, it is also predicted to be anomalous.

8 Conclusion

We have proposed a formal semantic analysis of two tunes in English discourse. We argued for modelling focus and tune jointly: both the *placement* of the stress and its *type* influence which (underspecified) presupposition gets triggered. The most coherent interpretation of this presupposition in context then determines whether the focal element (determined by where the stress is placed) can be *interesting* (and so felicitous), or not. We do not provide a model theory of interest, nor even a definition of it. Rather, we impose a necessary condition on it—it must not be entailed by given information, but it must be coherently related to it.

While there is lots of ambiguity (and confusion, even!) in mapping a raw acoustic signal to a specific intonation contour (Calhoun 2007), we postulated no ambiguity in the mapping from a specific intonation contour to its meaning representation. Its distinct pragmatic interpretations in distinct contexts are then determined entirely by how discourse coherence interacts with linguistic and non-linguistic content, as required for analysing other linguistic phenomena such as anaphora, elided constructions and presuppositions (Hobbs 1985, Kehler 2002, Asher & Lascarides 1998). The fall–rise contour is often a vehicle for conveying content indirectly, but what exactly is implicated varies radically from one context to another. We captured that variation by making intonational form underspecify certain semantic elements, and then capturing how those elements are resolved to specific values via reasoning about discourse coherence and its interaction with compositional and lexical semantics, real world knowledge and cognitive states—in particular, the distinct perlocutionary consequences we associated with the two distinct tunes.

This is just the first step towards a more comprehensive account of intonation that exploits coherence, cognitive modelling and the difference between presupposed and proffered content. We feel that combining focus with tune is the right direction, but that to make real progress, the field of formal semantics needs to meet the challenge of accounting for the solid empirical evidence for the existence of a non-discrete *spectrum* of possible tunes with varying meaning (Ladd 1980, Calhoun 2007).

References

Asher, Nicholas & Alex Lascarides. 1998. The semantics and pragmatics of presupposition. *Journal of Semantics* 15(3). 239–300.

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

Asher, Nicholas & Alex Lascarides. 2008. Commitments, beliefs and intentions in dialogue. In *12th semdial workshop on the semantics and pragmatics of dialogue*, .

Asher, Nicholas & Alex Lascarides. 2011. Reasoning dynamically about what one says. *Synthese* 183(1). 5–31.

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

Beaver, David. 1997. Presupposition. In J. van Benthen & A. ter Meulen (eds.), *The handbook of logic and language*, Elsevier.

Beaver, David I & Brady Z Clark. 2009. *Sense and sensitivity: How focus determines meaning*. Wiley–Blackwell.

Bolinger, Dwight. 1972. Accent is predictable (if you're a mind-reader). *Language* 633–644.

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

Bolinger, Dwight. 1985. Two views of accent. *Journal of Linguistics* 21(01). 79–123.

Büring, Daniel. 2006. Focus projection and default prominence. In Valéria Molnár & Susanne Winkler (eds.), *The architecture of focus*, 321–346.

Büring, Daniel. 2007. Semantics, intonation and information structure. *The Oxford handbook of linguistic interfaces* .

Calhoun, Sasha. 2007. *Information structure and the prosodic structure of english: A probabilistic relationship*: University of Edinburgh dissertation.

Carston, Robin. 1996. Metalinguistic negation and echoic use. *Journal of Pragmatics* 25. 309–330.

Dryer, Matthew S. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26(4). 475–523.

Geurts, Bart & Rob van der Sandt. 2004a. Interpreting focus. *Theoretical Linguistics* 30. 1–44.

Geurts, Bart & Rob van der Sandt. 2004b. Interpreting focus again. *Theoretical Linguistics* 30(2-3). 149–161.

Groenendijk, Jeroen, Martin Stokhof & Frank Veltman. 1996. Coreference and modality. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 179–216. Oxford: Blackwell.

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

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19(02). 377–417.

Halliday, Michael AK. 1967. Notes on transitivity and theme in english: Part 2. *Journal of Linguistics* 3(02). 199–244.

Hamblin, Charles. 1970. *Fallacies*. Metheun.

Hobbs, Jerry. 1985. On the coherence and structure of discourse. Tech. Rep. CSLI-85-37 Center for the Study of Language and Information, Stanford University.

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

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. MIT Press.

Kehler, Andrew. 2002. *Coherence, reference and the theory of grammar*. CSLI Publications, Cambridge University Press.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and philosophy* 12(5). 607–653.

Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. In Joachim Jacobs (ed.), *Informationsstruktur und grammatik*, 17–53. Westdeutscher Verlag.

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

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

Liberman, Mark & Ivan Sag. 1974. Prosodic form and discourse function. In *10th regional meeting of the chicago linguistics society*, 416–427.

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

Reese, Brian. 2007. *Bias in questions*: Department of Philosophy, University of Texas at Austin dissertation.

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Rooth, Mats. 1999. Association with focus or association with presupposition? *Focus: Linguistic, cognitive, and computational perspectives* 232–244.

Šafářová, Marie. 2005. The semantics of rising intonation and interrogatives. In E. Maier, C. Bary & J. Huitink (eds.), *Proceedings of sinn und bedeutung 9*, 355–369.

van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4). 333–377.

Schlöder, Julian J. 2015. A formal semantics of the final rise. In *Student session of the european summer school in logic, language and information 2015*, .

Schlöder, Julian J & Raquel Fernández. 2015. Pragmatic rejection. In *Proceedings of the 11th conference on computational semantics*, 250–260.

Schlöder, Julian J & Alex Lascarides. 2015. Interpreting english pitch contours in context. In *19th semdial workshop on the semantics and pragmatics of dialogue*, 131–139.

Schlöder, Julian J, Antoine Venant & Nicholas Asher. 2017. Aligning intentions: Acceptance and rejection in dialogue. In Robert Truswell (ed.), *Sinn und bedeutung 21*, to appear.

Schwarzschild, Roger. 1999. Givenness, avoidf and other constraints on the placement of accent. *Natural language semantics* 7(2). 141–177.

Silverman, Kim EA, Mary E Beckman, John F Pitrelli, Mari Ostendorf, Colin W Wightman, Patti Price, Janet B Pierrehumbert & Julia Hirschberg. 1992. ToBI: a standard for labeling english prosody. In *2nd international conference on spoken language processing*, .

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

Vallduví, Enric & Elisabet Engdahl. 1996. The linguistic realisation of information packaging. *Linguistics* 34. 459–519.

Walker, Marilyn A. 1996. Inferring acceptance and rejection in dialogue by default rules of inference. *Language and Speech* 39(2-3). 265–304.

Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language* 747–776.

Ward, Gregory & Julia Hirschberg. 1988. Intonation and propositional attitude: the pragmatics of L*+ H LH%. In *Proceedings of the fifth eastern states conference on linguistics*, 512–522. Ohio State University Press Columbus, OH.

Weaver, Warren & Claude Shannon. 1963. *The mathematical theory of communication*. University of Illinois Press.