UNIVERSITY OF CALIFORNIA
SANTA CRUZ

GREAT INTONATIONS

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Kelsey N. Kraus

June 2018

The Dissertation of Kelsey N. Kraus
is approved:

Associate Professor Pranav Anand, Chair

Professor Donka Farkas

Associate Professor Grant McGuire

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Contents

List of Figures

List of Tables

Abstract

Great Intonations

by

Kelsey N. Kraus

This dissertation seeks to address the following question: can we use prosodic meaning as a metric to inform the meaning of other semantically underspecified elements? This question is asked here with respect to a class of elements called discourse particles, which speakers use to structure a conversation in an expectationally and emotionally informative way. In this work, I propose to capture the effects of both prosody and intonation in an updated view of the Table Model of discourse representation (Gunlogson 2001, 2008, Farkas & Bruce 2010), wherein pragmatic effects are entered into the conversational scorecard by way of a speaker's Discourse Commitments. This view is underlain by a causal model framework of discourse structure (Halpern & Pearl 2005), which is based off of a probabilistic notion that I define as *Expectation*. Once these tools have been established, I show how this framework can be used to talk in an informed way about Mirativity (DeLancey 1998, Aikenvald 2012). In doing this, I identify three discourse particles and one intonational contour in English that make up a part of the mirative system of this language. Broadening the reach, this work also picks out two strategies of discourse navigation in German whose effects can also be explained by making reference to speaker expectations in a discourse: the modal particles *ja* and *doch*. Here, the latter constitutes an example of a mirative marker in German, whose discourse effect picks out violated expectations of the speaker, while the former is an example of a particle that signals the status quo. This dissertation also seeks to experimentally validate the claims made about the interactions between prosody and discourse particle interactions. Is introduces a series of experiments designed to test the predictions that this new view on discourse management and coherence makes. In all, I argue that what these elements have in common is their ability to enrich a discourse with pragmatic material that contributes speaker-oriented commentary on a participant's beliefs and expectations in a conversation.

For Ralph Kraus and Edith Stein

My academic role models

who showed me how to be curious,

to listen,

and to approach everything with open eyes and an open heart.

Acknowledgments

It's difficult to imagine a world in which this dissertation was completed without the support of my advisor, Pranav Anand. I came into grad school thinking I would be a syntactician, having been seduced by the siren song of syntax. I met with Pranav at the end of my first year to talk about an idea I had for my first Qualifying Paper, which would constitute my offside QP (my plan was to get this out of the way first, so that I could spend the rest of my time doing research on German or Turkish syntax). I had recently read a paper about modal particle combinations in German, and I was sure that this author was wrong. I brought the paper with me to our meeting, along with a book I had picked up in the library, and declared that I would write my first QP on the semantics of modal particle combinations. Seeing as how German has more than 20 modal particles, I had reasonably narrowed my scope to what I assumed would be a manageable amount of work: four particles, *ja, doch, halt* and *eben*, and all of the possible pairs of combinations. This seemed straightforward enough. Our first meeting lasted two and a half hours, and when we were done, we were farther from an answer than we were when we started. After that, we decided that maybe just two of the particles and their combinations was a big enough project for a QP. At the next week's meeting (which also well overflowed our allotted time), we had, much to my dismay, narrowed the scope even further, and again, waded deeper into uncharted territory. It was decided that I would write only about *doch*. As it turns out, when one endeavors to write about modal particle interactions, one must have a complete understanding of what it is that modal particles actually contribute to a discourse. Those first few meetings remain some of the most exhilarating, exhausting, and exciting parts of this entire process, and throughout writing this, I came back to my notes of these first days often to make sure that the analysis we were building could account for all that we had discovered. Pranav has been an inspiring, encouraging, engaged, and thoughtful advisor, and can't imagine writing this document without his guidance and support. Thank you for your patience and enthusiasm

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Grant No. DGE 1339067. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

for this project, which has served in many instances over the past year and a half as motivation and inspiration.

As one may have deduced, I did not end up a syntactician. For this, I must apologize to Jorge Hankamer, an unofficial advisor, the chair of my second (syntax) QP, and my original motivation for pursuing a degree in Linguistics. Once, when I was working on a particularly tricky problem, I came to Jorge for advice. In his usual way, instead of giving me an answer, he gave me an avenue of inquiry. When I protested, he smiled and proceeded told me that in order to succeed in this field, you have to work to know thine enemy. This has proven to be very useful advice.

Early on, I got another piece of advice from Donka Farkas: always be charitable toward other people's work. This is something that I have tried, sometimes more successfully than others, to incorporate into this work and others—it turns out that this advice is hard to follow; it's much easier to handwave and not engage deeply with other work. I am thankful to Donka for her insights, her helpful feedback and thoughtful criticisms throughout this project, and for reminding me to present every analysis in its best light.

There are so many other people at UCSC who deserve credit for their help, advice, support, and guidance. I have enjoyed working with my other committee member Grant McGuire—I owe much of the design and analysis of the experimental portion of this work to his helpful suggestions and ideas. Also to Osa, for critical dog therapy breaks. To my pedagogical role models, Sandy Chung and Jim McCloskey, both for the seminars I have been able to take, as well as the courses I've taught with you, thank you for your intellectual rigor, and unique perspectives on teaching socratically and effectively to masses of undergraduates. Thanks too, to Adrian Brasoveanu for being his wonderful self.

I have had the immense pleasure of being surrounded by linguists who serve dual roles as both colleagues and friends. When I began the program, I looked to Mark Norris for advice in how to balance work and life, and he reminded me that when you need to sleep, you should sleep. Bern Samko and Karl DeVries pattern together, mostly for great conversation, movies, and for baseball stats. To Karen Duek, for breaks for cake, walks to the woods, bottles of wine, and the occasional impromptu therapy session. To Nick Kalivoda, Filippa Lindahl, and Erik Zyman: tack så mycket för syntax och vänskap. To the entire Bellik Family—Jenny and Ozan for being sympathetic ears, great friends and excellent sounding boards, and Isaac and Lizzie for asking the tough questions (“Kelsey, you’re still a human being, right?”) and for teaching me the sound that baguettes make (*crack*). To Margaret for sharing an office and for allowing me to unload my daily anxieties, and to Dan Silverman, who is always willing open his home as a quiet place to work or as a place for luxuriating. To my extended office-mates Jed Pizarro-Guevara, Netta Ben-Meir, and Nick Van Handel for breaks from writing, keeping me updated in department gossip, and for critically important discussions of current television. To those who have been with me along the way: Nate Clair, Scarlett Clothier-Goldschmidt, Mansi Desai, Lisa Hofmann, Hitomi Hiriyama, Ben Meriçli, Chelsea Miller, Maho Morimoto, Jason Ostrove: I owe you the sincerest gratitude.

Thank you to my Konstanz Family, who so graciously welcomed me into their community, and who I had to leave much too soon. To Maribel Romero, María Biezma, and Regine Eckerdt, for making my transition smooth, and for valuable comments and insights on the work that is presented here. To Erlinde Meertens and Ramona Wallner, for tolerating my office pop-ins, for allowing me to make just one more recording, and for helping me slide in with the “in” crowd of Konstanz. To Natasha Korotkova and Igor Yannovic, for hikes, excursions, reflections on life and work, and reminders that we can do linguistics without having it be the only thing we do. Finally, thanks to Andrea Beltrama, for meditations of life, on semantics, on basketball, and on finding happiness; for Kaffee Zeit while in Konstanz and for e-Cafes in the interim.

There is a group of people that I lump together because I associate them with a place that has in part accounted for my continued sanity during the past five years. Linghaus has always been a home base and a safe haven, despite the ants, the rat, and the lack of central heating. There are four core Linghausers (core only by way of having actually lived there), and three others who I associate very strongly with this place. First, thank you to Jeff Adler, Sophie Bassett, Steven Foley, and Tom Roberts for being people I looked forward to coming home to every day. Thank you for staying up late to drink beer, eat cookie dough and do chin-ups, for wine and help with troubleshooting the water heater, for cooking together, for game nights, for marathon seminar paper/QP writing sessions, for banjo time, and for feeling at home. I share

this same fondness and association with Deniz Rudin, a constant presence at Linghaus, and presence throughout my entire graduate career. Along with the others, he has been one of my closest confidants. As cohort-mates with similar research interests and extracirricular activities, most of my Santa Cruz memories have been made at least in part with Deniz. Kelsey Sasaki, thank you for being the Department Dad to my Department Mom—I think we've done a great job of keeping this ship afloat. And to Heather Purchase, for reminding me to always take time for myself.

Thank you to the people who have supported me from afar–Hannah Cory, Nina Carson, Theresa Wright, Carlie Amacher and Abby Flowers for being a cheering section when I needed it, and a shoulder to cry on. To Angelica Ramirez, one of the most thoughtful and giving people I have the pleasure to know. And of course to my parents, Paul and Carla Kraus, for their unwavering support and love.

Lastly, I want to thank Brianna Kaufmann and Paul Willis, my partners in crime (and travel). Brianna, thank you for the tough love through hard times, for helping me to evaluate my choices and for supporting me unquestioningly through my trials and successes. Thank you for spending six weeks in a car with me driving across the country, for being a home away from home, and for your continued friendship these past five years. Paul, thank you for being a friend, a partner, an advisor, a teacher, an in-house statistician, a counselor, and the supportive foundation off of which I have built this dissertation. This degree is as much mine as it is yours.

Part I

Markers of Discourse: Intonation, Particles, and Expectation

Prologue

An introduction to the puzzle

There is a curious scene in the first season of the HBO series *The Wire*. In it, detectives Bunk and McNulty are investigating a cold case homicide at an apartment building. As the scene unfolds, Bunk and McNulty learn that the detectives who initially worked the case have botched the investigation, and by using crime scene photos and re-enacting the course of events, they are able to paint a clearer picture of the night of the crime. Bunk lays out photographs of where the body was positioned in the kitchen, while McNulty examines and questions the same photos and descriptions. When they find a discrepancy, they note it aloud, and the other responds, sometimes with surprise, sometimes in disgust, and other times with a hint of cynicism. They quickly realize that they are dealing with a case of premeditated murder, and they work backwards to try to create a timeline of events. Their frustration turns to celebration when they realize that the killer was not in the apartment at all, but rather, shot the victim through a closed window. They go outside to confirm their suspicion, and sure enough, find a shell casing in the grass a few feet away. They've cracked the case wide open.[1]

The curious thing about this scene, which is over five minutes long, is that all of this information is conveyed by the detectives through repeated use of a single word. The richness in their emotional and expressive performance is not constrained, even though the literal content of their lines are. Despite their limited vocabulary, their utterances are calculated, unambiguous, and meaningful, and it is clear by their tone when they are frustrated or when they can't believe

[1] A link to this clip is here:
`https://www.youtube.com/watch?v=1lElf7D-An8`
A precautionary note to sensitive viewers—this *is* an HBO show, and as such, contains some strong language.

their eyes. By varying the performance of their single, shared line, the two are able to exchange ideas in a meaningful way to themselves as well as their audience, and they do this by changing only their prosody and intonation.

While this is perhaps an extreme example of the role that prosody plays in a discourse, it does emphasize the importance that it has in guiding our understanding of a conversation. Prosody can not only help diagnose when an utterance is an assertion, a question, or an exclamative, it can also allow listeners to draw conclusions about a speaker's emotional state and their expectations given the current conversational backdrop. The pragmatic meaning component that prosody contributes to an utterance is not trivial; by manipulating pitch, melody, or vocal quality, a speaker adds layers of meaning beyond what is contributed from pure syntactic or semantic form. This is oftentimes done for a particular pragmatic effect. If a speaker wants to position herself as being authoritative and committed to the content of her utterance, she might choose to use a neutral, falling contour. But if she chooses to use a rising contour on the same string, she comes off as more hesitant she is not proposing to settle an issue, but rather, deferring to another participant's authority (Gunlogson 2008).

(1) a. Tom watches the West Wing. ↓ — Falling declarative[2]

b. Tom watches the West Wing? ↑ — Rising declarative

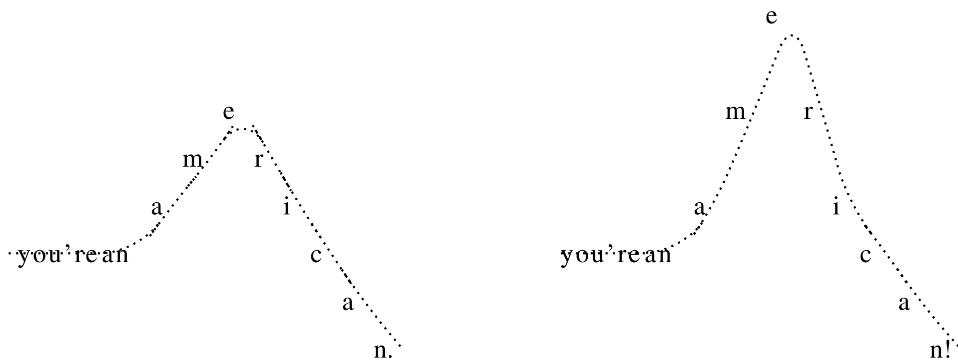
The difference between the two utterances above stems from different abstract prosodic templates: impressionistically, a final rise on a declarative utterance has the effect of speaker uncertainty (1b), while a falling final contour does not (1a).

English relies heavily on prosody and intonation to impart information about where a speaker's places herself and her beliefs in the current conversation. If she is excited, a speaker might log this with a high pitch excursion from the normal baseline. The shape of this contour is almost identical to the shape of a falling contour, which is represented impressionistically in the diagrams in (2). While neutral utterances have a characteristic rise and fall utterance-finally

[2] I will use a combination of punctuation and arrows to represent the different contours throughout. A '.' (sometimes followed with a '↓' for rhetorical clarity) is to be taken as an utterance with neutral prosody. What will come to be known as 'Excited' prosody is represented with an '!' and a fall '↓'. Rising intonation will be secondarily expressed with an '↑'. I borrow this notation from Farkas & Roelofsen (2017), and add another dimension to show surprise or exclamation with a rightward-falling arrow, ↘.

(2a), the steeper rise of (2b) is an extra indication that points toward her positive stance toward the proposition.

(2)



a. You're an American.↓ b. You're an American!↓

This deviation in (2b) from a speaker's normal vocal range is meaningful. It allows listeners to infer how the speaker views her contribution, and has an effect on how another participant may react to this utterance. In this case, manipulating the height of the position of main sentential stress correlates with a speaker's excitement: the higher the peak, the more intense a listener will perceive the speaker's excitement to be.

As opposed to many focus or topic marking strategies employed by various languages, the prosodic contours that will be of interest in this dissertation are much more idiomatic in their use. In the work that has followed from Rooth (1985), much attention has been paid to the shift in semantic interpretation that certain intonational prominence can have on elements in a conversation. This is the case in (3), where focusing TOM in the examples in (3a-c) is a licit response to the question posed, while putting focus on other elements seems wrong:

(3) Who was gone last weekend?

a. TOM$_F$ went to Canada.

b. Sophie told me that TOM$_F$ was.

c. I'm not sure, but I think Sophie drove TOM$_F$ somewhere.

d. # Tom went to CANADA$_F$.

e. # SOPHIE$_F$ told me that Tom was.

f. # I'm not sure, but I think SOPHIE$_F$ drove Tom somewhere.

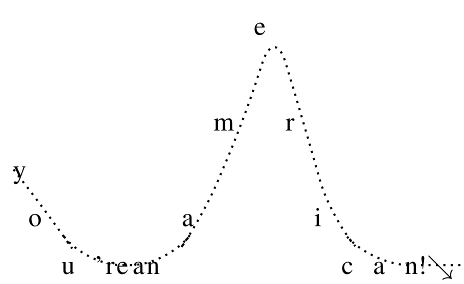
The contours that I am concerned with here differ from focus marking in a number of ways. Perhaps the clearest difference lies in the way they are marked on the phonetic string: focus marking picks out a relevant semantic contrast, and regardless of the way in which a question is answered, will appear on the same element. This is the case in (3). Prosodic contours are a bit more formulaic—they attach to prominent *positions* in a structure, notably, positions where primary and secondary sentential stress are anchored across an utterance.

The English Surprise-Redundancy Contour is a tune template that both conforms to a regular stress-based template, while also imparting a recognizable pragmatic function. This tune is idiomatic, allowing a speaker to indicate to their listener that there is something in the conversation that is being overlooked or not taken into account. It's a contour that has a reminding feel to it, telling the hearer that they "should have known" some fact or course of events. Taking the same string as in (2), adding the Surprise-Redundancy Contour is a way for Sophie's annoyance to be logged in conversation, alongside her cooperative response to Tom's question. Her overall falling contour is augmented in an informative way. Not only is she answering, she is indicating that she finds her answer redundant.

(4) *Tom and Sophie are talking about visa regulations. Sophie tells him that for visits to Europe, US passport holders don't need a visa if their trip is shorter than three months. Tom isn't getting it. He asks:*

Tom: So do I need a visa for Estonia?

Sophie: (I already told you:)



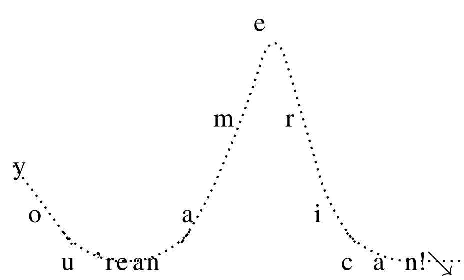
You're an American!↘

The shape of this contour is not limited to this particular utterance: Sophie could have responded using the same prosodic representation with *Estonia's in Europe!*↘ or *You don't need a visa!*↘ or even simply *No!*↘. The pragmatic contribution of this contour is clear regardless.

This same abstract template, which anchors to primary and secondary stress positions in a larger prosodic phrase, can also be used to indicate a speaker's own surprise at learning a new fact. The difference here is in the performance of the contour. Whereas in (4), the speaker's overall pitch was low, the pitch in (5) is significantly higher, leading a listener to conclude that the speaker is surprised in a positive way:

(5) *Sophie is traveling in Mongolia. She hasn't come across another English speaker in a while. Sitting in a restaurant, she overhears a familiar accent. When she realizes this, she turns to the voice and exclaims:*

Sophie:



You're an American!

What these few examples show is that prosodic meaning is incredibly rich. At the very least, overall pitch trajectory, prominence relations, and relative pitch height are all factors in determining a speaker's intent and her contribution when she offers up a proposition for negotiation. Each of these pieces has the potential to change the interpretation of an utterance in context. One way to get at their contributions is by holding all but one variable constant. In that way, as with the single difference between (4) and (5), or the change that shifts interpretation between (2a) and (2b), these pragmatic differences are allowed to surface.

Once we have established a meaning for particular prosodic contours, these too can be used as knobs to turn while fine-tuning the contribution of other semantically bleached content.

One very clear way of using this to our advantage is by using prosody as a diagnostic for discourse particle meaning. What is the pragmatic effect of using *oh* as an answer in (6a), for example, and how does it differ from (6b)?

(6) **Sophie:** We're out of flour.

a. **Tom:** Oh. ↓

b. **Tom:** Oh? ↑

In the face of no other evidence, one might be inclined to say that *oh* in (6a) marks acceptance: Tom appears to accept and view the content of Sophie's utterance as uncontroversial. But the example in (6b) suggests the opposite: Tom uses rising intonation to call the assertion into question rather than to accept it.

So what is the contribution of *oh*? And more importantly, how can we separate out intonational meaning from discourse particle meaning? Discourse particles are inherently flexible, and appear alongside a wide range of illocutionary types with a broad spectrum of functional interpretations. A goal of this dissertation is to argue for the compositionality of prosodic meaning and discourse particle meaning. Apparent conflicting uses of particles like those in (6) arise from the additive nature of the meaning contributed by intonation and the actual contribution of the particle itself. Under this view, the meaning of a discourse particle is quite semantically bleached. Their richness in interpretation comes from the varying prosodic contours that they can host. This work takes this approach to understanding the effect of discourse particles and asks what prosody can tell us about meaning the contribution of the particles themselves.

In what follows, I take an approach to formalizing the contribution of discourse particles and intonation in terms of speaker expectation. Recall the two mini dialogues above. In each case, the particle and prosody combination betrays some fact about Tom's expectational state. In (6a), *oh* seems to commit him to the fact that his expectations have been violated, though subtly. Perhaps he and Sophie were about to make cookies, only to learn that they were out of flour. He had built one expectation of how things would progress, but then to realized that he needed to re-evaluate his belief structure. In (6b), Tom is also faced with a situation in which his expectations have been violated. In this case, he has yet to restructure any beliefs.

Rather, he indicates that while he expected one course of events, Sophie apparently expects an alternative.

Framing the contributions of discourse particles and intonation in terms of expectation has two clear advantages: it can easily represent two very basic knowledge states. Expected information are facts and events that are a matter of course. Unexpected information deviates from that. It is an indication of surprise or revelation or realization. Expectations can stay the same, or they can be updated and revised. With the constant receipt and exchange of information, it comes as no shock that languages have innovated "shorthand" ways of contributing information about a speaker's expectational state. This is what I propose the pragmatic discourse effect of discourse particles and intonation is: commentary on a speaker's current belief state in the context.

In particular, marking *violated* expectations is another theme that winds its way through this dissertation. While the stated goal is to use different prosodic contours to help diagnose combined particle and contour meaning contributions, this is done in order to tease apart the meanings of particular mirative strategies. Starting with English, I identify the particles *oh, huh* and *what* as mirative strategies, as well as the Surprise Redundancy Contour. The angles taken here are both theoretical as well as experimental, adding evidence to the claim that the contribution of intonation and particle meaning is additive.

The focus then turns to German: are similar mirative strategies found in this language as well? Can the same intonation + particle outlook be extended to the data here? Does framing the contributions of these elements in terms of expectations extend past just English? In particular, I look at two German modal particles, *doch* and *ja*, which have both received wide attention and varied treatment in the literature, and both of which have been said to play very similar pragmatic roles. The difference, I argue, can be reduced to the kind of expectation they commit a speaker to. The core difference is apparent in a minimal pair like (7):

(7) **Sophie:** I'll wait to start the movie until Jeff gets here.

a. **Tom:** *Jeff hat* ***ja*** *keine Interesse.*
Jeff has *ja* no interest
(*We all know*) Jeff isn't interested.

b. **Tom:** *Jeff hat **doch** keine Interesse.*
Jeff has *doch* no interest
(*Obviously*) Jeff isn't interested.

The difference in *ja* and *doch* is subtle, but ultimately based on the assumption of different sets of expectations: *ja* is a particle that marks shared expectations, while *doch* is a mirative strategy, marking a speaker's violated expectations given the conversational context.

The dissertation is divided into three main parts, each containing two chapters. The first part is a general overview of all of the moving parts. Chapter 1 introduces discourse markers, and identifies the core group of markers to be discussed in this work. It also introduces a background on intonational meaning. Chapter 2 is an engaged discussion on themes ranging from mirativity, expectation, and evidence. This chapter also introduces the Table model of discourse management (Farkas & Bruce 2010), as well as a probabilistic approach toward modeling belief, based off of expectation.

The second part is divided into a theoretical chapter and an experimental chapter. Chapter §3 is a theoretical look at the English discourse particles *oh, huh* and *what*, and their interaction with Neutral, Excited and Surprise Redundancy prosodic contexts. It outlines in detail the conditions of use for each particle and contour combination, and proposes pragmatic representations that extend the Farkas & Bruce (2010) Table Model of discourse. While previous accounts of English discourse particles show an impressive range of flexibility in what they can express, no research to date looks at the compositionality of discourse particle meaning and intonational meaning. This chapter aims to address this. Chapter §4 puts the predictions made in the previous chapter to the test. Using a new methodology created for testing participants' intuitions when faced with emotive utterances, the experiments here probe participants' perception with respect to pragmatically enriched speech. It ultimately shows a fine grained sensitivity in unbiased participants to the same patterns hypothesized in Chapter §3, lending credence to the idea that the pragmatic effects of discourse particles and prosodic contours are both active and additive.

The final part consists of two case studies in German, which attempt to provide an updated understanding of the modal particle system. Framing the particles in terms of a speaker's expectations, Chapter §5 argues for a treatment of *doch* that views the particle as committing the speaker to expectations for one outcome, while indicating that the addressee has conflicting expectations. From this, *doch* can be viewed as a mirative marker. Chapter §6 argues for the opposite treatment of *ja*, instead viewing it as marking uncontroversiality: both the speaker and the hearer are on the same page. Overall, a picture emerges where the semantic representation of a discourse is enriched by many different kinds of pragmatic material, but which all contribute similar content: speaker-oriented commentary on a participant's beliefs and expectations in a discourse.

Chapter 1

An introduction to the moving parts

This chapter is meant to be a collection of themes that will play prominent roles in the chapters to come. It is a hodgepodge of introductions, relevant assumptions, and background ideas that I use to ground the work that will be integral in the following chapters. Some sections focus more on the semantic underpinnings I assume, while others touch on phonetic and phonological concepts that are relevant to a study of prosody. Each section is ultimately relevant for an engaged look at the way markers of surprise are conventionalized across languages, and how calculations of speaker expectation plays a role in their interpretation.

1.1 Flavoring a discourse: linguistic forms and interpreting intent

The goal of this section is to narrow the field. What is meant, exactly, when we talk about discourse markers? Even within a single language, this category is generally quite ill-defined. Take the table in (1.1) below. The discourse markers in English alone that have been defined in various works range from interjections and honorifics to contrastive elements and hearsay adverbials, which presents us with a large set of very different types of content, all subsumed under the same general heading:

Table 1.1: Types of English Discourse Markers

Type	Example	
a. Deference Marker	*Sir, Ma'am*	Fraser (1990)
b. Hearsay Marker	*certainly, reportedly*	Papafragou et al. (2007)
c. Contrastive Discourse Marker	*but, so, however*	Heritage (1984), Toosarvandani (2013)
d. Discourse Management Marker	*See, As I said...*	Schiffrin (1987)
e. Interjections	*Oh! Gosh! Wow! Oops!*	Heritage (1984), Fox Tree & Schrock (1999)
f. Intonation	Excited/Rising intonation	Schiffrin (1987), Heritage (1984)
g. Pragmatic Marker	*ok, well, so*	Fraser (1990)

In addition to making for a broad class of discourse markers, many of these elements live dual lives. Some of them, like *but*, function both as a semantic operator that contrasts two propositions (8), as well as a marker of utterance-initial discrepancies between a speaker and an addressee (9):

(8) Sophie wants beef *but* Jeff wants shrimp.

(9) **Tom**: Jeff wants shrimp.
Sophie: *But* I'm allergic.

This is similar to the French particle *enfin*, as identified by Hölker (1990), which can be interpreted in (10) for its literal content (with a temporal ordering meaning of *at last*—he was the last to arrive), but also as a pragmatic marker, indicating that the speaker is coming to the end of the list that she has started:

(10) *À cette réunion sont venus Pierre, Paul, Jacques et enfin Marcel.*
at this reunion are come Pierre, Paul, Jacques and ENFIN Marcel
At this reunion, Peter, Paul, Jacques and *finally/at last* Marcel came.

Other pragmatic content can be variably expressed with extra phonetic content, or simply by marked intonation structures. In a situation where Tom and Sophie are discussing Jeff's vacation plans, they can have the following exchange. Notice that Sophie can register her surprise both with a particle, as in (11a), or purely with intonation, as in (11b):

(11) **Sophie:** I wonder where Jeff has decided to travel to.

Tom: He told me yesterday that he's booked flights to Japan.

Sophie:

a. Oh! He's always wanted to go there!

b. He's always wanted to go there!

Again, marking a discourse as surprising in these two ways is not English-specific. If Sophie spoke German, she could respond in either of the ways in (12), with much the same pragmatic effect as the English equivalents. (12a) marks surprise with intonation, while (12b) with the discourse particle *ja*.

(12) a. **Sophie:** *Er wollte immer dahin!*
he wanted always there
He always wanted to go there!

b. **Sophie:** *Er wollte* ***ja*** *immer dahin!*
he wanted JA always there
(As we know), he's always wanted to go there![1]

In both (11) and (12), the propositional meaning of Sophie's utterance does not change given the presence or absence of English *oh* or German *ja*, nor does it change when she modifies her prosody.

Clearly the discourse navigating strategies outlined above are not restricted to English. But widening the scope of investigation does not ease the burden of what exactly defines a discourse marker. In part, this is complicated by the vast number of terms that have been used to describe these elements cross-linguistically. The literature on Dutch prefers to refer to these discourse-structuring units as *discourse particles* (Aijmer 2002, Davidson-Nielson 1996), German, Indonesian, and Chinese have *modal particles* (Abraham 1991, Mei Lee Wong 1993), and Italian, Romanian and French are interested in their languages' *pragmatic markers* (Ghezzi & Molinelli 2014, Livescu 2014, Hölker 1990). But there is a reason that these elements have been identified cross-linguistically as belonging to some class outside basic lexical and semantic

[1]Particular non-neutral intonation on (12a) as well as (11b) is necessary for these to be licit responses. This important piece of the puzzle will be set aside until later chapters.

categories. For most of the elements outlined so far, the most basic property that they all share is that they function as pragmatic cues that speakers and addressees use to help integrate the larger proposition into a discourse context.

It is not my aim here to reclassify what is meant by the term ‘discourse marker’ in every language; that is a monstrous task and a tangled web. But understanding some of the content that they can express can help to narrow down to a more manageable investigation. In this work, I want to focus in on a handful of discourse markers that have to do with the expression of surprise. On the surface, they seem like an odd bunch: What does German *doch* have to do with the English Surprise Redundancy Contour? And how can we reconcile the effects of what appear to be the English interjections *oh* or *huh* with the pragmatic effect of the German modal particle *ja*? And given their obvious distributional differences, what can we say about how they form a cohesive class of elements?

The core of the analysis and investigation here hinges on the idea that these markers of surprise are unified by their ability to express a speaker’s *expectations* about a discourse context. In what follows, I suggest that German *doch* and the English Surprise Redundancy contour express the same *kind* of speaker expectation, namely, expectation violation. And whereas *oh* and *huh* in English comment on a specific way in which a speaker expectations did not meet reality, German *ja* comments on how a speaker’s expectations *have* been met. What emerges is an updated view of tracking a speaker’s commitments in a discourse through their own eyes: commitment can be framed in terms of what a speaker expects (or doesn’t) given a discourse context, and how that interacts with the kind of commitment that is ultimately interpreted.

1.1.1 Three discourse particles, two modal particles, one contour

Though both languages have many ways of marking surprise and violated expectations, I limit this investigation to three English discourse particles (*oh, huh* and *what*), two German modal particles (*ja* and *doch*), and one intonational contour (the Surprise Redundancy Contour, (SRC)). On the surface, this may seem strange, given the very obvious differences in the ways the respective languages categorize and integrate these particles into speech, and the fact that one of the strategies at issue has no inherent phonologic content. In this section, I wish to lay out an argument for why these particular expressions of surprise are more deserving of similar treatment than of different.

First, I want to make clear the terminology I use throughout. Though their uses are similar, I use two distinct terms to distinguish the language-specific markers discussed here. This is done in part as a way to connect this work to previous literature on these topics in their respective languages. I take the term **discourse particle** to refer to the English-specific discourse structuring units of interest, while **modal particles** refer to the same class of German elements. These terms are both ways of referring to subsets of the larger class of pragmatic discourse markers. When I use the more general term *discourse marker*, I assume that the prosodic marker is subsumed under this heading as well.

What I refer to as the English discourse particles *oh, huh* and *what* are more commonly thought of in classical grammars and in much of the linguistic literature as interjections. This is partially based on the observation that they can occur on their own as a response to an utterance or event:

(13) a. **Tom:** It's happy hour at the Red Room tonight.
Sophie: Oh! / Huh!

b. *Tom, assuming he is home alone, walks in to the kitchen to see Jeff at the table:*
Oh. / Huh.

But they are clearly different from other kinds of interjections, as evidenced by (14):

(14) *Jeff punches Steven in the arm. Steven responds:*

a. Ow! / Hey! / Fuck!

b. #Huh! / #Oh! / #What![2]

Cuenca (2013) notes that these interjection-like elements are extremely context dependent, and are often used in conjunction with utterances that further refine their meaning. In a way that pure interjections cannot, elements like *oh* or *huh* can be anaphoric to an utterance that further explains their discourse function. Schiffrin (1987) refers to this distinction as

[2]Note here that it is not the fragment question reading of *what* that is intended here.

forward-looking and backward-looking uses of particles. I prefer to distinguish these different uses as *linking particles* and *detached particles*. Interjection uses of the particles are detached uses: they react to a previous speaker's utterance, but they do not make an inferential connection between that utterance and a following one. In these cases, the particle's meaning seems to stand alone as an independent discourse contribution, much like independent propositions might.[3] This is best illustrated with an example, as in (15).

(15) **Colonist:** It's very strange. This evening I saw two lanterns hanging in the Old North Church.

Paul Revere: Huh. The British are coming!

This is a run-of-the-mill, normal interaction. But while this is a licit use of *huh*, the particle is a detached use, rather than a linking particle use. It is used register Paul Revere's reaction and understanding of what the observant colonist has just told him. It is a strange thing that two lanterns would be hanging in the Old North Church. Crucially, the particle response does not directly invoke his realization that the British are coming. In order for him to felicitously announce that there are invading troops, he must have undergone some other cognitive process. The particle utterance and his follow up are two separate discourse moves. Only once he has remembered his instructions (*one if by land, two if by sea*) and his mental state is updated, can he felicitously announce the invading Redcoats. The particle here is detached from the rest of the semantic content, and functions as an interjection and a following utterance. These are clearly two disjoint conversational moves which happen in quick succession.

Detached particles are not the particles of interest in this work. Rather, the focus is on linking particles, which take the contribution of the particle as well as preceding context and a follow up utterances into account, In (16), Paul Revere has a different reaction to the colonist's observation—he uses *oh* as a way to connect the information he hears from the colonist with an inference he is able to draw from it:

(16) Oh! The British are coming!

[3]This is of course, not assuming that the particles in these cases contribute actual propositional meaning. I leave fine-grained details of how content is placed into the common ground for these types of utterances to future work.

The link between the particle and utterance is much clearer, as the realization that the two lanterns hanging in the Old North Church is immediately registered as a predetermined signal for action. Paul Revere uses *oh* to both register the significance of the colonist's statement, as well as mark his surprise. The follow up utterance is then understood as an explanation of this surprise. The discourse particle and the utterance in this case clearly refer to the same event, which triggers a response from Paul Revere.

One way to tease apart a detached use from a linking use of a particle is to notice what the particle responds to, and what the follow up utterance responds to. In detached cases, the particle responds to a previous utterance or event, but the follow up does not. One diagnostic that pulls apart this difference is to paraphrase an explanatory relation. Detached particle uses will seem like infelicitous responses with the paraphrase, while linked particle uses will not be:

(17) **Sophie:** The server is down.

a. # What. *The reason I said that is because* no it's not.

b. Oh. *The reason I said that is because* no it's not.

This type of explanatory connection between particles and the statements that further refine their contributions is one that is not present for interjections and detached particle uses, and one which inches particles like *oh, huh* and *what* in these contexts closer to a grouping with German modal particles. Unlike English discourse particles, modal particles always occur utterance-medially, which link them inherently to the content of an utterance. This contrasts with English particles which can have a dual use as a detached and linking particle, partially due to their ability to occur at the juncture of two separate utterances. But the functions of modal particles and discourse particles are much the same. Both the German and the English utterances in (18-19) convey the attitude of the speaker toward the utterance and the discourse context. Interestingly, English can also use prosody to register this surprise, using the Surprise Redundancy Contour:

(18) (*points*:) Oh! There's a cat!

(19) *(points:) Da ist* ***doch*** *eine Katze!*
there is DOCH a cat

(*To my surprise*) there's a cat!

(20) (points*:) There's a cat!*↘
L* H*L%

Though one should hesitate to call these utterances direct translations, the particles *oh* and *doch* do appear to contribute similar speaker attitudes, namely, that the speaker is surprised about the presence of the cat. In the case of (20), the speaker manipulates stress placement and pitch accenting to highlight the fact that this sighting is surprising in much the same way that the overt particles do. In all three cases, the particles as well as the contour seem to operate above the literal semantic content of the utterance, commenting instead on the speaker's view of the discourse. Cuenca (2013) observes that this is one of the hallmarks of discourse markers more generally in her attempts to establish a tripartate classification of discourse marker subtypes. Though her goals are different, her description of *modal meanings* inherent to some particle subtypes is relevant here: the contributions of many discourse markers operate above the level of the speech act.

(21) **Modal meanings** are inherently interactional in that they put forward the attitude, knowledge, or stance of the speaker with respect to what [has] been said, or to the hearer. Pragmatic connectives and, especially, modal markers are typical vehicles of modal meanings such as agreement, disagreement, or emphasis.

Cuenca (2013), p.207

These modal meanings are contributions that enhance the effect of a given speech act. A speaker uttering *There's a cat* in a neutral environment makes an assertion. Adding *oh* does not change the fact that this is an assertion, but rather, adds the extra inference that the content expressed in this utterance was somehow unexpected. The same is true of *doch* in (19). The utterance at its core still has the force of an assertion, but it has been modified to represent how the speaker has integrated this into her view of the discourse situation.

Waltereit (2001) makes a similar observation: German modal particles are generally discussed as a unique linguistic innovation, as they are extremely common and integral parts of both written and spoken forms of the language, but are relatively uncommon in other languages.

Yet it is not the case that other languages do not have strategies for expressing the same things that modal particles express in German. In Portuguese or Italian, this can be found in some uses of the diminutive. In French, you find modal particle-like meanings in some uses of the perfect tense (Waltereit 2001). For English, there is significant overlap with discourse particles, and a very rich system of intonational patterns that contribute similar effects.

In sum, this section is a starting point for an analysis that treats the contributions of English *oh, huh* and *what* on par with the German modal particles *ja* and *doch*. Though it does not comment explicitly on the surprise elements that each particle contributes (discussions of these can be found later in Chapters §3, 5, and 6), it does make a case for a similar treatment of English discourse particles and German modal particles. It argues, too, that pragmatically marking surprise can be extended to the intonational realm as well. Given this, we proceed in our pursuit of an expectation-based view of discourse markers.

1.2 Intonation and meaning

Discourse markers function outside of the traditional semantic representation of discourse. As the previous sections have argued, these elements are largely pragmatic, and interpreted broadly as tracking discourse participants' attitudes, expectations, and evaluations of propositions in a conversation. But discourse particles are not the only elements that contribute necessary information to participants navigating a discourse. The performance of an utterance also contributes crucial pragmatic information. This means that in order to really understand the contributions of pragmatic discourse markers, we must also understand the contributions of prosodic tunes. That is to say, understanding the contributions of discourse markers *requires* a detailed understanding of intonation as well.

Attempting to arrive at a meaning for a discourse particle without regard for its actual acoustic manifestation is a losing battle. So much of the meaning of discourse particles is relayed in the performance of the utterance itself. Because they both operate at the pragmatic level, the relation between the text and the tune is very close. The following is an attempt to ground the current work in a prosodic theory outside of a semantic one, in which prosodic and phonetic generalizations can help inform the pragmatic theory we build from it.

1.2.1 Some background on prosodic theories

A semantic or pragmatic investigation of intonation must be sensitive to the phonetics and phonology of the languages involved. Much of the work on labeling intonation and pieces of prosodic contours is done in the Tones and Break Indices (ToBI) framework developed by Pierrehumbert & Hirschberg (1990) and Silverman et al. (1992). This system is an attempt at standardizing the analysis of speech prosody such that labeling conventions are non-language specific and can be easily adapted to the characteristics of individual languages without losing crucial phonetic detail.

The ToBI system developed out of the Autosegmental-Metrical tradition of phonological theory. It seeks to broadly characterize sentential melodies in terms of peaks and valleys corresponding to high and low pitches of an intonational contour. In classic metrical phonological theory, high and low pitches are obscured, collapsed into a single category of *strong* positions in an utterance, which are anchored to sentence accents and stressed syllables of content words. Prominence is determined in hierarchical, tree-like structures which branch in a binary fashion according to strong and weak relative positions. Primary sentential stress is assigned to a single strong position that is not dominated by any weak nodes. Figure 1.1 shows a sample metrical tree based on strong and weak sentential accents.

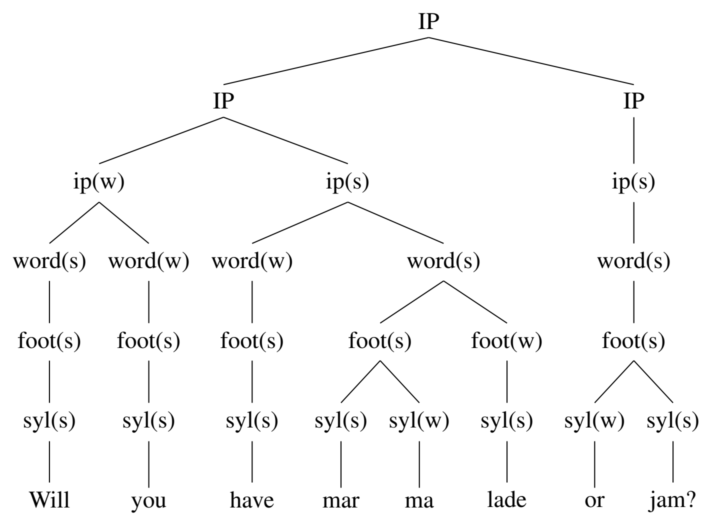


Figure 1.1: *Will you have marmalade, or jam?* A metrical tree, showing Intonational Phrases (IP), intermediate intonational phrases (ip), words, metrical feet and their corresponding weak (w) and strong (s) syllables.

The entire tree represents a single Intonational Phrase (IP) made up of two smaller phrases. In an alternative question like *Will you have marmalade or jam?*, there are traditionally assumed to be two IPs that combine to form the overall, canonical rising tune on the first alternate and falling accent on the second. Pitch rises and falls in this schema are not immediately apparent. Instead, prominence in phrases is determined by relations between adjacent syllables with abstract *weak* (w) and *strong* (s) features. Strong features are calculated for stressed syllables on lexical words, and labels that are assigned to abstract functional phrases (*foot, word, intermediate phrase*) are calculated based on intuitions about sentential stress and accent placement. The functions of these trees are to show prominence relations, but they do not necessarily clearly recreate the information carried by intonational contours.

Though I will not comment extensively on the syntactico-prosodic theory underlying these distinctions, I will say the following. IP boundaries are assumed to align with *some* syntactic constituent, and are assumed to occur after major syntactic constituents. This means that in the base case, one is unlikely to find an IP boundary between a preposition and the noun in a prepositional phrase like *by the sea* in English, or as *unter den Linden*, 'under the Linden trees' in German. These constituents are assumed to phrase together. But timing and phrasing in speech is not exact, especially when dealing with naturally occurring speech. Thus, while syntactic phrasing and intonational phrasing in the base case seem to pattern together, this is not always the case.

Metrical grids have a less formally abstract hierarchical notation, and instead are based on representing individual segmental prominence in a phrase. Figure 1.2 takes the same sentence and annotates it using this system, stacking x's on metrically strong phrasal positions and allowing no strong positions to have identical prominence.

This annotation system is more visually representative of the fundamental frequency (f0) pitch track of an utterance. The rises and falls of the metrical stacks can be thought of as approximating a pitch track. But much of the intricacies of the actual speech sound are still obscured.

			x				
			x				x
			x		x		x
x	x	x	x	x	x	x	x
Will	you	have	mar	ma	lade,	or	jam?

Figure 1.2: Sample metrical grid analysis showing prominence relations

The ToBI annotation system takes the building blocks of autosegmental-metrical analyses and pairs them with actual speech recordings. ToBI analyses annotate features of the speech sound based on pitch fluctuations, which are identified and labeled as sequences of varying high (H) and low (L) tones (Pierrehumbert 1980). This approach is an attempt to represent metrical analyses of speech in a phonetically grounded way. The system looks at utterance-specific analyses of tone height, which then may be interpreted in relation to tone heights within the same utterance. This adds a level of dynamic analysis of tonal prominence that was previously obscured. High tones can be maintained across or within prosodic phrases, or they can fall. Low tones can rise and then fall again. This dimension of movement improves upon previous analyses, which focus only on strong and weak positions within a phrase.

The abstract hierarchical information previously encoded by trees and grids is also incorporated into ToBI analyses. Prosodic phrase boundaries and intonational phrase information is preserved on an abstract tier of *break indices*, which measures the association strength of adjacent words from 0 (strong phonetic association) to 4 (clear phonetic break). These roughly correspond with prosodic phrasing: larger measures indicate clearer prosodic phrase boundaries, and smaller ones mark word or segment-level intonational boundaries.

In addition to a tier of break indices, ToBI analyses annotate three more parallel tiers of information. Peaks and valleys in the fundamental frequency of an utterance correspond to H and L tone annotations, respectively, which are recorded on the *tone* tier. An *orthographic* tier pairs time-course data of speech production with words, and the final *miscellaneous* tier annotates aspects of speech production not characterized as informationally part of the prosodic structure (coughs, disfluencies, etc...). Many of the finer grained labeling conventions used are still subjective; break index tiers are impressionistic measures of phonetic association, and may be influenced by particular theories of prosodic phrasing. Overall, the ToBI framework provides a phonetically motivated way to begin analyses of intonation.

The following example is taken from the Guidelines for ToBI Labelling (Beckman & Elam 1997: p. 9), and shows some of the conventions used when marking pitch dynamics and prosodic phrase annotations:

(22)

Will		you		have		mar ma lade,		or		jam?	
	1		1		1		3		1		4
						L* H-				H* L%	

The numbered break index tier shows prosodic phrasing levels. A mark of 4 on this tier represents a pause or the end of a prosodic phrase. In this case, 4 appears utterance finally. Smaller, intermediate phrases are are marked by a 3, such as the space between the first and second conjuncts of the alternative question. 1 indicates high strength of association between linearly adjacent words. The tones tier tracks the relative height of the prosodic contour through the utterance, and uses diacritics to mark pitch accents (*), phrase accents (-) and boundary tones (%). The utterance in (22) has two intermediate intonational phrases corresponding to each conjunct. The first has a low pith accent (L*) that rises to a high phrase accent (H-), and the second begins with a high pitch accent (H*) that falls to a low boundary tone (L%). The utterance has an overall rising contour, which then falls in the second disjunct of the question. Figure 1.3 shows a complete analysis of the sentence in (22) by ToBI standards, produced and annotated by the author:

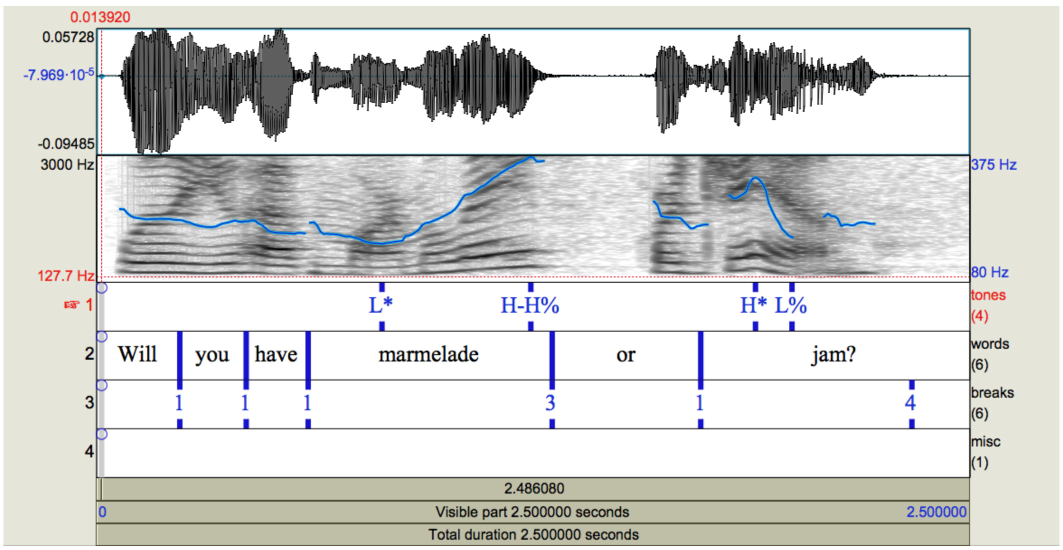


Figure 1.3: ToBI annotation of the sentence *Will you have marmelade, or jam?*

In what follows, we adopt a ToBI analysis of prosodic contours based on the autosegmental-metrical framework. Terms such as *final rise* and *final fall* are used throughout, but will generally be associated with their corresponding ToBI labels. Finally, as prosodic contours must generally anchor to prominent intonational positions in an utterance, explanations of contours will also include information about how the overall contour may be shaped with respect to the syntactic characteristics of a phrase.

1.2.2 Pieces of prosodic contours: rising and falling tones

German and English have very similar patterns of prosody and intonation. Both have rhythmically governed stress distribution, and in matrix clauses, align primary phrasal stress on the rightmost argument of the verb. New and focused information can shift this placement, but in both languages this is done in very similar and predictable ways. The distribution of content words throughout an utterances creates a series of rises and falls of high and low pitches which constitute the melody of the utterance. In both languages, declarative sentences have the force of assertions, and are characterized by what I will refer to as *neutral* prosody, which is a high tone that falls to a low boundary tone (H* L%) phrase-finally (Liberman & Prince 1977, Féry 1993, Hayes 1995). This I assume is characteristic of entire utterances.

(23) a. The cat is in the garden.
H* L%

b. *Die Katze ist im Garten.*
H* L%
The cat is in the garden.

Final falling boundary tones are also very commonly the default utterance-level intonation in constituent questions. Bartels (1999) assumes that the neutral falling contour on wh-questions is indicative of the speaker's overall intent: everything in wh-question that is not the wh-phrase itself is assumed to be old or presupposed content. In such questions, some focus on the wh-phrase is expected, but an overall falling contour phrase finally is the default. In an utterance like (24), the speaker seeks additional information, signaled by tonal prominence on the wh-phrase utterance-initially. The overall prosodic shape of a constituent question bears striking resemblance to ordinary declaratives with neutral prosody. Additional prominence is only calculated on the new, alternative-generating wh-phrase:[4]

(24)

[What kind of hummingbirds] drink from your feeder?
H* L% H*-L%

[4]The corresponding audio files, produced by the author, are available at the following link. Note that (25) is not assumed here to act as an answer to the question posed in (24), but rather, as a intonational minimal pair with the constituent question. https://www.dropbox.com/sh/ig9buw9jbavv6c6/AACampiJfPUC-c9YmsVn1PUXa?dl=0

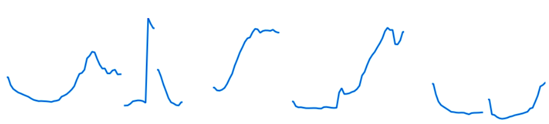
(25) 

Anna’s hummingbirds drink from your feeder.
H*-L%

Rising final tones are also found in German and English on both declarative and interrogative sentence types. They are in some sense the converse of falling tones, as they begin on a low L* pitch accent and rise to a high H% boundary tone. These tones are the default for polar questions, and have been generally described to express some sense of incompleteness (Féry 1993).

Recent experimental work on English by Jeong & Potts (2016) indicates that rising tunes can signal that a speaker’s authority over the proposition in question is lower for falling tunes. Rising tunes lead a listener to infer that a speaker wishes to be less specific about her contribution, and also that a speaker is signaling her politeness and positive stance toward a listener. This approach to rising and falling tunes utterance-finally is unifying in its attempts to extract overall pragmatic meaning from pieces of a tonal contour, independent of sentence type. For rising declaratives in particular, this can help to explain the commonalities between “uptalk” situations, as in (26), and questioning uses of final rises. The connection lies in the shared inferences a listener can make about a speaker’s knowledge state:

(26) *On a nature walk through a natural park*:



My name’s Melissa, I’ll be your tour guide.
L*H% L*H%

(27) a. Did Lars build a birdhouse?

H* L%

b. Lars built a birdhouse?

H* L%

Though rises and falls can be pieces of larger intonational contours, the assumption here is that there are canonical neutral falling and neutral rising contours that operate on the level of the utterance. Absent of any context, these contours in both languages appear to nicely align with particular illocutionary types: assertions generally have falling contours, while interrogatives are characterized by rising intonation. As we will see in the next sections, this is not as straightforward as it seems on the surface.

1.3 Variations in prosodic meaning

In her overview article on intonational meaning, Prieto (2015) discusses three properties of intonational structure which have been more or less taken as given. First, work in the realm of intonational prosody must have a clear *duality of structure*, and must make clear differences in phonological and phonetic levels of intonational analysis. Gussenhoven (2002) categorizes these two levels as the *tamed* (phonological) and *untamed* (phonetic) differences in a speech stream. The tamed half is made up the the distinctive tonal morphemes that compose in different ways to create linguistic meaning. The untamed half is where phonetic variation is measured, which is gradient and meta-communicative.

Yet the lines here are blurred. This is in part due to a second factor, namely, context dependency. Actual prosody in the wild does not easily lend itself to generalizations made in the lab. There is not a clear one-to-one mapping of intonational contour to speech act type in (28), all of which have overall rising contours, but whose contexts can alter the intended meaning:

(28) a. *Tom walks in and sees Sophie, who he hasn't seen since yesterday:*
Tom: You got a haircut? Observation

b. *Sophie asks Tom to guess what is different about her appearance:*
Tom: You got a haircut? Guess

c. *Tom, a movie producer, has told Sophie not to change her appearance for continuity while shooting. She comes to the studio the next day with her long hair in a bob. Tom is very annoyed:*
Tom: You got a haircut? Threatening assertion

d. *Tom is asking Sophie if anything is new. He asks:*

#Tom: You got a haircut?

Ideally, intonational morphemes like H* or L% could be assigned inherent meanings that would remain constant regardless of the content of an utterance. This is done in works like Pierrehumbert (1980), Pierrehumbert & Hirschberg (1990), Hirschberg (1993), Féry (1993), Bartels (1999), among others, each coming to different conclusions about which intonational pieces contribute bits of meaning to an utterance. But this is also mired in controversy, as no one can seem to come to an agreement about even the most basic of meanings. Speaker assertiveness, something that many believe is contributed by intonation, is for Pierrehumbert & Hirschberg (1990) a property of a dynamic contour, while for Bartels (1999), this is attributed to a particular phrasal tone. Or, is speaker assertiveness a feature of entire tonal contours? This compositionality vs. holistic view of tonal melodies is the third part of intonational meaning.

In this work, I take the holistic stance, arguing that the meaning of complex tunes cannot be completely broken down into component parts. This is in part based on the work of Goodhue et al. (2015), who argue that breaking down complex intonational contours into smaller intonational morphemes distorts the larger picture that is gained from looking at the contour in its entirety. On analogy with morphologically complex lexical items, they argue that focusing too intently on one piece of an intonational contour is akin to examining only a prefix or a suffix for the category membership of the word. While a prefix can betray details about larger grammatical category membership, it often falls short of characterizing a word's meaning as a whole. They present several English contours that are all characterized by a final *rising* pitch contour, but which contribute distinct discourse effects that need not rely on a lack of authority or incompleteness toward the proposition in question. All are assumed to be functions that take a proposition, check felicity conditions, and return that proposition (p. 313-314).

The English Contradiction Contour (H*+L L*-H%) is a complex sequence of tones, beginning with a high tone that falls to a low nuclear pitch accent, and then rises again phrase finally. Its effect is much like the name suggests, and it seems to imply to the addressee something like "Didn't you know??":

(29) **Tom:** Jeff is trying out for basketball this year.

Sophie: Jeff can't play basketball!
H* L L* L-H%

Though this has a high rising terminal, Ladd (1980) and Bartels (1999) both agree that declarative sentence forms with this contour have the illocutionary force of an assertion. What the contour seems to do is require contextually salient evidence for $\neg p$. Sag & Lieberman (1975) introduce intonational contours with these very specific meanings as a kind of intonational idiom (Wennerstrom 2001), as their melody alone seem to convey particular speech acts. The contradiction contour in its entirety, though it asserts the content of the proposition, seems to have the illocutionary force of a contradiction.

The Rise-Fall-Rise (L*+H L-H%) contour is another complex contour which ends with a final rise, but which asserts the content of *p*. This contour rises on a focused constituent with main sentential stress, then falls again before rising utterance finally. Again, while it has the illocutionary force of an assertion, the contour is used to also insinuate alternatives to *p*. Attention is drawn to the fact that the speaker *did not* assert the potential alternatives, allowing the hearer to make their own inferences as to why not (Ladd 1980).

(30) A: I bought a lemon tart for the party.

B: Lars likes car- rot cake.
L*+H L-H%

If the Rise-Fall-Rise contour is uttered with a higher overall pitch range, the effect shifts. Bartels (1999) argues that this higher overall pitch range implies a lack of speaker commitment, and does not have the basic force of an assertion. However, it also does not have the force of a question either. This contour takes a proposition *p* and conveys that it is unexpected in the face of evidence for *p*, or in the face of an interlocutor's commitment to *p*.

(31) *Lars is known to hate all root vegetables. B observes him eating carrot cake at a party:*

B: Lars likes car- rot cake!
L*+H L-H%

All of these contours work on declarative sentence types, and all involve final rises, yet their functions are much more complex than the basic meaning of a final rise. This is of course not to say that contours are non-compositional. Instead, it is intended to show that prosodic contours cannot easily be broken down into component parts. At least in these cases, the tones and the trajectories of pitches do not seem to be additive. Though these contours have been specific to English, the same patterns can be observed in German as well (c.f. Féry 1993, 2011, Turco et al. 2012: u.a.). I leave the compositionality question of intonation aside, and proceed in the rest of this section to talk about a biological rational for perceiving and producing contrasts in intonational meaning.

1.3.1 The Effort Code

Intonation has two crucial aspects: pitch movement throughout an utterance, and the relative prominence, or pitch height, that a segment is given. Pitch movement refers to the dynamic rise and fall of the fundamental frequency throughout an utterance. Height is a relative measure throughout an utterance of the peaks and valleys of the f0. But height can also serve as an absolute measure as well. Consider the impressionistic difference between the contours on the B and B′ responses to A's question in (32):

(32) A: What animals do you see out the window?
B: Cats.
B′: Cats!

In terms of pitch movements, the utterances are quite similar. Both appear to have falling, contours with similar trajectories, yet the B′ response has higher overall pitch, which a listener in turn perceives as informationally significant:

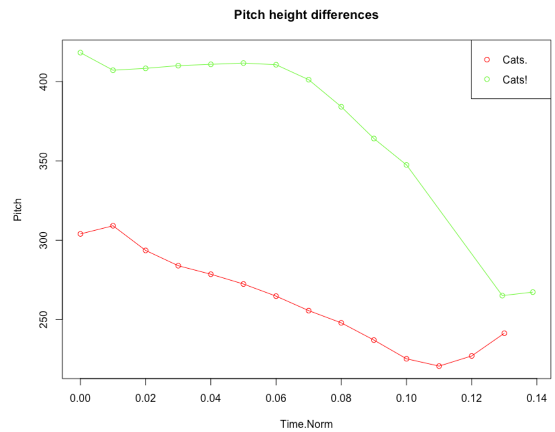


Figure 1.4: Pitch height differences between neutral *Cats.* and excited *Cats!*

Both of the contours can be described as a high pitch accent falling to a low boundary tone. Figure (1.4) shows the author's production of the B and B′ responses in (32) with normalized times. Both utterances do indeed involve a final fall, and do so in approximately equal time spans. But the utterances differ in at least two significant ways. First, while both fall to approximately the same frequency level, the slope of the fall is significantly greater in the highly emotive B′ *Cats!* than in the same answer with neutral prosody. Second, and perhaps more strikingly, the starting pitch height in the exclamative utterance is approximately 100Hz higher than the neutral utterance. Impressionistically, this makes sense. Speakers tend to use higher pitched utterances when they are excited, and the extra expenditure marks the information as more highly salient. We predict from this that relative tonal height to a baseline f0 can have a particular pragmatic effect when paired with an utterance.

Gussenhoven (2002) argues for communicative power in intonational meaning along biological grounds, following Ohala (1983)'s Frequency Code. Speakers are universally equipped with the basic fundamentals of producing language. Speech production requires the physical articulators, energy expenditure to produce sounds, and the ability to parcel out that energy into discernible chunks synced up with the normal intake and expulsion of breath. All of these are essential in producing a speech signal. For the purposes here, the energy that a speaker requires to produce a particular sound or gesture is particularly relevant. The articulatory effort expended by speakers increases a speaker's precision of articulation. More effort is indicative of a greater concern for the way in which the speaker conveys her message.

The Effort Code assumes that speakers are concerned with the manner in which their message is conveyed. Greater precision in presenting oneself will require more effort placed into speech production, which in turn is realized as greater articulatory precision. This comes across in noticable deviations from neutral prosodic contours (Gussenhoven 2002). By using a larger pitch range, speakers can signal emphatic meaning above the semantic content of their utterances. In Germanic languages, this is grammaticalized in part as the expression of focus, and also through deaccenting not-at-issue content (Swerts et al. 2002).

The Effort Code is often also seen as imparting speaker affect. A concrete example comes from child-directed speech. Speakers often are perceived as more friendly when they are speaking with children, which is the product of higher pitch and more precise gestural movements. The speaker comes off as more helpful and more approachable as a result of a sharper definition of these tonal contours (Uldall 1960).

Since pitch varies from males to females as a function of vocal fold thickness and length as well as various indexical properties, we must speak here in relative terms when it comes to fluctuations in a speaker's pitch. This means that though individual speakers can have absolute height ranges for normal and emotive pitches and contours, there is no one value or range that characterizes high or low pitches. Instead, listeners must adapt to a particular speaker's mode of speaking, and infer meaningful excursions from knowledge of their normal range (Gussenhoven 2002). Taking this into account, listeners are extremely skilled at quickly accommodating a speaker's speech style, and even after a short period of exposure, can pick out prominent positions and content (Reitveld & Gussenhoven 1985).

Hirschberg (2002) explains the Effort Code in terms of Gricean Maxims in an effort to link the biological functions of speech production to actual theories of speech communication. She paraphrases the Effort Code in terms of a Maxim of Emphasis: "Try to make informationally important parts of your speech intonationally prominent" (p. 3). Thus, particular pragmatic effects that we might infer from relative pitch height or contour shape can be comparable to conversational implicatures. Focused elements are given prominence to signal to the listener that the element is informationally significant. Deaccenting has a similar effect: it signals to a listener that this information is informationally old.

In the same way that speakers might violate Grice's Maxims of conversation, one might expect to be able to violate this proposed biological maxim as well. Deaccenting informationally relevant information to achieve a particular goal is one way to test this. One can imagine that a perturbed mother whose child has skipped class might ask this child where they were all afternoon. He might answer with a dismissive "Oh nowhere" and follow it with a (highly) deaccented "just at the movies with some friends." Deaccenting here, in the son's mind, signals to his mother that this information should be regarded as unimportant, even though it's exactly the important information the mother seeks.

In light of the previous sections, likening this to conversational implicatures has a certain appeal: intonational contours in English occur with a wide variety of sentence and illocutionary types. Ascribing them meanings that can be defeasible in context can derive the pragmatic effects that arise in the right contexts, and explain how they do not arise elsewhere.

1.4 Chapter Summary

This chapter has been a mixed bag of topics that will appear throughout this work. The central theme of this thesis necessitates it. Assuming that discourse particles interact with prosodic meaning requires engagement with linguistic areas that for the large part, do not frequently interact with one another. This chapter first touched on prosodic discourse markers as a class, and what it is that characterizes these pieces of the grammar. In doing this, we restrict the scope of the particles that this work engages with: English *oh, huh* and *what*, German *doch* and *ja*, and the English Surprise Redundancy Contour. We find that these particles and contours share that they are optional, act above the propositional level, are sensitive to discourse context, and do not change the truth conditions of an utterance. These markers operate on the pragmatic level to help structure coherent discourses.

In addition, this chapter introduces the topic of intonational meaning, and some of the prosodic underpinnings that are assumed when talking about intonation. Being able to talk about the acoustic representation of particular tunes is an important foundation for later chapters, which rely on a knowledge of the ToBI annotation system of pitch and accenting. This chapter adopts the assumption that while intonational contours may be broken up into high and low pitches, and rises and falls between pitches, the system is not clearly compositional. Instead, I argue for a more holistic approach to modeling intonational meaning. Neutral sentence intonation in English is characterized by an overall falling contour, while excitement can be characterized by this same contour, but with a higher overall pitch excursion. I analyze the meaning of prosodic tunes at the utterance level, a level of granularity that can capture broad scale intonational differences without resorting to attributing meaning to single tones or pitches within an utterance.

The following chapter tackles another one of the building blocks needed for this inquiry into discourse particle meaning: surprise. In particular, this ties in with the notion of (violated) expectation, an ingredient I argue is crucial in understanding discourse particle, as well as intonational meaning.

Chapter 2

Evidence, Expectation and Surprise

As seen in the previous chapter, much of the discussion of pragmatic discourse markers and prosodic contours introduced in this work deal with the notion of violated expectations. This chapter explores the idea of expectation violation and surprise, and in particular, how these concepts relate to what a speaker believes and knows in a context. The first part of this chapter is an attempt to situate this work in the preexisting literature on expectation and epistemic uncertainty, while focusing narrowly on the expression of surprise. The second half of this chapter looks at the application of surprise in the grammar, and how mirative strategies can be formulated in terms of violated expectations.

2.1 Belief, credence, expectation

Before we get into the notion of expectation, I want to briefly talk about some of the assumptions that I make about belief. What kind of state is a belief state? This is of course not a new question, with scores of linguists and philosophers debating this territory for years. In what follows, I assume that it should be comprised of a few things, taken from various scholarly pursuits on the topic (see, among others, Hintikka 1969, Kamp 1981, Beaver 2001, Halpern & Pearl 2005, Yalcin 2010). First, it contains the facts that a speaker knows are true in the world (their epistemic state). In addition, it can include the inferences that have been drawn about the current world and potential future states of the world. In setting up a belief state, we need some way of categorizing or ordering these propositions and eventualities in such a way as to filter out those that you know, those that you believe, and those that are likely, given what you know and believe. In effect, we need to also be able to talk about a speaker's expectations in a given situation, We want to be able to talk about uncertain beliefs.

For instance, you can believe that it will rain today without being entirely sure that this will turn out to be true. In this case, it seems that belief can be structured in terms of credence levels, or *expectations*: you expect with some high probability that it will rain. Rational agents hold coherent, yet uncertain doxastic states all the time, which suggest a need for a way of talking about non-categorical belief:

(33) *You are a burglar doing a thorough casing of a house. You've been outside for hours and haven't seen signs of movement inside the entire time:*
I believe the house is empty, but it might not be.

(34) *Ann is casing the house with you*:
Ann believes the house is empty, but it might not be.

(35) I believe, but I am not certain, that we met on a bus.

Beddor & Goldstein (2017), among many others (for example Kratzer 1991, Halpern 2005, Yalcin 2010, Lassiter 2017), are some of those who propose a solution to the uncertain belief problem. They argue that belief is not certainty, but rather, a probability value above a certain threshold. While it is not the current enterprise to pinpoint this exact threshold, from this, it seems reasonable to talk about belief in terms of expectations of a particular outcome given observations about the world.

Farkas & Roelofsen (2017) base their account on credence levels for particular propositions. I propose to talk about expectation in terms of relative credence levels. A speaker's expectation for a proposition is based off of her (non-categorical) belief calculation for that proposition. The choice here is largely aesthetic, and reflects the general trend in the literature on mirativity where the focus is less on traditional notions of commitment, and more on a speaker's expectations and how they align with a discourse context. Thus, I repackage Farkas &

Roelofsen (2017)'s notion of credence levels into a shorthand I term *expectation*.[1] Expectation in this sense is a measure of relative credence levels.

A speaker's expectation for a proposition (or event denoted by a proposition) is built off of her credence level in that proposition. But her credence level for a proposition is conditioned off of various other propositions in the context set. If a speaker believes a proposition q, we can say that her expectation in q is relatively high, but this expectation is not context free: it is relative to the rest of the propositions in her context set. So regardless of what her expectations are toward a particular proposition, whether she expects an eventuality or it violates her expectations, this must be calculated relative to her credence level toward other propositions in the context set. We can base an agent's expectation for a proposition q on the joint probability of that proposition and other ps in the context set. Expectation is high for q when all ps in the context set have a high joint probability with q. On the other hand, a single low joint probability between q and any context-set p should be enough to lower the expectation in q. This suggests the following definition for a speaker's expectation for a proposition q:

(36) Expectation for q by a speaker α:

$$\mathrm{Exp}_{\alpha}(q) = \min_{p \in CS} P_{\alpha}(p, q)$$

We can say that a speaker α believes or expects q if $\mathrm{Exp}_{\alpha}(q) \approx 1$. A speaker's subjective epistemic bias can then be very easily represented as an inequality in expectations for a proposition's truth. We can say that $\mathrm{Exp}_{\alpha}\ (\varphi) > \mathrm{Exp}_{\alpha}\ (\neg\varphi)$ is a way of talking about a speaker's non-categorical belief in φ, which is formulated in terms of their expectations. This will ultimately prove to be a useful way of talking about surprise, which in the very general case, will turn out to be an instance of a speaker's violated expectations. We will return to this notion more in the next section, as well as sections §2.2.4 and 2.4.

[1] Expectation, as I use it here, is a technical term, and should not carry the same intuitions and presuppositions that might go along with the use of the word *expect*.

2.2 Refining the role of expectation

Much of the data to be discussed in the coming chapters depend on certain assumptions about how discourses are structured. In particular, the data depend on how individuals in a discourse navigate and update their beliefs and views of a conversation through mutual exchange of information. Semantic models of discourse generally attempt to capture aspects of dialogue that are part of the lexical or compositional meaning of the utterances, as well as those parts that are left unsaid. This means that many models of discourse generally subscribe to a notion of the conversational common ground. They may also want to be able to talk about the commitments made and held by each participant, and how these commitments can shape the upcoming discourse moves.

Such approaches focus on explicit contributions to a discourse, and the strategies that conversational agents use to navigate through a conversation. Building a common ground is like merging distinct networks of reasoning collected from each participant into a single coherent representation of the discourse. But being able to do this warrants a closer look at an individuals' own network of knowledge, belief, and expectations about the current state of the conversation. This section aims in part to expand on the role of expectation in a discourse, particularly when it comes to expressing surprise at some turn of events.

Talking about surprise as it relates to expectations must make reference to the changes that can be made to an agent's internal belief state after the receipt of new information. An agent's belief state is not rigid: beliefs and expectations are constantly changing and updating over time as experiences and observations about the world grow. Given this, a model of real-world learning patterns should be able to represent changes that take place without disrupting the entire system.

Perhaps you have built up the following set of beliefs and inferences from observation: You see Sophie walk into Mission Hill Creamery, a business renowned for their ice cream. While she is inside, she interacts with the person behind the counter, selects a flavor, asks for sprinkles, pays for the treat, and proceeds to eat it. From this, you can form a set of expectations about Sophie and her ice cream habits based on information, beliefs and inferences. Seeing her at the ice cream shop might allow you to form a high expectation that she likes ice cream. Seeing her choose sprinkles might lead you to believe that she also likes sprinkles. You know that she has been in an ice cream shop, and you have seen her pay for and eat her treat, which, given the evidence you have access to, might be enough for you to form the belief that Sophie likes ice cream. While this is something that you don't know for sure, your causal knowledge allows you to conclude with fairly high expectations that once verified, it will turn out to be true.

But though these expectations may have been formed from solid reasoning based on knowledge, belief, and stereotypical ways in which the world works, the *expectations* themselves you have formed can be challenged. In the cases that I focus on in this work, this expectation violation leads to the expression of surprise.

Going back to our toy scenario, suppose that it was Jeff who observed Sophie's ice cream habits. Later, he and Tom discuss where they should meet Sophie for dessert. Jeff suggests Mission Hill Creamery. Tom, knowing that Sophie is allergic to dairy, can respond in surprise, and Jeff, having seen Sophie at the ice cream shop, can also express his surprise at this new information:

(37) **Jeff:** Let's meet Sophie at Mission Hill.
Tom: (But) Sophie's lactose intolerant![2]
Jeff: (But) I saw her there last week!

Tom and Jeff's expectational models differ here, but are not irreconcilable. Tom, knowing of Sophie's lactose intolerance, believes that meeting at an ice cream shop would exclude Sophie from a sweet treat, while Jeff has seen her eating what he had assumed to be frozen dairy. What happens, then, when both turn out to hold valid inferences? The truth of the matter could be that Sophie had ordered a lemon sorbet, avoiding the question of dairy entirely. In both cases, Jeff and Tom must update their doxastic states. Tom, whose expectations based on Sophie's ability to digest dairy were correct, must somehow integrate the fact that there are non-dairy options at an ice cream shop. Jeff, who had seen Sophie at an ice cream shop and

[2]While including the counterexpectational *but* is natural here, it is not necessary with the right prosody. I set the issue of this marker aside here, and instead refer to Toosarvandani (2013) for a discussion of this.

had assumed she could eat dairy, must reevaluate his inference network in order to allow for her to be eating non-dairy frozen desserts. In both cases, the agents' belief states must be flexible enough to accommodate this new information, and must be able to reflect this update from a previous to a current doxastic state.

2.2.1 Approaches to modeling violated expectation

Returning back to expressions of surprise, how is it that we can represent the violated expectations that we see in examples like (37)? Intuitively, it seems that this could be done by comparing a speaker's doxastic state before and after learning some new piece of information. Surprise could be modeled as trying to integrate a proposition that is not predicted by entailment relations given the information in a speaker's prior doxastic state. We want to be able to capture the surprise that Tom or Jeff may experience in (37), and for a more run-of-the-mill situation of surprise, as in (38):

(38) *The speaker, having no expectation for rain, steps outside a windowless building and sees that it's raining:*
I'm surprised that it's raining.

One way to represent this surprise could be by constructing an analysis based on a counterfactual reasoning strategy. An analysis of this sort would be akin to taking the union of speaker's current doxastic state and a counterfactual doxastic state, in order to model a speaker's current (expectation violating) belief about p, while also making reference to a counterfactual situation in which $\neg p$ were true, in effect, imagining how it might be if it were not raining, while also being aware that it is.[3] Imagining surprise in this way is not entirely different from Beltrama & Hanink (2017)'s *belief widening*, an extension of the Lasersohn (1999) idea of a proposition's pragmatic halo. While their analysis is specific to the English mirative marker *like*, the goal of their proposal is similar: we must somehow capture the contribution of surprise to a discourse while also widening the contextual restrictions relative to an agent's doxastic

[3]Thanks to Maribel Romero, p.c., for discussion of this alternative.

state. What we need here is a way to talk about a speaker's subjective expectations, and what it means when these expectations are violated.

For the situations in (37) and (38), the speaker receives a piece of information that in some way violates an expectation (or network of expectations) that they have formed. Trying to integrate this information as it is causes a clash, and some belief restructuring must happen. In the case of (38), perhaps the speaker went to work in the morning, and it was sunny. In the evening, when she is finished, unbeknownst to her, the weather has changed dramatically. She walks outside with one doxastic state, not predicting that it will rain, and is faced with a new fact that she must update her beliefs with. The surprise here could be characterized by widening the reasonable worlds of evaluation to those which were previously excluded.

Yet formulating violated expectations in terms of comparing a previous and a current doxastic state after learning something new could also run into an issue. Learning a new fact might not be the surprising element in your reasoning system. Rather, your surprise could come from implications that follow from it, given the way that you have assumed the likelihood of particular outcomes based on the information that you have. That is, there are particular cases of surprise where your doxastic state *did* entail all of the necessary information to formulate a particular expectation, but your evidence favored another outcome more highly. In these cases, it's not about not being able to integrate a new fact into your pre-existing belief network. Instead, the expectations that you have calculated relative to various facts and beliefs lead you to expect one outcome over another.

Assume for a moment that you are named Deniz, and you are part of a family that generally names children after other members of the family. Your doxastic state includes variables that you take into account when forming expectations about how you were named. It includes background knowledge like the fact that you are male, or knowing what all the names of your family members are. It also includes variables that relate to background knowledge, such as knowing about this naming practice, believing certain things about how your name relates to the names of your relatives, and perhaps knowing that your name is gender neutral. From these facts, you can form sets of expectations. Perhaps because you are male, you assume that is is the case that you have been named after a male relative. You know that your father was close with his brother, who shares your name. You also know that you had an aunt with your name, but know nothing more. From these two paths, you might infer that being named after an uncle is a more likely outcome. With no further information, you set up the expectation that you were named after your father's brother, even though another valid inferential path is available. At a family gathering, your mother reveals the information that you were in fact named after the aunt, contrary to your expectations. You don't need to integrate a new proposition or even widen your doxastic set— your information was not faulty, nor was it the case that your network of reasoning is somehow flawed. Rather, your internal calculus about the likelihood of who you were named after and the actual reality of the situation turned out not to accord with your expectations.

In such a situation, you can reasonably be surprised by a fact that your model of the world already predicts. You can even utter something like one of the alternatives in (39), indicating that surprise:

(39) a. I'm surprised you named me Deniz!

b. You named me Deniz!

With certain situations of surprise, it may be difficult, though not impossible to continue with a counterfactual or doxastic state-widening analysis. I take a different route here, and instead look at surprise in terms of causal reasoning. Because surprise in many cases involves reasoning from a cause to an outcome, I choose to extend the ideas of causal modeling in order to conceptualize expectation and expectation violation (Halpern & Pearl 2005, Baldi & Itti 2010, Halpern & Pucella 2007). The choice here is at its core functional: surprise and expectation violation in many varied cases involve chains of reasoning from knowledge and beliefs to expectations about the world. Using this type of reasoning system has the advantage of making these causal connections very clear.

2.2.2 An interlude: Surprise as abductive reasoning?

Before we delve into a causal model solution for modeling surprise, it is important to pause and reflect on other avenues that might lend themselves to capturing the contribution of surprise in a conversational scenario. Because surprise and violated expectations in many cases do involve chains of reasoning and belief revision, one might conclude that expressions of surprise might be modeled as a way of abductive reasoning. When a speaker is surprised, they tend to look for a reason for that surprise. Abductive reasoning takes an effect, and tries to use general knowledge about the world to reason back to its cause. If I walk into a room and see a spilled cup of coffee and a very enthusiastic golden retriever, I might reason abductively about the cause of the spilled coffee. Since dogs are generally unaware of their surroundings (especially very excited ones), I conclude that the golden retriever is the cause of the spilled coffee.

Abductive reasoning infers a resulting event or outcome from the most likely or "best-fit" explanation for that outcome (Peirce 1955, Winans 2016). Reasoning in this manner allows an inference α to be an explanation of a particular outcome β, without necessitating that the premise (β) entail the conclusion (α). Rather, the conclusion is simply a likely cause for the premise. Methods of abductive reasoning can be valid ways of determining a cause, but they are not *logically* valid. Below is a template outlining reasoning of this sort, as well as an example of abduction, taken from Winans (2016):

(40) **Abductive reasoning (taken from Peirce 1955)**

The surprising fact *P* is observed.

But if *Q* were true, *P* would be a matter of course.

Hence, there is reason to suspect that *Q* is true. *(Winans 2016: p.40-41)*

(41) **Abduction**

The coffee is spilled.	*P*
When golden retrievers are excited, they knock things over.	If *Q*, then *P*.
Therefore, the dog knocked over the coffee.	Therefore, *Q*.

While abductive inferences may turn out to be actual explanations for a result state, it is more helpful to think about this kind of reasoning as proposing a hypothesis for the cause of an event or observation. As hypotheses can turn out to be validated or invalidated, so can the inferential causes of a situation of abductive reasoning. From this, we could formulate a theory of violated expectations based on of abductive reasoning. Since this reasoning builds in the potential for the premise to not entail the conclusion, surprise could be captured in just those cases where inferring a cause from a resulting effect turn out to **not** model the real world. In other words, we express surprise in situations where we havereasoned abductively, and we later are shown that this reasoning strategy did not actually reflect reality. This could be the case for a situation like (42):

(42) *You read the weather forecast in the morning and see that it calls for sun today. You assume that this means it won't rain, so when you walk outside and see that it's raining, you are surprised:*
(I'm surprised) It's raining!

Here, your faulty pattern of reasoning could be captured abductively, as in (43):

(43)	The weather forecast calls for sun.	P
	If it rains, it is not sunny.	If Q, then $\neg P$
	Therefore, it won't rain.	Therefore, $\neg Q$

Expectation violation is then calculated by your initial expectations being based on a faulty premise—you expect there to be sun, which you have assumed precludes the possibility of rain. When it does indeed rain, you are surprised: your pattern of reasoning did not predict this outcome. You had assumed a false conclusion.

But violated expectations are not always straightforward. In a situation like (44), there is no evidence that you have done any previous abductive reasoning at all. You have not assumed a false conclusion, since you have assumed no conclusion at all, which fails to predict a surprised reaction:

(44) *Speaker, who has not read the weather forecast, walks out of a windowless building to see that the streets are wet:*
It was raining!

Here, a speaker's surprised response is not due to a false hypothesis made through abductive reasoning. Instead, the speaker has simply reasoned abductively to infer that the streets are wet as a result of the rain. You have drawn an inference about a plausible explanation for a surprising event, and you express violated expectations *at that inference*, not at the fact that the inference is faulty.

When the surprising observation is removed from consideration, all the ways in which you came to learn that proposition are also presumably removed as well. If you have not observed the fact that the streets are wet, you cannot be surprised about it having rained. If you remove the expectation-violating event, you also remove the reasoning for why the event would be surprising in the first place. For the case in (44), uttering p would at once introduce the surprising outcome, but also say something about the calculation of its cause. This is not an inference that we draw in unbiased cases of surprise like in (44).

Given these instances of expectation violation, it is clear that this calculation should be made relative to how your network of beliefs has been structured. Another case where relying on an abductive-style reasoning strategy fails is in (45):

(45) *You are a meteorologist, and you have you been in the lab watching your instruments all day. The pressure has been steadily rising in the barometer, and you have been watching this change happen. Though you trust your instruments, and you know the weather patterns, this change is still surprising to you. You know that this behavior in your tools means that there is a very high likelihood that it will rain later in the day. When it does indeed start to rain a few hours later, even in the face of all the evidence you have, you can still utter:*

a. I'm surprised it's raining!

b. I can't believe it's raining!

Here, it's not the case that your surprise is the proposition that it is raining. Based on the evidence that you have gathered throughout the day, this outcome should be predicted. Yet somehow, you are still able to be surprised. In this case, your surprise is not at a proposition, or even at a result state; it is quite simply at the particular chain of events that lead to this outcome.[4]

[4]An interesting real world example of the kind of situation in (45) was recently observed in the author's social media feed. A friend posted the following update: "Having an I-don't-know-why-I'm-surprised-when-I-saw-this-coming kind of day" (23 April, 2018).

2.2.3 A simple causal model for surprise

The introduction to Halpern & Pearl (2005)'s approach to causal models includes the following line: "Whenever we undertake to explain a set of events that unfold in a specific scenario, the explanation produced must acknowledge the actual cause of those events" (p. 194). The goal here is to make the connections between related events explicit. When reasoning about the world, there are many variables that we must consider all at once in order to come to the clearest picture. If we can make the patterns of causation explicit between these variables, we end up with a way of modeling direct and indirect outcomes that a particular agent's beliefs and knowledge predict. Put simply, if we can make a speaker's doxastic state and subjective epistemic bias explicit, we have a way of formally modeling expectation and surprise.

Before this goes much further, I want to make clear my goals in formulating surprise in this manner. First, this way of modeling surprise does not propose to contribute one way or another to the discussion about how to model subjective epistemic bias; there are a variety of proposals for how to best model this (e.g. a pair of a set of worlds and a probability measure, a set of probability measures, a probabilistic causal model, etc.—see Halpern (2005) for an overview). Though there is active debate in the linguistic literature about which of these different representations is best suited to the treatment of the semantics of epistemic language (see Lassiter 2017, Rudin 2017b: for an overview of this debate), the goal of this work is much smaller. I aim to provide a way of looking at expectation and subjective epistemic uncertainty in order to capture generalizations about particular natural language phenomena. Thus, to a large extent, the system that I use here is arbitrary, and generalizable to any theorist's framework of choice. My proposal does not hinge on the fine details of formal epistemology, and as such it would be a waste of our time to wade too deep into the weeds on this issue. What is important for our purposes is that all formal models of subjective epistemic bias provide a well-defined notion of expectation (see Halpern (2005), Ch. 4).

I analyze what we've been calling "surprise" simply as a form of asymmetric expectation. In the most basic case, $\neg p$ is surprising to the extent that the expectation for p outweighs the expectation for $\neg p$: $\text{Exp}(p) > \text{Exp}(\neg p)$. The more a speaker is subjectively biased toward a proposition—the more they expect it—the higher their expectation for it is. The formal notion of expectation will vary in its details depending on one's preferences for how to represent subjective epistemic bias, but not in ways that have ramifications for such simple formulas.

A simple way of representing a speaker's reasoning strategy is by integrating subjective bias with causal models (see Halpern & Pearl 2005, Baldi & Itti 2010, Schultz 2011, Lauer & Nadathur 2017: for more explicit notions of causal models than will be introduced here). Thinking about surprising events as reasoning chains can be easily represented graphically, as drawing lines and arrows from one state to another can make explicit the (sometimes subjective) causal links a speaker makes between propositions.

Causal models are generally used to show how certain variables can be arranged in order to model the likelihood of one event being the cause for another. How likely is it, for example, that a lit match is the cause of a forest fire, versus a strike of lightening? Which assumptions can we take for granted (exogenous variables), and which do we determine to be causally dependent on others (endogenous variables)? A speaker's knowledge and expectations can be modeled in this way as well.

The underlying reason for favoring a causal approach to belief structuring is the appeal toward being able to clearly reason about uncertainty. It assumes that the world can be described by random variables, but that some of this randomness can be predicted: some of these random variables are causally linked. Halpern & Pearl (2005) approach these connections by proposing a model of belief based on sets of *structural equations*, which represent the causal paths that can serve as explanations for a particular observation. These equations show the links between a set of variables, laying out various paths of inference that allow an individual to reason from one state to another. This is distinct from reasoning from causes to effects: focusing on the paths that may lead to a particular result state actually zooms out to the level of observation and asks, what can we infer about the interpretational possibilities and links between a set of facts, their connections, and the outcomes?

Imagine that we are back in the room with the excitable golden retriever and the spilled cup of coffee. How did the current state of affairs come to be? We must take all plausible factors into account. We must ask whether the coffee cup is realistically within reach of the dog, whether a door or a window might be open, or how much coffee was in the cup. More interestingly, we want to identify the set of variables that we hope to use to build causal relationships. In our case, these are things like the presence of the energetic dog, the possibility of the wind knocking over the cup, or perhaps the knowledge that someone else may have been in the room.

Most views of causally modeling the world assume that there is a relevant split between random variables. Exogenous variables are those that can be taken for granted (like whether the coffee cup is within reach, or whether the door is open). These are factors about the situation which are external to the model itself. Those that are more relevant to building causal relationships are endogenous variables, which are causally dependent on others (such as whether it was the dog or the wind that knocked over the coffee).

Halpern & Pearl (2005) formalize the relationships within a causal model as the relationship between a pair $(\mathcal{S}, \mathcal{F})$, where $\mathcal{S}$ is a *signature*, or list of endogenous and exogenous variables in the model, and $\mathcal{F}$ is a set of structural equations that relates the possible values of the variables. (46) outlines their other basic assumptions:

(46) a. A signature $\mathcal{S}$ is a tuple made up of $(\mathcal{U}, \mathcal{V}, \mathcal{R})$, where:
   i. $\mathcal{U}$ is a set of exogenous variables
   ii. $\mathcal{V}$ is a set of endogenous variables
   iii. $\mathcal{R}$ associates with every variable $Y \in \mathcal{U} \cup \mathcal{V}$ a non-empty set $\mathcal{R}(Y)$ of possible values for Y

b. $\mathcal{F}$ associates with each endogenous variable $X \in \mathcal{V}$ a function denoted F_X, such that:
   i. $F_X : (\times_{U \in \mathcal{U}} \mathcal{R}(U)) \times (\times_{Y \in \mathcal{V} - \{X\}}) \mathcal{R}(Y) \rightarrow \mathcal{R}(X)$
   ii. which is to say that F_X determines the value of X, given the values of all the other variables in $\mathcal{U} \cup \mathcal{V}$

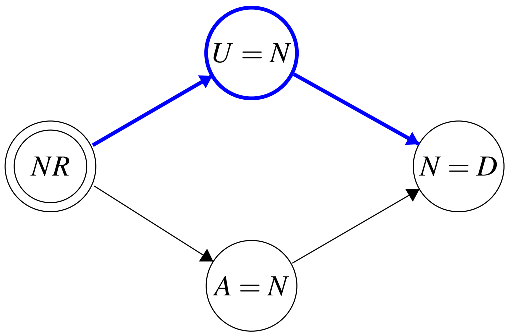
Halpern & Pearl (2005) also assume that these causal modals are acyclic: given a context and a unique setting for the exogenous variables in $\mathcal{U}$ and that there is a total ordering over endogenous variables, there is a unique solution for all equations. That is to say, there are causal paths from each node to others, and these paths are ordered.

Given a model and relations between exogenous and endogenous variables, patterns of causation can then be represented by underlying prior probabilities, which determines the likelihood of one causal pathway being a true cause of another. These connections can be represented graphically, with arrows between nodes in a graph showing the direction of causal relationships.

Returning back to expectation, it seems that causally modeling belief structures in this way builds in this notion for free. Expectation represents an individual's prior belief in a particular path of causation. It also allows for multiple live options for possible inference paths, inherently taking a wide view of a belief network. Surprise is a straightforward, contextually determined state that arises from a causal model structure. It arises in those cases where one setting of exogenous variable has set a particular scene, and given this, weights assigned to the endogenous variables produce a single path forward with a higher likelihood than the rest. Expectations are violated when the actual world lies unexpectedly along a different path. Given a network of related events, surprise is a the effect of a low-weighted causation path turning out to be the actual cause of a series of events.

Recall the naming scenario of surprise in (39). Deniz has assumed that he was named for his beloved uncle, not for his aunt, whom he knows nothing about. In this toy example, we can restrict the exogenous variables to just the fact that he knows that he has been named after a family member. He then has two causal paths of reasoning represented by two endogenous variables: either he was named for his uncle, or he was named for his aunt. We can represent this in the graph below. In both scenarios, the outcome, his name being Deniz, is a result of being named after one of the relatives, but the expectation for one path representing the actual world (marked in blue for ease of identification) is greater than the other:

When Deniz learns that he has been named after the aunt, his surprise is not due to the fact that his reasoning chain was somehow inadequate, but rather, that his subjective preference for a particular pattern of reasoning was wrong. Learning that he was named for the aunt did not require him to radically restructure his network of subjective epistemic beliefs; instead, in he



NR	Know you are named for a relative
$U{=}N$	Named after uncle
$A{=}N$	Named after aunt
$N{=}D$	Know you are named Deniz after a relative

Figure 2.1: A toy model of reasoning with surprise

needed to re-evaluate inferences paths he had previously predicted to favor another accessible, but less expected outcome.

In approaches to modeling language, talking about absolute probability calculations does not make much sense. Speakers do not have access to an exact numerical value for the paths in their causal networks. But modeling languages in probabilistic terms is not about the numbers themselves, but about how the mathematical concept of probability is a reasonable starting point for other types of reasoning strategies. This is the position of Yalcin (2010), Lassiter (2011), and Lassiter (2017), who argue that reasoning in natural language systems can be achieved through the same assumptions that underly statistical Bayesian inference. I take much simpler stance, and argue that for pragmatically modeling violated expectations, simple relative probabilities is enough.

With this toy version of a causal modeling framework, violated expectations can be looked at in terms of an unexpected observation given a prior expectation above a contextually defined threshold. Violated expectations arise from deviations from the expected value to the observed value. It is what happens when a subjective bias for one outcome over another turns out to not represent the observed state of the real world.

2.2.4 The threshold of surprise

Another aspect of surprise is the degree to which one's expectations must be violated in order to feel surprised. If two outcomes are equally likely to happen, and there is no outside bias for either to happen over the other, it seems odd for a speaker to express surprise when she learns of the outcome. Imagine that Sophie watching a foot race between Tom and Jeff. She is unaware of either of their athletic abilities, and as far as she is concerned, both have an equal chance of winning. When the race is over and Jeff is the winner, it is odd for her to utter (47):

(47) #I can't believe that Jeff won!

But perhaps Sophie knows that Tom was a sprinter in high school, and competed at the national level in the 200 meters. Jeff, on the other hand, never competed in sports. In this case, when the two are up against each other in a race, Sophie expects Tom to win, given the evidence she has that he is an elite runner. With these facts and expectations in place, if Jeff wins the race, it becomes completely normal for Sophie to utter (47), repeated in (48):

(48) I can't believe that Jeff won!

Here, the threshold for surprise must be fairly high, or at least above chance. Other situations may set the threshold of surprise at the level of contextual stereotypicality. Suppose you walk into your office and see a bakery box on your desk. Though you have no prior calculated expectations about the presence of donuts on your desk, the box is still surprising: it is a deviation from the norm. In the usual case, your will walk into your office and it will be as you left it. But suppose that every day after this, you receive donuts on your desk. You will likely continue to be surprised by this for a while, but at some point, that surprise will fade. Your calculation of what is 'normal' will begin to include a box of donuts on your desk. Thus the threshold for surprise is both context dependent and malleable as well. Things that are at first surprising can cease to be so: surprise is conditioned on a context and subjective biases.

Though the two examples above are fairly simple, it is clear that setting the threshold of surprise is extremely context dependent. One way of thinking about how surprise might be calculated appeals to probabilistic notions. For the purposes here, this need not involve exact numerical calculations of a conversational agent's prior expectation and their updated doxastic state, given a new piece of information. This seems artificial, as it is not the way in which the typical agent goes about calculating her expectations about the world. Though we may be rational agents, our Bayesian reasoning is not as precise as a statistical model, and is influenced by factors that are at times non-rational. But probabilistically modeling a speaker's change in doxastic state after learning a surprising new fact does have advantages.

In the Bayesian tradition, we can model expectation as the relation between a speaker's prior beliefs and their updated posterior beliefs, given a piece of new information (see Halpern & Pearl 2005, Baldi & Itti 2010, Halpern & Pucella 2007, Yalcin 2010, Lassiter 2012: for fully specified Bayesian probabilistic models.). A speaker's prior doxastic state can be determined by a probability distribution over a set of possible worlds in a hypothesis space. A speaker assigns some probability to a part under the curve where they expect the real world lies, and this is updated when observations are made. 'Surprise' can be measured by how much the updated belief distribution has had to change relative to the the prior one.

For example, imagine that two speakers, Tom and Jeff, are talking about Sophie's ice cream habits again. Recall that Tom, but not Jeff, knows that Sophie has recently had to give up ice cream due to lactose intolerance. Tom can relay this information to Jeff, and Jeff's surprise is relative to how likely he had expected Tom's contribution to be. Since we know that Jeff was surprised by this, we can assume he had assigned a high probability to a proposition with conflicting entailments.

Formally measuring surprise in a fully specified system amounts to measuring how much an observation deviates from the subjective bias of the speaker. If Jeff, for example, also knows that Sophie is lactose intolerant, he will not be surprised about the information that Tom relays to him, and the posterior distribution of his belief state (which can be calculated from the likelihood of the model to include the real world and the new observation) will not deviate much from the prior. Jeff will integrate this information into his belief set, and all will be as expected. But if Jeff did not know this, it represents a change from his prior beliefs. If that change falls outside of the expected distribution, we categorize this information as surprising.

But while these notions are useful for actual, measurable observations of surprise, a pure Bayesian theory of surprise and expectation violation is difficult to use to model non-experimental cases of expectation violation and belief revision. When Jeff is surprised by the information he receives, it is not clear how the weights of expected values should change to represent this, or even what exactly the weight assigned to this particular outcome was in the first place.

Determining a threshold for surprise should also include reference to a degree of surprise. While this is outside the immediate scope of the project, it is important to note that languages have ways of expressing different amounts of surprise. This can range from slight surprise (perhaps only a slight deviation from an expected outcome), as well as moderate to high degrees of surprise. In English, this can be marked by prosody, discourse markers, or particular turns of phrase:

(49) *Speaker opens a box to reveal a puppy inside:*

a. Huh.

b. Wow!

c. Oh my God!

Returning to the question of the threshold or degree for surprise: can we calculate it? The answer here is not easily. Surprise not only depends on context, but also on the doxastic state of the surprisee. Luckily, this is not my enterprise, and it is not crucial to the phenomena I describe later. I offer no clear answer on this matter, and instead assume that a speaker's threshold of surprise relies crucially on contextual information.

2.3 The table model of discourse management

Probabilistically modeling a speaker's expectation is only one part of representing discourses that involve surprise. Conversation is of course a two-way street, and must be able to capture not only the expectations of particular discourse participants, but also the contributions and commitments that agents make to the common ground. The following is an outline of a model of discourse that will be used in coming chapters to capture the relations between commitments speakers make during the course of a conversation, which includes the expectations that they make salient through language use.

Recall Tom and Jeff, who are discussing Sophie's ice cream habits. Once more, Tom, but not Jeff, knows that Sophie no longer eats dairy due to her lactose intolerance, and as a result, doesn't go to Mission Hill Creamery anymore. Tom tells this to to Jeff, and if Jeff accepts the contribution without hesitation, this information can become common ground. There are many ways to go about capturing this information in a model of discourse.

Farkas & Bruce (2010)'s Table model of discourse is a system developed to model commitments, salience, and common ground management for a given discourse structure *K*. This system developed out of the works of Gunlogson (2001), which draws heavily on the ideas introduced by Stalnaker (1978)'s work on common ground management. The Table model grows out of the desire to decompose the act of assertion into specific effects that it has on a discourse that do not become part of the common ground, and also grows out of the desire to unify the way in which declaratives and interrogatives are formally represented. The original Stalnakerian model comprises a Context Set and a Common ground, which represents the shared public commitments of discourse participants. Farkas & Bruce (2010) propose the addition of a stack of at-issue propositions called the *Table*, sets of *Discourse Commitments* indexed to particular speakers in the conversation, and a *Projected Set* of future common grounds that are accessible from the current contextual state. The act of assertion can then be defined by the effects that it has on these three elements of the model.
In summary, the important components of the Table model are presented in (219):

(50) a. **Table:** s stack of issues modeling salience in the current context

b. **Common Ground (*cg*):** the set of all propositions that all members of the conversation are publicly committed to during the course of the conversation

c. **Discourse Commitments (DCs):** for every participant *x*, the set of propositions that *x* has publicly committed to, but which are not yet part of the CG

d. **Projected Set (*ps*):** the set of propositions currently under consideration for addition into the *cg*

e. **Context Set:** the set of all worlds that are compatible with the propositions in the Common Ground

Participants in a conversation move forward through a shared commitment to growing the common ground, and thereby shrinking the Context set. This is done when speakers raise issues (sets of propositions), and place them on the Table for consideration. Issues are resolved by clearing them from the Table.

Assertion in this model is an act by a speaker that raises an Issue, $\{p\}$, which places that issue on the table, and also updates a the speaker's Discourse Commitments with p. Farkas & Bruce (2010) define this in terms of a function from an input context K_i, in which an author a asserts a sentence s that denotes a proposition p, to an output context K_o:

(51) $Assert(a, p, K_i) = K_o$ such that: *Based on Farkas & Bruce (2010), p. 92*

a. $\mathrm{DC}_{a,o}$ = DCa,i $\cup$ $\{p\}$

b. $\mathrm{Table}_o = \mathrm{Table}_i + \{p\}$

c. $ps_o = ps + \{\mathrm{cg}_i + \mathrm{p}\}$

The table in (53a) represents the state of the discourse before a conversation begins. There is nothing in either Tom or Jeff's discourse commitments (DCs), and the table is free of issues. When Tom raises the issue in (52), the context is updated in the following ways, schematized in (53):

(52) **Tom:** Sophie can't go to Mission Hill Creamery anymore.

The context is updated with his DCs, the issue is placed on the table, and the projected set (ps) shows us how the common ground could be updated, should Jeff accept Tom's proposal. Here, s_0 represents the initial state of the common ground.[5]

[5]This initial state need not be empty. It is likely full of previous issues that Tom and Jeff have raised and agreed on. This is also a simplification of each participants' discourse commitments, as we can assume that Tom and Jeff are not totally in agreement about everything up until this point. There is if course some simplification here when the initial discourse context in (53a) is presented with completely empty DCs. Actual discourses do require some enrichment.

(53) a. Context before utterance:

DC$_T$	**Table**	**DC$_J$**

$cg : s_0, ps = \{s_0\}$

b. **Tom:** Sophie can't go to Mission Hill Creamery anymore. $= p$

c. Update context with p:

DC$_T$	**Table**	**DC$_J$**
p	$\{p\}$	

$cg : s_0 = s_1, ps = \{\{s_0 \cup p\}\}$

If Jeff accepts this proposal, p would be cleared from the table and added to the common ground, and p would become a joint discourse commitment shared by Tom and Jeff. If, however, he rejects Tom's proposal, the common ground is not updated, and Jeff's DC's are in turn updated with $\neg p$. This is the effect of an assertion in general: a proposal to add the content of an utterance to the common ground, and public commitment to the proposition.

In cooperative conversations, assertion also requires speaker commitment to the truth of the utterance. When a speaker asserts a sentence, she must at the very least act as though she believes in the content of her utterance. But what does the particular intonation of an utterance contribute? This is a controversial area. Whereas propositional content that is added to the table is a proposal that must be agreed upon by other discourse participants, the fine-grained pragmatic contributions of an utterance's performance and the context that it has been performed in are left out of this model.

Going back to our interaction involving Jeff, Tom, Sophie, and the ice cream shop, suppose that neither Jeff nor Tom is aware of Sophie's recent diagnosis of lactose intolerance. The three have planned to meet at the shop, but then Tom receives a message from Sophie, informing him of this change. If we use this updated context, the Tabletop model of discourse as it stands cannot capture the pragmatic calculations that Jeff is forced to compute to make sense of Tom's utterance. In the presence of Jeff, Tom can utter (54), indicating his surprise at receiving this message, while also putting this proposition on the table for Jeff to consider:

(54) **Tom:** Sophie can't go to Mission Hill Creamery anymore! $= p$
Extra inference: *Tom is surprised that p*

If we plug the propositional content of (54) into the model of discourse, we get the same calculation as we do in (53). Crucially, we don't capture the surprise that Tom expresses with this utterance, which is integral to Jeff's understanding of how it fits into the larger context. The intonation is working on the assertive speech act itself here, fine-tuning the exact sincerity contribution that Tom makes. By manipulating the pitch and melody of his voice, Tom pragmatically indicates something that the semantics of the utterance doesn't, which he makes clear to his addressee; he modifies the sincerity conditions on his utterance. This is an important piece of information that the model as it stands does not capture.

In general, when a speaker uses a falling tone, characteristic of unmarked declarative utterances, she presents herself as believing the content of her utterance. If she believes the content of her utterance, she must also *expect* her utterance to be true as well. In an utterance like (54), the speaker's tone is also overall falling, but the final fall has a higher pitch excursion than normal, signaling to the listener information both that the speaker is asserting this proposition, as well as expressing surprise about it.[6] That surprise is speaker-oriented, and it can't be questioned by the listener. Given Tom's reaction in (54), it would be infelicitous for Jeff to respond targeting Tom's surprise:

(55) **Jeff:** #You're not surprised to learn that.

Given this minimal pair, one can conclude that intonation in English is a way for the speaker to comment on their privately held beliefs and expectations. If we assume that intonation is another factor to consider in our game of discourse management, we should also be able to represent this in the model. The particular intonational contours that we use are a way for us to comment on our expectations for how a conversation should proceed, and how we are modifying the commitments we make in a discourse. Discourse commitments can be calculated from multiple different pieces of the performance of an utterance.

[6]See Chapter (1.2.2) and (1.3) for more on the contributions of intonational contours.

One level of a speaker's discourse commitments comes from the act of uttering semantic content itself. This equates to placing a proposition on the table, and marking the proposition down in a speaker's DC list. Another way material can be added to a speaker's list of DCs is through the particular prosody of their utterance. Here we must make clear that this content is not added to the table; instead, it is logged directly in the discourse commitments of the speaker. It is not a proposal to be accepted, rejected, or challenged by other participants in the conversation. Rather, it forms a part of the not-at-issue content of an utterance, as commentary that the speaker makes public, but non-negotiable by the other players.

If we add these expectations to a speaker's DC list, we can assume that an utterance level, final falling tune commits the speaker to having a high expectation of truth toward their utterance:

(56) The neutral H*L% contour adds the following to the speaker's DC:
$\text{Exp}_{spkr}(p) \approx 1$

The difference between (53) and (54) lies in the expectation calculations that are registered in the DCs of the speaker. Adding this information further specifies how a speaker expects her contribution to a conversation to be integrated into a discourse, and how she might signal this with her intonation. Information contributed by intonation is added to that participant's list, and integrates a speaker's comments about propositions into the greater model of discourse: addressees are able to see and compute a speaker's beliefs or expectations surrounding a particular proposition, but they need not commit to them themselves.

In this work, I will use this model of discourse to represent how a speaker integrates expectations into a discourse context. In a way, this is a conscious choice to pick a straightforward model of discourse and add a level of complexity. This could have been done to fit with any model of choice. The main message here is that integrating a speaker's expectations about propositions in their immediate discourse environment is an integral part of understanding the pragmatic interpretation of intonation. As we will see in the chapters to follow, I argue that it is crucial to understanding the contributions of discourse particles as well. The remaining sections of this chapter focus on how to fit expectation into the larger picture, which in this case, concerns itself a great deal with mirativity and evidentiality.

2.4 Violating expectation

2.4.1 Evidence and sourcehood in theories of discourse

Many current theories of discourse are based on notions of evidence and sourcehood, with particular interest paid to which participants in a conversation can be considered the source of information (see, among others, Gunlogson 2001, 2008, Farkas & Roelofsen 2012, Northrup 2014). "Source" here refers to a participant who has committed to particular propositions, and on particular evidence. In languages that have overt evidential marking, this distinction is crucial. A language might use one mark for information the speaker personally has evidence for (therefore fully committing a speaker to the truth of a proposition), and another to mark information that has been gained second-hand, distancing the speaker from the truth or falsity of it. Languages like English can mark evidence by way of sensory verb constructions (57a), adverbial modification (57b), or parenthetical "illocutionary evidential" strategies (57c) (Matthewson et al. 2007), which are verbs that introduce speech act operators suggesting a speaker's evidence:

(57) a. I **heard** Jeff's in town today. *Sensory verb*

b. Tom **apparently** broke his laptop. *Adverbial modification*

c. Sophie's home now, **I gather**. *Illocutionary evidential*

Using these strategies of evidence marking allows a speaker to commit to a weaker version of a potential proposition: they indicate that the speaker is not the original source of this information. Such utterances are contingent on another source for validation—as a result, they behave differently from propositions a speaker has fully committed to. Northrup (2014) notes that assertions that indicate this kind of evidence marking are resistant to direct contradiction, though they can be negotiated. In these instances, speakers seem to be committing to both a proposition and the kind of evidence they have. The evidence component is to some extent not-at-issue.[7]:

[7]Not all subtypes of evidential strategies are subject to this restriction. Murray (2017) shows that epistemic

(58) a. A: I heard Jeff's in town today.

B: #No, you saw him. / #No, he isn't. / # No you didn't. / ✓ Who told you?

b. A: Tom apparently broke his laptop.

B: #No, you watched him do it. / #No he didn't. / ✓ Weren't you there?

c. A: Sophie's home now, I gather.

B: # No, you saw her walk in. / #No she isn't. /✓ I heard footsteps, too.

Northrup (2014) draws a parallel between commitment and evidence in these cases, specifically, that commitments made by participants in a discourse are underwritten by a default evidential base. This source of evidence is known to the speaker alone, and is based on her experiences and her beliefs. She can use this base to lend credence to her discourse contributions. Speakers appeal to a default base when they have enough evidence to fully commit to their discourse contributions, conditioning non-evidential marked utterances. When the speaker is not sure of the truth of the utterance, or she does not have enough evidence to fully commit to it, she defaults to a dependent base, which contains commitments made by other discourse participants.

A dependent or 'weak' base indicates that a speaker has insufficient evidence off of which to base her claims. The speaker may defer to the addressee's authority, but need not. Crucially, when a speaker uses a weak base, she indicates that her contribution cannot be entered into the discourse record without reference to someone else's commitments.

When issues are raised in a conversation, they come underwritten by a particular conversational participant (see Gunlogson 2008, Roelofsen & Farkas 2015). When a speaker accepts a proposition contributed by another actor, it creates an imbalance in commitments: the speaker's commitment to that proposition is dependent on her interlocutor's commitment. Though this is a gross oversimplification of Northrup's theory, it does underline the fact that our

evidentials can be targeted as part of the at-issue content in Cheyenne, much like the English modal 'could' in an epistemic context. Though the modal invites the same inference of lack of evidence, the modality can be directly targeted as at-issue content:

(1) Tom: (*hears knock at the door in Santa Cruz, CA*) That could be Jeff!
Sophie: No, (you know) Jeff's in Cabo.

We set these instances aside here.

evidence for a particular proposition is a way of managing the commitments made by participants in a discourse. On this view, we might ask how expectations relate to an agent's evidence in a conversation. In particular, how does evidence interact with the expectations a participants has about the state of a discourse? To answer this, we must have a better notion of what separates evidence from expectation.

2.4.2 Integrating expectation into discourse representation

Marking expectation, and in particular, violated expectation, can be used as a grammatical device to underline a particular pragmatic effect a speaker wants to make with her utterance. Many languages have grammaticalized the notion of surprise or expectation violation through particular marking on the verb, on nominals, or as freestanding morphemes in an utterance. Broadly, these strategies are subsumed under the term *mirativity*. Miratives are also often grouped together with evidential marking, which begs the question: are miratives sensitive to a speaker's evidence, or their expectations?

Miratives are generally thought to grammatically encode a discourse participant's epistemic state at the time of utterance. Aikenvald (2012) shows through a detailed look at many languages and a number language families that time after time, mirative markers overtly map a speaker's surprised reaction toward a proposition to the current state of the common ground. As these markers are audible parts of the speech stream, these reactions are recorded in the discourse in real time. Mirative-marked utterances are marked discourse moves: by adding emotive content to an utterance, a speaker stacks a pragmatic level of communication on top of the literal semantic contribution. The default state of the discourse assumes that there is some baseline range of expectation that most utterances will fall into. Speakers are not expressing violated expectations with every discourse move. Miratives are a way for a speaker to convey that some proposition or eventuality falls outside of that normal baseline range; English mirative strategies express these ideas by way of discourse particles and intonation. Consider the difference between (59a) and (59b):

(59) a. Jeff ate the whole carton of yogurt.

b. Wow! Jeff ate the whole carton of yogurt!

(59a) is a neutral assertion which is not informationally out of the ordinary. Assuming neutral intonation, we cannot read much into any pragmatic-level contribution of the speaker. (59b) on the other hand is much more expressive. While the speaker conveys the same message as in (59a), the addressee computes an extra pragmatic effect: the speaker is surprised by the content of their utterance. This is a mirative strategy in English.

There is a tendency in the literature to classify miratives as a subclass of evidentiality. But though they seem connected on the surface, mirativity and evidentiality are distinct in a number of ways. Miratives can mark surprise, unprepared mind, expectation violation or newly received knowledge in the languages that they are attested in, but they are unified by their ability to convey new or unexpected information (Aikenvald 2012). Evidential systems are used in many languages to mark ways that a speaker has *come to know* a piece of information. Turkish, for example, marks whether the speaker was the original source of some piece of information, or whether the information was gained second-hand. Other languages, like Pomo, have sensory evidentials that distinguish visual evidence from evidence gained through other senses (Mithun 2001). Evidentials are a way to mark *how a speaker has gained her evidence* for a particular proposition. Mirativity marks *how a speaker integrates new knowledge with old knowledge* in a discourse, and how that may affect the expectations she (and possibly her interlocutors) have built. As in (59b), a speaker need not specify her evidence source in order to express her surprise.

2.5 Expectation: mirativity, evidence and exclamations

2.5.1 Expectation and Evidence

In many languages, mirativity is marked in the same way that evidentials are, even though speakers agree that the uses are completely distinct. Turkish is an example of this type of system. The language has a default past tense marker that commits the speaker to having direct evidence for her proposition (*-dI*) as well as a a past tense marker *-mIş* that marks the speaker as having indirect evidence (Göksel & Kerslake 2005).[8] But this indirect past tense marker can also surface as a mirative in contexts where an evidence source is obvious to all parties.

(60) a. *Paul parti-ye gel-iyor-du.*
Paul party.DAT come.PROG.PAST
Paul came to the party. *(I saw him.)*

b. *Paul parti-ye gel-iyor-**muş**.*
Paul party.DAT come.PROG.EVID
Paul came to the party *(I gather.)*

In (60), *-mIş* comments on the fact that a speaker has gained her evidence second hand. Yet a mirative interpretation of the utterance in (61) is only licensed when the speaker has not expected her addressee to be home early. When the addressee walks in unannounced, the speaker uses the indirect evidence marker *-mIş* to indicate her surprise, even though she has direct visual evidence of the events before her.

(61) *Erken gel-**miş**-sin!*
early come.EVID 2SG
(I'm surprised that) you're here early! Meriçli (2016), p. 18

The evidential-mirative system also overlaps in Cheyenne in much the same way we find in Turkish. The marker *neho* is both a narrative evidential, appearing as a part of a story in the remote past, in which the speaker indicates that she has indirect evidence for the proposition. As a mirative marker, it indicates surprise, regardless of whether the speaker has direct or indirect evidence for a proposition.

(62) a. *É- x- hoo'kohó- **neho**.*
3sg REM. PAST rain NARR.SG.INAN
It rained long ago, it is told.[9] Narrative Evidential

b. *É- hoo'kohó- **neho**.*
3sg rain NARR.SG.INAN
It's raining! Mirative

[8]Here I follow the convention set by many grammars of Turkish by marking vowel-harmonic sensitive vowels with 'I', meant to represent an underspecification of backness.

[9]Glosses: 3sg = 3rd singular, REM. PAST = Remote Past, NARR.SG.INAN = Narrative singular inanimate

This pattern in seen in many languages throughout the world that have small evidential systems, where firsthand knowledge is distinguished from non-witnessed or hearsay forms, and in systems where these distinctions are expressed by three or fewer forms (Aikenvald 2004). The same patterns can be found in Cuzco Quechua, (Faller 2006), Lahsa Tibetan (DeLancey 1998), and Gitiksan (Peterson 2015), to name a few. In the languages that have this pattern, the indirect evidential is always the form which can give rise to mirative readings (mirative forms are not marked with eyewitness evidential marking), and when used as miratives, are subject to a *recency restriction*, a general requirement that miratives comment on something that a speaker has just experienced (Rett & Murray 2013). Given this, there are two clear approaches to capturing the double duty that these markers do in these systems.

Meriçli (2016) proposes one strategy for interpreting these mirative evidentials: as evidentials whose mirative use comes about as a form of flouting the Gricean Maxim of Quality. Using data from Turkish, he argues that the suffix *-mIş* picks out a particular class of indirect evidence, such that it relates the speaker's own evidence source to the very best world-compatible evidence. Indirect evidentiality is encoded as the result if the speaker's evidence being at best second-best. This approach assumes a constrained epistemic modal base, which is a function that assigns to every possible world w the set of propositions that the speaker knows in w and considers evidence for p:

(63) $f_s(w,p) = \{q : s \text{ knows } q \text{ and considers } q \text{ evidence for } p \text{ in } w\}$

This modal base is intersected with the set of all general world knowledge $\mathcal{G}(w)$ with the at-issue proposition p, resulting in the set of all inferences that might arise from the utterance of p in w: $I(w,p)$. He also assumes a partially ordered set of all possible evidence for the at-issue proposition p (64a), which is intersected with I (64b), resulting in the set of all inferences that might arise from the utterance of p in w:

(64) a. $f_0(w,p) = \{q_{best}, q_{best-1}, q_{best-2}, \ldots\}$, $q_{best} > q_{best-1} > q_{best-2} > \ldots$

b. $f_c(w,p) = f_0(w,p) \cup I(w,p)$

Once an evidential base, a partial ordering over this base, and the set of general world knowledge is included in a speaker's view of the discourse context, the indirect evidential has the denotation in (65), and is licensed in contexts where the speaker's evidence is not the best evidence to support an at-issue proposition p:

(65) $[\![-mI\c{s}\,]\!]^{s,w} = \lambda p.[\forall r \in f_s(w,p)\exists q \in f_c(w,p)\ s.t.\ q > r]$

The mirative evidential here is framed in terms of a speaker's evidence. It says that for all propositions r, there is some other proposition q in the set of world-accessible knowledge such that q is better evidence for p than r. When mirative readings arise, they do so as pragmatic implicature in cases where the speaker and the hearer both have access to the speaker's direct evidence source. This is what happens in situations where a speaker uses *-mIş* when they are surprised, even though they have direct, immediate, and visual confirmation of their evidence, as in (66).

(66) *Sophie does not expect Tom until 5pm. At 3pm, Tom walks in the door. Sophie exclaims:*

*Erken gel-**miş**-sin!*
early come.EVID 2SG

✓ (I'm surprised that) you're here early! Meriçli (2016), p. 18

\# Apparently you came early!

Using *-mIş* keys the listener in to the fact that the speaker has flouted the Gricean Maxim of Quality. This pragmatic calculation leads to the inference of violated expectations, and the mirative use arises contextually. In effect, a speaker communicates that her expectations were violated to the extent that she cannot even believe her own evidence. This dual use as evidential and mirative is used to argue for an updated take on indirect evidentiality. Indirect evidentials express that "in relation to the very best world-compatible evidence, the speaker's own evidence is at best second-best" (43). The indirect evidential used in situations of surprise allow the speaker to exploit a particular marker as a way to modify the speech act to reflect her own emotional state.

But this is not a unique fact about Turkish. English can also overtly flout the Maxim of Quality in order to downplay a speaker's evidence source. While English does not overtly mark evidence in the same way that other languages do, it does have paraphrastic and adverbial forms that allow the listener to infer that a speaker's knowledge source is not first hand. These can be used sincerely, or as a way to mark surprise in the face of direct, visual confirmation. The examples in (67) could be used if a speaker has second hand knowledge of an addressee's arrival, or in cases of surprise, mirroring situations like (66) for Turkish.

(67) a. *Apparently* you're here early!

b. *I guess* you're here early!

There is a second way that one could go about interpreting these dual-use mirative evidentials. Rett & Murray (2013) argue for Cheyenne that the marker is inherently a marker of a speaker's expectations, and indirect evidential uses of the marker arise in very particular contexts. This argument is built off of the observation that mirative uses of mirative evidentials always occur with respect to the speaker very recently having learned some proposition p. This is true of the Turkish example (66) above, where the mirative reading is interpreted in the present tense, and evidential as in the past. In Cheyenne, narrative evidential readings arise when the marker is paired with past tense, whereas mirative readings do not.

To do this, they assume that mirative evidentials relate the at issue proposition p to a contextually salient set E of epistemically accessible propositions, which they formulate in terms of expectations. Expectation states are anchored to a salient individual x and time t, E_x^t. This base contains the set of propositions x knows and believes about the past, present and future. These propositions are constrained to those that have been assigned a prior probability. They also assume a recency restriction, which characterizes the illocutionary relation between uttering p and whether mirativity or narrative evidentiality is interpreted:

(68) **The Recency Restriction**: For an event e_s of a speaker i uttering a form with at-issue content p, and for the event e_l of i learning that p, e_s satisfies the recency restriction iff $e_s \in \text{TARGET}(e_l)$.

In addition, they assume that the mirative evidential contributes not-at-issue illocutionary content, which is not assumed when the marker is being used as the narrative evidential. With these elements in play, they can then derive the different polysemous readings of the Cheyenne mirative evidential:

(69) a. *É-* *x-* *hoo'kohó-* ***neho***.
3sg REM. PAST rain NARR.SG.INAN
It rained long ago, it is told.

b. *É-* *hoo'kohó-* ***neho***.
3sg rain NARR.SG.INAN
It's raining! Mirative

In cases where the speaker intends to use the evidential, as in (69a), the speech act e_s is much later than the event of learning that p, and the recency restriction is not in effect. In these cases, the relevant expectations E are valued at the larger community's expectations at some past time (marked by the remote past *x*-), hence the narration use. In the cases where the marker is used as a mirative (69b), the speaker's expectations are the ones that are relevant, and the recency restriction is satisfied: uttering p proposes to add it to the common ground as well as marks it as information that was not previously known.

Rather than a division between evidence source and expectations, the differences observed between indirect evidentials and mirative uses in Cheyenne lie in how expectations are indexed from utterance to utterance in a discourse. If a proposition is new to the speaker, one set of speaker-oriented expectations is activated, and the mirative interpretation is available. If the at-issue proposition is assume to already have some prior expectation in the larger discourse context, the recency restriction is switched off, and the result is an indirect evidential reading.

As with the Turkish data, we again see an interplay between evidence and expectation. In Cheyenne, the system is modeled off of an underlying notion of expectation, with evidence being a contextually relevant flip of a switch—evidence is determined on either the speaker's, or on a community-wide basis of expectations. For Meriçli and Turkish, interpretation is inherently dependent on a speaker's evidence, which can lead to inferences of a speaker's expectations.

Looking at the mirative evidential systems in both of these languages, the motivation for wanting to either derive evidence from expectation, or expectation from evidence is clear. Deriving mirative readings from underlying evidence sources can compute subjective epistemic biases from the evidence that a speaker might have for her claims. The converse derives indirect evidence from a base that is calculated on the salient expectations in the context. If the expectation base is anchored to the speaker, the utterance is interpreted as a mirative. If it is not, it is interpreted as being indexed to the community, and we derive an indirect evidential basis of knowledge.

But looking at these mirative evidential systems in these ways, where one is inherently built off of the other, conflates how a speaker comes to know a piece of information with how the speaker reacts to the receipt of that information in context. In those cases where a speaker uses an mirative evidential in a mirative context, it is not clear that it retains any of its evidential character at all. Notice that when the surprise content of *-mIş* is called into question in an otherwise mirative context, the listener is forced to accommodate an evidential reading of the suffix, and the mirative reading is suppressed:

(70) *Erken gel-miş-sin, ama şaşır-ma-dı-m.*
early come-EVID-2SG but surprise-NEG-PAST-1SG

a. *Evidential:* You were apparently here early, but I wasn't surprised!

b. *Mirative:* # You're here early! But I wasn't surprised.[10]

In the interest of separating the contributions of miratives from the contributions of evidentials, I go the way of Rett & Murray (2013), arguing that at their core, miratives are not inherently derived from evidential meaning, but rather, are a reaction to a speaker's expectations. I remain agnostic about the exact treatment of mirative evidentials in general, but submit that approaches that rely on a speaker's expectations rather than evidence can be easily extended to capture mirative meanings in languages that do not have a collapsed system of mirative and evidential marking.

While Turkish and Cheyenne's mirative evidentials may prove evidence for a dual treatment of evidence and expectation, Magar, a Sino-Tibetan language of Nepal has distinct mirative and evidential markings. Such a language distinguishes evidence and expectation in a way that the collapsed systems cannot.

Magar has a three-way evidential system which overtly marks inferred (*-sa*) and reported (*-ta*) information (Aikenvald 2004). The mirative is realized as the suffix *-le* on a complex

[10]Judgements here are from Jorge Hankamer, Ben Meriçli, Ozan Bellik, p.c.

copular verb form, which attaches to a nominalized predicate, a strategy common on many languages in this family (Grunow-Hårsta 2007). Directly witnessed evidence in Magar need not be marked. In (71a), the speaker reacts with surprise to directly witnessed information, marking this with *-le*. This contrasts with (71b), where the speaker uses *-sa* to mark that she has inferred something about herself (in this case, that she has eaten meat that she shouldn't have), while using the mirative *-le* to further indicate her surprise[11]:

(71) a. *boi-e chitua-ke ŋap-o* ***le***.
father.ERG leopard.DAT shoot.NMLZ **MIR**
Father shot the leopard! Aikenvald (2012), p. 441

b. *ŋai i-din-cʌ sya ŋa-jya-o* ***le****-sa-ŋ*
1sg.ERG PROX.type.ATT meat 1pron.eat.NMLZ **MIR**.INFR.1PRON
Apparently I have eaten this type of meat! Aikenvald (2012), p. 441

The lack of evidential marking in (71a) is telling: Magar's system clearly differentiates how information was gained from how a speaker has integrated that information into her current state. Here, expectations and evidence are clearly distinct: the expression of a speaker's expectations (*-le*) need not comment on a speaker's source of evidence. You can have one without the other.

Stepping back, what can we say concretely about mirativity at this point? The answer here must keep the general descriptions of the two phenomena in mind. Evidential systems are concerned with tracking information sources throughout a discourse as they relate to the speaker. Put simply, they focus on whether a speaker's evidence for a claim is the best evidence that is available in the context. They can indicate whether the speaker witnessed the event they are communicating or if they have inferred it in some other way. They can indicate whether a proposition should be treated as a well known fact or as something that could use more investigating. Evidentials encode what a speaker's basis is for a claim. Miratives, on the other hand, help navigate a speaker's receipt of information into a wider communicative context.

[11]The examples in (71) use the following abbreviations in their glosses: ERG = Ergative, DAT = Dative, NMLZ = Nominalizer, MIR = Mirative, PROX = Proximity, 1PRON = 1st person Pronoun, ATT = Attributive, INFR = Inferred.

2.5.2 A comment on exclamatives

Much of this chapter has concerned itself with expectations and what happens when a speaker's expectations have been violated. A natural question that arises is whether miratives are related in any way to exclamatives. One very clear difference between the two is in the syntactic and semantic environments that exclamatives are restricted to, which contrast with the apparent freedom in the kinds of sentence types that miratives are found in. In English and in German, exclamatives are restricted to occurring alongside a wh-element, in an inversion context, or as a nominal expression, and they must comment on a salient degree:

(72) English:

a. What a beautiful cat! *Wh-exclamative*

b. (Wow,) does that cat like tuna! *Inversion*

c. The fur on that cat! *Nominal*

(73) German:

a. *Was für einen leckeren Kuchen!* *Wh-exclamative*
What a delicious cake!

b. *Ist die Katze dick!* *Inversion*
Is THAT cat fat!

c. *Der Duft, den sie trägt!* *Nominal*
The scent that she's wearing!

Exclamatives comment on gradable properties, and expectation violation happens when a relevant threshold of expectation is exceeded. Thus, if there is no set standard of how beautiful a cat might be, or how tasty a cake is, the exclamatives in (72) and (73) are infelicitous. But if we can interpret an implicit comparison to some stereotypical notion—cats usually like tuna, but this one does so to an exceptional degree—we can also frame the contribution of an exclamative in terms of some degree beyond the speaker's expectations.

Unlike exclamatives, miratives need not be restricted to particular forms or semantic content that can diagnose gradability. In (74a-b), we assume that our speaker walks outside with no prior expectation of rain at all. While we might assume that this is an exclamative utterance, other exclamative forms are not licit in this context. The wh-exclamative and inverted exclamative (74c-d) are infelicitous, as there is no gradable standard to compare them to:

(74) a. It's raining!

b. Wow! It's raining!

c. #How it rains! / #What rain!

d. #Is it raining![12]

Looking at exclamatives, the natural question is to what extent one could derive the fact that their syntax and semantics give rise to this notion of violated expectations. For miratives, surprise comes from their lexical semantics. For exclamatives, surprise is calculated from the compositional semantics. What I will argue is that in attempting to understand these questions, one major factor that needs to be taken into account is intonation. In what sense do these exclamatives have similar contours to mirative contours? However, it is far less clear to me what the intonation contours of exclamatives are, and trying to understand them would be a dissertation unto itself. But taking this into account, what I will say is that what is said about mirative contours and the calculation of expectation in the coming chapters could very well be profitably extended to exclamatives as well.

2.6 What kind of content is mirative content?

Being able to talk about mirativity as its own class of elements raises the issue of what role these markers might have in the larger discourse context. Their contribution to a discourse is clearly pragmatic, which raises the following set of questions: do miratives modify the speech act? Can their content be at-issue? Or do they really arise from conventional implicatures?

Zooming back in to the scope of the mirative markers in this work, we can look at how English *oh, huh* and *what*, the SRC, and the German particles *doch* and *ja* make their contributions available in various ways. One question to ask is at what level of communication

[12] These are of course made better in a context where we can set up some notion of gradable contrast, such as if we had expected a light rain, but it is instead pouring.

these elements operate. If the mirative contribution is constant across declaratives, interrogatives, and imperatives, we might wonder whether the markers operate on speech act content or illocutionary content more generally. A larger question then arises: how do we distinguish the two?

Rett (2017)'s detailed description of emotive markers attempts to draw a clear divide between different types of not-at-issue content, focusing specifically on the difference between descriptive and illocutionary content, represented in the figure below:

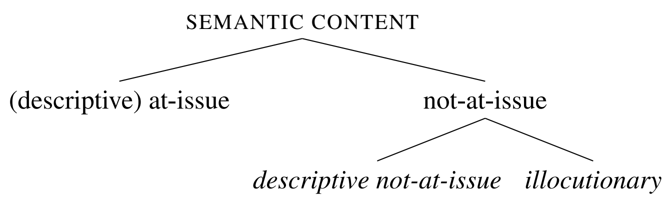


Figure 2.2: Kinds of semantic content

She argues that emotive markers, unlike other types of not-at-issue content encoders, contribute information about how the speaker intends the utterance to be understood, and are thus illocutionary markers. Descriptive not-at-issue markers are those that, as their name suggests, are part of the descriptive content of an utterance, but which nevertheless contribute not at-issue content. This is like the content that constructions like appositives contribute, as in (75):

(75) Jeff, an former phonologist, lives in Japan.

While the fact that Jeff is a former phonologist is not at-issue, it is still part of the projected content of the utterance by way of the speaker introducing it as given. Crucially, maintaining that Jeff is a former phonologist does not comment on a speaker's *attitude* toward the content of any part of the proposition. Teasing apart the difference between descriptive and illocutionary not-at-issue content is ultimately drawn from three observations that distinguish between illocutionary content and "canonical not-at-issue content encoders." Rett assumes the following are true about emotive markers, which she takes to be elements that encode illocutionary not-at-issue content:

(76) **Emotive markers** encode:

a. the speaker's attitude;

b. towards some propositional content of the utterance they are encoded in;

c. in not-at-issue content, i.e. in content that can't be directly denied; be targeted by truth-conditional operators; or address the QUD. *Rett (2017), p. 2*

How do these generalizations jibe with the mirative strategies under investigation in this work? Part of this claim about illocutionary not-at-issue content lies in the fact that emotive markers are comments on propositional content, and cannot target sub-propositional content. Additionally, they inherently comment on a speaker's stance toward a particular topic, suggesting that they operate on an illocutionary level. This initial split is motivated in part by the observations that illocutionary not-at-issue content does not update the projected set. Rather, emotive markers contribute content to a speaker's Discourse Commitments in a Gunlogson (2001)-style table model of discourse. I argue that this is the case for *oh, huh*, *what*, the SRC, as well as the German particles *ja* and *doch*: their contribution is not-at-issue illocutionary content.

Though I hesitate to broadly categorize all miratives as emotive markers, it is clear that the bifurcation of not-at-issue content into descriptive and illocutionary content is a meaningful distinction. For one, it allows for the contribution of mirative markers to be placed in the Discourse Commitments of the speaker without committing to them in the same way as proffered content of an utterance. These contributions also differs from the projected content of (77a), the content of the appositive in (77b), or the expressive use of *that idiot* in (77c), which would be part of the projected set in a Farkas & Bruce (2010) model of discourse:

(77) a. Jeff's workouts usually include pull-ups.

descriptive not-at-issue content: Jeff has/does workouts.

b. Sophie, a skilled banjoist, has recently taken up guitar.

descriptive not-at-issue content: Sophie is a skilled banjoist.

c. I can't believe that idiot Tom forgot to come to trivia tonight.

descriptive not-at-issue content: Tom is an idiot.[13]

[13] Rett (2017) notes that some Expressives, like *idiot* here, may in fact function as emotive markers, or at least,

Rett (2017) also notes that a subtle distinction exists between descriptive and illocutionary not-at-issue content with respect to Moore's Paradox. Elements that contribute illocutionary content (such as the sincerity conditions of particular speech acts) contribute a particular infelicity when denied, as in (78):

(78) It's raining, #but I don't believe it's raining.

The oddness here is not due to actual contradiction—we calculate infelicity in such utterances based on a speaker's asserting some proposition *p*, and then asserting that she does not believe this proposition, which does not deny the *content* of *p*, but rather, its *sincerity conditions* (Searle 1969). These Moore's Paradox situations are different than denying other types of not-at-issue content, such as denying the content of an appositive, which results in an actual contradiction (77b):

(79) Sophie, a skilled banjoist, has recently taken up guitar. #⊥ Sophie can't play the banjo, though.

Rett (2017) reports that denying the content of miratives in Moore's Paradox scenarios pattern like (78), which indicates that they are illocutionary not-at-issue content. She notes that in English, particular mirative intonational contours have a Moore's Paradox-like effect when their pragmatic contribution is denied, as in the pragmatically odd (80). The contribution of *oh* cannot be targeted in this way either (81). And in German, when the contribution of either the modal particles *ja* or *doch* is contradicted, it sets up the same sort of infelicity, (82-83):

(80) John arrived on time! ...#I'm not surprised, I knew he'd be punctual. *p.11*

(81) Oh, there's no flour ...#But I didn't think there would be any.

(82) *Ich habe die Tür* ***ja*** *zugemacht. ...# Aber das hast du nicht gewusst.*
I have the door *ja*.MIR closed but that have you not known.
(We both know that) I shut the door. ...# But you didn't know that.

(83) *Tom hat* **doch** *keine Kinder. ...# Aber das hättest du nicht wissen können.*
Tom has *doch*.MIR no children but that have you not know could

contribute illocutionary not-at-issue content. In a case like (77c), the attitude expressed by the speaker does not target the entire propositional content of the utterance, only a portion of it.

Tom obviously doesn't have any kids. ...# But you couldn't have known that.

Given these observations, I assume that the mirative content of the strategies under scrutiny in this work contribute illocutionary not-at-issue content.

One potential question could be raised with respect to what kind of content mirative evidentials contribute. These markers have a context-dependent dual use, marking both indirect evidence for an utterance, as well as a mirative marker. Evidential content has been argued by Faller (2006) and Matthewson et al. (2007) to be conventional implicature, or in the terms described above, descriptive not-at-issue content. This has the potential to clash with the proposal that miratives contribute content on the illocutionary level. But the polysemous nature of these makers should not be conflated; in situations where the marker is used as a mirative in a direct evidence context, the mirative meaning does not also comment on the speaker's evidence for the proposition, but rather, just their attitude toward the propositional content. The example in (84), repeated from above, shows this. In a direct evidence context, only the evidential interpretation is available:

(84) *Sophie does not expect Tom until 5pm. At 3pm, Tom walks in the door Sophie exclaims:*

Erken gel-miş-sin, ama şaşır-ma-dı-m.
early come-EVID-2SG but surprise-NEG-PAST-1SG

a. *Evidential:* You were apparently here early, but I wasn't surprised!

b. *Mirative:* # You're here early! But I wasn't surprised.

In (84a), which attempts to set up a Moore's Paradox situation, speakers note that it is interpretable, but only under an evidential interpretation of the suffix. Denying the content of the mirative is is infelicitous; the context must be such that the speaker had some indirect evidence for the addressee's arrival, and was not surprised by that. If the mirative reading is forced like in a context from (61), the *-mIş* can be contradicted, but sets up a paradox—why mark that you are surprised, and then deny this condition? In Turkish, there are two distinct uses of the mirative evidential, and their distributions are clearly distinguishable: it is not simultaneously a mirative and a marker of indirect evidentiality. The divide between descriptive and illocutionary not-at-issue meaning between evidentials and miratives can still stand. With this, I assume that the miratives I deal with in this work contribute illocutionary not-at-issue content.

2.7 Summary

The purpose of this chapter was twofold. First, it introduced a background on some theoretical notions of belief structuring and discourse modeling. Expectations are anchored to an individual, and they can be navigated through the lens of a discourse and discourse commitments. This is done in part by way of individual speakers' contributions to the Table, as well as their personal discourse commitments, which are not available to be accepted or rejected by other agents in the context. In addition to the traditional Table model of discourse in the style of Gunlogson (2001) and Farkas & Bruce (2010), this chapter proposed that intonation contributes a speaker's expectations about a particular proposition, and should be added to the model as an update to a speaker's Discourse Commitment list. Doing this makes the expectations of the speaker clear in a discourse, as particularities of intonation are registered as comments on a proposition, and not in play for questioning, negation or discussion by interlocutors.

Second, this chapter was a way to begin a conversation about mirativity, and more specifically, about what role expectation plays in discourse navigation. As expectation is a crucial component of the meaning of mirative markers, it was important to outline just what it is that mirative meanings contribute, and just as important to make clear what they do not. In particular, this chapter discussed the connection between expectation and evidentiality, ultimately drawing a line between how a speaker *comes to know* some piece information, as opposed to how they *integrate knowledge* into their base.

This chapter concludes with a discussion of how to characterize mirativity in terms of the content that it expresses. Ultimately, it concludes that mirativity contributes illocutionary not-at-issue content, much like Rett (2017)'s characterization of emotive markers. In all, this chapter made the case for the following three things: a) Miratives are markers that express a speaker's violated expectations in a discourse and how this information is integrated into their view of the current context, b) Miratives, like intonation, contribute illocutionary not-at-issue content, and c) In a tabletop model of discourse navigation, the contributions of intonation as well as of miratives are registered in the speaker's Discourse Commitment list, which is neither public, nor projected content.

In the chapters to come, the ideas that have been introduced here will be further refined, with particular emphasis on contours and mirative markers specific to English and German. The following chapter begins with a discussion of three English discourse particles and three English intonational patterns, which serve as fodder for a continued discussion of the themes introduced in the previous pages.

Chapter 3

Discourse particles and Intonation in English

The pragmatic discourse markers that will be important in this work (discussed in Chapter §1), have the following four things in common:

(85) a. They are always optional in an utterance

b. They contribute illocutionary not-at-issue content

c. They comment on things in the discourse context above the propositional level

d. They take the utterance context into account, not just an uttered proposition

This chapter will focus on three English discourse particles that fall under this umbrella. In doing so, it will explore their relationships with prosody, looking specifically at how they compose to contribute precise pragmatic meanings.

Intonational meaning is a widely researched topic. The differences between the particle responses in (86), for example, can be reduced to the differences in a speaker's intonation:

(86) **Sophie:** I think Jeff went to campus this morning.

Tom:

a. Oh. (He forgot his lunch.)

b. Oh? (He forgot his lunch?)

(86a) is a relatively neutral statement that appears to accept the issue raised in A's utterance and adding follow up information, perhaps after noticing that Jeff's lunch box is still on the counter. (86b) does not accept this issue, and with the follow-up, raises an alternative issue: If Jeff were on campus, why would he have left his lunch?

But which meaning component is contributed by the particle, and which by intonation? In (86a), Tom seems to accept the content of Sophie's statement with the particle. Perhaps *oh* signals acceptance. Yet in (86b), the intonation on the particle leads us to believe that the speaker cannot have accepted the issue raised by Sophie. Instead, Tom challenges the issue raised by Sophie. We might be inclined to write this lack of acceptance off as interference from the rising contour; the lack of commitment to the content of an issue raised with the rising contour might override any commitment that the particle inherently contributes. But then why use the particle at all? Impressionistically, it doesn't seem to contribute much to (86b). Under these off-the-cuff interpretations, it would appear that the only difference between the two responses is the intonational meaning. There must be something more complex going on if, as we assume, *oh*'s contribution is meaningful.

This chapter begins to investigate the individual contributions of intonation and discourse particles to a discourse. Starting from intonational meaning, we will work our way through to discourse particle meaning by way of examining how prosody interacts with the commentary that is provided by discourse particles. This chapter takes an additive approach to particle and prosodic meaning, assuming that the contributions of each are decomposable into component parts. In approaching this problem in this way, we are able to disentangle two types of pragmatic meaning that come packaged as a single speech unit.

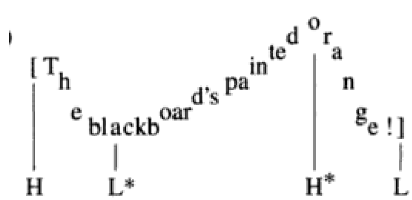
The organization of this chapter is as follows. Section §3.1 outlines a few utterance-level prosodic contours with their standard pragmatic meanings. Section §3.2 introduces three discourse particles in English: *oh, huh* and *what* and the complexities surrounding their interpretation. Section §3.3 tackles the largest issue of this chapter, and is divided into two parts. The first portion takes a look at the interactions of *oh, huh* and *what* with each of the prosodic contours of interest. This is not a trivial issue; pragmatic discourse markers are extremely context dependent, and seem to change their meaning from utterance to utterance. The picture is made cloudier still as particle, contour, prejacent, and context are all variables that if manipulated, could change pragmatic interpretation.

Second, and perhaps more importantly, this chapter situates these English particles into the existing literature on mirativity. I argue here that discourse particle meaning is grounded in speaker expectations about the discourse, and that their licit use depends on contextually-conditioned events. English also has a way of intonationally marking mirativity: the surprise-redundancy contour. Here, I show that intonation, too, can betray a speaker's expectations, and can contribute precise pragmatic meaning to an utterance. In all, this chapter serves as part of the wider goal of teasing apart particle meaning from prosodic meaning, producing what we hope to be a clearer picture of both.

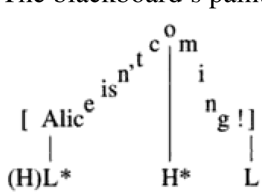
3.1 English Intonational Contours

3.1.1 Surprise-Redundancy Contour

Sag & Lieberman (1975) identify the surprise-redundancy contour (SRC) as an utterance-level prosodic tune that carries pragmatic information about speaker attitude. Their classification assumes that the contour itself is a minimal unit of meaning consisting of a low pitch associated with the primary sentential accent, and contrasting low pitch on the utterance's second most prominent syllable. Hayes (1995) notes that this tune is a particularly reliable diagnostic for locating secondary sentential stress: the a low (L*) pitch aligns with the strongest stress left of the main stressed syllable (H*), as in (87) below:

(87) a. 

The blackboard's painted orange!↘

b. 

Alice isn't coming!↘ *Hayes 1995, p. 16-18*

Ladd (2008) further specifies the prosodic shape of the contour, noting that smaller intonational phrases may not have an initial high tone ((H-)L* H*-L%), and compound phrases may have an extra high tone by virtue of normal English compounding processes ((H)-L* **H**...H*-L%). Sag & Lieberman (1975) originally identify the contour as expressing a speaker's surprise at a proposition or event in the world, much as a speaker might do in (87a) upon walking into a classroom with a brightly painted chalkboard.

The contour is also used in cases where the speaker believes the hearer "should have known" some proposition p, as in (87b). If we have established that Alice is currently vacationing in Bermuda, we don't expect her to come to our party in Santa Cruz. Bringing up the fact that Alice isn't around might surprise an addressee, who thinks that Alice's whereabouts is common knowledge. Using the SRC in the 'redundancy' case is akin to overtly questioning your addressee's information state: "Why are you asking? Isn't it clear? Alice isn't coming!" Ladd (2008) puts it another way: The surprise-redundancy contour "expresses [...] the view that one's interlocutor should have already known what one is saying" (120).

The SRC is an utterance-level phenomenon, and does not associate with a particular phrase type or lexical category. It instead aligns with prominent positions and breaks of an utterance. Though it requires two stressed tones, the contour can also appear on prosodic phrases with just a single prominent position, or even a single syllable. To do this, the vowel is lengthened, giving the effect of two syllables. It might be helpful to think of this as the "duh" contour:

(88) Mother: Did you brush your teeth?
Child: Du- uh!
L* H*-L%

(89) Mother: Don't forget to floss!
Child: Mo- om! (*especially petulant here*)
L* H*-L%

Wh-questions can also align with the SRC, especially those that require the addressee to provide explicit explanation of a previous utterance or an event, or those that are used rhetorically (Bartels 1999: p. 173):

(90) A: I don't want to eat this grilled cheese.

B: Then why did you order it?↘
L* H* L-L%

(91) A: I'm going to use cookie cutters to make cute tea sandwiches.

B: And what are you going to do with the crusts?↘
L* H* L-L%

The use of discourse-navigating markers in (90-91) is not insignificant; both *then* and *and* mark a explicit reference to a salient theme in the common ground. Bartels notes that this contour on the corresponding assertive versions of these utterances promotes the redundancy or "should have known" inference. While Wh-questions posed with the SRC are still inquisitive, they do not necessarily indicate deference to the addressee's information for the enrichment of the speaker's knowledge state. Rather, the use of the SRC is an attempt by the speaker to point to something about the hearer's actions or utterances in previous discourse that they are not currently taking into account. With the SRC, the speaker brings up an issue that she believes the hearer knows, or should be able to infer, but has not taken into account in a salient conversational move. Thus, the domain of the SRC is not limited to declaratives, as the final falling boundary tone might suggest in more conservative frameworks. Rather, it has particular pragmatic effects that arise in tandem with a variety of utterance types.

3.1.2 Rise-Fall-Rise Contour

Similar to the surprise-redundancy contour, the Rise-Fall-Rise, or "incredulity contour" is an utterance level phenomenon spanning multiple prosodic phrases. Ladd (1980) identifies this complex contour as a low pitch accent rising to a high tone, which falls, only to rise again utterance-finally: L*+H L-H%. Breaking down this contour into its component parts, it is not unreasonable to think that the rising utterance-final boundary tone might indicate non-assertiveness. Ward & Hirschberg (1985) propose that the rise-fall-rise contour implicates varying degrees of speaker (un)certainty: a speaker might use it to signal appropriateness of the utterance to the current discourse, uncertainty about how well an utterance fits into a contextually given scale, and the uncertainty of the utterance choice as a value on a particular scale. Pragmatic effects of politeness, irony or deference may also arise as a result of the rise-fall-rise contour (p. 765):

(92) A: You don't know any Estonian music!

B: I know Tormis' "Curse Upon Iron."↗
L* H L- H%

(93) A: Do you think we can borrow this ping pong equipment?

B: I'm pretty sure it's allowed.↗
L* H L-H%

In the Ward & Hirschberg (1985) system, B's answer in (92) implies B's uncertainty as to whether his answer is appropriate; Veljõ Tormis' piece is arranged by an Estonian composer, but it is sung in Finnish. The utterance asserts its content, but allows for the possibility that the addressee might reject it. In (93), *pretty sure* indicates that the speaker's certainty about the matter is scalar. She insinuates the alternative that they may not be allowed to play, but ultimately asserts the opposite. 'Incredulity' readings (whose contours have a higher overall pitch duration, but are otherwise identical in tone patterns) arise from scales implicitly invoked by an addressee in response to a previous utterance (Chen 2005: p. 115). If B assumes in (93) that A is questioning him for uttering a statement to the effect of *Here are some ping pong paddles to use*, his pitch might rise higher than his normal range in response, further expressing his surprise that his conversational partner might second-guess his judgment.

In the coming sections, it will be important to distinguish this tune from the SRC, as both are complex, operate on the level of the utterance, and have similar pragmatic effects. One prominent difference between the two is that the rise-fall-rise, as the name suggests, rises utterance-finally. I will not comment explicitly on the effects of final rising tunes (see Farkas & Roelofsen 2017, Jeong 2017, Rudin 2017a: for discussion). Rather, I point this out in order to note that caution must be taken in replicating the contours as they are intended, especially in the absence of explicit audio stimuli.

3.1.3 Excited Contour

"Excited" intonation is an intensification of the canonical high to falling boundary tone, (L)H*L%, coupled with higher than average pitch excursions on the H* peak. On paper, it is difficult to distinguish this from a normal utterance final fall, as ToBI markings do not distinguish between pitch height or gestural slope across prosodic phrases, let alone utterances. The effect of this extended pitch height on the last high tone in an utterance is interpreted as an intentional expression of speaker emotion: the speaker is trying to convey her excitement or positivity toward something in the discourse by manipulating pitch. In (94), B′ finds the prospect of cats outside her window new and surprising:

(94) A: What do you see out the window?

a. B: Cats.

b. B′: Cats!

The (a) response to A's question is more "normal" or default in the sense that it does not ask the addressee to compute any other pragmatic information. If information is new to a speaker, they can choose to accommodate this newness, and mark it with default tonal prominence. Extreme tonal height has the pragmatic effect of marking the utterance in question as new and informationally important. The response in (94b) chooses to highlight this newness by using a higher pitch. But though this contour indicates surprising content to the speaker, the content must be surprising in a positive or beneficial way, reflected in the choice to call this the *excited* contour. If A and B are in a tent and they hear a rustling outside, B might be surprised to see a bear, but as she knows that a bear poses a threat while camping, her surprise cannot be conveyed with the same contour as it is in (94). Her reaction is negatively surprising: a near outside the tent is not positive or beneficial to their situation.

(95) A: What do you see out the window?

B: *(with excited contour)* # A bear!

Note, too, that "newness" need not be anchored to the speaker. Speakers use higher pitch ranges when talking to children about things that they perceive the child as finding exciting or new, but which is known to them already. Take (96), in which a parent's feigned surprise has a similar high pitch excursion. The H* peak on *found* has the effect of imitating the child's excitement:

(96) *Small child, triumphantly holds up an Easter Egg.*

Parent: You found it!
H* L-L%

In a small corpus study of data gathered from a single speaker, utterances annotated as having the excited contour have an average pitch height maximum of 240.667 Hz, as opposed to an average pitch maximum of 142.155 Hz in utterances annotated as "neutral." While this is a small sample (n=21 for neutral utterances, n=20 for excited utterances), these preliminary results show a significant difference in pitch height maxima for these two contour types (f-value = 139.26, $p < 0.0001$***).

This is coupled by the fact that both utterances fall to similar f0 levels utterance finally. While mean pitch heights for final boundary tones in the neutral annotations were slightly lower than those with the excited contours (neutral means = 102.274 Hz, excited means = 104.823 Hz), this difference was well below significance (f-value = 0.517, p = .47662). This suggests that this pitch height excursion is a distinguishing factor in the pragmatic effect found here.

In what follows, I will use an exclamation mark '!' to orthographically indicate this excited intonation. Note that this contour is distinct from two contours with similar features: the calling contour (H* !H%) and what I identify as mirative vowel lengthening. The calling contour, sometimes known as vocative chant, is just what it sounds like–the pattern of intonation used when calling for someone who is perhaps out of range (Ladd 1980). It begins on a high tone and then is downstepped to a mid-level tone, indicated by the additional ToBI diacritic '!'. This contour can be made to sound scolding as well by lowering the overall pitch of the utterance. The following contexts distinguish the two levels of this contour:

(97) Calling contour:

a. *...Happy birthday dear Ste- ven,* ...
H* H!%

b. *Mother calling a child to dinner:*

So- phie!
H* H!%

(98) Low calling contour:[1]

a. *Mother, scolding child*:

Je- ff. (You know better than that.)
H* !H%

b. *Jeff, responding, slightly defensive:*

Wha-t! (I didn't do anything!) *said like "wu-ut"*
H*!H%

Mirative vowel lengthening is a process that happens in circumstances of high speaker surprise or violated expectations. It is characterized by an extended lengthening of a stressed syllable in a prosodic phrase, and in written communication, is often characterized by repeated instances of the lengthened vowel, where more vowels are representative of longer vowel duration:

(99) **Sophie**: We just won airline tickets to Sweden!

Jeff: Whaaaat!

(100) *Tom comes home from vacation to find the apartment ransacked*:

a. Hooooooly shit

b. Hooooooly shiiiiiit.

c. Hooooooolyyyy shiiiiiit.[2]

[1]Note the downstepped pitch on monosyllables is realized as a lengthening of the vowel to make two syllables. The downstepping can then be applied.

[2]The prosodic restrictions here have not been fully explored, but it seems as though the first element with intonational prominence of a prosodic phrase must be lengthened, and lengthening following peaks are optional. In (100c), *shit* carries the most intonational prominence, but it cannot be the only syllable to receive mirative lengthening. Note how in the variants below, when the first syllable is not lengthened, the utterance is infelicitous:

(1) a. *Holyyyy shit.
b. *Holy shiiiiiit.
c. *Holyyyy shiiiiiit.

This mirative lengthening can be a positive response (99), as well as negative reaction to a proposition or event (100). While this contour is indeed licit on discourse particles in English, it is not the exclamative contour of interest; the exclamative contour relevant here may only respond in a way that presents the speaker as evaluating the content in a positive matter. I leave investigation of this to future work.

3.2 English particles

3.2.1 Discourse particles and prosodic phrasing

In broad strokes, discourse particles in English are a way for a speaker to integrate facts about her epistemic state into a conversation. But it is not only particles that have this effect: prosodic contours, too, contribute similar pragmatic effects to an utterance. Adding to the complexity, discourse particles can be variably accentuated and prosodically phrased, each giving rise to a particular discourse effect. Schiffrin (1987) shows that some utterance-initial discourse particles in English occur with their own intonational contour, and are prosodically independent elements of an utterance. Alternately, they can be integrated into the contour of the sentence as a whole, and prosodically parsed as part of the following phrase. The following prosodic trees show a rough sketch of the differences here, with the major deviations occurring in how the particles have been integrated into the leftmost intonational phrases in each tree:

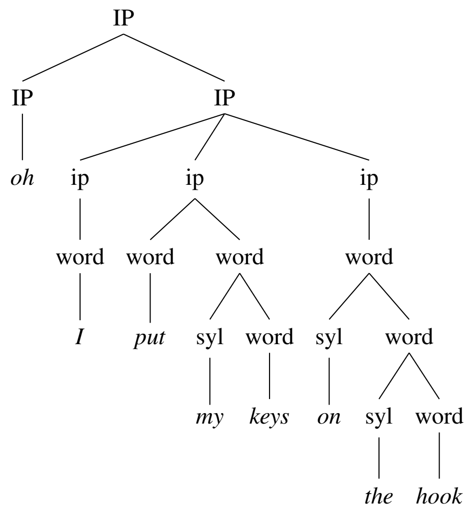


Figure 3.1: Independent discourse particle

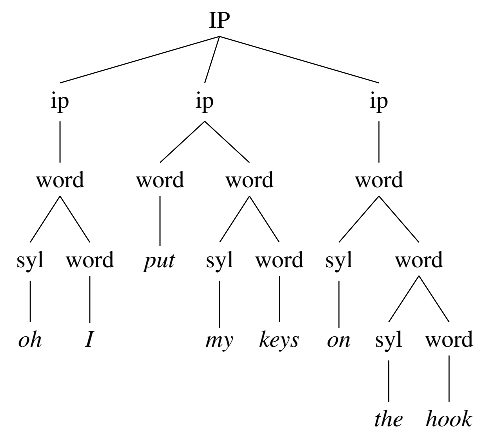


Figure 3.2: Integrated discourse particle

(101) Oh! I put my keys on the hook last night. Independent

(102) Oh I put my keys on the hook last night. Integrated

In (101), both the particle *oh* and the following utterance occupy their own prosodic phrases. (102) integrates the particle into the overall melody of the utterance, with *oh* cliticizing to the prosodic phrase containing the subject and the verb. Though overlapping for the most part, the difference in prosodic phrasing is able to diagnose at least one case where the integrated particle is interpreted differently than the independent particle. If Jeff were to walk around a corner and see his keys hanging on a hook, without already being in the act of looking for the keys, he can utter (101), but (102) is odd. Completely out of the blue, only the independent prosodically phrased particle and utterance is licit. In contrast, if someone asks Jeff where his keys are, he can respond with either (101) or (102). (101) has the effect of positioning his utterance as relevant to the hearer's needs. Incorporating the particle as in (102) has a more reassuring effect, indicating that he is not worried about the location of the keys. In this work, we will discuss only the prosodically independent English discourse particles, but with the same forward-looking restriction as discussed in section §1.1.1. I leave the analysis of prosodically incorporated particles and their contributions to future work.

3.2.2 English *oh*

Many descriptions of the discourse particle *oh* rely on interpreting a speaker's emotion. Heritage (1984) and Schiffrin (1987) note that many uses of this particle are elicited in situations where the speaker has had a strong emotional reaction to some event, be it positive or negative (anger, surprise, excitement, sadness or fear are a few examples). On an abstract level, the particle broadly marks some change of state that has occurred, whether in the speaker's reasoning or in their accessible information. (103) shows this particle in what Heritage (1984) calls an interjection use of *oh*, whereas (104) is its discourse particle use:

(103) *Speaker walks into a room full of snakes:*
Oh!

(104) *Speaker receives news that the room full of snakes is to be incinerated:*
Oh! That's great news!

But the two are not unrelated: as discussed in Chapter §2, surprisal can be captured in terms of violated expectations, which mark a change in a speaker's biases or reasoning from a previous state. Other emotions can also be thought of in terms of a change of state: anger, for example, can be framed as a negative change in speaker expectations.

Aijmer (2002) takes a wider view of *oh*, and suggests that this particle can indicate a change of state on the part of the speaker as well as the addressee's representations of the discourse. When using *oh*, a speaker not only indicates her own change of state, in doing so, she is also priming the listener for an upcoming change to their own current discourse model.

Fox Tree & Schrock (1999) find that this is precisely what happens in experiments that seek to pinpoint an agent's perception of a discourse. One reason a speaker may use *oh* is to facilitate some aspect of comprehension on the part of their addressee. They find that the presence of *oh* in audio stimuli sped up a listener's perception of what they heard next when compared to similar stimuli without the discourse particle. From this, they conclude that the speaker uses *oh* to mark both her own change-of-state as well as to inform an addressee that they should also prime themselves for potential unexpected information.

Gunlogson (2008) notices that *oh* can be used to diagnose speaker commitment, but only when paired with a neutral falling contour (H*-L%). A neutral *oh* response to a previous utterance has a similar effect to the response particle *yes*, committing the speaker to the content of the previous proposition. But unlike *yes*, *oh* signals that this is new information to the speaker:

(105) A: They quarantined that room.

a. B: Yes. *commits to A, A can already be known to B*

b. B: Oh. *commits to A, A is not already known to B*

In both cases, the response particles seem to commit the speaker to the content of A's utterance. Yet changing the prosody on the particle changes this effect of commitment for *oh*. *Yes* with rising intonation commits the speaker to the content of A's utterance, but has an extra pragmatic effect that appears to question the relevance of the previous discourse contribution (for an in-depth discussion of the varied contributions of types of rising intonation, see Jeong 2017). The same rise on *oh* does not commit B to A's utterance:

(106) A: They quarantined that room.

a. B: Yes? *B is committed to content of A*

b. B: Oh? *B is* ***not*** *committed to content of A*

For *oh* in (105b) and (106b), the particle itself does not contribute commitment to a salient proposition, but rather, this commitment is contributed by the intonation. The contribution of the particle must be something else. This is particularly clear in out-of-the-blue cases, where the speaker reacts to some environmental factor, rather than a previous utterance:

(107) *Speaker walks past a charred building:*

Oh! They burned the snakes!

Here, *oh* with excited intonation highlights a speaker's surprise about some event. While a follow-up utterance may provide some proposition to commit to, the particle in itself does not. Walking past the charred building and expressing surprise at this with *oh* simply conveys a speaker's violated expectations. In keeping with Heritage (1984), Fox Tree & Schrock (1999) and Aijmer (2002)'s ideas of a salient 'change-of-state' as a core meaning component of *oh*, I propose that the particle requires an inequality in terms of a speaker's expectations and the information that they have been presented with. *Oh* comments on a speaker's expectations, allowing the hearer to infer that some information did not accord with how the speaker imagined the conversation would progress, as outlined in (108):

(108) *oh* is anaphoric to a proposition p salient in the discourse s.t.:

a. $oh(p)$ adds the following to the speaker's DCs:
$\mathrm{Exp}_{spkr}(p) < \mathrm{Exp}_{spkr}(\neg p)$

For this particle, a speaker presents herself as having a higher expectation for the complement of a proposition than for the proposition itself. Formulating the contribution of this particle in terms of expectations allows the contribution of *oh* operate on the same level as the contributions of intonational contours: since the meanings of these particles can be influenced and shaped by their prosodic environment, capturing their discourse effects in similar terms is beneficial for separating the contribution of one from the contribution of the other.

This characterization of *oh* also places its contribution into a speaker's discourse commitment list as a part of a discourse, but not on the table, subject to public scrutiny. While these expectations are made publicly available (an addressee can easily compute a speaker's expectations from her performance of an utterance), they are not directly challengeable. As in (109), it is odd for A to challenge the pragmatic contribution of B's discourse particle response:

(109) A: The University has banned reptile rooms.
B: Oh.
A: #That's not true. You expected that.

Another advantage of formulating the contribution of *oh* in terms of expectations is that it casts this particle as a mirative marker. Mirativity relies on somehow challenging the expectations of a speaker; when a speaker uses *oh*, she is communicating just this to her partner. She had expected one outcome, and is surprised to find that it is not the case. In (109), perhaps the speaker is only mildly surprised. Yet, with excited intonation, the mirative function of this particle is magnified: the strategies have an additive effect. In an utterance like (107), the speaker effectively augments her degree of surprise by using an excited contour. The interactions of *oh* with other contours will be addressed further in Section §3.3, as being aware of the functions of each, as well as restrictions imposed by the context, will be crucial to understanding their overall discourse function.

3.2.3 English *huh*

While a fair amount of work has been done on *oh*, a similar particle *huh* has not gained as much attention. In many cases, this particle patterns alongside *oh*, occurring as apparent synonyms in similar environments. In this sense, it seems plausible that *huh* might pattern alongside *oh* as a particle that indicates a change-of-state, along the lines of Fox Tree & Schrock (1999) and Aijmer (2002):

(110) *Speaker walks into a room full of snakes:*
Huh!

(111) *Speaker receives news that the room full of snakes is to be incinerated:*
Huh! That's great news!

Given this, one might be tempted to propose that *oh* and *huh* contribute the same thing to a discourse. But in particular cases, *huh* patterns differently than *oh*. One place in particular where the two particles pull apart are in cases where two participants in a discourse have equal information. In (112a), B can use *oh*, although she overtly indicates that she is already aware of the content of A's utterance. In this case, it seems that the violated expectations she refers to is the level of the speech act itself: she is surprised that her speaker (re-)raises an issue that she assumes to be common ground. This is not so for *huh*:

(112) A: They quarantined that room.

a. B: Oh. I know.

b. B: #Huh. I know.

Huh's behavior also pulls apart from *oh*'s in the Gunlogsonian test for commitment to a previous utterance. Whereas *oh* with neutral, falling intonation seems to commit the speaker to the content of a previous utterance, *huh* requires something more. In (113a), B can make her *oh* claim without going to check on the status of the room in question. But uttering (113b) is odd, especially if B has not gotten up, opened the door, and verified the claims that A has made:

(113) A: That room is full of snakes.

a. Oh. No it isn't. (It just looks like that.)

b. #Huh. No it isn't.

For (113b) to be licit, some accommodation on the part of the speaker must be made. This is distinct from the use of *oh*, where the continuation utterance seems to form an explanatory relationship with the use of the discourse particle. That is absent in (113b)—the particle and the continuation are two distinct conversational moves. In contrast to *oh*, *huh* must acknowledge and at least temporarily accept the previous proposition. Of course, revisions can be made, but they come at the cost of striking the accepted content from the discourse record.

Clearly, though, *huh* has much in common with *oh*. In keeping with the characterization of *oh*, I propose a formulation of *huh*'s discourse contribution in much the same way. *Huh* also requires a contextual inequality in terms of a speaker's expectations and the information that they are currently faced with. The difference here is that *huh* requires a further restriction on the speaker's expectations:

(114) *huh*(p) adds the following to the speaker's *DC*s:

$$\text{Exp}_{spkr}(p) < \text{Exp}_{spkr}(\neg p) \wedge \text{Exp}_{spkr}(p) > 0$$

While, like *oh*, *huh* requires that the speaker have expected $\neg p$ over p, this particle is not silent about the speaker's relative expectations about p. This is particularly clear with contextual differences in cases of speaker surprise. In these cases, the performance of the particle shifts from a neutral, falling contour to an excited contour (a higher falling contour, with an overall higher f0). In a completely out of the blue scenario of surprise where the speaker does not suspect anything out of the ordinary, *oh* is licit, while *huh* is not:

(115) *Speaker, oblivious, rounds the corner and is hit with confetti:*

a. Oh! I didn't expect that!

b. #Huh! I didn't expect that!

But suppose the speaker and her colleagues occasionally surprise each other with practical jokes. The speaker, though still surprised, in this case has some inkling that a surprise is possible, though doesn't know when it will come. In this case, when she is hit with confetti, *huh* is licit:

(116) *Speaker rounds the corner and is hit with confetti:*

a. Oh! I didn't expect that!

b. Huh! I didn't expect that!

Formulating the contribution of *huh* in terms of violated expectations also has the advantage of placing this particle in with other markers of mirativity. This is a welcome categorization: *huh* communicates to an addressee that the speaker must shift her expectations in order to compensate for a new piece of information. This will be addressed in more detail in Section §3.3.

3.2.4 English *what*

The particle *what* has a sort of split status for a prosodic discourse marker. While *oh* and *huh* are able to occur in many different prosodic environments, *what* is more restricted. One obvious explanation stems from the semantic meaning and function that is available to describe the phonologically identical wh-word. But despite this overlap, there is reason to categorize this word as a particle as well. For one, it has functions, both syntactically and semantically, that the interrogative use does not.

The particle in question occurs as a reaction to a previous utterance, much in the way that *oh* or *huh* do. It is distinct from the wh-determiner *what* in part due to its out-of-the-blue use in contexts such as (117) below, or the non-repetition surprise use of the particle in (118):

(117) *Speaker is listening to the lottery numbers be announced on TV. She realizes she's holding the winning ticket:*
What! / What?! (I won!)

(118) A: I just got a new haircut!
B: What! Who from?

One might be tempted to analyze these occurrences of *what* as root sluices (see Ginzburg & Sag 2000: for a discussion of this phenomenon). But doing this begs the question of what exactly the content of the ellipsis site would be, if these were indeed wh-remnants of sluices. In neither case does the speaker seem to require an overt antecedent to the sluice. While (117) could be a rhetorical use of *what* to implore the addressee (here, the TV announcer) to repeat their utterance, it is more likely an expression of surprise. This is not the environment of a sluice remnant. (118) is especially telling, as the particle use of *what* is followed by an actual root sluice. In these two examples, *what* is not a marker that asks for more information about a topic, nor does it expressly ask for clarification. Just like with *oh* and *huh*, *what* is used to betray a speaker's violated expectations. It is another mirative strategy in English.

The question at present is to distinguish just how the contribution of this mirative particle differs from the contributions of *oh* and *huh*. Right off the bat, it is interesting to notice that in contrast with the two other particles, *what* does not naturally occur with either the neutral falling contour from Section §1.2.2 or the excited contour outlined in Section §3.1.3: *what* has a more limited distribution than the other two particles.

While this particle can occur with a falling contour, it seems like a much more marked move than when the other two particles appear in this prosodic environment. Marking *what* with falling intonation seems to necessitate a very low pitch excursion, as well as an extended pitch duration on the nucleus.[3] This can be elicited in situations like (119) below:

[3]These are impressionistic judgments; the requisite data has not been collected to verify these claims entirely, which I leave at present to future work.

(119) **Sophie**: You're late again.

Jeff: What. My car broke down.

These falling contour on *what* is pragmatically as well as intonationally distinct from identical utterances with *oh* and *huh*. Falling contours with *what* always come off as defensive, as if the speaker is threatened by the characterization of his own actions by someone else, and is responding to veiled criticism. This particular pitch contour must reference the speaker, and is always construed as negative. As this is distinct from the falling contour used most naturally on the other two particles under discussion here, I leave an in depth discussion these uses of *what* aside.

What also does not fare well with the excited contour. While it can occur in cases of surprise like in (117) and (118), the contour is not the same excited contour of interest that we find with the particles *oh* and *huh*. Rather, the nucleus of the particle is lengthened, and is likely an instance of the mirative lengthening contour, discussed briefly in section §3.1.3.

Given this, I propose that the meaning of *what* in its discourse particle form is simple expectation violation of a speaker, calculated in relation to a salient proposition p. This differs from both *oh* and *huh*, which express levels of disbelief based on calculated expectations for *another* proposition to be true. *What* simply encodes uncertainty about an event or proposition.

(120) *what*(p) adds the following to a speaker's DCs:

$\text{Exp}_{spkr}(p) < 1$

Formulating the effect of this particle in such simple terms of speaker uncertainty can help explain why it is not licensed in neutral falling or excited contexts. The neutral falling contour has the pragmatic effect of an assertion, which positions the speaker as being authoritative over the prejacent. For *oh* and *huh*, the speaker indicates that they had previously thought $\neg p$, and must now find a way to reconcile p and $\neg p$. With *what*, the speaker is only responding to a single proposition. Because they are expressing surprise at this proposition, it would be odd to assert this too—it would create an inference of both uncertainty and authority.

Similarly, the excited contour anchors elements to a speaker's excitement or perceived informational importance. Marking something as informationally important allows for the inference that a speaker has some sort of prior knowledge or authority for this proposition. In this way, excited contours are ruled out with *what* in the same way that falling ones are. *What*, like *huh*, accounts for situations in which the speaker knows that her expectations about a proposition have been violated. The difference is that with *huh*, the speaker encodes another proposition that she finds more likely.

What's more limited use as a discourse particle may be a remnant of the semantic content that is attached to the phonologically identical wh-word. In that use, the element is used only in syntactically [+wh] contexts, which in itself is a limited set of sentence types. It seems that this particle is only licit with the SRC in the restricted inventory of prosodic contours under discussion in this work. For the purposes here, we will only deal with *what* in SRC contexts.

3.3 Discourse particles, discourse contexts, and the relation of text to tune

While previous accounts of English discourse particles show an impressive range of flexibility in what they can express, no research to date looks at the compositionality of discourse particle meaning and intonational meaning. But while it is important to look at these different particle and contour combinations, it is perhaps more informative to look at the kinds of discourse contexts that these particles and contours can be use in. In an effort to sharpen the picture even further of the function of discourse particles and prosodic interaction, I outline a number of contexts that turn out to have an interesting effect on a particle performance's felicity.

Doing this constitutes another manipulation in the search for these particles' pragmatic contribution to an utterance. By pinpointing specific discourse contexts and holding them constant, we are able to see what other kinds of extra-linguistic factors these pragmatic markers are sensitive to. The felicity of these particles is not limited to their combinability with particular contours; they interact with the overall structure of the discourse as well.

By doing this, we can more accurately narrow down where differences in meaning or interpretation stem from. If a particle and contour pair is infelicitous in one context, we can form a hypothesis about the cause of this change. If we have held everything constant save for one element, and in doing this, observe a measurable change, we can reasonably assume that this change is meaningful. Perhaps felicity conditions on one prosodic contour clash with the discourse assumptions that have been made common ground, but are mollified with a change in the prosodic contour. In contrast, if a contour is held constant across a single context and particles show varying degrees of felicity, we can conclude that any observed difference must be inherent in the meaning of the particles themselves.

The following eight mini-discourse contexts will be used in what follows to disambiguate the pragmatic contributions of the particles *oh, huh*, and *what* when paired with the neutral, excited and surprise-redundancy contours. What all of these contexts have in common is the notion of expectation violation. Use of these particles must happen in a situation where surprise plays some role in an utterance's interpretation.

(121) **Discourse contexts for diagnosing changes in particle/contour felicity**

a. **Sudden realization:** Out-of-the-blue contexts. The speaker has just come to know a fact, either by direct observation, or from inferred reasoning strategies.

EXAMPLE: *The speaker walks outside from a windowless building, having no previous knowledge of the weather*:

(Oh/huh/what.) It's raining!

b. **Implied Speaker responsibility:** Another conversational agent raises an issue to the speaker in which it is implied that the speaker is responsible for, or should be responsible for the resolution of the issue.

EXAMPLE: *A is making cookies and is searching for an ingredient in the cupboard*:

A: We're out of flour.

B: (Oh/huh/what.) You didn't put it on the list.

c. **Eavesdropping:** The speaker comments on the actions of an agent who is not in the current discourse.

EXAMPLE: *Speaker is watching a movie, when one of the characters contradicts a plot point*:

(Oh/huh/what.) That was clearly Julia Roberts.

d. **Accept a fact:** The speaker fully acknowledges and commits to a preceding utterance that they previously indicated uncertainty toward.

EXAMPLE: *The speaker is on a quiz show, and buzzes in to respond to a question. Upon hearing the correct answer*:

(Oh/huh/what.) I was right.

e. **Contradict a statement:** The speaker directly contradicts a preceding utterance, p vs $\neg p$

EXAMPLE: *A asserts that Sandy is from Nebraska.*

B: (Oh/huh/what.) No she isn't.

f. **Correct a fact:** The speaker offers alternative information to show that a previous proposition should be accepted as false.

EXAMPLE: *A asserts that Sandy is from Nebraska.*

B: (Oh/huh/what.) She's from California.

g. **High situational gravity:** The discourse context is construed as highly unfavorable for one or more participant.

EXAMPLE: *A has just confessed that she has cancer.*

B: (Oh/huh/what.) That's awful.

h. **Low situational gravity:** The discourse context is mildly unfavorable for one or more participant

EXAMPLE: *A has just spilled coffee on her shoes.*

B: (Oh/huh/what.) That's awful.

In what follows, I rely heavily on the notions of expectation violation and surprise as outlined in Chapter §2. I mean this in terms of a speaker's subjective epistemic uncertainty given her representation of the discourse context and the common ground. What is important here is that epistemic likelihood drives the interpretation of mirative strategies, both in terms of discourse particles and prosodic contours. Expectation in this sense represents the speaker's own internal degree of belief or disbelief of an issue on the table or in a discourse context.

In keeping with this, there are many varied contexts in which a speaker might find herself in a state of surprise. Violated expectations are not always violated in the same way. The number of contexts here is to a large extent arbitrary, but does serve the larger immediate purpose of successfully disambiguating between the three particles under scrutiny here. This is of course also language-specific—these contexts are a breakdown of different scenarios where a speaker of *English* might find herself using a mirative strategy. These same contexts may not induce the same effects in other languages. I take this breakdown of expectation-violating contexts as a starting point for understanding these English mirative markers. Table 3.1 is a preview of what is to come, and outlines the patterns and contexts to be discussed in the following sections.

Table 3.1: Utterance contexts and contours for the particles *oh*, *huh* and *what*

Neutral contour	Example	*oh* + H*L%	*huh* + H*L%	*what* + H*L%
Sudden Realization	(124)	✓	✓	*
Implied Speaker Responsibility	(125)	✓	✓	*
Eavesdropping	(126)	✓	✓	*
Accept a fact	(127)	*	*	*
Contradict a statement	(131)	✓	*	*
Correct a fact	(138)	✓	*	*
Solidarity, situational gravity: high	(132)	✓	*	*
Solidarity, situational gravity: low	(133)	✓	✓	*
Excited contour	Example	*oh* + ↑ H*L%	*huh* + ↑ H*L%	*what* + ↑ H*L%
Sudden Realization	(145)	✓	✓	*
Implied Speaker Responsibility	(146)	✓	✓	*
Eavesdropping	(147)	✓	✓	*
Accept a fact	(154)	✓	✓	*
Contradict a statement	(149)	✓	*	*
Correct a fact	(148)	✓	*	*
Solidarity, situational gravity: high	(150)	*	*	*
Solidarity, situational gravity: low	(151)	*	*	*
Surprise-redundancy contour	Example	*oh* + SRC	*huh* + SRC	*what* + SRC
Sudden Realization	(167)	*	✓	✓
Implied Speaker Responsibility	(157)	✓	✓	✓
Eavesdropping	(158)	✓	✓	✓
Accept a fact	(159)	✓	✓	✓
Contradict a statement	(162)	*/?	✓	✓
Correct a fact	(163)	✓	✓	✓
Solidarity, situational gravity: high	(160)	✓	✓	✓
Solidarity, situational gravity: low	(161)	✓	✓	✓

In broad strokes, a few patterns about the distribution of the particles show subset and superset relations between particular particle and contour pairs emerging, as well as blanket restrictions on other particle and contour combinations. As outlined in the previous section, the table lists *what* as disallowed across all sentence contexts when paired with either the neutral falling contour or the excited contour. These restrictions are lifted under the SRC, and the particle with this contour is licit in all scenarios. When paired with either the neutral falling or the excited contour, *huh* appears in a proper subset of the environments that license *oh*. But this generalization is reversed when *oh* and *huh* appear with the SRC. In all these cases, it is a combination of the pragmatic contribution attributed to a particular particle and the contribution of a particular prosodic contour that can derive the patterns seen in the table above. The following three sections break this down, and show just how these patterns emerge in particular discourse contexts.

3.3.1 Text and Tunes: Neutral Falling H*L%, *oh*, and *huh*

The apparent synonymy enjoyed by *oh* and *huh* in many of the neutral contexts presented in this work so far suggest that the particles are extremely similar. This is to a large extent true. In part, this can be attributed to the types of expectation-violating contexts that the particles have been presented in so far. Recall that *oh* and *huh* differ only very slightly in the expectations that are placed in the discourse commitments of the speaker:

(122) *oh* and *huh* are anaphoric to a proposition p salient in the discourse s.t.:

a. *oh*(p) adds the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(p) < \text{Exp}_{spkr}(\neg p)$

b. *huh*(p) adds the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(p) < \text{Exp}_{spkr}(\neg p) \wedge \text{Exp}_{spkr}(p) > 0$

The main difference between the two is the extra pragmatic effect that *huh* expresses. When using this particle, while the speaker is surprised, she conveys that she did not rule out entirely the situation that she currently finds herself in.

The effect of both of these particles is calculated in large part in conjunction with pragmatic effects of intonation, a statement that is to be taken literally: particles must be expressed with some kind of audible tune, and the effects of both are additive. In what I assume is the base case, neutral, falling contours on discourse particles have the effect of an assertion—a property of an intonational contour, not of a particular sentence form. I assume that neutral, falling contours assert the content of the proposition they take as their prejacent. In this sense, I follow Farkas & Roelofsen (2017) and Malamud & Stephenson (2015) in assuming that a falling declarative asserts speaker belief in *p*, which is contributed by the contour:

(123) The neutral falling H*L% adds the following to the speaker's *DC*s:
$\text{Exp}_{spkr}(\text{p}) \approx 1$

For the most part, *oh* and *huh* pattern in very similar ways, and are available in overlapping contexts. Both can be used with the neutral contour to indicate sudden realization (124), implied speaker responsibility (125), and eavesdropping situations (126). They are both also disallowed in cases of fact acceptance, as in (127):

(124) *Speaker walks out of a building, with no expectation of the weather.*

a. Oh. It's raining.

b. Huh. It's raining. *Sudden Realization*

(125) A: There's no baking powder!

a. B: Oh. You didn't put it on the list. *Implied speaker Responsibility*

b. B: Huh. You didn't put it on the list.

(126) A, *Character speaking to other characters in a film*: It's Tess Ocean!

a. B, *Moviegoer, observing the film:*
Oh. It's clearly Julia Roberts. *Eavesdropping*

b. B: Huh. It's clearly Julia Roberts.

(127) A, (*Quizmaster*): What's the capital of Delaware? *Accept a fact*
B: Dover.
A: It is Dover.

a. B: #Oh. I was right.

b. B: #Huh. I was right.

But for how similar *oh* and *huh* seem, there are places where their distinct contributions become clear. The Gunlogsonian test for commitment with the neutral falling contour appears to extend to *huh* in cases where the speaker has less evidence or a lower expectation from A's claim (128). But in cases where a speaker has equal to or more evidence, or a pre-established high expectation for the truth of a proposition, *huh* is no longer a licit response, let alone a diagnostic for commitment (129):

(128) A: The server's down.

a. B: Oh. (I didn't know that.)

b. B: Huh. (I didn't know that.)

(129) A: The server's down.

a. B: Oh. I know. *(Gunlogson 2008: p. 11)*

b. B: # Huh. I know.

In (129), the speaker sets up an inconsistent set of discourse commitments if she uses *huh* to mark her surprise. In this context, B already knows the content of A's utterance to be true. It is not this information that is surprising. Rather, the particle reacts to the speech event itself. B is surprised that she is being informed of something that she has reason to believe is common ground. When she uses *huh* to indicate this, she is communicating the following things about her doxastic state:

(130) a. A expected B to <u>not</u> make the speech event "The server is down" [$\text{Exp}_A(p) < \text{Exp}_A(\neg p)$]

b. B holds a non-zero expectation for the possibility that A will make the speech event "The server is down."

But while this is a consistent belief state to hold, it is a confusing thing to convey to a speech partner. It is akin to telling them, "I didn't expect for you to perform that speech act, but also, I did expect that the chance that you *would* was non-zero." When surprise is directed at the level of illocution, *huh* is a strange pragmatic choice. This seems to be a special case of fact contradiction contexts, as in (131), where a speaker attempts to use *huh* in an instance where they mean to directly contradict the previous statement:

(131) A: The server's down.

a. B: Oh. No it isn't. (It just looks like that).

b. B: # Huh. No it isn't.

Note that there is a reading of (131b) that is felicitous, but it requires some accommodation on the part of the listener. With *oh*, the speaker can respond without looking up from what she is doing: this is a fact that she knows to be true. For *huh*, the only way this is felicitous is if the speaker utters the discourse particle, gets up and checks the status of the server, and reports back. Only then can B correct with the follow-up utterance. This does not constitute an explanatory relationship between discourse particle and utterance continuation–the two are distinct conversational moves. In contrast to *oh*, *huh* <u>must</u> acknowledge and at least temporarily accept the previous proposition. Revisions can be subsequently made, but the come at the cost of striking the accepted content from the discourse record.

Restricting the utterance context of *huh* and *oh* to only falling intonation introduces a subset-superset relation of acceptability and usage. In all cases where *huh* is licensed, *oh* is licensed as well, yet *oh* is much more permissive. Take, for instance, the gradable situational gravity measures introduced by (132-133):

(132) ***High*** *situational gravity*

A: I just found out I have cancer.

a. B: Oh. That's awful.

b. : # Huh. That's awful.

(133) ***Low*** *situational gravity*

A: I spilled coffee on my new shoes.

a. B: Oh. That's awful.

b. B: Huh. That's awful.

In cases of high situational gravity, B can easily respond with *oh* and convey the weight of the information just relayed to him. With *huh*, the speaker comes off as insincere, as if he is not taking the seriousness of the situation into account. In Brown & Levinson (1987) terms, this discourse move is both face-threatening and uncooperative. But in cases where the situational gravity is not as high, the *huh* utterances become much better. In (133), as in previous examples, B′ appears to use *huh* to accept (in some way or another) the contribution of A, and then explain the relationship between his surprise and his continuation. The *huh* cases are more restricted cases of *oh*-marked utterances. This is of course clear in the formulations of *oh* and *huh* in (122a) and (122b): *huh* has an added restriction that there must be some expectation already calculated for a given surprising event or proposition.

This predicts that *oh* may be used in cases of extreme surprise or with unexpected events, events in which the speaker may have zero expectation for p. For *huh*, while the context does favor expectation for $\neg p$ over p, it still requires p to be an available option for the speaker's chain of reasoning.

Both particles are different mirative marking strategies, which rely on the expectations of the speaker. In most of the contexts discussed here, the surprising element is an issue raised and placed on the table by an interlocutor. But as we have just seen, out of the blue cases of surprise, as well as some correction cases, target a larger event. The restrictions on each particle determine their felicity in particular contexts. What they have in common is the ability to make clear a speaker's expectations surrounding certain discourse moves, introducing their own subjective bias for these claims.

As discussed briefly in Section (3.2), I proposed that these markers mirror the contributions of intonation by adding expectations to a speaker's discourse commitment list. In doing so, the speaker makes these commitments known to their audience without subjecting these commitments to the scrutiny of the discourse community at large. The contributions of these markers are not-at-issue. Placing these commitments in a table model of discourse is a convenient way of keeping track of issues raised, common ground material, and conversational participants' subjective epistemic states. Discourse moves are recorded in this updated table model. For particle only responses like (128), a speaker's discourse commitments are updated as in (134):

(134) A: The server is down.

B: Huh.

a. Context after A's utterance:

$\mathbf{DC}_A$	**Table**	$\mathbf{DC}_B$
p, $\text{Exp}_A\ (p) \approx 1$	$\{p\}$	

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

b. B utters *huh*

c. Context after B utters *huh*:

$\mathbf{DC}_A$	**Table**	$\mathbf{DC}_B$
p, , $\text{Exp}_A\ (p) \approx 1$	$\{p\}$	$\text{Exp}_B(p) < \text{Exp}_B(\neg p)$ $\wedge\ \text{Exp}_B(p) > 0$

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

With no clear commitment one way or another, the speaker and hearer may choose to take this particle response as committing B to p, thus clearing the scoreboard. This reasoning is due in part to the contribution of the prosody. In our case in (134), a neutral falling contour on the particle allows A to infer that B has accepted their contribution. In a subsequent step, this can grow the common ground. When B responds with the particle and a following utterance, the particle puts the speaker's expectations into their discourse commitments, and puts the following utterance on the table. Pragmatic reasoning again takes the entire utterance and performance into account to determine whether B's contribution was an accepting or rejecting move.

Here, I modify the table model to include the stack of propositions on the table that have yet to be accepted or rejected. In (135), neither p nor q has been explicitly accepted by both participants. For the sake of simplicity, I assume that only the most recent proposition in the stack, q is represented in the projected set, though there may be other approaches to representing this state of a discourse.

(135) A: The bank is closed today. = p

B: Oh. I need to deposit a check. = q

a. Context after A's utterance: $\Rightarrow$

$\mathbf{DC}_A$	**Table**	$\mathbf{DC}_B$
p, $\text{Exp}_A\ (p) \approx 1$	$\{p\}$	

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

b. B utters *oh* + q

c. Context after B utters *oh*+q:

$\mathbf{DC}_A$	**Table**	$\mathbf{DC}_B$
p, $\text{Exp}_A\ (p) \approx 1$	{q} ——— {p}	q, $\text{Exp}_B(q) \approx 1$, $[\text{Exp}_B(p) < \text{Exp}_B(\neg p)]$

$cg : s_0 = s_1,\ ps = \{\{s_0 \cup q\}\}$

B's utterance here is not a clearly accepting or rejecting move. As *oh* does not supply this confirmation, and B's following utterance leaves the option open, both p and q are issues up for discussion on the table. In the response case of (135), we can assume that the falling contour on both the particle and the utterance lead the listener to infer that their issue had been resolved by the speaker. At this point, p can be cleared from the table into the common ground.

I assume that the individual discourse commitments of a particular speaker are related by conjunction. Pragmatic meanings contributed by particles and contours are calculated additively by the expressions in a participant's discourse commitments. The individual contributions of *oh* and *huh*, coupled with the contributions of the neutral falling contour distinguish the varying levels of appropriateness in high situational gravity cases, as in (132), repeated below as (136).

(136) ***High*** *situational gravity*

A: I just found out I have cancer.

a. B: Oh. That's awful.

b. B: # Huh. That's awful.

The pragmatic effect of a falling contour on *huh* in cases of high situational gravity has the following effect: falling H*L% prosody on both the particle and the continuation presents the speaker as expecting their utterance to be true, and has the conventional discourse effect of an assertion. The bare meaning component of *huh* expresses an expectation for $\neg p$, while indicating that p might also be a possibility. Though B′ indicates to A that he expected the opposite of her proposition to be true, namely, *It's not the case that A has cancer*, he also indicates that he held a non-zero expectation that p might be the case. Though this is a pragmatic inference, it is strong, especially in the face of the alternative *oh*. B$'$ here signals to his interlocutor that he already assumed that A's chance of getting cancer was a computable, real world option. With *oh*, the speaker's expectation of p could realistically have been zero. B$'$'s response in (132) comes off as insincere, as he doesn't take the gravity of the situation into account.

With the contextual seriousness lowered, *huh* again becomes an option for expressions of surprise:

(137) ***Low** situational gravity*

A: I spilled coffee on my new shoes.

a. B: Oh. That's awful.

b. B$'$: Huh. That's awful.

When the stakes are lower, and all that has happened is that A has spilled coffee on her new shoes, both particles are equally permissible. In this case, it is not so socially taboo to assume that there is a non-zero possibility of a coffee-spilling event. Including that calculation in a discourse particle selection does not induce an ordering with pragmatic repercussions.

This distinction between non-zero and no expectation can also explain the oddness introduced by fact correction scenarios, as in (138):

(138) A: Sandy is from Nebraska.

a. B: Oh. She's from California.
   H*-L% H*-L%

b. B: # Huh. She's from California.

In both (131) and (138), the B utterances disagree with A's proffered content, in the former, by offering explicit contradiction, and the latter, by offering up conflicting information that B knows. In both of these scenarios, *oh* is allowable, while *huh* is not. Knowing a fact q about the world implies high expectation for it to be true, meaning that [$\text{Exp}_{spkr}(q) \approx 1$]. If someone utters p, a fact that would make q false, a speaker can signal this discrepancy with *oh*; based on their belief in q, they have no reason to believe p to be true. When B knows that Sandy is from California, he has reason to believe that she is not from other places: being from California implies that she is not also from Nebraska. With this information in B's discourse commitments, but not actively on the Table, the listener can infer the following:

(139) *Sandy is from Nebraska* = p *Sandy is from California* = q

a. Context after A's utterance: $\Rightarrow$

$\textbf{DC}_A$	**Table**	$\textbf{DC}_B$
p, $\text{Exp}_A(p) \approx 1$	$\{p\}$	

$cg: s_0,\ ps = \{\{s_0 \cup p\}\}$

b. B utters *huh* + q

c. Context after B utters *huh* + q:

$\textbf{DC}_A$	**Table**	$\textbf{DC}_B$
p, $\text{Exp}_A(p) \approx 1$	{q} / {p}	q, $\text{Exp}_B(q) \approx 1$, $[\text{Exp}_B(p) < \text{Exp}_B(\neg p) \wedge \text{Exp}_B(p) > 0]$

$cg: s_0 = s_1,\ ps = \{\{s_0 \cup q\}\}$

i. **Falling contour:** $\text{Exp}_B(q) \approx 1$

ii. ***huh*:** $\text{Exp}_B(p) < \text{Exp}_B(\neg p) \wedge \text{Exp}_B(p) > 0$

iii. **Pragmatic inconsistency:** $q \rightarrow \neg p$, so $\neg p = 1$. Since q and p are contrary, one cannot commit to q being true and to p being possible.

Huh's added restriction that the expectation for p be non-zero leads to an inconsistent pragmatic calculation. If B knows his information is correct, there is no way that he can felicitously signal that with *huh*. But he can with *oh*:

(140) *Sandy is from Nebraska* = p *Sandy is from California* = q

a. **Falling contour:** $\text{Exp}_B(\text{q}) \approx 1$

b. ***oh*:** $\text{Exp}_B(p) < \text{Exp}_B(\neg p)$

c. **Pragmatic consistency:** $q \rightarrow \neg p$, so $\neg p = 1$. B must have expected p to be 0.

Notice how *huh* again becomes licit when B's uncertainty is put back into play:

(141) A: Sandy is from Nebraska.

a. B: Oh. I thought she was from California.

b. B: Huh. I thought she was from California.

In those cases where B indicates overtly that he is not sure about the contribution, *huh* again becomes licit. Here, the semantic content of his utterance allows for felicitous interpretation of *huh*.

The following table summarizes the relevant points from this section: The conditions that license *huh* are a subset of the conditions that license *oh* in neutral, falling contexts.

Table 3.2: Table of usage for neutral falling *oh* and *huh*

	Example	*oh* + H*L%	*huh* + H*L%
Sudden Realization	(124)	✓	✓
Implied Speaker Responsibility	(125)	✓	✓
Eavesdropping	(126)	✓	✓
Accept a fact	(127)	*	*
Contradict a statement	(131)	✓	*
Correct a fact	(138)	✓	*
Solidarity, situational gravity: high	(132)	✓	*
Solidarity, situational gravity: low	(133)	✓	✓

3.3.2 Text and Tunes: Excited *oh* and *huh*

As particles that inherently comment on a speaker's surprise, we expect both *oh* and *huh* to be able to appear with the excited contour.[4] This is the case with the particles in (142):

[4]As indicated earlier in Section §3.2.4, *what* is disallowed with this contour for independent reasons.

(142) *A teacher hands back an exam to a student, who has received a higher grade than expected:*

a. Oh!

b. Huh!

The excited contour is not in itself a mirative strategy. While it often combines other strategies that include surprise, it need not. Rather, the contour indicates newness, and positive speaker reaction to this newness. This, too, can be defined in terms of a speaker's expectations, As the excited contour still asserts the content of an utterance, the speaker indicates that they have a high expectation for the truth of that proposition. This positivity is framed in (143) in terms of a speaker's desires: a speaker desires that the content of proposition p they have asserted to be true over $\neg p$.

(143) The excited H*-L% adds the following to the speaker's *DC*s:

$\mathrm{Exp}_{spkr}(p) \approx 1 \wedge \mathrm{Boul}_{spkr}(p) > \mathrm{Boul}_{spkr}(\neg p)$

With the Excited contour, we see again that while *oh* and *huh* can be uttered felicitously, *what* is disallowed[5].

(144) A: Marcie's going to have a baby!

a. B: Oh! How wonderful!

b. B$'$: Huh! How wonderful!

c. B$''$: # What! How wonderful!

In many contexts, the same subset-superset relation between the particles holds. In sudden realization, speaker responsibility, eavesdropping, fact correction, and contradiction cases, excited intonation does not change the felicity of the particle in a particular context, only the implied speaker affect:

(145) a. Oh! It's raining! Sudden Realization

b. Huh! It' raining!

[5] As with the falling contour, we turn away quasi-vocative, speaker-accusation readings of *what*.

(146) A: There's no baking powder! Speaker Responsibility

a. B: Oh! You didn't put it on the list!

b. B: Huh! You didn't put it on the list!

(147) A, *Character speaking to other characters in a film:* It's Tess Ocean! Eavesdropping

a. B, *Moviegoer, observing the film:* Oh! It's clearly Julia Roberts!

b. B: Huh! It's clearly Julia Roberts!

(148) A: Sandy is from Nebraska. Fact Correction

a. B: Oh! She's from California!

b. B: # Huh! She's from California!

(149) A: The server's down. Contradiction

a. B: Oh! no it isn't! (It just looks like that).

b. B′: # Huh! No it isn't!

All of the grammaticality judgments here with the excitement contour track the judgments reported in §3.3.1. Indicating positive-valence or novel information in the surprise-inducing contexts above does not change the felicity of B's response. They all pragmatically signal the speaker's emotional state with respect to the current discourse context. In clearly negative contexts, the added excitement does register a pragmatic effect. For contexts with high or low situational gravity, the excited contour is infelicitous, though there does seem to be a slight difference in acceptability. The excited contour pragmatically signals to the listener that the speaker is reacting to new and surprising information, which in cases where something unpleasant has happened, is odd. This is not so much a restriction on the particle, but on the particle and contour combination:

(150) ***High*** *situational gravity*

A: I just found out I have cancer.

a. B: # Oh! That's awful!

b. B: # Huh! That's awful!

(151) ***Low** situational gravity*

A: I spilled coffee on my new shoes.

a. B: $^{?}$/# Oh! That's awful!

b. B: $^{?}$/# Huh!. That's awful!

In cases of low situational gravity, the oddness that is calculated can be explained by conflicting pragmatic inferences: the restrictions on the discourse particles assume that events that signal speaker expectation violation should be felicitous with *oh* and *huh*. This is certainly one context where a surprise response is licit. For both of the situational gravity cases, the pragmatic calculations are much the same. The contributions of the discourse particle and the intonational contour are logged in the speaker's discourse commitments list, and their conjunction leads to confusion, given the conversational backdrop.

(152) *A has cancer* = *p* *That's awful* = *q*

a. Context after A's utterance: $\Rightarrow$

DC$_A$	**Table**	**DC$_B$**
p, $\text{Exp}_A(p) \approx 1$	$\{p\}$	

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

b. B utters *oh!*

c. Context after B utters *oh*:

DC$_A$	**Table**	**DC$_B$**
p, $\text{Exp}_A(p) \approx 1$	{p}	$\text{Exp}_B(p) \approx 1 \wedge \text{Boul}_B(p) > \text{Boul}_B(\neg p)$, $\text{Exp}_B(p) < \text{Exp}_B(p)$

$cg : s_0 = s_1,\ ps = \{\{s_0 \cup p\}\}$

i. **Excited contour:** $\text{Exp}_B(p) \wedge \text{Boul}_B(p) > \text{Boul}_B(\neg p)$:
B indicates positive stance toward p

ii. ***oh*:** $\text{Exp}_B(p) < \text{Exp}_B(\neg p)$

iii. **Pragmatic inconsistency:** B indicates that the desire for p is greater than $\neg p$ with this contour, As p is A's news that she has cancer, this is not an appropriate response.

Hearing that someone has spilled coffee on themselves is a similar for the low situational gravity cases. Using the excited contour gives the impression that the speaker is being insincere. While this is new information, the contour is conventionally used to indicate strong, positive association between a speaker and new information. Using the excited contour with these discourse particles leads A in (151) to think that B somehow is excited about her misfortune. Varying the contour in this context results in infelicity.

Placing a speaker's expectations within the table model of discourse but not directly on the table itself is one way of accommodating a speaker's attitude toward a discourse move. With intonation, like with the discourse particles, this information is placed in the discourse commitments for reasons like those in (153): bringing a contour's pragmatic contribution into direct question is both odd and uncooperative.

(153) A: *(with excited contour)* We just got a puppy!
B: # That's not true. You're upset.

Using *oh* and *huh* with the excited contour also leads to felicitous situations where neutral particle and contour combinations in certain contexts were previously judged as pragmatically odd. In cases of fact correction like (154), the contour on the particle signals that the speaker has judged some part of the preceding discourse as new or surprising:

(154) A, *Quizmaster:* What's the capital of Delaware?
B: Dover.
A: It is Dover.
a. B: Oh! I was right!
b. B: Huh! I was right!

In both cases, the discourse particles' meanings can pragmatically accommodate a speaker's reaction to new information. Though B's response is uttered with falling intonation, his follow up in response to the quizmaster's confirmation accommodates the fact that though he positioned himself as being authoritative about his previous utterance, he did in fact require some confirmation. Using the excited contour on the discourse particles indicates his context has been updated, thus leading the hearer to infer that he had not excluded the fact that his initial utterance may be wrong. The listener is invited to infer that though the speaker positioned himself as knowing a fact, *oh* and *huh* signal that the speaker had some expectation, however low, that his answer may be incorrect.

The mirative component of both *huh* and *oh* is pragmatically restricted by the excitement contour. In these cases, the degree to which a speaker's expectations may be violated must be of a positive valence. This effect of the exclamative contour shapes the interpretation of the discourse particles, showing a more restricted distribution in the various environments introduced by examination of these particles in neutral falling contexts. Table 3.3 schematizes these observations, and shows again that the meaning of *oh* is present in the meaning of *huh*.

Table 3.3: Table of usage for excited intonation and *oh* and *huh*

	Example	*oh* + $\uparrow$ H*L%	*huh* + $\uparrow$ H*L%
Sudden Realization	(145)	✓	✓
Implied Speaker Responsibility	(146)	✓	✓
Eavesdropping	(147)	✓	✓
Accept a fact	(154)	✓	✓
Contradict a statement	(149)	✓	*
Correct a fact	(148)	✓	*
Solidarity, situational gravity: high	(150)	*	*
Solidarity, situational gravity: low	(151)	*	*

3.3.3 Text and Tunes: Surprise-Redundancy, *oh*, *huh* and *what*

The previous two sections showed the overlap in meaning between *oh* and *huh*, as well as their composition with two phonetically similar but pragmatically distinct falling contours. In this section, pragmatic effects of the surprise-redundancy contour (SRC) create a split between the conditions of use of *huh* and *oh*. But this contour reveals an interesting fact about the discourse particle use of *what*. In conjunction with the SRC, *what* and *huh*'s conditions of use completely overlap.

The SRC is more complex than other final falls, requiring at least a rise-fall pattern with two different prominent tone heights, (H) L* H*-L%. The pragmatic effect of the SRC is complex as well–it indicates that the speaker believes the listener "should have known" some salient proposition in the discourse (Hayes 1995). This contour bakes in a mirative meaning itself, anchoring itself to speaker expectations about a participant in the discourse. Formally, I represent the pragmatic contribution of the SRC to the discourse in (155):

(155) The English Surprise Redundancy Contour (H) L* H*-L% is anaphoric to a salient proposition or event p in a discourse context C, and is admissible for discourse-salient participants x when

a. q is the proposition expressed by the speaker (uttered content or the presuppositions introduced by a question),

b. add the following to the speaker's *DC*s:
$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

This formalization does not presume to break down the prosodic components of the SRC into meaningful segments corresponding to particular phonetic features. Rather, it assigns a conventional pragmatic effect to this phrase-level melody. In doing so, it anchors speaker expectation to a proposition q, while expressing the speaker's secondary expectation that for anyone in the discourse context, given q, they should have very low, if not zero expectation for p. In other words, the speaker expresses surprise that the conversational participants have deduced p, given what she expects them to have predicted by knowing q. The speaker does not expect p, and assumes no one else should either.

A small note on the notation used here to show the relationship between p and q is in order. The formulation $\text{Exp}_x(p|q) \approx 0$ could refer to one of two things. This notion is used both to notate conditional probability, as well as likelihood ratios. Here, this should be understood in the likelihood function use for the simple fact that we are not making predictions about some future outcome or state of the discourse. Rather, this contour is used only when an observation has been made, and a speaker is attempting to find an explanation for this utterance that does not accord with her view of the discourse. In effect, the SRC says that given the data you have been presented with, the only way it can be explained is if expectations are set up in the way described in (156). Given the facts that we have and the observations that have been made, what is the likelihood that they turn out to be true?[6]

Just like the other mirative strategies, I assume here that using the SRC is a strategy that places the speaker's current epistemic state into active view. By placing it in the discourse commitments of the speaker, but not on the table itself, the speaker signals to the listener that this is not at-issue content. Like with the Excited contour, the content of the SRC is something that the speaker commits to:

(156) A, *(with src:)* This cheddar cheese is orange!↘

B: # That's not true. You aren't surprised at all.

The SRC can combine with all three discourse particles of interest, and in particular, is licensed by *huh* and *what* in all eight contexts. In many of these contexts, this contour can combine with *oh* as well. So as not to confuse this with a rising declarative, in the examples below, I use an exclamation mark followed by a question mark '!?' to represent the SRC. Often, the SRC on the particle can also co-occur with a final rise or rising declarative question intonation which I indicate with a '?' when appropriate:

[6] I make the choice to use the likelihood ratio for the relevant notion of violated expectations for a few reasons here. Perhaps the most salient is the basic idea that if you use simple probability, it will typically not be the case that average propositions and their negations will have roughly equal probability. This is based off of the observation that for a proposition p to hold, all of its entailments must hold as well. But for $\neg p$ to hold, just one of the set of entailments that make p true can be false. Given this, it will tend to be the case that out of p and $\neg p$, one of these will have an intrinsically higher base probability than the other. But though this is a valid observation, this fact should not have an impact on the calculation of surprise.

For example, we can say that the proposition *Jeff's plane comes in tonight* is more likely to be true than *Jeff's plane comes in tonight at 10:45pm*. This is true on the surface, because we have added an entailment to the latter context—they differ in terms of prior probability. But in a a context where both of these are surprising, they don't differ in inherent surprisal; prior probability by itself shouldn't be responsible for our notion of violated expectations.

If we think about surprise in terms of joint probability, conditional probability, or even just raw probability, we miss the relevant idea of surprise because we cannot relativize this to a particular context. We want to get rid of the notion that surprise can be calculated in a context-free environment, which is what we would predict if we used any of the measures mentioned above. If we use a likelihood ratio of the joint probability of the antecedent and the nat-issue proposition, normalized against the probability of the antecedent and the at-issue proposition ($\frac{P(p,q)}{P(p)\times P(q)}$), we get rid of the question of the sensitivity to the number of entailments that are carried either by the antecedent (p) or the at-issue proposition (q). Likelihood ratios are a good way of achieving the normalization that you need for context sensitivity by unlinking the violated expectations mechanism from raw probability. Thanks to Paul Willis and Pranav Anand for discussion on this topic.

(157) A: There's no baking powder! Speaker Responsibility

a. B: Oh!? You didn't put it on the list!?

b. B: Huh!? You didn't put it on the list!?

c. B: What!? You didn't put it on the list!?

(158) A, *Character speaking to other characters in a film:* It's Tess Ocean! Eavesdropping

a. B, *Moviegoer, observing the film:* Oh!? It's clearly Julia Roberts!?

b. B: Huh!? It's clearly Julia Roberts!?

c. B: What!? It's clearly Julia Roberts!?

(159) A, *Quizmaster:* What's the capital of Delaware? Fact Acceptance

B: Dover.

A: It is Dover.

a. B: Oh!? I was right!?

b. B: Huh!? I was right!?

c. B: What!? I was right!?

(160) A: I just found out I have cancer. **High** situational gravity

a. B: Oh!? That's awful!?

b. B: Huh!? That's awful!?

c. B: What!? That's awful!?

(161) A: I spilled coffee on my new shoes. **Low** situational gravity

a. B: Oh!? That's awful!?

b. B: Huh!? That's awful!?

c. B: What!? That's awful!?

Each particle interacts with the SRC in a very subtle way. With *oh*, there is a pragmatic inference that the addressee's utterance is very speaker-new. The speaker is taken off guard; they have not previously entertained the proposition in question. This nicely falls out from *oh*'s characterization in (122a), as no imposed floor on a speaker's expectation for a preceding utterance allows for a genuine violation of expectations.

Like *oh*, *what* and *huh* also express heightened levels of disbelief when paired with the SRC. Another contribution of this contour comes at the perlocutionary level: the contour not only expresses a speaker's expectations, but also persuades the addressee to react. This contour is inherently a request for the addressee to bolster their claims before the speaker will commit to their discourse move. This effect is strong—so strong that it is strange for a discourse to end with an utterance that carries the SRC. The pressure is high for the addressee to respond.

As alluded to in Table 3.1, both *huh* and *what* fare well in all situations with the SRC. Both particles bake in the explicit possibility that the speaker's expectations are violable. Though as speaker may believe strongly in one thing, these particles indicate that this belief is not absolute. The difference is that *huh*, the speaker explicitly encodes that she expects that an alternate proposition is more likely.

The prosodic environment imposed by the SRC effectively flips some of the patterns of grammaticality form the three particles involved. Whereas composing *huh* with either the neutral or the excited contours revealed that this particle occurred in a subset of the contexts where *oh* could occur, the findings are switched with the SRC. *Oh* becomes the particle with the more limited distribution, which is a subset of the contexts in which *huh* can appear. One place where the particles pull apart is in fact correction and contradiction contexts:

(162) A: Sandy is from Nebraska. Fact Correction

a. B: Oh!? She's from California!?

b. B: Huh!? She's from California!?

c. B: What!? She's from California!?

(163) A: The server's down. Contradiction

a. B: #/? Oh!? No it isn't!? (It just looks like that).

b. B: Huh!? No it isn't!?

c. B: What!? No it isn't!?

One first thing to observe is that none of the B responses completely settle the issue at hand. Rather, they invite A to weigh in on the matter again, whether it be to provide counter-evidence, or to commit to the proffered alternative. This is a pragmatic effect from the SRC, calculated when the speaker projects their own relatively high expectations about the prejacent and the expectations generated about other participants' beliefs about the alternative. The SRC must comment on something mutually manifest. In the case that the speaker offers an alternative to *p* that their conversational participant couldn't have been aware of, the result is an incoherent discourse state, regardless of the particle used:

(164) *Sophie and Tom are supposed to pick Jeff up from the airport at 2pm. Tom, but not Sophie, receives a flight alert, telling him that Jeff will be delayed until 4pm. Later, they are talking about when the need to leave to pick Jeff up*:
Sophie: I can be ready to leave for the airport at 1:30pm.
Tom: # Oh/Huh/What? Her flight doesn't come in until 4!

Though the intuition is not as sharp as in other cases, the felicity of *oh* is questionable in cases of direct contradiction. Recall though, the difference between *oh* and *huh* in neutral contexts. There, *huh* was illicit, but in an interesting way. It necessitated extra accommodation on the part of the speaker in order for it to be interpretable. In point of fact, *huh* responses in those cases constituted two conversational moves: a discourse particle response, an intermediate verification step, and then a second corrective conversational update. A similar thing happens with SRC *oh* responses to contradiction environments. The (a) response in (163) is only licit if *oh* is one discourse move, followed by the speaker going to check the status of the server, and then issuing the corrective follow-up utterance.

So what is the reason for *oh*'s infelicity here? The answer ties back to the SRC's discourse management strategy that allows a speaker to assert an alternative proposition while implicitly requesting the addressee's feedback on it. The alternative that the speaker offers must be an alternative that the listener could arrive at by carefully considering the common ground and the projected set. *Huh* and *what* are no problem because their meanings bake in the fact that there is a non-zero chance that an alternative explanation is available. Cases of contradiction or fact correction are not purely contradiction because these mirative particles assume that however unlikely p may be in the speaker's estimation, it is still a viable alternative.

Oh on the other hand, has no requirement that the speaker's expectation of the prejacent be greater than zero. *Oh*'s freedom in this case is a weak restriction: because there are alternative particles that require uncertainty about a proposition, using a particle that does not confronts the listener with a potential Maxim of Quantity violation. Why would the speaker have used *oh* when *huh* was available? Perhaps it is because they want to indicate that they expect p to be 0 (from 163):

(165) *The server is down* = p *The server is not down* = q

a. Context after A's utterance: $\Rightarrow$

DC$_A$	**Table**	**DC**$_B$
p, $\text{Exp}_A(p) \approx 1$	$\{p\}$	

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

b. B utters *oh* + q

c. Context after B utters *oh* + q:

DC$_A$	**Table**	**DC**$_B$
p, $\text{Exp}_A(p) \approx 1$	$\frac{\{q\}}{\{p\}}$	q, $[\text{Exp}_B(p) < \text{Exp}_B(\neg p)$ $\wedge\ \text{Exp}_B(p) > 0]$, $[\text{Exp}_{A,B}(p\|q) \approx 0]$

$cg : s_0 = s_1,\ ps = \{\{s_0 \cup q\}\}$

i. **SRC:** $\text{Exp}_B(q) \approx 1\ \wedge \forall x \in DC\ [\text{Exp}_{A,B}(p|q) \approx 0]$

ii. ***oh*:** $\text{Exp}_B(p) < \text{Exp}_B(\neg p)$

iii. **Pragmatic consistency (?):** $q \rightarrow \neg p$. Speaker believes q = 1, infer $\neg p$ = 1. There is no restriction that p be 0.
BUT: Pragmatic competition with *huh* can lead a listener to wonder why *oh* was used.

If the speaker wants to directly contradict p, then the speaker must have evidence against p. But the SRC still allows for the possibility that the speaker's expectations might be violated, which sets up a (defeasible) inconsistency: a speaker can theoretically fully expect p

to be false (from the requirements of *oh*) but when they do this, they cannot also expect it to be *almost always* be false (from the SRC). Impressionistically, this appears to be a by-speaker pragmatic calculation in need of empirical validation. While some argue that *oh* with the SRC in contradiction cases is pragmatically fine, the same individuals do agree that there is a cline of acceptability in the following responses to the same utterance, with the direct contradiction cases more pragmatically odd:[7]

(166) A: Sandy is from Nebraska.

a. B: ??Oh?↘ No she isn't! She's from California!↘

b. B: ?Oh?!↘ She's from California!↘

c. B: Oh?!↘ She has a California driver's license?!↘

With *huh* and *what*, the listener accommodates the fact that they are not being directly contradicted: the speaker's expectation for p was crucially greater than zero to begin with, leaving room for the possibility that their expectation of q, however strong, must not be certain. Though the SRC signals to the listener that by the speaker's estimation, they should have known an alternative proposition q, it invites further commentary from the listener by leaving the door open for them to weigh in on the matter.

A similar but more robust pattern of acceptability emerges from cases of sudden realization:

(167) a. #Oh?!↘ It's raining?!↘ Sudden Realization

b. Huh?!↘ It' raining?!↘

c. What?!↘ It's raining?!↘

If a speaker walks out from a windowless building, having no expectation for it to be raining, it is strange for her to utter (167a). In these cases, especially when the speaker is talking to herself, this seems pragmatically odd. But when sudden realization cases involve another participant, *oh* gets better:

[7]Thanks to Margaret Kroll, Deniz Rudin, Erik Zyman, Tom Roberts, Maziar Toosarvandani, Jim McCloskey and Pranav Anand for an extended hallway discussion of this.

(168) *A and B are solving math problems. A turns to write one on the board and picks up the chalk in her left hand*:

B: Oh?!↘ You're left-handed?!↘

When the speaker is addressing herself, the Quantity implicature arises from the interaction of the SRC on the particle, the requirements of *oh*, and the pragmatic discourse effect of a rising declarative. The listener assumes that the speaker must have had some reason for using *oh* when *huh* was available as well. They deduce that the speaker expects q to be the case:

(169) *The fact that it's raining = p* *It is raining = q*

a. Context after B utters *oh* + q:

$\mathbf{DC}_A$	**Table**
q, $[\text{Exp}_A(q) \approx 1$	{q}
$\wedge\ \text{Exp}_{A,B}(p\|q) \approx 0]$	{p}

$cg = \emptyset$, $ps = \{\{p\}, \{q\}\}$

i. **SRC:** $\text{Exp}_B(q) \approx 1 \wedge \text{Exp}_B(p|q) \approx 0$

ii. ***oh*:** $\text{Exp}_B(p) < \text{Exp}_B(\neg p)$

iii. **Pragmatic inconsistency:** $q = \neg p$. Speaker believes $q = 1$, infer $\neg p = 1$.
There is no restriction that p be 0.
The speaker is both the addressee and the source of q; they hold one belief about the actual facts in the world, p, and another about their expectations, $q = \neg p$.

When p turns out to be true, the free variable in the definition of the SRC allows for the speaker to hold a belief about q approximately 1, while also also strongly expecting that p is not the case. But the speaker has also just observed the fact that it is raining. Using a rising declarative has the pragmatic discourse effect of indicating the speaker cannot commit to p, the highlighted alternative. But facts about the world prove otherwise. The speaker has seen the rain, and knows p is true. Her internal state is in conflict: she cannot know for sure that it's raining (p=1), expect that it's not ($q = \neg p$=1), and expect that given q, the expectation for p should be 0. In this case, there appears to be such a thing as too much mirativity.[8]

[8]Note that in this section, as opposed to the previous sections, expectations surrounding propositions which are

The table below summarizes the conclusions drawn from this section, which discuss the SRC in conjunction with the mirative markers *oh, huh* and *what.*

Table 3.4: Table of usage for Surprise-Redundancy Contour and *oh*, *huh* and *what*

	Example	*oh* + SRC	*huh* + SRC	*what* + SRC
Sudden Realization	(167)	*	✓	✓
Implied Speaker Responsibility	(157)	✓	✓	✓
Eavesdropping	(158)	✓	✓	✓
Accept a fact	(159)	✓	✓	✓
Contradict a statement	(162)	*/?	✓	✓
Correct a fact	(163)	✓	✓	✓
Solidarity, situational gravity: high	(160)	✓	✓	✓
Solidarity, situational gravity: low	(161)	✓	✓	✓

3.4 Taking stock

The previous sections have argued for the following interpretations of the English mirative discourse particles *oh, huh* and *what*, as well as for the pragmatic discourse effects of the neutral falling H*L% contour, the excited contour, and the surprise-redundancy contour.

The falling and excitement contours share a core effect of speaker authority over a proposition. When pitch excursions deviate from a speaker's normal range, this signals informational significance. The excitement contour is an exaggeration of the H*L% contour which marks information as new, surprising, and markedly positive. Both assert their propositional content.

≈ 1 are treated in two different ways. The neutral fall and excited contours both take this to mean that when a speaker presents herself as believing p, we can make the assumption that she *believes* p and further, that $p = 1$. With the SRC, this is not one of the added effects. Here, ≈ 1 really does indicate that the belief is open to revision. This is in part due to the intuition that the SRC, as opposed to the other two contours, does seem to be asking the listener to weigh in on the contribution bearing the contour. If a listener does not respond to an utterance directed at them with the SRC, the discourse seems to stay unresolved. Although I leave an in-depth analysis of this to future work, I suspect that this effect arises in part from the calculation of the entirety of the SRC, which includes reasoning about the likelihoods of both the speaker and the hearer's expectations.

(170) The falling H*L% adds the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(p) \approx 1$

(171) The excited H*-L% with a high pitch contour indicates

$\text{Exp}_{spkr}(p) \approx 1 \wedge \text{Boul}_{spkr}(p) > \text{Boul}_{spkr}(\neg p)$

Section §3.3.3 argues that the pragmatic effect of the SRC in English can be defined as in (172):

(172) The English Surprise Redundancy Contour (H) L* H*-L% is anaphoric to a salient proposition or event p in a discourse context C, and is admissible for discourse-salient participants x when

a. q is the proposition expressed by the speaker (uttered content or the presuppositions introduced by a question),

b. add the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

I argue that the SRC in English is a mirative strategy which signals a speaker's expectations at the time of utterance relative to the expectations of their discourse partners. When a speaker uses this contour, she presents herself as expecting q, while also indicating that for some salient participant, given q, they should infer that p is not likely the case. This has an extra perlocutionary effect as well, which puts pressure on the listener to respond to the utterance with an explanation or elaboration on their previous utterance, suspending resolution of an issue on the table.

The particles *oh* and *huh* are very tightly connected. Sections §3.2.2 and 3.2.3 propose the following pragmatic contributions for *oh* and *huh*, respectively. The meaning of *oh* is baked into the contribution of *huh*.

(173) *oh* and *huh* are anaphoric to a proposition p salient in the discourse s.t.:

a. *oh*(p) adds the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(p) < \text{Exp}_{spkr}(\neg p)$

b. *huh*(p) adds the following to the speaker's *DC*s:

$\text{Exp}_{spkr}(p) < \text{Exp}_{spkr}(\neg p) \wedge \text{Exp}_{spkr}(p) > 0$

Like the SRC, I argue that these particles have a mirative component. This pragmatic effect is to encode the speaker's epistemic state at the time of utterance by means of tracking expectations throughout the discourse. *What* also signals mirativity, but in a slightly different manner:

(174) *what* is anaphoric to a proposition p salient in the discourse, and *what*(p) adds the following to the speaker's *DC*s:
$\text{Exp}_{spkr}(p) < 1$

Instead of encoding speaker expectation about a proposition and its complement, this particle relies on inherent uncertainty of a speaker in relation to an uttered proposition p. This accounts for why it is permissible in SRC contexts, but not in falling or exclamative contexts, which require some inference of speaker authority.

In sum, this chapter is both an argument for the recognition of English mirative strategies, as well as a petition for an updated look at the contribution of discourse particles and prosodic meaning to a discourse. Discourse particles contribute information about how a speaker has structured her expectations, and how those expectations have come to be met or violated. Intonation plays an extremely similar role. While some utterance-level contours contribute pragmatic effects of authoritativeness or excitement, others have a more complex contribution. Intonational contours can also be markers of mirativity in English. Viewing these contours in terms of a speaker's expectations about a discourse has the advantage of pinpointing which contours convey authoritativeness by virtue of assigning the truth of a proposition a high expected value. It can also capture the effects of contours that express violated expectations in view of a discourse context. In all, reshaping the way we view intonation and discourse marker meaning creates a system designed from the same basic pieces. This gives us insight into a speaker's internal belief structure, a view of the conversational structure, and a way to talk about the interactions between the text of a sentence and its tune.

Chapter 4

Experimental Approaches to Intonational Meaning

4.1 Background

Much of work done on the interaction between semantic meaning and prosodic information falls into one of two categories. The first category, with arguably the most research available, focuses on word-level focal accenting and the various focus, topic, and deaccenting patterns found throughout the world's languages. Much of this research stems from Rooth (1985)'s observations that focal accenting on particular constituents can change the truth conditions of the utterance. Different focal prominence changes the interpretation of the otherwise identical string of words in (175):

(175) a. Jeff only gave [onions]$_F$ to Sophie.

b. Jeff only gave onions to [Sophie]$_F$.

It is trivial to construct scenarios where (175a) is true and (175b) is false. Imagine Jeff is handing out vegetables. He can give onions to Sophie, and hand Tom and Heather carrots and rutabagas and both sentences are true. But if Jeff also hands Sophie a carrot, (175a) is all of the sudden odd. Suppose that he takes back the carrot, and he hands Tom as well as Sophie an onion. Suddenly, (175b) is not longer a licit characterization of the state of the affairs. Focus sensitivity and the problems and puzzles related to it has been discussed extensively from both theoretical (Horn 1969, Rooth 1985, Beaver & Clark 2008, von Stechow 1996, Pierrehumbert & Hirschberg 1990, Erteschik-Shir 1997: u.a.) and experimental angles (Cooper et al. 1985, Hirschberg 1993, Krahmer & Swerts 2001, Xu & Xu 2005, Katz & Selkirk 2011: u.a.). Each of these avenues of inquiry share the common understanding that focus marking contributes a crucial communicative function to an utterance. Work in this area is particularly relevant in relating how the focused or topic marked element can be integrated into the information structure of the utterance as a whole, and work in this realm has spawned much work on alternatives and alternative semantics. Focus marking can cue a listener to which pieces of the utterance are new, relevant, backgrounded, or contrasted with something else in the discourse. All of these strategies can be integral to understanding the propositional meaning of an utterance, and to the interpretation of an utterance in context.

The second avenue of inquiry, in which semantic meaning crosses paths with prosodic interpretation occurs at the utterance level. Instead of pinpointing particular elements within an utterance that may be sensitive to particular focus operators, zooming out to the sentence level examines how the prosodic packaging of an entire utterance may change its interpretation. In particular, much of this line of research looks at how particular sentence-level intonation patterns interact with the illocutionary force of an utterance, especially between utterances with similar syntactic realizations (see: Baker & Bradlow 2009, Bell et al. 2009, Kaiser 2015, Ouyang & Kaiser 2015). Neutral, falling intonation on the utterance in (176a) has the effect of an assertion. Yet the same string of words with a rising contour can have the effect of an interrogative, positioning the speaker as uncertain about the propositional content of her utterance:

(176) a. You're late. *Neutral, falling*

b. You're late? *Rising*

Here, the theoretical landscape is again incredibly rich. Stemming from the works of Austin (1962) and Searle (1969) on speech acts, many researchers have studied the connection between the compositional or propositional meaning of an utterance and a speaker's intention behind that utterance. In particular, much work has been done on the interplay between sentence types and prosodic speech act operators, which modify the canonical illocutionary force of a sentence (see, among others, Bartels 1999, Farkas & Bruce 2010, Condoravdi & Lauer 2012, Malamud & Stephenson 2015, Krifka 2015, Farkas & Roelofsen 2017). In this realm, much of the literature focuses on the interaction of declarative and interrogative sentence types with rising and falling intonation contours. But while these issues have generated many papers, discussions and disagreements in the literature as to the correct interpretation of intonation and literal meaning, there are very few investigations into how speakers actually use and perceive these sentence types and illocutionary force-modifying prosodic contours.

Recent studies by Kaiser (2015), Jeong (2017), and Jeong & Potts (2016) are some of the first to ask the question of how slight differences in intonation might signal different affective meanings in speech perception. Jeong (2017) investigates the differences in interpretation that arise when final contours on polar questions are manipulated (falling, level, and rising). Not only does this pull apart meaning differences that can be attributed to the steepness of a rise, Jeong's work also probes a further layer of meaning, eliciting participants' judgements on relative speaker authority and attitude. Both of these are assumed to be inferred by a hearer based on an intonational pattern in context.

Ouyang et al. (2017) show using spoken data gathered from interactive gameplay experiments, that discourse expectations are rapidly updated based on the way that information is prosodically introduced. In a cooperative gameplay scenario, higher acoustic prominence signals particular information to the hearer, which affects how she is able to update her expectations in context. The prominence that she assigns to particular constituents in her response is indicative of this last expectational update: updates that were expected receive baseline levels of prominence, while responses that indicate unexpectedness deviate from this baseline.

Along these lines, the following studies are a further attempt to tap in to listeners' perceptions of particular prosodic contours. Here, another layer of complexity is added in: how can we use assumptions we already have about prosodic contours to test the meaning contributions of discourse particles? The three particles *oh, huh,* and *what* are further put to the test. We ask how we can use the meanings of specific, known prosodic contours to diagnose and pull apart the differences between discourse particles in English. Results suggest that much like the findings of Jeong (2017), licit interpretation of intonation and of discourse particles is closely correlated with the overall discourse context.

4.2 Pieces of an experimental approach

Previous chapters have argued for a compositional account of discourse particles and intonation, based off of fine differences from introspective judgements. The basic intuition hypothesizes that if intonation carries one layer of meaning and discourse particles another, adding their respective effects together should approximate their natural language use. Put concretely, we expect that *oh* with a falling contour is a composition of two pieces: the pragmatic effect of *oh* and the pragmatic effect of a falling contour. Taking a compositional approach means that we hypothesize that pragmatic effect of intonation and discourse particles is **not** an interaction above and beyond each element's effect in isolation. Rather, their effects are additive. I have argued in Chapter §3 that this is a key factor for understanding the contribution of discourse particles, which are inherently semantically underspecified. The performance of these utterances is integral in their licit interpretation. Understanding the contribution of discourse particles must take their bare meaning and use other contextual factors to enhance it. The richness in interpretation that discourse particles allow is not due to the interpretation of the particle in isolation, but rather, to the culmination of prosodic features and semantic or pragmatic effects that are present in the discourse context as a whole. Most of the integration of these particles into a licit discourse is made possible (or impossible) by a speaker's use of particular prosodic contours.

So far, we have limited ourselves to particles and utterances which share a common prosodic shape. Put concretely, particles with neutral contours have only been analyzed in the context of follow-up utterances that are also produced with neutral, falling contours. But intuitively, this is not representative of speech production as a whole. Tonal melodies on particles do not have a strict identity constraint which mandates that the following utterance mirror the melody of the particle. In many cases, mismatching the contour present on the particle and on the utterance shifts the interpretation of a speaker's expectations in a discourse. The examples in (177) show this:

(177) *The speaker walks into Tom's garden:*

a. Huh! Tom made a birdhouse? — Excited + Rising Contour

b. Huh? Tom made a birdhouse! — SRC + Excited Contour

In a neutral context, both of these utterances are licit. The first registers the speaker's surprise at receiving new information, marked on *huh*.[1] The rising contour on the continuation of the utterance proposes that the speaker is not quite willing to commit to this *p*, licensing the inference that she may have reason to believe some alternative. In the case in (177b), the speaker first signals her violated expectations, and then uses a higher pitch contour as a strategy to explain her confusion. Excited intonation highlights the new information she has just received as expectation-violating.

The following sections report on experiments which were run as a first attempt at understanding how particle+contour tokens can drive a listener's interpretation and comprehension of a discourse. Experiment I reports on one aspect of this, taking the idea of manipulating particle and contour pairs to its logical end. it asks whether all particle and contour combinations (*oh, huh, what* x NEUTRAL, EXCITED, SRC) can find a pragmatic interpretation in conjunction with *utterances* carrying NEUTRAL, EXCITED, or SRC contours. Experiments II and III focus on particular aspects related to the outcomes on Experiment I. Experiment II tests the hypothesis that the English Surprise Redundancy Contour is really a conflation of two distinct pragmatic effects whose shape is identical, but whose f0 determines the relevant "surprise" or "redundancy" reading. Experiment III further hones in on particular findings of Experiments I and II. Using the same "surprise" and "redundancy" productions from Experiment II, it introduces explicit contexts to further strengthen the claims of Experiment II, while adding an explicit control for participants' expectations and accommodation strategies that were overlooked in Experiment I.

4.2.1 Experimental intent

Prosodic judgments, especially those that have been presented in the preceding sections, are extremely subtle. Slight variations in pitch accent, peak duration or the amount of

[1]Notice that this need not take an uttered proposition as its prejacent; it can occur upon simply realizing (by inference, or visual evidence) that Tom made a birdhouse.

silence left between one prosodic phrase to the next can all have an effect on the perception of an utterance, and can easily shift an otherwise discourse-coherent utterance in the direction of complete infelicity. The pilot study presented here is an attempt to do two things. First, it pioneers a new methodology for studying semantic effects of intonational variation. While there are many studies on the semantics of focus, givenness, backgrounding and lexical prominence throughout a variety of languages, very little work has been done on experimentally investigating the pragmatic effects of utterance-level prosody. Running this experiment was an attempt in itself to present a method of empirically testing claims about intonational meaning. Second, this experiment is meant to serve the more immediate goal of attempting to tease apart discourse particle meaning from intonational meaning in English. This study probes the discourse relations introduced by the interaction between prosody and particles.

Previous chapters have touched on the fact that in English, it is possible for prosodic contours on particles to not completely "match" the contours found on the utterances that they occur with. In many cases, mismatches are perfectly acceptable, as in (178a), but in cases like (178b), mismatches produce a jarring pragmatic clash between what a listener expects and what a speaker produces:

(178) *Tom walks into the kitchen and hears popping. Rounding the corner, he spots Jeff:*

a. Oh! Jeff's making popcorn.

b. #Oh. Jeff's making popcorn!

This section attempts to systematically test a listener's perception of a range of contours, particles and following utterances in order to back up the claims presented above. In short, this chapter attempts to bolster the claims of the previous chapter experimentally: if prosodic contours on discourse particles are a part of overall speaker propositional attitudes, they should be subject to global discourse coherence phenomena. In addition, their effects should be additive.

This study involves felicity judgements elicited from participants about audio stimuli, but also asks for participants to match corresponding emotions to audio files they have heard. In a two step process, listeners must access information about their immediate responses regarding an audio file and in a series of follow up questions, and reflect on the reasons for their initial classification. Section §4.2.2.1 explains this setup in more detail.

This study puts the observations made in previous sections about the distribution and grammaticality of the discourse particles *oh, huh* and *what* to the test. In particular, it tests the following hypotheses:

(179) Hypotheses about particle and contour meaning:

a. **Hypothesis 1:** The discourse particles *oh, huh*, and *what* have distinct meanings.

b. **Hypothesis 2:** NEUTRAL, EXCITED, and SRC contours have distinct meanings.

(180) Hypotheses about felicity and acceptability

a. **Hypothesis 3:** There should be a cline of acceptability with the particle *what* between the three contours. In particular, *what* + SRC should rank more natural than *what* paired with either the NEUTRAL or the EXCITED contour.

b. **Hypothesis 4:** Based on the intuition that *what* should be disallowed with the NEUTRAL contour, we expect *oh* + NEUTRAL and *huh* + NEUTRAL to be perceived as more natural than *what* + NEUTRAL.

c. **Hypothesis 5:** Based on the intuition that *what* should be disallowed with the EXCITED contour, we expect *oh* + EXCITED and *huh* + EXCITED to be perceived as more natural than *what* + EXCITED.

(181) Hypotheses about emotion responses

a. **Hypothesis 6:** There should be an overall interaction between PARTICLE and PARTICLE CONTOUR, matching intuitive judgments from the previous chapter.

b. **Hypothesis 7:** Tokens where either the particle or the utterance contour is NON NEUTRAL are predicted to be categorized as NON-NEUTRAL on a higher average than tokens where both contours are NEUTRAL.

c. **Hypothesis 8:** A particle with an EXCITED contour should align with *Excited* or *Surprise* emotions; A particle with an SRC contour should align with *Confused* or *Surprise* emotions.

This is an ambitious list for any experimental design. The pilot study that follows is an attempt at answering each of these hypotheses by looking at responses gathered from a set of particle and contour interactions.

4.2.2 Methods

4.2.2.1 Stimuli and Design

The design for this experiment was ambitious, seeking to explore the interactions between PARTICLES, PARTICLE CONTOURS, and SENTENCE CONTOURS. Each of these factors contained three levels: PARTICLES = {*oh, huh, what*}, PARTICLE CONTOUR = {*neutral, excited,* SRC}, SENTENCE CONTOUR = {*neutral, excited,* SRC}. Fully crossing these results in a 3x3x3 = 27 conditions. A sample item set is below, where punctuation roughly represents intonational contours. Here, '.' represents falling "neutral prosody" on a particle or a sentence (H*L%), '!' represents a high rise-fall pattern (H*L%, EXCITED), and '?↘' indicates a low to rising and final falling contour ((H) L* H*-L%, SRC):

(182) PARTICLE x PARTICLE CONTOUR x SENTENCE CONTOUR

a. {Oh./Oh!/Oh?↘} Another pile of papers to grade. [NEUTRAL]

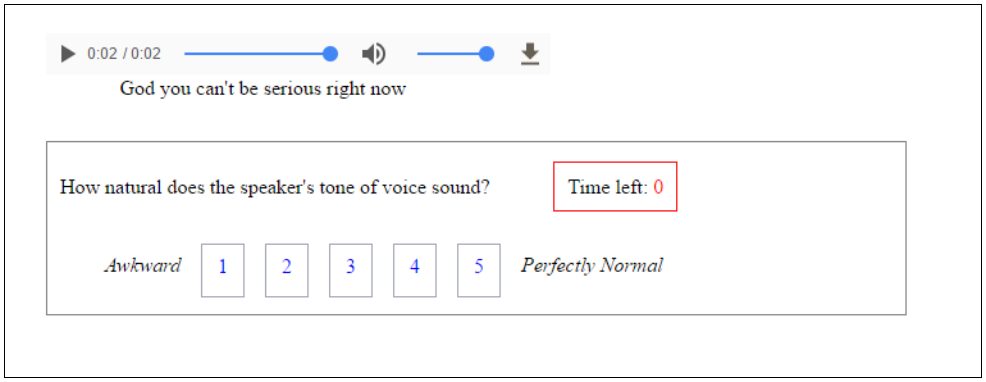
b. {Oh./Oh!/Oh?↘} A package arrived for you! [EXCITED]

c. {Oh./Oh!/Oh?↘} I thought apples were in the sunflower family?↘ [SRC]

Participants were solicited using Amazon's Mechanical Turk service, and compensated for their participation. 670 participants took part in this study. Geographical regions were restricted to the United States and Canada in order to control for English native speakers and dialectal variation in prosodic usage that arise in other English speaking communities. Each participant listened to two random, non-identical audio clips and were asked to complete four short ratings questions for each recording.

Participants were presented with an audio file on their computer screen using a modified version of the Ibexfarm experimental platform:

Figure 4.1: Audio presentation: experimental view after clip has played



After listening to the clip, an orthographic representation of the clip would appear on the screen without punctuation so as not to bias participants toward one or another interpretation. Participants were then given five seconds to perform a rating task, wherein they rated the naturalness of the utterance on a scale from one to five, with endpoints "Awkward" (1) and "Perfectly Normal" (5).[2]

After participants had made their selection, they were presented with three more questions and the option to replay the stimuli. The first was a forced-choice judgment task, which required a selection of an emotion that most closely correlated with the speaker's perceived attitude (*Angry, Excited, Surprised, Confused, Neutral*). The second question asked them to rate the intensity of the emotion they had chosen (under the assumption that *Neutral* contours would rate as low intensity, and more emotive contours would be rated as higher intensity). The third was a free-response question, asking them to explain, if present, the relation between the emotion expressed and the sentence uttered.[3] Participants were able to revise all of their responses with the exception of the first naturalness rating. All answer revisions, as well as time stamps for mouse clicks, were recorded. The full text of these questions, as well as a visual

[2]These Likert scale ratings were chosen from smaller-scale ratings tasks which tested the effectiveness of different labeling schemes on Likert scale questions. Identical data run with Likert scale ratings of "Very odd" (1) to "Perfectly Normal" were not as effective at gathering reliable data.

[3]Under the assumption that the prosodic contour on the discourse particle would trigger an evaluation by the participant of the speaker's overall emotional state, the researchers were interested in those cases where the discourse particles were expected to sound incoherent. This question was an attempt at honing in on a participant's given naturalness rating by forcing them to explain the coherence relation between the particle and the following utterance.

representation of this experimental setup can be found in Appendix B.

4.2.2.2 Stimuli creation

Audio files were recorded in a soundproof booth using the built-in microphone of a 2012 Macbook Pro. One male, linguistically informed native speaker of English was recruited to produce all discourse particles and utterances in question. The speaker was coached by the investigator in order to produce perceptually distinct contours for each of the particle type and utterance type conditions. In addition to knowing the approximate shape of the contours in question, the speaker was given contextual information that was intended to help produce the sentences in a natural environment. All of these environments were constructed in order to keep the speaker's tone of voice and other potential voice quality aspects of the productions constant.The speaker was given a list of utterances, all beginning with a particle followed by one of the twelve sentences used as carrier phrases. All initial sentence/particle pairs were designed to sound as natural as possible, with contours produced on the particles being mirrored in the corresponding sentence. This was done in order to have a baseline rating for subsequent particle and sentence pairs: all but one of the initial stimuli were expected to rate high on the naturalness scale by participants.[4]

Particles were analyzed in Praat to ensure that each had comparable contour shapes and peak heights. Contours and peak heights of each particle were plotted against each other after being time-normalized. Tokens were selected based partially on their effectiveness of displaying the contour in question, and partially on their similarities in relative peak heights and time-course measurements of the other particles in each contour group set. Figures 4.2-4.4 track the fundamental frequency of each particle and contour pair:

[4]The neutral final fall contour on the particle *what* was the only particle + sentence pair of the original stimuli that was assumed to be infelicitous. This assumption was based on the findings summarized in table (3.4), which shows that a neutral contour on *what* that does not indicate speaker rudeness or defensiveness (assumed by the researchers to be a different contour) is infelicitous.

Figure 4.2: NEUTRAL particle contours

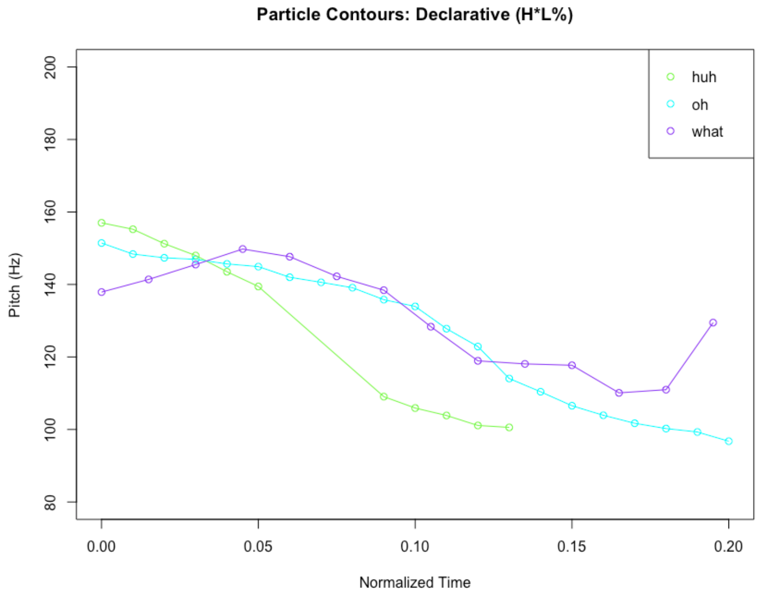


Figure 4.3: EXCITED particle contours

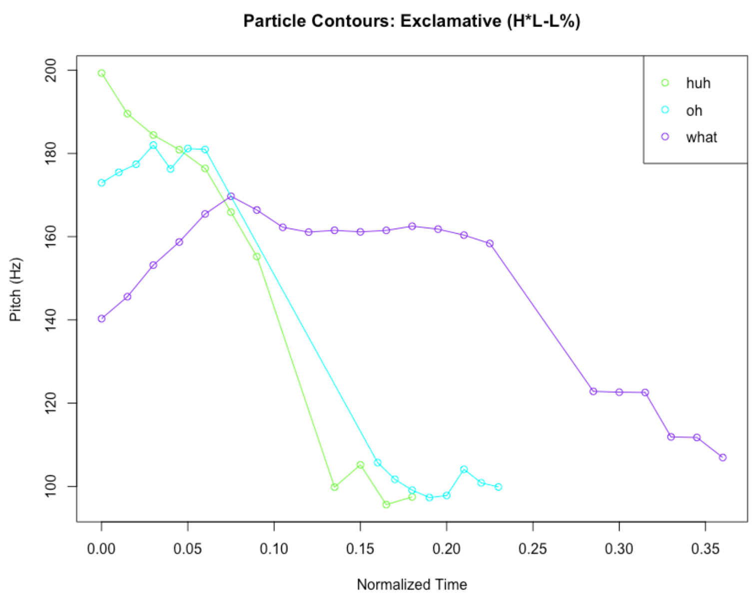
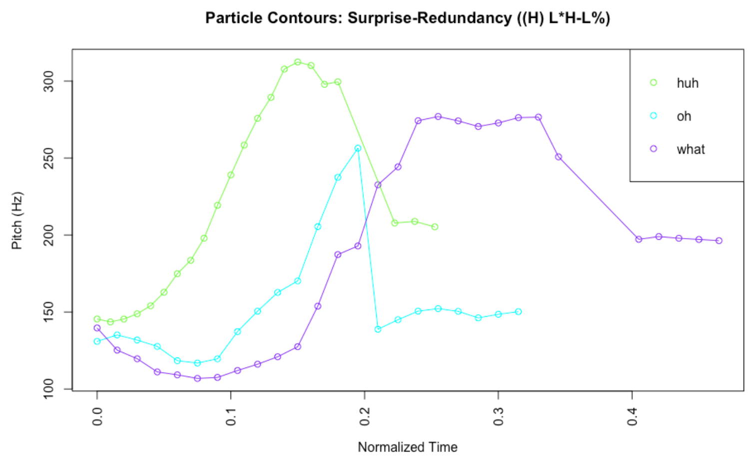


Figure 4.4: SRC particle contours



Particles and sentences were then cut and spliced together to produce a full factorial design of 81 individual sentences (9 distinct particle + contour combinations x 9 distinct sentences). Sentence conditions were collapsed by contour type, leaving 27 actual conditions of PARTICLE x PARTICLE CONTOUR x SENTENCE CONTOUR. A full list of these particles and their carrier sentences, as well as a link to the audio files, can be found in Appendix A.

4.2.3 Results

4.2.3.1 Analysis

Each of the 670 participants listened to two recordings, yielding a total of 1340 observations. Filtering by Native Speaker, excluding trials that took participants longer than 6 seconds to answer, and three individual utterance types yielded a total of 956 utterances that will be used in the subsequent analysis.[5]

[5] The three utterances to be excluded from this analysis had two features in common that the researchers felt should warrant complete exclusion. First, post-hoc acoustic analysis of the sentences revealed that the contours present on these tokens could not be entirely ruled out as separate prosodic contours. Second, each of the sentences in question also employed focal accenting of particular phrases above the presumed prosodic contour operating on the utterance as a whole, creating an unintended pragmatic effect with some particle and contour pairs. This effect was made clear in free response answers from participants. These sentences were then cut from the results so as to not introduce an unintended pragmatic confound.

4.2.3.2 Naturalness ratings data

A two-way analysis of variance was conducted on the influence of PARTICLE and PARTICLE CONTOUR on RATING, the perceived naturalness rating of audio recordings. PARTICLE consisted of three levels (*oh, huh, what*), as did PARTICLE CONTOUR (*neutral, excited,* SRC). The analysis found both main effects to be significant (Hypotheses 1+2), but their interaction was not. The factor PARTICLE indicated a significant difference between *oh, huh*, and *what* [F(2.947) = 18.62, p < 0.001***], and PARTICLE CONTOUR showed significant distinctions contributed by the three contour types [F(2,947=14.49), p < 0.001***]. The interaction between the independent variables did not reach significance [F(4.947)=1.57, p = 0.27]. The following table reports the Ratings means and standard deviations of the particle and particle contour data[6]:

PARTICLE	PARTICLE CONTOUR	Count	Mean	SD
oh	decl	107	4.018	1.0043
	SRC	111	4.028	0.9437
	exclam	85	3.800	1.0328
huh	decl	120	3.789	1.0626
	SRC	114	3.783	1.0682
	exclam	116	3.328	1.1555
what	decl	85	3.458	1.1167
	SRC	107	3.776	1.2231
	exclam	111	3.099	1.2930

Table 4.1: Means and standard errors for Naturalness ratings, based on 1-5 Likert scale judgments

Planned comparisons within PARTICLE do reveal a significant, three-way distinction in ratings between *oh, huh* and *what* (Hypothesis 1). The particles *what* and *huh* show a highly

[6]A mistake in the randomization algorithm caused some tokens to be presented more frequently than others, as indicated by the non-identical counts of observations here.

significant difference [t=-2.3, p < 0.001***], with *what* receiving lower overall naturalness ratings irrespective of contour type. *Oh* is rated as significantly less natural than *huh* overall [t=-4, p > 0.001***], which is expected given the variability in felicity between the three contour types with this particle. The standard for significance between *what* and *oh* is lower, but still enough to reject the null hypothesis [t=-2.3, p = 0.02*], and to assert that *oh* is deemed more natural than *what*. This aligns with the impressionistic investigation conducted in Chapter §(3), in which *what* was assumed to be licit only in a subset of the prosodic contexts that could license *oh* and *huh*.

The factor PARTICLE TYPE also has significance between levels (Hypothesis 2). The NEUTRAL contour is rated significantly more natural than the EXCITED contour [t=4.2, p < 0.001***], while the SRC also rates distinctly higher on the naturalness scale than the EXCITED contour [t=5.4, p < 0.001***]. NEUTRAL falling and SRC contours did not reach significance against one another in terms of naturalness ratings, but SRC did have a slight tendency to be rated more natural [t=1.2, p=0.2]. This is not expected, but perhaps unsurprising, as both the SRC and the NEUTRAL contours also tended to be able to host more particles in more pragmatic contexts than the EXCITED contour (see Chapter §3.2, Figures 3.2-3.4).

Planned comparisons involving the interactions between independent variables should track the judgments laid out in Chapter §3. In particular, a strong contrast between the naturalness ratings of *what* + NEUTRAL and *what* + EXCITED when rated against *what* + SRC is expected. This is partially borne out (Hypothesis 3). Particle and contour combinations *what* + NEUTRAL and *what* + EXCITED are not expected to rate differently on the naturalness scale, as they are both infelicitous pairings. Their significance is scant, a finding that is predicted by impressionistic judgments that *what* does not allow for these contours at all [t=2.1, p=0.04*]. *What* with the SRC is, as expected, rated significantly higher than *what* + EXCITED [t=3.9, p < 0.001***], but it does not reach significance (though it is just out of range) against utterances marked with *what* + NEUTRAL [t=1.9, p=0.06].

Two-sample t-tests reveal that Hypotheses 4 is borne out: both *oh* + NEUTRAL and *huh* + NEUTRAL are perceived as more natural than *what* + NEUTRAL ([t=2.1, p = 0.03*] for the former, [t=3.8, p < 0.001***] for the latter). Hypothesis 5 is partially validated: *huh* + EXCITED contour scores higher than *what* + EXCITED [t=4.2, $p < 0.001$***], but there is no significance with *oh* and *what* in the same contour environment [t=1.4, p=0.2]. This could be due in part to discrepancies in utterance productions in the excited case; see Section §4.2.4.1 for discussion of this matter.

4.2.3.3 Emotions categorization data

A multinomial logistic regression model was fitted to the categorical response data using the *lmtest* statistical package in R. Using maximum likelihood techniques, the probability of each of the Emotions rating possibilities (*Surprise, Anger, Confusion* or *Excitement*) against the *Neutral* emotion reference level was estimated for each of the PARTICLE CONTOUR x SENTENCE CONTOUR pairs. Items with the NEUTRAL x NEUTRAL particle and sentence contour combination served as baselines for comparison against each other particle and sentence contour combinations. Here, the emotions rating data do not take particular particles into account: all data reported here is across PARTICLE TYPE.

The data used in constructing this model was based first off of initial naturalness ratings for each token. Only ratings for utterances that were judged as within ± 2 standard deviations from the mean were included. This was to ensure the most accurate representation of the ratings data in conjunction with the emotions categorization data. If a participant rated an utterance as very low on the naturalness scale, her perception of what the emotion that they have categorized will be affected. In such cases, there is a significant chance that the emotion rating ascribed to it is a best guess, or worse, a random guess. This method of subsetting the data is in an effort to catch such behaviors and rule out these outliers.

Table 4.2 shows data for comparisons of particle + sentence contour pairs and their likelihood of being rated *Surprise* against the *Neutral* emotion baseline.

		Estimate (β)	Odds Ratio	z-score	p-value	
Baseline: neutral x neutral		-0.25490	0.77500	-1.6525	0.2867	
Particle Contour	SRC	1.98224	7.25898	4.3847	1.16e-5	***
	EXCITED	2.50618	12.25801	5.10019	3.393e-7	***
Sentence Contour	SRC	1.38329	3.988802	3.28265	0.001	**
	EXCITED	1.78636	5.96767	3.90483	9.42933e-5	***
Particle Contour x Sentence Contour	SRC x SRC	1.50445	4.50170	1.30224	0.1928	
	EXCITED x SRC	0.25744	1.29362	0.21896	0.8266	
	SRC x EXCITED	-0.6061	0.50631	-0.86355	0.3878	
	EXCITED x EXCITED	-0.75411	0.47043	-0.87701	0.3805	

Table 4.2: Multinomial regression coefficients for Response = *Surprise*

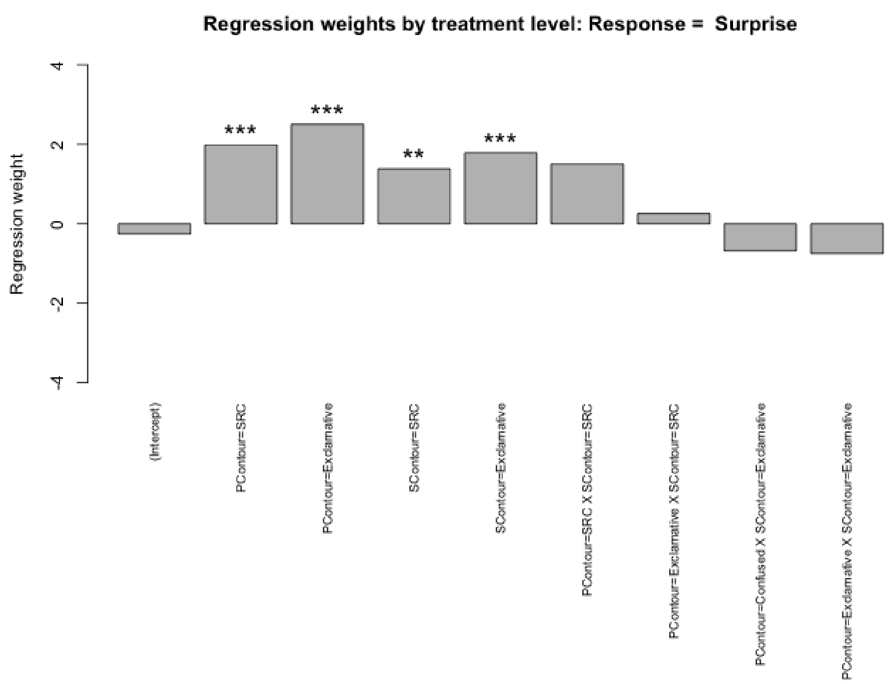


Figure 4.5: Logistic model results for *Surprised* Response

Compared against NEUTRAL x NEUTRAL reference level, particles occurring with the SRC and with the EXCITED contours were highly likely to be categorized as *Surprise* [SRC: $p < 0.001$***, EXCITED: $p < 0.001$***] when paired with an utterance with a NEUTRAL sentence contour. This effect is shown when SENTENCE CONTOUR is manipulated as well, with SRC and EXCITED contours indicating very high statistical preference for a *Surprise* categorization when crossed with a particle with a NEUTRAL contour [SRC: $p < 0.001$***, EXCITED: $p = 0.001$**]. There was no significant interaction of PARTICLE CONTOUR and SENTENCE CONTOUR.

Table 4.3 and Figure 4.6 compare the likelihood of a *Confused* rating against the same *Neutral* reference level:

		Estimate (β)	**Odds Ratio**	**z-score**	**p-value**	
Baseline: neutral x neutral		-0.64434	0.5250	-2.3911	0.0168	*
Particle Contour	SRC	0.9629	2.6193	1.7926	0.073	
	EXCITED	0.2387	1.2696	0.3412	0.7330	
Sentence Contour	SRC	2.1382	8.4843	4.9857	6.174e-7	***
	EXCITED	1.3375	3.8093	2.6224	0.0087	**
Particle Contour x Sentence Contour	SRC x SRC	0.58768	1.7998	0.4885	0.6252	
	EXCITED x SRC	0.6656	1.9456	0.5117	0.6089	
	SRC x EXCITED	-0.4034	0.66806	-0.4517	0.6515	
	EXCITED x EXCITED	0.86013	2.3635	-0.4517	0.6515	

Table 4.3: Multinomial regression coefficients for Response = *Confused*

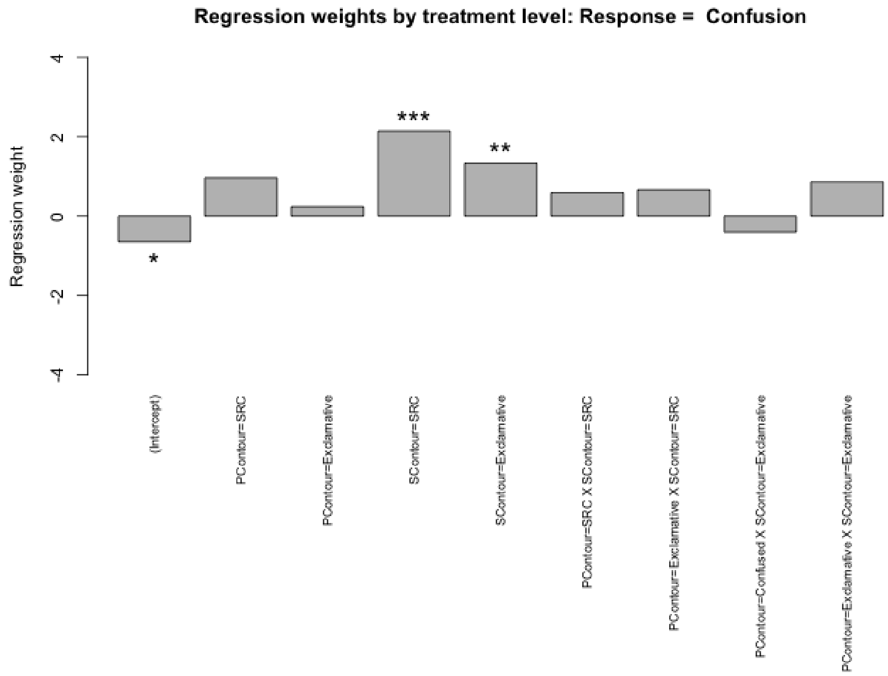


Figure 4.6: Logistic model results for *Confused* Response

Baseline PARTICLE CONTOUR and SENTENCE CONTOUR responses with neutral contours show significant negative preference for a *Confused* categorization [p = 0.0168*]. There was no effect of PARTICLE CONTOUR on emotions categorization, but SENTENCE CONTOUR manipulations did reach significance. Utterances with NEUTRAL contours on particles are statistically likely to be rated as *Confused* when SENTENCE CONTOUR is either SRC or EXCITED [SRC: p < 0.001***, EXCITED: p < 0.0087**]. There was no interaction effect for the categorization of utterances as *Confused* with any permutation of PARTICLE CONTOUR and SENTENCE CONTOUR.

The following data in Table 4.4 and Figure 4.7 compare a categorization of *Anger* against the baseline *Neutral* categorization for PARTICLE CONTOUR x SENTENCE CONTOUR emotions.

		Estimate (β)	Odds Ratio	z-score	p-value	
Baseline: neutral x neutral		0.02468	1.02498	0.11104	0.9115	
Particle Contour	SRC	-0.02349	0.97571	-0.04495	0.9641	
	EXCITED	1.16175	5.04051	3.246	0.0012	**
Sentence Contour	SRC	0.06225	1.06423	0.13163	0.8952	
	EXCITED	-0.02475	0.9755	-0.04524	0.9639	
Particle Contour x Sentence Contour	SRC x SRC	-0.0618	0.94007	-0.0393	0.9686	
	EXCITED x SRC	0.2416	1.2733	0.1931	0.8469	
	SRC x EXCITED	-0.6688	0.5123	-0.5867	0.5574	
	EXCITED x EXCITED	-0.2311	0.7937	-0.2416	0.8091	

Table 4.4: Multinomial regression coefficients for Response = *Anger*

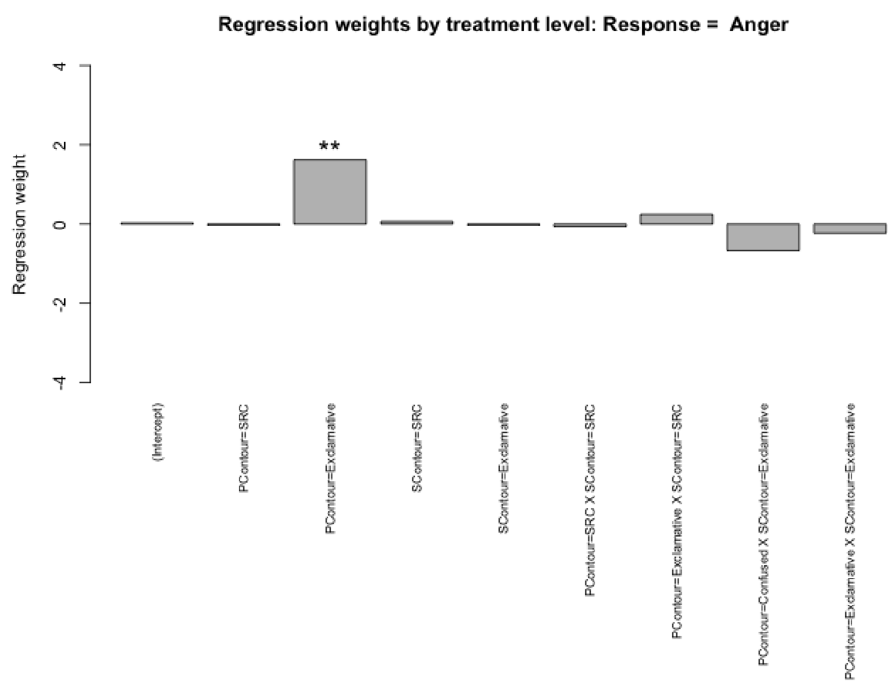


Figure 4.7: Logistic model results for *Anger* Response

There is only one manipulation within the model parameters which showed significance for a token to be categorized as *Anger* over the baseline. Particles that carry the EXCITED contour but which are paired with NEUTRAL sentence contours are significantly more likely to be categorized as *Anger* over *Neutral* [p= 0.0012**].

Table 4.5 and Figure 4.8 summarize the comparison of the *Neutral* baseline to the *Excited* emotion categorization:

		Estimate (β)	**Odds Ratio**	**z-score**	**p-value**	
Baseline: neutral x neutral		-2.5908	0.0750	-4.327	1.51e-5	***
Particle Contour	SRC	0.5116	1.6679	0.42	0.6745	
	EXCITED	2.4085	11.1167	2.8282	0.0048	**
Sentence Contour	SRC	1.2915	3.6382	1.4598	0.1443	
	EXCITED	3.1503	23.345	4.2212	2.343e-5	***
Particle Contour x Sentence Contour	SRC x SRC	0.7881	2.1992	0.3987	0.6901	
	EXCITED x SRC	0.2772	1.3195	0.1789	0.85796	
	SRC x EXCITED	0.4864	1.6272	0.3438	0.7295	
	EXCITED x EXCITED	-0.9755	0.3770	-0.8554	0.3923	

Table 4.5: Multinomial regression coefficients for Response = *Excited*

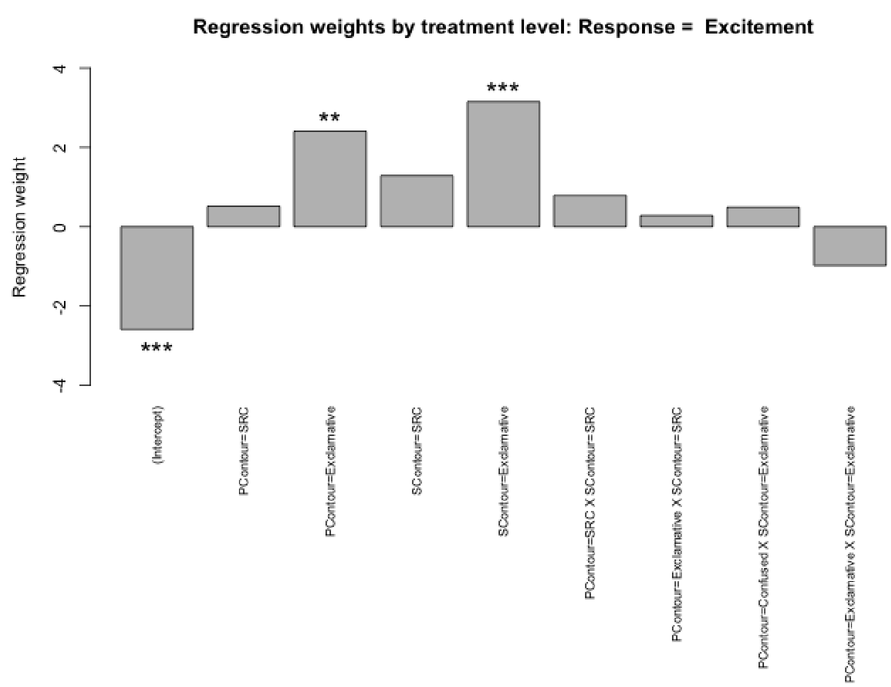


Figure 4.8: Logistic model results for *Excited* Response

The baseline *Neutral* emotion was rated significantly lower than chance when pitted against the *Excited* categorization: participants far prefer to rate NEUTRAL x NEUTRAL utterances as *Neutral* over *Excited* [$p < 0.001$***]. Being paired with an EXCITED contour significantly increased the likelihood of being rated *Excited*; if a PARTICLE CONTOUR is EXCITED, it is significantly more likely to be judged as *Excited* [$p = 0.0048$**]. Similarly, if SENTENCE CONTOUR was EXCITED, the entire token was more likely to be categorized as *Excited* than to the baseline [$p < 0.001$***].

4.2.4 Discussion

4.2.4.1 Unanticipated issues in speech production

Preliminary analyses of the EXCITED tokens of *what*, seen here in Figure 4.3, were judged as representing similar peak height and target fundamental frequency value to the EXCITED tokens of *oh* and *huh*. However, post-hoc analysis of the data did not yield the expected judgments for items containing this particle and contour combination, much to the surprise of the researchers. Further scrutiny revealed that the excited contour on *what* used in the experiment actually carried a distinct contour from the other two particles. As Figure 4.3 shows, the pitch on *what* begins at a similar hight to the other excited tokens, but the pitch accent on the vowel extends approximately 0.2 seconds before making the final fall to heights on par with other excited tokens. This vowel length is approximately 0.15 seconds longer than the vowel length durations in the *oh* and *huh* tokens.

The researchers originally assumed that this was an ordinary deviation stemming from larger syllable weight leading to longer production time of the *what* token in general. As opposed to *oh* /oʊ/ and *huh*, /hʌ̃/, *what*, /wʌt/ contains both an onset and a coda, and is a closed, heavy syllable. However, the delay on the nucleus, seen in Figure 4.3 does point to a difference in the overall gesture involved in producing the sound. As such, this token cannot be compared to those utterances which do not display a similar extended pitch duration. While this may not seem relevant to a semantic account of the meanings of these contours, Gussenhoven (1984) shows that differences in vowel height and duration change the timing and alignment of the pitch accent, and can have pragmatic effects that listeners can readily distinguish. In particular, extended pitch duration before the final fall on a vowel can be taken to indicate that the utterance is in some way significant or out of the ordinary. Though similar, this is distinct from the meaning ascribed to "excited" or high rise to falling contours, which marks an utterance or phrase as new, positive, and informationally distinct. This distinction was discussed more in depth in Chapter §1.3.

In the face of this, new particle and contour pairs were synthesized, taking identical pitch tracks from a canonical SRC utterance and replacing all tokens of *oh, huh* and *what* with identical pitch features that were then time normalized. This resulted in unnatural sounding particles, which, when paired with unmanipulated sentence contours, came across as highly artificial. In light of this, and in light of the task, which asked participants for naturalness ratings, the decision was made to go ahead with the naturally produced particle and contour combinations, while keeping this potential confounding factor in mind. As a result, this extra dimension of meaning may have interacted with the results as they relate to the theoretical predictions made in Chapter §3.2, but ultimately does not hinder the progress that this investigation has made.

4.2.4.2 Testing intuitions against ratings data

In general, naturalness ratings were able to back up the intuitive judgments that were presented in previous sections of this work. *Oh, huh,* and *what* were given variable acceptability ratings despite occurring in identical prosodic environments with (close to) identical contours aiding in the audio presentation of the particles. This suggests that the motivation for characterizing these particles as distinct discourse navigation strategies is strong. In terms of acceptability, *huh* seems to be the most permissive, adapting relatively well across the board to all prosodic contours. Though *oh*'s mean naturalness ratings rank higher than *huh*, adjusted scores reveal that in many cases, this does not equate to significance. *What* is rated the least "natural" of the three particles, but this comes as no surprise: an assumption going in to this experiment was that *what*'s distribution with the prosodic contour inventory was much more limited than the other two particles. This appears to have been validated here. We can reject the null hypothesis in terms of Hypothesis 1.

Varying acceptability in naturalness ratings between prosodic contours also indicates that participants in this study are indeed able to perceive and identify three distinct prosodic contours. In addition, the lack of an interaction between PARTICLE and PARTICLE CONTOUR main effects suggests that, as predicted, the effects introduced by each may be additive.

The data show a clear split between the contour types and naturalness ratings: both SRC and NEUTRAL contours tend to be rated more natural than utterances with EXCITED contours. Based off of the data in Chapter §3, we might expect a clear ordering from most to least permissive in terms of naturalness, based partially off of the ungrammaticality of *what* in NEUTRAL falling and EXCITED contexts. However, the data here must be taken with a grain of salt, as Section §4.2.4.1 notes, due to the unanticipated interpretation attributed to *what* + EXCITED utterances. From this we learn that based on this experiment, we cannot reject the null hypothesis that there is no interaction between particles and particle contours (Hypothesis 6).

The predictions of Hypotheses 4 and 5 were very specific: *oh* and *huh* combined with NEUTRAL would score as more natural than *what* + NEUTRAL, and similarly, *oh* and *huh* combined EXCITED would score higher than *what* + EXCITED. These were partially validated. Lack of significance can be attributed here to a number of factors, but perhaps the two most worthy of follow-up experiments deal with the particles and their complex patterns of context dependence. Unlike the investigation in Chapter §3, this experiment failed to control for context type when pairing particles and contours with utterances. Not controlling for this means that the entire possibility space has not been covered, leaving room for contexts which have not been included to potentially skew the results toward a particular result. Future iterations of this experiment take these contexts into account.

On the one hand, lack of a clear contextual background was an intentional design plan by the researchers: allowing free interpretation of the audio presentations was meant elicit unbiased responses from participants, as well as reduce the complexity level of the entire experiment. But this was not without a calculated tradeoff. Lack of context for stimuli made interpretations of the relation between the discourse particle and the utterance much freer. This introduced unanticipated readings of these particles, and in one case, lead the researchers to exclude an entire particle from the results.[7]

The second factor that could have contributed to the relatively high naturalness ratings given to many of these utterances also stems from their manner of presentation. Though participants were asked to consider the discourse particle and judge its naturalness based on its coherence in relation to the following utterance, many participants were able to attribute "natural" readings to some of the expected-infelicitous tokens in an unanticipated way. In particular, the researchers expected strong unnaturalness ratings for any utterance with *what* and neutral prosody as the discourse particle in question. Though this was partially borne out, some of these particle+contour and sentence+contour pairs were rated much higher than expected. In particular, the item in (183) was rated particularly high, despite the prediction that this would be infelicitous:

(183) # What. I'm on the phone.
H*-L% H*-L%

Analysis of the free response answers to the explanatory relation between the particle and the utterance can begin to explain why this might be: participants repeatedly interpreted this utterance as responding to two distinct acts or utterances. The intended reading of (183) took the speaker to be conversing with a single participant, this his particle and his follow up should have both been addressed to the same individual. Over half of the participants indicated that the interpretation given to this utterance was licit if the speaker meant the particle and the sentence to be two distinct conversational turns. The first, marked by the particle, is a response to a discourse participant on the other line of the phone call. The second reacts to another speaker present in the room who was interrupting the call. In this case, though the particle and utterance referred to two distinct discourse contexts, they were judged felicitous. This again is a fault in the presentation of these utterances in isolation with no clearly defined or disambiguating

[7] Originally, *God* was included as another discourse particle variant expected to have distinct but overlapping discourse navigation strategies as the other three particles. However, a closer look at the free response questions in the experiment revealed that participants were consistently interpreting *God* as a vocative , and not as an actual discourse particle (something akin to "God? It's me, Margaret." These are included in Appendix A). This may have been due to the presentation of the clip's text after the initial playing of the audio–capitalization of 'God' as the first word in the sentence may have had the unintended consequence of invoking reference to the deity. Lowercase orthographic representation may have suppressed this interpretation.

contexts. With this is mind, no conclusive generalizations can be drawn, and Hypotheses 4 and 5 need to be revisited.

4.2.4.3 Testing Emotions Intuitions

The experimental setup was designed to test two things: the first goal was to verify intuitive judgments about the acceptability of various discourse particles and the contours that they can occur with. The second was to probe discourse coherence relations with these particle+contour combinations and sentence contours. We make two predictions about these combinations, summarized in Hypotheses 7 and 8 above. The first relates to the interpretation of the entire utterance when a discourse particle is uttered with neutral prosody. Will the entire utterance be judged as neutral, or will the interpretation shift to the more expressive tune? The second is the converse. We ask what the interpretation of an utterance should be when a sentence paired with an emphatic discourse particle is not itself expressive.

The first hypothesis concerns itself with general pragmatic effects introduced by non-Neutral prosodic contours. It predicts that if a particle or an utterance is paired with the SRC or the EXCITED contour, it should more likely be rated as indicative of a non-Neutral emotion that utterances performed with entirely neutral falling contours. This turns out to be partially true. Because these effects are assumed to be additive, we see no significance across the board in ratings for the interactions. Within main effects, the generalization in Hypothesis 7 holds for cases where the rating was *Surprise*–both SRC and EXCITED contours on either PARTICLE CONTOUR or SENTENCE CONTOUR shift the likelihood for the emotion rating strongly in the direction of a *Surprise* categorization.

The prediction only holds for SENTENCE CONTOUR manipulations of *Confused* categorizations: it is no more likely for an utterance to be categorized as *Confused* over *Neutral* when when only the particle carries a non-Neutral contour.

The data are a bit different for *Anger* and *Excited* emotions categories. *Anger* turns out to be a significant category label only when an EXCITED contour on a particle has been crossed with a NEUTRAL sentence contour. In all other cases, choosing an *Anger* response is almost exactly as likely as choosing *Neutral*. The *Excited* category is the only one where the baseline is significantly skewed away from an *Excited* categorization. Hypothesis 7 is only partially borne out here: while SRC contours do not effect bias toward an *Excited* rating, EXCITED contours are highly correlated with an *Excited* rating.

While Hypothesis 7 cannot be entirely confirmed, the data here provide more evidence for confirmation of Hypothesis 2: differences in Emotions ratings across PARTICLE CONTOUR and SENTENCE CONTOUR lead more credence to the idea that these three contours do indeed have distinct pragmatic effects.

Hypothesis 8 is a bit harder to quantify, but the data here are suggestive of a larger pattern. Because the multinomial model only compares single emotions against the *Neutral* baseline, the ratings data as they are presented say nothing about the likelihood that a token is rated *Surprise* over *Excited*, or *Confused* over *Excited*; the model only predicts that a token is more likely to be categorized as *Surprised* or *Excited* or *Confused*. With this in mind, Hypothesis 8 has (conservative) positive evidence: The *Excited* category does show a preference for a particle carrying the EXCITED contour to be rated as such, and the same can be said for the *Surprise* category. Particles with EXCITED are significantly likely to be rated as *Surprise* or *Excited*. Interestingly, they are also likely to be categorized as Angry.

For the SRC, the evidence is slightly different: participants prefer to rate particles carrying this contour as *Surprise*, but less likely as *Confused*. This is a potential pragmatic effect introduced by the competing interrogative contour, not present in this study, but distinct enough for participants to notice a difference. The pragmatic effects of these contours are key. This finding suggests that the characterization of the SRC in (155) is on the right track, as it does not signal Confusion on the part of the speaker, but rather, *Surprise* brought about by violated expectations.

4.3 Experiment 2

The purpose of this set of experiments is to follow up on a hypothesis that stems from some of the results gathered from Experiment 1. The SRC is described as having two distinct semantic meanings: one of genuine surprise, as when a speaker encounters a situation that she did not expect, and one of redundancy, in which the speaker seems to chastise a salient figure in the discourse, insinuating that they "should have known" some relevant piece of information. These readings seem very pragmatically distinct, yet there is no solution offered as to why these interpretations may both arise from a single contour. Both readings clearly have to do with a speaker's violated expectations, but the question is how the two distinct interpretations come about: how does one receive a positive construal, while the other is viewed as a negative reaction? The hypothesis here is that this difference at least partially due to a speaker's relative f0 height. When the SRC has a higher overall pitch, it pragmatically shifts the contour's interpretation toward the positive, "surprise" reading. When its overall pitch excursions are low, this shifts the interpretation of the SRC toward the "redundancy" reading. The aim of this experiment is to test whether f0 height can serve as a basis for these two readings of the SRC. We split the SRC in two, giving rise to what we term SRC-HIGH and SRC-LOW:

(184) a. **SRC-HIGH**: a positively skewed SRC, "surprise" reading.

b. **SRC-LOW**: a negatively skewed SRC, "redundancy" reading.

The following study examines the perceptual effect of these two performances of the SRC. In order to investigate this, two separate tasks were performed with identical stimuli. To first establish a baseline for perception, participants were given a simple ABX discrimination task where their task was to match one speaker's production of a particular contour to another. Once a sufficient perceptual difference was established, a second study used the same tokens to elicit emotion categorizations from participants. Utterances with higher overall pitch excursions are hypothesized to align with "surprised" interpretations of the contour, which are construed as positive, while utterances with a lower relative pitch are hypothesized to be aligned with a "redundancy" interpretation, which are construed in a more negative fashion.

4.3.1 Methods

4.3.1.1 Stimuli and Design

This experiment is an attempt to tease apart the differences between NEUTRAL, SRC-HIGH and SRC-LOW contours. In this, there is one overarching question, with two crucially important sub-questions that we must answer:

(185) **Q1:** Is there a perceptual difference between the SRC-LOW and the SRC-HIGH contours?

a. **Q2:** Is the NEUTRAL contour perceived as distinct from the SRC-HIGH?

b. **Q3:** Is the NEUTRAL contour perceived as distinct from the SRC-LOW?

The design presents three audio tokens in a modified ABX-fashion. ABX tasks are usually designed as a matching-to-sample task; participants listen to two stimuli, A and B, and determine which is closer to a sample stimulus, X. In most cases, X is identical to either A or B. In this study, participants were asked to match prosodic contours across speakers. Instead of identical stimuli, the participant hears two stimuli A and B spoken by one voice, and is asked to match the contour that is the *most similar* to a different voice's contour, in this case, the X stimulus.

Participants heard two stimuli, A and B, spoken by the same speaker, varying only in their prosody, and separated by 500ms. After the stimulus B and another 500ms, the third token X was spoken by a distinct speaker, using one of the two contours heard previously by the participants. Participants are then asked which of the first two clips sound the most similar to the sample, or 'X' clip. The order of the A and the B clips were randomized and balanced across presentations. All possible arrangements of stimuli (3 stimuli types x 2 {AB, BA} orderings x 2 'X' possibilities (A,B) = 12) were included in the experiment in an effort to offset temporal ordering and recency effects that ABX tasks impose on the listener. The following is a sample presentation of an audio stimulus:[8]

(186) **Listener hears Speaker 1 (female)**:

(A) She climbed a mountain with a guide dog. + NEUTRAL

(B) She climbed a mountain with a guide dog. + SRC-LOW

Listener hears Speaker 2 (male):

(X) She climbed a mountain with a guide dog + *either* NEUTRAL *or* SRC-LOW

[8]A version of this task can be found at `people.ucsc.edu/~knkraus/ABX/main.html`.

Possible presentations of 2 stimuli (4):

a. A—B—X_A

b. A—B—X_B

c. B—A—X_A

d. B—A—X_B

Participants were solicited using Amazon's Mechanical Turk service and compensated for their participation. 48 participants (34 males, 14 females, aged 18-57) took part in this study, and participants were geographically restricted to the United States for reasons of dialectal intonational coherency. One participant's data was excluded due to their reported non-native English fluency. Each participant listened to 72 random, non-identical presentations of items and made a forced choice decision after each event. Participants were given the option of re-listening to the stimuli one time. All mouse clicks were time-stamp recorded. A list of the stimuli that were used is available in Appendix B.

4.3.1.2 Stimuli Creation

Norming study

This study uses audio recordings presented to participants to test their perception of "surprise" and "redundancy" readings of the SRC. In order not to bias speakers by introducing semantic content that might prefer a surprised or a redundancy reading of an utterance, a norming study was performed in which participants were asked to rate 72 utterances for their emotive content. 38 participants (15 males, 23 females, age range 18-65) were recruited via social media and took the experiment on the Ibex experimental platform. Their task was visually presented, and given a sentence, they were asked to make a forced-choice decision about emotion invoked, with choices of *Neutral, Positive* or *Negative*.

The original data were coded with expected answers (Negative = -1, Neutral = 0, Positive = 1), and the gathered data were given an ordinal interpretation on the same scale based on participants' responses. These were then aggregated by subject and item in a contingency table to compare expected vs. observed values in an ordinal interpretation of the data. Using Cramer's V, which compares the cells in a contingency table to one of a random table, the 24 sentences that were rated as being most *Neutral* from the set were pulled to form the items for the second portion of the experiment.

ABX Experiment

The results of the norming study were used as the base sentences for the ABX portion of the experiment. All audio files were recorded in a soundproof booth using the built-in speakers of a MacBook Pro. One male and one female native speaker of American English were recruited to produce the utterances for the discourse contexts, with the male speaker's voice being used as the sample, and the female voice used as the match candidates.

For target utterances, the speakers were given utterance contexts and a list of target items. Each of these target sentences included a disambiguating context, which were designed in order to bias a neutral, surprise, or redundancy reading of the particular token. Utterances that were produced with a neutral contour were given with no context, and the voice talents were instructed to read these as regular assertions. Surprise contexts were constructed using semantically biased language which skewed productions toward a positive reaction: these contexts included tokens such as *excited, surprised, thrilled,* and *pleased*. Redundancy contexts were elicited with negatively emotionally skewed language, whee the voice talent was prompted to produce these utterances as if they were frustrated, annoyed, angry or confused. The speakers were recorded separately so as not to influence one individual's speech style by the other.

Figure 4.9 shows the mean differences between each speaker's SRC-HIGH and SRC-LOW productions, along with standard errors. Both plots show the average height of the L*, H* and L% tones measured in semitones. These productions are significant both within speakers (each speaker produced contours whose L* to H* pitch excursions were significantly distinct from each other), and across genders (female and male voices realize the SRC-HIGH and SRC-LOW at different pitch ranges). Table 4.6 outlines these results in more detail.

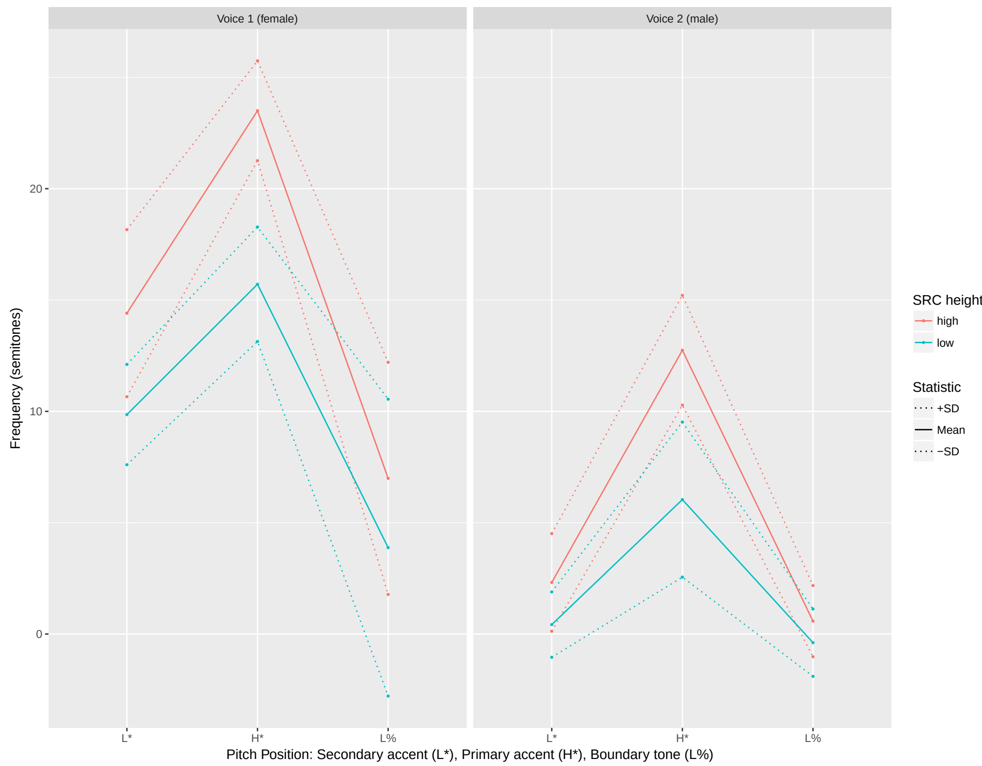


Figure 4.9: Simplified SRC contours: differences in HIGH and LOW productions within speakers

The following table reports the values plotted in Figure 4.9 above. Values for neutral productions are not plotted, but Primary accent and Boundary Tone f0 heights are reported. As the neutral utterances do not have the characteristic L* to H* rise, Secondary accent measurements were not recorded.

		Secondary Accent L*		Primary Accent H*		Boundary Tone L%	
		mean	SD	mean	SD	mean	SD
Voice 1 (female)	SRC-High	14.409	3.752	23.501	2.242	6.992	5.213
	SRC-Low	9.855	2.252	15.709	2.57	3.88	6.66
	Neutral	–	–	17.722	2.324	4.93	6.29
Voice 2 (male)	SRC-High	2.319	2.191	12.745	2.464	0.579	1.6
	SRC-Low	0.424	1.468	6.038	3.483	-0.387	1.513
	Neutral	–	–	8.968	3.031	-0.203	1.629

Table 4.6: Means and standard errors for Pitch Accents and Boundary Tones for female and male voice recordings, calculated in semitones

The effects of speaker and utterance type were tested by fitting a linear model to the data. Dependent variables were calculated on the basis of differences between rises and falls in the contour data. The following two tables show that the difference between the rise (L* to H* high pitch excursion) is significant across SRC-LOW and SRC-HIGH contour types, as well as the fall (H* to L% low pitch excursion) across SRC-LOW, SRC-HIGH, and Neutral contours. Note that measurements taken from Neutral contours represent the difference in rise from utterance initial f0 to main H* pitch accent.

		Estimate (β)	Std. Error	t-value	p-value	
Baseline:	SRC-High	9.092	0.696	13.071	<0.001	***
L* to H*	SRC-Low	-3.238	0.983	-3.291	0.00127	**
excursion	Neutral	-3.109	0.983	-3.160	0.00194	**

Table 4.7: Linear regression coefficients for differences in f0 excursion from L* to H* across SRC-HIGH, SRC-LOW, and NEUTRAL contours, collapsed across Female and Male Voice type

There was no significance within VOICE for comparisons between L* and H*, and H* and L% (SRC-HIGH, p = 0.177; SRC-LOW, p = 0.26; NEUTRAL, p = 0.609). This is as expected: we expect female and males to have different absolute values for pitch accents and boundary tones, but not significantly different excursion slopes. This data supports the position that female and male speakers employ roughly the same pitch modification techniques in order to produce similar contours.

The fall from H* to the L% boundary tone is also significant across production types. Here there is a slight effect of Voice Type, but only in the SRC-HIGH condition (SRC-HIGH, p = 0.0225*; SRC-LOW, p = 0.512; NEUTRAL, p = 0.716).

		Estimate (β)	**Std. Error**	**t-value**	**p-value**	
Baseline:	SRC-High	16.509	0.986	16.742	<0.001	***
H* to L%	SRC-Low	-4.68	1.395	-3.356	0.00102	**
excursion	Neutral	-3.716	1.395	-2.665	0.0086	**

Table 4.8: Linear regression coefficients for differences in f0 excursion from H* to L% across SRC-HIGH, SRC-LOW, and NEUTRAL contours, collapsed across Female and Male Voice type

Finally, we look at the differences between the L* and L% values to check whether these positions might have an impact on the way speakers perceive the two contours. However, we find no significant difference between SRC-HIGH and SRC-LOW productions based on the difference between their L* and L% values. NEUTRAL has been left out of this model, as the relevant L* measurements are not part of this contour's shape.

		Estimate (β)	Std. Error	t-value	p-value	
Baseline:	SRC-High	7.417	1.0	7.417	<0.001	***
L* to L%	SRC-Low	-1.442	1.414	-1.019	0.31	
difference	Neutral	–	–	–	–	

Table 4.9: Linear regression coefficients for differences in f0 values from L* to L% across SRC-HIGH and SRC-LOW, collapsed across Female and Male Voice type

What we see here is an overall difference in production between the SRC-High, SRC-Low, and Neutral contours. Here, the difference was measured solely based on the difference in relative f0 height as it moved from prominent positions throughout the utterance. While these measurements do turn out to be significant in their pitch excursion differences, this is likely not the only signal that listeners perceive when hearing these contours. I will return to this discussion in Section §4.4.3.

4.3.2 Results

Of the 48 participants in this study, four participants were excluded. One participant's data was dropped after learning that they were not a native speaker of English. The other three participants were dropped for a combination of reasons. A first pass inspection revealed that accuracies for these participants were well outside of 2 standard deviations from the mean accuracy; participants overall were able to match the predicted stimulus to the sample 74.6% of the time. Across conditions (3: NEUT—SRC-HIGH, NEUT—SRC-LOW, SRC-HIGH—SRC-LOW), participants' mean accuracy was also quite high:

	NEUT—SRC-LOW	NEUT—SRC-HIGH	SRC-LOW—SRC-HIGH
mean	.688	.789	.761
chance	*.500*	*.500*	*.500*
P1	.208	.250	.125
P2	.375	.167	.167
P3	.333	.250	.250

Table 4.10: Mean accuracy ratings for responses by presentation type, collapsed across order, compared against three dropped participants P1, P2, P3

Given these, the three participants that were excluded were done so on the basis of their significant deviation from mean accuracy scores, as well as their large divergence from the chance response rate (50%). This would seem to indicate signal detection, followed by a decision strategy that deviated from the given on-screen instructions.

Results were analyzed using a signal detection theory of perception. Participants are actively involved in the experimental process, and as such, their perception of experimental stimuli will bias their decision strategy in ways that are not immediately clear to the experimenter. In order to smooth out these biases, we attempt to separate the subjects' inherent ability to discriminate between A and B stimuli when matching to a sample from any implicit or explicit bias that may have been introduced into the experiment.

For each of the three types of comparisons (NEUT—SRC-LOW, NEUT—SRC-HIGH, SRC-LOW—SRC-HIGH), subjects' results were converted to d' by treating the ABX on par with a roving design same-different task. The reason for this is to separate out sensitivity from possible bias toward one or another of the stimulus categories. For the analysis, one category in each pair was treated as containing a signal and the other as containing no signal. Hit rates and False Alarm rates were calculated against these labelings of the trials, and d' calculated for each subject from those rates. The results are used to verify how reliably subjects can distinguish the three stimulus categories. d' values greater than 1 are taken (somewhat arbitrarily) as an indicator of reliable sensitivity, since d'-s greater than this value imply a bimodal stimulus distribution in perception.

Subjects' results within each comparison type were compared to a d' of 1 using one-sample t-tests with a 2 alternative match-to-sample differencing model (MacMillan & Creelman 2005). The results show sensitivity significantly greater than d' = 1 in all three comparisons.

	Mean	t-value	p-value	
NEUT—SRC-LOW	2.156	8.218	4.854e-10	***
NEUT—SRC-HIGH	3.164	9.991	1.508e-12	***
SRC -LOW—SRC-HIGH	2.856	9.282	1.246e-12	***

Table 4.11: Means, t-values and p-values for comparison types NEUT—SRC-LOW, NEUT—SRC-HIGH and SRC-LOW—SRC-HIGH

4.3.3 Discussion

Participants displayed a very clear ability to discriminate between all three contours tested in this experiment based on the perceived emotions of the speakers at the time of utterance. While this experiment does not say anything definitive about the actual emotional responses that participants may or may not assign to the stimuli, it does provide evidence in favor of a more general claim that hypotheses about a semantico-pragmatic meaning distinction between these contours are built upon. In this study, participants were very accurately able to match fine-grained differences in production both across speakers and across genders. This finding is significant for a number of reasons. First, it shows that on a very basic level, there is a perceptual difference between the three stimuli types. Listeners are not only able to distinguish the SRC from neutral, non-emotional speech, they are also able to separate out the hypothesized difference between the SRC with high pitch excursions from those with lower overall fundamental frequency, minima and maxima. Even without knowing what kind of pragmatic meaning is being assigned to these stimuli, it is a positive result that participants can detect the differences in signals with relative ease. The hypothesized difference between the SRC-HIGH and the SRC-LOW is perceptible, and subjects show a sensitivity to this distinction.

Second, this finding is further strengthened by the fact that participants were sensitive to these prosodic distinctions from one speaker to another, as well as between speakers of different genders. Participants are able to quickly perceive and distinguish between similar and different contours that occur in different pitch ranges and with distinct voices. Furthermore, these results show that participants are able to both perceive a difference between a single speaker's vocal range (shown by their ability to discriminate between a single speaker's A and B presentations), as well as match relative f0 height contrasts to a distinct voice. This funding is indicative of a fine-grained capacity to match features of one production to another distinct production. Adding to the complexity of the data, participants were also able to discriminate between contours whose shape was identical, but whose height had been manipulated.

In sum, the results of this experiment point strongly to the conclusion that participants can not only reliably discriminate between SRC-High, SRC-Low and Neutral contours, they can also discriminate across voices, accurately matching female productions to male productions. Experiment 3 is a follow up to this experiment, which takes one step further: given a context, can participants discriminate between the emotive component of the SRC-High and the SRC-Low in order to match an appropriate contour to a context?

4.4 Experiment 3: Extending the ABX results

Experiment 3 builds off of the results of Experiment 2, which was crucial in determining whether a difference in f0 height could explain the difference between "surprise" and "redundancy" readings of the SRC. The results of Experiment 2 show that participants are able to reliably discriminate between high f0 and low f0 productions, but the design does not allow for claims to be made about the pragmatic effect of the contour. Experiment 3 takes this established perceptual difference and asks whether participants are able to reliably match SRC-High productions to contexts that bias a "surprise" reading, and SRC-Low productions to contexts that bias a "redundancy" reading.

4.4.1 Methods

4.4.1.1 Stimuli and Design

This experiment attempts to tease apart the "Surprise" and "Redundancy" pragmatic interpretations of the SRC. This is done by manipulating the contextual factors that would lead a participant to choose one of either the SRC-HIGH or the SRC-LOW as a match to a particular context. We expect the distribution outlined below:

	Surprise Context	**Redundancy Context**
SRC-HIGH	✓	✗
SRC-LOW	✗	✓

Table 4.12: Expected distribution of CONTOUR+CONTEXT categorizations: SRC-HIGH should pattern with *Surprise* contexts, SRC-LOW with *Redundancy* contexts.

This experiment seeks to address the following questions:

(187) **Q4**: When faced with a context biased toward a **surprised** interpretation, is an audio follow up with an SRC-HIGH contour preferred over SRC-LOW?

(188) **Q5**: When faced with a context biased toward a **redundancy** interpretation, is an audio follow up with an SRC-LOW contour preferred over SRC-HIGH?

The design is modified Two Alternative Forced-Choice task (2AFC). In this task, participants are given a context line-by-line which forms the background scenario for the audio evaluation task. Once they have read the scenario, two clips (one SRC-HIGH, one SRC-LOW) are played sequentially. Once they have played, participants are forced to evaluate which of the clips (A or B) is the best fit for the scene they have just read. This task deviates slightly from the standard 2AFC method in that it asks participants to not only choose between tokens that differ in terms of prosodic contour, but also in terms of voice: A and B tokens alternate between male and female speakers.

The design crossed factors CONTEXT (2: *surprise, redundancy*), VOICE (2: *female,*

male) and ORDER (2: SRC-LOW, SRC-HIGH) for a total of 2 x 2 x 2 = 8 conditions. The following is a sample item set of the context, target utterance, and audio order presentation:

(189)

Context 1: SURPRISE	Context 2: REDUNDANCY
<ACTOR> enters the kitchen.	<ACTOR> enters the kitchen.
Breakfast has already been made.	Breakfast has already been made.
<ACTOR> is delighted.	<ACTOR> is annoyed.

Target: She spread butter on the sourdough.
Heard twice, with manipulations of **Presentation order + Voice (F, M)**

a. SRC-HIGH$_{\mathbf{F}}$ – SRC-LOW$_{\mathbf{M}}$

b. SRC-HIGH$_{\mathbf{M}}$ – SRC-LOW$_{\mathbf{F}}$

c. SRC-LOW$_{\mathbf{F}}$ – SRC-HIGH$_{\mathbf{M}}$

d. SRC-LOW$_{\mathbf{M}}$ – SRC-HIGH$_{\mathbf{F}}$

After reading scenarios, participants heard two stimuli, A and B, separated by 1000ms. These audio clips consisted of identical strings of words, but varied whether the voice was female (F) or male (M), and the presentation order of the voices. Items were counterbalanced across participants, such that each participant saw the same number of *Surprise* and *Redundancy* contexts, as well as response orders and trials where the female or the male voice corresponded to the correct answer.

Participants were solicited using Amazon's Mechanical Turk service and compensated for their participation. 24 native speakers of American English (19 male, 5 female, aged 21-55) took part in this study, with participants geographically restricted to the United States for reasons of dialectal intonational coherency. Each participant listened to 24 random, non-identical presentations of the experimental items and made a forced choice decision after each event. Selections were logged and choices were time stamp recorded.

4.4.1.2 Stimuli Creation

Productions from one female and one male speaker of American English were used. Audio stimuli were identical to those used in Experiment 2. This was chosen for continuity, and because these stimuli had already undergone rigorous featural analyses showing that SRC-HIGH and SRC-LOW productions were in fact different in terms of relative height both within and across female and male voices. See Section §4.3.1.2 for detailed discussion.

4.4.2 Results

The following table shows the raw results from all participants with responses grouped by Context (*Surprise, Redundancy*) and collapsed by Voice Type (*female, male*). As expected, *Surprise* contexts bias the listener toward choosing the SRC-HIGH token over the SRC-LOW token (Surprised context: SRC-HIGH = 93.75%, SRC-LOW = 6.25%). Accordingly, we see *Redundancy* contexts provide a bias for SRC-LOW contour responses (Redundancy context: SRC-LOW = 85.42%, SRC-HIGH = 14.58%).

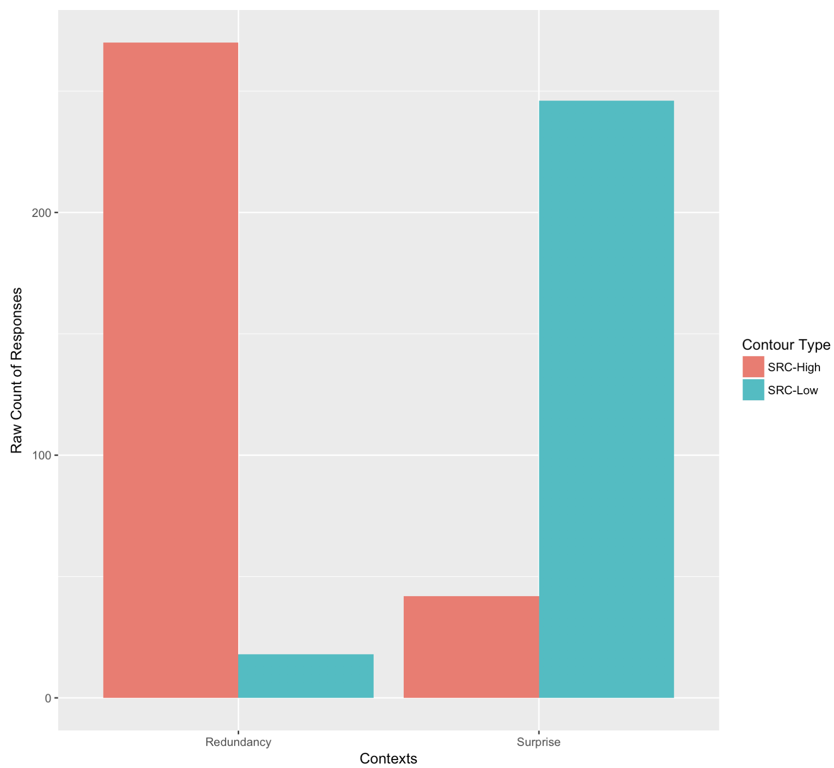


Figure 4.10: Responses for SRC-High and SRC-Low categorizations based on contexts coded for *Surprise* or *Redundancy*

Though simply looking at percent correct statistics in this case shows a relatively clear pattern of classification, speakers' correct identification percentages were also subjected to a d' analysis. Applying a signal detection model to the data, we look at the z-transformed probability of correct identification on a per-participant level MacMillan & Creelman (2005). The figure below shows the distribution that we see for the data gathered. Points in this figure represent a single participants' Z-scores for their two decision strategies. The x-axis identifies the relationship between a participant's classification of a sound file as SRC-HIGH when the scenario was a *Redundancy* context. The Y-axis identifies this relationship between a participant classifying a sound file as SRC-HIGH when they had been exposed to a *Surprise* context.

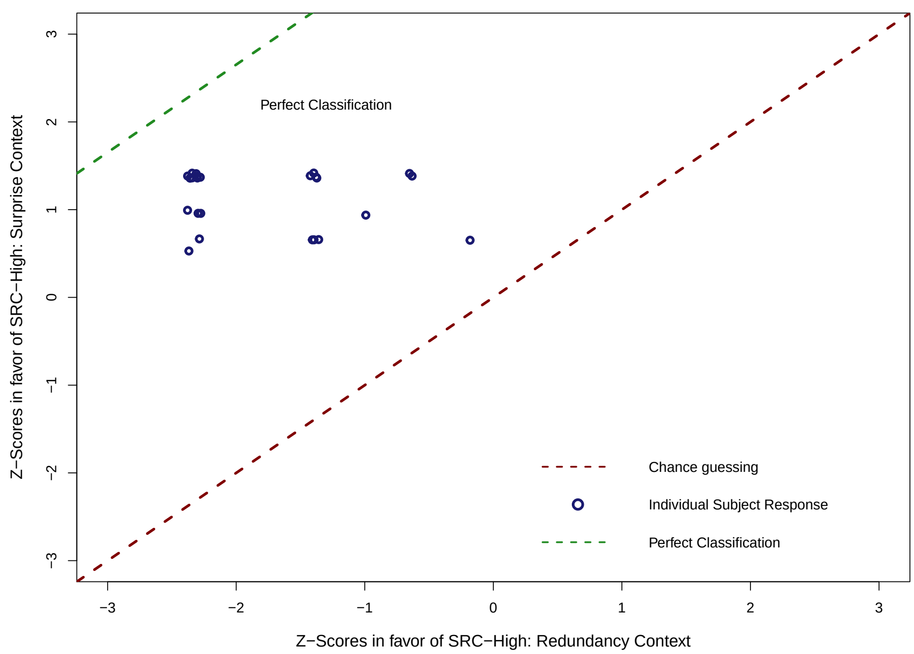


Figure 4.11: Decision space for participants' classification of SRC-HIGH in *Surprise* (y-axis) and *Redundancy* (x-axis) contexts

Z-scores on either side of the dashed line reveal that participants are performing at a level different than chance. In the figure above, participants are reliably doing so for both contexts. When the participant has been presented with a *Surprise* context, they show a clear bias for choosing the response with a SRC-HIGH contour. Similarly, when they have been shown a *Redundancy* context, they are biased toward *not* choosing the response with a SRC-HIGH contour (i.e., they chose the SRC-LOW contour).

A probit logistic regression model was fitted to the categorical response data in R. Table 4.13 summarizes the results.

	Estimate (β)	**Std. Error**	**z-value**	Pr(> |z|)	
Intercept	1.07196	0.09139	11.73	< 2e-16	***
Context (Surprise = 0, Redundancy = 1)	-2.55009	0.14481	-17.61	< 2e-16	***

Table 4.13: Summary of Probit Logistic Regression for SRC-HIGH Response

In Table 4.13, we look at preference for an SRC-HIGH response relative to the predictor variable CONTEXT, with *Surprise* contexts coded as 0 and *Redundancy* contexts coded as 1. The Intercept value was significant with a beta weight of 1.07, indicating a consistent preference for SRC-HIGH responses in the *Surprise* context, favored 6:1 relative to SRC-LOW responses. The value of the predictor CONTEXT was also significant with a beta weight of -2.55, showing that the response pattern was consistently reversed in *Redundancy* contexts: there, SRC-LOW responses were favored 13:1 over SRC-HIGH responses. The significant p-values of both of these model variables demonstrate that subjects on the whole responded consistently and categorically, matching SRC-HIGH responses with surprise and SRC-LOW responses with redundancy.

4.4.3 Discussion

The purpose of this experiment was to build off of the findings from Experiment 2, which showed that listeners are able to easily distinguish productions of the SRC with a high overall f0 from SRC productions with a lower overall f0, as well as categorize these as distinct from a neutral, falling prosody. The previous study was aimed at a listener's perception, and simply asked whether they were able to correctly categorize distinct performances of contours with identical shapes. The current experiment takes the findings from Experiment 2 and asks a further question: is there a pragmatic difference in perception between the SRC-HIGH and SRC-LOW contours?

In order to target the hypothesized pragmatic differences, contexts were created to set up a scene in which an actor would have an emotional response to a salient addressee, event, or situation. Though all of the audio tokens had been semantically normalized for use in a previous experiment (see Section §4.3.1.2 for details), Experiment 3 sought to reintroduce bias through contextual conditioning factors. We hypothesize that SRC-HIGH productions will be matched with contexts that are emotively *positive*; these contexts were constructed using positive emotive language (*surprised, pleased, excited*). SRC-LOW contours are thought to appear with more *negative*-oriented emotions. Contexts that bias a *Redundancy* reading of the SRC contained language like *annoyed, exasperated, irritated.*

Results from this experiment show that participants prefer to match SRC-HIGH-productions to contexts that are semantically biased toward surprised or excited emotions, while they very consistently match SRC-LOW productions with contexts that favor frustrated or annoyed emotions. This lends credence to the hypothesis that the high and low SRC productions are both perceptually and pragmatically distinct. Coupled with the results from Experiment 2, which showed that participants are very reliably able to discriminate between high and low productions of the SRC across voices, this result is welcome, as it stands as further evidence that relative f0 is a distinguishing factor between "surprise" and "redundancy" interpretations of the SRC.

This contour in itself is an interesting case study in the larger framework of intonational mirative strategies. Previous work has always identified the SRC as a contour with two distinct pragmatic meanings: surprise or an inference that someone in the context *should have known* some fact. Both of these comment on violated expectations, but until now, no research has commented on how the two readings come about. From the investigations here, we have the beginnings of an answer. Both readings, though composed off of the same abstract (H)L* H* L% template, interact with other acoustic factors in a compositional way. When speakers use the SRC with a high relative f0, the underlying prosodic template contributes a basic notion of violated expectations, which is modified (at the very least) by pragmatic cues associated with relative pitch height.

Given this, we must also be aware of other paralinguistic factors that listeners tune into when distinguishing the pragmatic effect of these contours. It is possible that listeners were not focused as much on pitch height or pitch movements during the task, but rather, tuned into a speaker's timing, loudness, voice quality, or other suprasegmental properties of the speech stream (nasality, breathiness, voice on/offsets, etc...). None of these things were controlled for in these tasks, setting up a potential confound in the data here. But though the data may be muddied by these other factors, we can think about the general connotative meanings that the two different SRC contours have and draw parallels to other places in the grammar that might lend credence to our hypothesis in another way. The SRC-High, like the Excited contour, is a contour that is characterized by *positive* surprise: it seems that in general, positively emotive utterances are associated with higher pitch, while lower pitched utterances have a tendency to be perceived as negative (specifically, as anger or frustration). What we might instead want to claim is that the SRC is a templatic contour, whose interpretation of general violated expectations is modulated by fluctuations in pitch. We have identified to general pitch heights, High and Low, which seem to be able to characterize this expression of violated expectations as either "positive and surprising" or "negative and redundant." One can conceive of a follow up experiment in which multiple pitch levels are tested for their perceptible difference, as well as for their differences in pragmatic meaning.

In light of this, and given the discussion above, it seems reasonable to conclude from the data presented in Experiments 2 and 3 that at the very least, one of these distinguishing factors in distinguishing the High from the Low SRC readings is a speaker's relative pitch height.

4.5 Outlook and future progress

The research program outlined in the previous sections is still in its beginning stages. All three experiments explored new empirical terrain with methodologies that were created to suit the particular questions being asked. In a sense, these studies were in themselves experiments in testing the efficacy of new linguistic methodologies. Though there have been many previous studies that experimentally test semantic and pragmatic ideas, the methodology here broadens the reach to show one way that the pragmatics of intonation can also be included in these efforts.

These methodologies are not without avenues for improvement, and the analyses here do not, in their current form, take advantage of all of the analytical angles that are offered by the data. In particular, none of the three experiments above look at reaction time data that was recorded for each participant's responses. In part, this was because in most cases, the patterns observed appeared to be explained simply and effectively with the forced choice responses. But looking at patterns of reaction time data may lead to further insights into a participant's behavior, and perhaps into the difficulty of the tasks as a whole.

In particular, looking at reaction time data might help to explain the results gained in Experiment 1, which asked participants to give an answer on a scale from 1-5 regarding an utterance's *naturalness*. Utterances where contours present on discourse particles do not set up a pragmatic clash between contours on a follow up utterance are predicted to receive higher ratings than those where there is a clash. In this situation, we predict that utterances that are rated higher on the Naturalness scale will have faster reaction times than utterances rated lower on the Naturalness scale. We interpret a pause on the part of the participant as indicative of a clash in information processing: something between hearing one particle+contour and then a different contour with a follow up utterance should cause a processing slowdown. On the other hand, particle and sentence contours that cohere should be interpreted quickly.

Zooming out, looking at how intonation interacts with other semantic and pragmatic content is a clear path for discovery, part of which is currently being explored in work related to Jeong & Potts (2016) and Jeong (2017) on utterances with different types of rising intonation in English. It also lends itself well to other languages that are intonationally structured like English (such as German or Dutch), or which have clear prosodic patterns with distinct meanings (as in Brazilian Portuguese, French or Spanish).

One clear parallel to the English discourse particle system presented in both the current and previous chapters is the German modal particle system. German, like English, has seen extensive work done in both the semantic literature on modal particles, and in the literature on language-specific intonation and prosody. But there little to no research that explicitly links the two. Doing this with the methodologies presented above could lead to new insights about modal particle meaning, and could help to tease apart pragmatic discourse effects that come about through uses of modal particles from those that arise from prosodic means.

Phonetically and prosodically grounded work in intonational semantics and pragmatics ultimately puts armchair theoretic judgments to the test: they ask whether intuitions about felicity and theoretical conceptualizations of the contributions of a particles and contours can be verified experimentally. If they can, then our theories can (for now) rest easy. If not, this is also positive for our pragmatic theories of language understanding, as it shows that we must approach the problem from another angle. Either way, it is a step forward.

Part II

Modal particles and Intonation in German

Chapter 5

Reinterpreting the German modal particle system

German modal particles (MPs) as a class have received much more attention than English discourse particles. Accounts of their syntactic and semantic contributions to an utterance make up a substantial body of literature, and as a result, their treatment, meaning, function and contributions to an utterance have many varied explanations. Like the English discourse particles discussed in the previous chapters, these modal particles fall under the umbrella of prosodic discourse markers. The goal of this chapter is to present a unified picture of German and English discourse markers by framing these modal particles in terms of a speaker's expectations.

Modal particles tend to be left out of written varieties of German, yet like English, the spoken language is rife with them, and they are largely non-register specific. German has over twenty MPs, all which serve distinct semantico-pragmatic roles in an utterance. The following chapters looks at two German modal particles in particular, both of which encode forms of speaker expectation. But unlike the English particles from the previous chapters, I argue that only one of the German particles under discussion here encodes mirativity–*doch*. The particle *ja*, on the other hand, while still a marker of a speaker's expectations in a discourse, marks content as *unsurprising* in the eyes of the speaker and in view of the context, drawing attention to information that is accessible to both a speaker and a hearer. The two particles form an interesting test pair, especially in the face of some theories that attempt to capture the meaning of one in the meaning of the other.

Understanding the effects of German intonational patterns is also crucial to determining just what the contributions of each MP is in a context. But due to the full integration of MPs into the center of a clause, German presents an more nuanced prosodic picture than English: MPs and intonation are intricately bound in a way that the English particles are not. Because of this, it is even more important to attempt to separate out intonational meaning from MP meaning. This chapter will explore the connections that German modal particles have to intonational contours, and the contributions of each in a discourse.

This chapter first outlines the German modal particle system and its defining properties. It then turns to a discussion of the modal particle *doch*. Here, I first present the pretheoretical empirical terrain, and then move on to discuss a treatment of this particle, which I assume is a marker of mirativity in German.

5.1 Properties of modal particles: a sketch

German has a particularly rich system of particles that contribute actively to discourse navigation. One subgroup of this system is the modal particles, which pattern independently of other particle or adverbial groups. Modal particles form a class based on shared syntactic and semantic properties and have the distinction of being semantically uninterpretable outside clear contextualization; their contribution appears to be entirely pragmatic. Kaufmann & Kaufmann (2012) refer to these as 'epistemic particles', as many take their basic function to be marking information as known or given in a context. However, their clearest and perhaps most basic function is to provide commentary on how a speaker has integrated (or plans to integrate) information into a discourse. (190b) contrasts with (190a) not in terms of the semantic content or propositional truth of the utterances, but by introducing the speaker's attitudinal stance in relation to the utterance's informativity:

(190) a. *Nach* HAMBURG *möchte Paul fahren.*
to Hamburg wants Paul go
Paul wants to go to Hamburg.

b. ***Ausgerechnet*** *nach* HAMBURG *möchte Paul fahren.*
of-all-things to Hamburg wants Paul go
Paul wants to go to HAMBURG, of all places! *(König 1991)*

Focus marking on HAMBURG in either utterance in (190) invites the listener to infer an alternative set of places that Paul may have been choosing between, a là Rooth (1992). But while the addition of the MP *ausgerechnet* does not change the truth value of the utterance, it does invite the listener to infer that this desire of Paul's frustrates or annoys the speaker, thus introducing her subjective bias into the discourse. Another kind of bias is calculated in (191b) with the MP *ja*:

(191) a. *Den Brief hat Paul am Freitag gehabt.*
the letter has Paul on Friday had
'Paul had the letter on Friday.'

b. *Den Brief hat Paul* ***ja*** *am Freitag gehabt.*
the letter has Paul *ja* on Friday had
'Paul clearly/obviously had the letter on Friday.' *(modified from Lindner, 1991)*

While both (191a-191b) commit the speaker to Paul having the letter on Friday, the MP provides extra commentary as it relates to the speaker's view of the common ground. Using *ja* signals to the listener that the speaker views this information as already known to everyone in context. In a neutral context and with neutral prosody, it seems to imply something along the lines of "As we all know...".

Previous studies of MPs have shown a range of properties common to this particular grammatical class. They include all of the properties of PDMs, as well as a few more that are specific to the workings of German. A partial list of all of the proposed properties is below, taken from Doherty (1982), Thurmair (1989), Abraham (1991), and Diewald (2007):

(192) Modal Particles:

a. are optional elements

b. cannot have their contribution negated

c. have sentence scope

d. occur only in the Mittelfeld

e. are uninflected and generally unaccentuated

f. cannot have their contribution be questioned

g. are dependent on sentence mood or type

h. are combinable with other modal particles

Some of the most common modal particles appear in Tables 5.1 and 5.2 below. In Table 5.2, these particles are grouped according to the descriptive generalizations argued for by Thurmair (1989), who proposes that in addition to the properties listed in (192), MPs can be classified in terms of their basic functions in a conversation.

Table 5.1: Common German Modal Particles

aber	*auch*	*bloß*	*denn*	***doch***	*eben*	*einfach*	*etwa*
halt	***ja***	*mal*	*nur*	*ruhig*	*schon*	*überhaupt*	*wohl*

Thurmair identifies four broad interpretational categories that modal particles appear to fall into, with respect to their general effect on a discourse. In her system, modal particles can be categorized by whether they provide *propositional evaluativity*, or commentary on a speaker's beliefs about a particular proposition. They can also be grouped by those that may only provide *illocutionary modification* in a particular clause type, either strengthening or weakening the effect of imperatives or deontically modalized utterances. Other particles must make *reference to a conversational partner* salient in the context, and others still require *reference to an antecendent* that is on the discourse record.

	aber	*auch*	*bloß*	*denn*	***doch***	*eben*	*einfach*	*etwa*	*halt*	***ja***	*mal*	*nur*	*ruhig*	*schon*	*überhaupt*	*wohl*
Prop. Evaluativity	✓	✓			✓	✓	✓	✓	✓	✓				✓	✓	
Illocut. Modification			✓							(✓)	✓	✓				✓
Partner Reference					✓							✓	✓			
Antecendent Reference		✓		✓		✓	✓	✓	✓					✓		

Table 5.2: Summary of Thurmair (1989)'s Modal Particle classifications based on their effect on an utterance

For Thurmair, these interpretational categories form the basis of her featural account of modal particles. Once a particle has been shown to possess certain discourse-level properties, their functions can then further restricted. Particles that are *propositionally evaluative* are composed of features that further specify what kind of evaluative statement they make. *Doch*, for example, comments on the fact that a proposition is *known* in a context, while also requiring *reference to a partner*. Once these general categories have been established, a particle is further specified for its individual features. In the case of *doch*, that additional feature is a *correction*. *Ja* also specifies that a proposition is *known*, and in particular contexts, can be a particle that *strengthens* the illocutionary type. While I will not engage too deeply with these classifications for each MP and what it means to be *known*, *corrected*, or *strengthened* in this system, these classifications provide great insight into the discourse navigatory power of MPs more generally. This featural breakdown is one way of systematically representing the way that speakers actively place their own expectations and beliefs about the conversation and their interlocutors into view.

Shifting back to properties of MPs, the first four common to all MPs make reference to their syntactic position within a clause. German is a split V2/SOV language, whose syntax traditionally divides the clause into a pre-verb 'Vorfeld' region, a 'Nachfeld' region following the finite verb and the 'Mittlefeld' between the inflected and finite verbs in matrix clauses. Though MPs are typically assumed to be some sort of adjunct or syntactic modifier, they are only allowed in this Mittelfeld region: they are ungrammatical pre-and post-verbal (Hentschel 1986). But this restriction aside, their placement in a clause between other constituents is relatively free. The only differences that arise are pragmatic. (193a-d) show that placement of MPs here is unrestricted, even able to occur on either side of the adverbial *gestern* 'yesterday'; The Mittelfeld in (193) is demarcated here by angle brackets.

(193) a. *Sophie hat* **<*doch/ja*** *gestern ihrer Tochter das versprochene Buch***>** *geschenkt.*
Sophie has *doch/ja* yesterday her daughter the promised book gave
'Sophie gave her daughter the book she was promised yesterday.'

b. *Sophie hat* **<***gestern* ***doch/ja*** *ihrer Tochter das versprochene Buch***>** *geschenkt.*
'Sophie (clearly) gave her daughter the book she was promised yesterday.'

c. *Sophie hat* **<***gestern ihrer Tochter* ***doch/ja*** *das versprochene Buch***>** *geschenkt.*

'Sophie (clearly) gave her daughter the book she was promised yesterday.'

d. *Sophie hat* **<** *gestern ihrer Tochter das versprochene Buch* ***doch/ja*** **>** *geschenkt.*
'Sophie (clearly) gave her daughter the book she was promised yesterday.'[1,2]

As soon as an MP is placed outside of the Mittelfeld, the utterance is uninterpretable as an utterance modified by an MP:

(194) a. **Sophie* ***doch/ja*** *hat* **<** *gestern ihrer Tochter das versprochene Buch* **>** *geschenkt.*
'Sophie (clearly) gave her daughter the book she was promised yesterday.'

b. *****Doch/Ja** *Sophie hat* **<** *gestern ihrer Tochter das versprochene Buch* **>** *geschenkt.*[3]

c. **Sophie hat* **<** *gestern ihrer Tochter das versprochene Buch* **>** *geschenkt* ***doch/ja***.

The semantic meaning or grammaticality of these sentences is not dependent on the presence of these modal particles, but each of them has a slightly different pragmatic interpretation. This interaction with focus and intonation will be addressed in a later section.

MPs are also not subject to negation, and in fact, are syntactically restricted in the presence of negation. Though MPs and negation operators frequently occur in the same utterance, both of which are assumed to operate on the sentence as a whole, Thurmair (1989) notes that the scope of negation must be narrower than the scope of an MP when they co-occur. In (195b), when *doch* scopes lower than negation, ungrammaticality results, though the word order is otherwise acceptable without *doch*:

(195) a. *Tom kann* ***doch*** *den Weg* ***nicht*** *finden.*
Tom can *doch* the way not find
Tom (clearly) can't find the way.

[1] All of these sentences have multiple interpretations based on stress/focus placement. One interpretation of these sentences is with no special focus or topic marking on elements that may have been picked out by *doch* or *ja*, which is the prosodic mapping that I am assuming. It is also possible for a sentence such as (193b) to have focus on DAUGHTER, but this is not necessary. Whether or not modal particles must lexically associate with focus will be addressed in a later section.

[2] Some argue that MPs may only be adjoined at the vP level, and that their seemingly free placement in a clause is a result of scrambling and topicalization operations of other constituents. Others maintain that MPs attach higher to TP or to CP. As the syntactic placement of these elements is not crucial to the analysis I propose, I do not take a stand here.

[3] There is a reading of (194b) that is grammatical, but it involves interpreting *doch* not as a MP, but as a complementizer meaning roughly 'however'. This is distinct from the MP use in both syntactic placement and semantic interpretation, and not considered part of the analysis presented here.

b. *Tom kann **nicht** den Weg (***doch**) finden.*

A similar pattern is seen when negation and *doch* are linearly adjacent, with either high (196) or low (197) placement of negation. When a MP is in the scope of negation, the sentence is uninterpretable:

(196) a. *Tom kann **doch nicht** den Weg finden.*
Tom can *doch* not the way find
Tom (clearly) can't find the way.

b. **Tom kann **nicht doch** den Weg finden.*

(197) a. *Tom kann den Weg **doch nicht** finden.*
Tom can the way *doch* not find
Tom (clearly) can't find the way.

b. **Tom kann den Weg **nicht doch** finden*.

MPs are not inflectable elements of the language, rejecting case marking entirely. With the exception of two particles, they are also unstressable. However, in both cases where particles are able to carry prosodic prominence, the distribution of the stressed particle is more restricted than its phonologically unstressed counterpart. Thurmair (1989) argues for the inclusion of these two stressed particles into the MP class, but she also notes that their accentuation seems to change their function from pure MP elements to particles that tend toward a scalar or polarity particle interpretation.

MPs can also be categorized based on their semantic similarities. Previous accounts have assumed that MPs comment on backgrounded material, either presuppositionally triggered, entailed, or conventionally implicated (Kaufmann & Kaufmann 2012). Leading to this claim is the observation that their content cannot be denied; being not-at-issue prevents this, as in (198). For many, the contribution of *doch* triggers an *uncontroversiality* presupposition, marking its prejacent as given or shared content, and further, insinuating that the addressee *should have known* some salient fact that they seem to be overlooking (see Lindner 1991, Karagjosova 2001, Grosz 2012, Kaufmann & Kaufmann 2012). This inference turns out to be not-at-issue, as evidenced by the infelicity of overtly denying it, as in (198):

(198) *B accuses A of not helping with a project. A points to a sentence written in her own handwriting*:

A: *Das habe ich* ***doch*** *geschrieben!*
that have I *doch* written

I wrote that!

B: # You didn't think I already knew that.

Though all examples thus far have only shown MPs modifying declarative sentence types, this is not a generalization that extends past particular MP requirements. MPs have been shown to appear in all sentence types. MPs like *mal* can occur in a question that serves as a polite request, as in (199), or *wohl*, which appears to betray a speaker's violated expectations:

(199) *Kannst du mir* ***mal*** *die Butter reichen?*
can you me MAL the butter pass
Could you pass me the butter?

(200) *Sophie and Hannah hear an investigator ask a leading question. Sophie to Hannah:*

Was meint sie ***wohl*** *damit?*
what means she *wohl* that-with

What (could/do you think) she means by that?

Though no particular clause type can be ruled out as a host to MPs, each MP individually has restrictions on what utterance types it can occur in. On the view that MPs encode expectations that a speaker has about a discourse context, this is not so much a feature of the grammar, but rather, a pragmatic calculation given factors introduced by clause type restrictions or general facts about the discourse. Ultimately, it is pragmatic conditions that rule out MPs in particular sentential environments.

Lastly, MPs can be defined in terms of their relation to other MPs. Stacking MPs (conjunction is strictly ungrammatical), as long as they are distinct particles, is considered a fairly regular process, and serializations of up to four particles are found in quite widespread use:

(201) *Ruf* ***doch***$_{MP1}$ ***ruhig***$_{MP2}$ ***mal***$_{MP3}$ *bei mir an!*
call DOCH RUHIG MAL by me on
Do feel free to just give me a call!

The combinations of MPs in itself an interesting puzzle, as there are complex rules governing the placement and ordering of the combinations. For now, the main focus of this inquiry will restrict itself to *ja* and *doch*. What follows is an attempt to accurately portray the landscape of use of these two particles.

Though Thurmair (1989) analyzes a wide range of particles with her closed class of attributes, this work will focus in particular on two particles, *doch* and *ja*. But following in the tradition of Thurmair, the overarching theme of this line of inquiry can be summarized in the following way:

> "With modal particles, the speaker gives the hearer clues to how he should interpret what has been said and where he should integrate it, as well clues about how the speaker views the [content of an] utterance, how the previous utterance was integrated [into his view of the conversation], or how he has assessed the expectations of the hearer. The hearer should take these clues into consideration in his reaction; in this way, the function of modal particles is partially to influence the conversational partner in a particular way." [4] (94)

This description of modal particles is particularly relevant in the wider scope of this project, as identical things need to be taken into consideration with respect to the contribution of intonation to a discourse. Thus, separating out the meaning of modal particles from the meaning that intonation contributes is an important task, but it is just as important to separate the meaning of particles from each other. As she proposes, MPs comment on a speaker's *expectations* about how a hearer will react in a conversation.

Doch and *ja* are two particles whose meanings and contexts of use are incredibly similar. Both seem to comment on the expectations of both the speaker and the addressee, and are extremely similar in their discourse functions. This closeness in meaning has lead many

[4]Original German: "Mit Modalpartikeln kann der Sprecher nämlich dem Gesprächsparnter Hinweise darauf geben, wie er das gesagte auffassen, und wo er es einordnen soll, sowie Hinweise darauf, wie der Sprecher die Äußerung bewertet, wie er die Vorgängeräußerung aufgenommen hat oder wie er das Wissen und die Erwartungen des Hörers einschätzt. Diese Hinweise soll der Hörer in seiner Reaktion berücksichtigen; auf diese Weise ist es auch eine Funktion der Modalpartikeln, den Gesprächspartner in einer bestimmten Richtung zu beeinflussen." (94)

researchers to examine their use in tandem, with some even proposing that the meaning of one includes the meaning of the other. This work attempts to contribute to this body of literature by looking at the particles and how they contribute to the commitments of a speaker in a table model of discourse. This is done by examining the expectations the particles place in a speaker's discourse commitments, and how contextual factors play a role in the felicitous interpretation of these speaker expectations. In doing this, I identify *doch* as a marker of mirativity, and show that while *ja* comments on a speaker's expectations, it differs from *doch* in that there is no assumption of expectation *violation*. The similarity between the two particles lies in both of their reference to the expectations of multiple conversational actors, not just the expectations of the speaker. The following section lays out the data for these two MPs.

In addition, I claim, following Lindner (1991) that German modal particles are operators on illocution types, and as AnderBois (2016) suggests for Yucatan Maya *bakáan*, that *doch* is an illocutionary mirative, which indicates a speaker's violated expectations relative to a context, a speech act, an event, or a proposition. Both *doch* and *ja* are markers that contribute particular conventional discourse effects.

5.2 The modal particle *doch*

Doch is the canonical example of a German modal particle, known both for its relative frequency in the spoken language, as well as the attention it has received in linguistic literature. Previous accounts have differed in terms of their treatment of the MP's core function; some argue that the particle is used to contradict a salient fact *p*, while others frame the contribution in terms a speaker's assumption that their addressee *should have known* this salient *p*. Others still present the particle as a way for the speaker to answer the current QUD in a particular way. I take a different tack. Given the variety of sentence types and discourse contexts that *doch* appears in, and coupled with the pragmatic subtleties that are introduced with *doch*, I propose an analysis that captures the contributions of the particle in terms of the *expectations* that a speaker has of herself and the participants in a discourse. In particular, *doch* marks a speaker's *violated* expectations given the current state of the discourse, and places these expectations in full view to their addressee as part of their discourse commitments. Using a modified Table model of discourse management in the style of Farkas & Bruce (2010), I propose that placing the pragmatic component of *doch* in a speaker's Discourse Commitments list is a way to make the particle's contribution public information, while maintaining the observation that the contribution is not-at-issue.

5.2.1 How many *doch*s?

The trouble with pinpointing the meaning component of *doch* is apparent from the start: it can be used as a polarity response particle (202a), as an accentuated sentence medial modal particle (202b) and as a sentence medial unstressed modal particle (202c):

(202) a. **Tom**: Jeff's not coming to the recital tomorrow. *Polarity Response*

Steven: DOCH.
Yes he is.

b. *Jeff tells Sophie that he has class and can't make it to her recital. Later, before the recital begins, Jeff walks in the door.*

Sophie: *Du bist* DOCH *gekommen!*
you are DOCH came

You came after all! *Stressed MP*

c. *Tom is hungry during the recital. He asks Steven if there is a concession stand. But Steven saw Tom pack his bag before they left.*

Steven: *Du hast* ***doch*** *Schokolade in der Tasche!*
you have *doch* chocolate in the bag

(But) you have chocolate in your bag! *Unstressed MP*

There are three main stances that propose to capture this apparent polysemy. The first is conservative, and proposes that all three particles should receive their own treatment. The second is minimalist, proposing that the particles are all derivable from a single underlying item. The third is the middle ground, and the route taken here. This path assumes that polarity response particle instances of *doch* form a distinct use, and should be considered separate from modal particle uses of *doch*. This view is based on a variety of factors that distinguish the particle on the grounds that it is both syntactically distinct and semantically more restricted than the modal particle uses.

The polarity particle **DOCH**, like French *si*, is part of a tripartite system of polarity responses in German. Utterances or questions with positive polarity are affirmed with one particle (*oui* in French, or JA in German), while a distinct particle is used in responses to utterances with overt negative polarity (*si* in French, **DOCH** in German). The polarity particle **DOCH** is disallowed as a response to a question whose highlighted alternative has positive polarity (203a), while it is allowed (and JA is not) when the highlighted alternative of a question is negative (203b):

(203) a. **Jeff**: *Haben wir Joghurt?*
have we yogurt
Do we have yogurt? p

Sophie: ***Ja*** *(haben wir). / #***DOCH**.
Yes, (we do).

b. **Jeff:** *Haben wir* ***kein*** *Joghurt?*
have we **no** yogurt
Do we **not** have yogurt?[5] $\neg p$

Sophie: **DOCH** *(haben wir). / #**Ja**.*
Yes (we do).

Syntactically, the polarity response **DOCH** does not carry the restriction that MPs do of appearing only in the Mittlefeld of a clause. MP uses of *doch*, both stressed and unstressed, must at the very least co-occur with an inflected verb. Second, the polarity particle always appears utterance-initially, and in its own prosodic phrase with utterance level prominence, setting it apart from MPs, which must be prosodically incorporated into a larger prosodic phrase. I set polarity response particle uses of **DOCH** aside in this work, and instead defer to Farkas (2010), Krifka (2013), Roelofsen & Farkas (2015) for a in-depth treatment of their semantics.

Taking the middle ground stance assumes that the idiosyncrasies of the unstressed *doch* and the stressed MP DOCH can be derived from a common source. This is the position I

[5] *Ja* here is ungrammatical on the intended reading where it asserts that there is in fact yogurt to be had. However, it is not completely ungrammatical as a response to the question, as Krifka (2013, 2015) shows. *Ja* can be used to agree with the highlighted proposition of the question, namely, that yes, there is **no** yogurt.

take here. I argue that the unstressed *doch* and DOCH share an identical core pragmatic contribution to an utterance, and are distinguished based on the extra accenting that is placed on the stressed DOCH. This is to say that the stressed and unstressed particles are not instances of two distinct lexical items, but rather, have the same underlying conventional discourse effect. I return to this distinction in Section §5.4.2. The discussion now turns to an empirical investigation of *doch*.

5.2.2 The Empirical Terrain

Most of the research on *doch* looks at how the particle can modify assertions, either occuring as answers to questions, or as responses to previous utterances. As a first pass distinction between DOCH and *doch*, perhaps the most obvious point of departure between the stressed and unstressed variants is DOCH's requirement for appearing alongside negation. Though this is reminiscent of the requirements of the polarity response particle, the restrictions are more lax with DOCH: the stressed MP can react to a salient item on the discourse record with negative polarity (204), or it can appear with negation in the same clause in order to contrast against a salient proposition (205):

(204) *Sophie and Steven both think that Tom* ***can't*** *dance. Later they observe him doing a perfect tango. Sophie to Steven:*

Tom kann DOCH *tanzen!*
Tom can DOCH dance
Tom CAN dance after all!

(205) *Sophie and Steven are pretty sure that Jeff is a good dancer. But later they see him flailing about on the dance floor. Steven to Sophie:*

Jeff kann DOCH ***nicht*** *tanzen.*
Jeff can DOCH not dance
It turns out Jeff CAN'T dance.

In both instances, the MP DOCH appears to be a signal by the speaker that some discourse expectations need to be revised. In (204), both participants in the conversation have expressed a high certainty that some proposition, $\neg p$, is true. Later, they realize that their expectations have been violated: p turns out to be true. They signal this discrepancy with DOCH, and by using the stressed variant, indicate that not only is there a contradiction, that contradiction was surprising. A similar thing happens in (205), but here, instead of expecting $\neg p$, the two believe p to be true. When they encounter the opposite to be true, the can also signal this with DOCH, but just as in English, they cannot do this without negation present in the prejacent. It seems that with DOCH, both contradiction to a salient p, as well as a speaker's calculation of violated expectations are in play.[6]

In addition to declaratives, the MP more generally is also allowed in a variety of other speech act and sentence types. The unstressed *doch* can also be used in imperatives, biased questions and wh-exclamatives, while the accented variant is licensed in imperatives, but has a more restricted distribution in exclamative and interrogative contexts.

The additional accentuation on stressed DOCH seems to play a part in its more restricted distribution. For the stressed DOCH to be licensed, it requires reference to two propositions or events that stand in clear, logical opposition to each other. As in (204-205), DOCH could be used to contradict a discourse-record proposition with positive polarity (p, DOCH($\neg p$)), or one with negative polarity ($\neg p$, DOCH(p)). Placing sentential prominence on DOCH appears to signal that the speaker's violated expectations lie in the way they had calculated the polarity of the utterance, which gives rise to an inference of direct contradiction in terms of the particle's discourse effect. A sentence may only be construed as having positive or negative polarity. On a violated-expectations account, stressing DOCH indicates that your expectations were violated, and that there is only one logical way to resolve your inference path. Because of this singular

[6] One might be tempted to treat all stressed variants of DOCH the same, lumping the polarity response particle use in with the MP, along the lines of Egg & Zimmermann (2016). Yet treating these two particles as identical items fails to explain at least two observations that seem to separate the two. First, there is a clear temporal contrast that arises between MP and polarity particle uses of DOCH. MP utterances can refer to an event far back on the discourse record just as (204–205) can; Sophie and Steven can have held these beliefs long before bringing them up again. In contrast, the response particle *must* respond to the immediately preceding utterance; it is infelicitous to utter **DOCH** or **DOCH nicht** in isolation as a response to either (204) or (205).

Egg & Zimmermann (2016) mention that one way the two particles could be united is through sentential ellipsis, where all but the particle "may be elided under verum focus" (p.231). But such an account raises serious questions about whether there is evidence or motivation for a syntactic movement of DOCH to a peripheral position which would license phrasal (CP or TP?) deletion, or on the other hand, whether verum focus is a licensing condition for this kind of sentential deletion at all. Given this, I opt for separate treatment of polarity response **DOCH** and MP DOCH.

option for revision, the stressed particle sets up the illusion of requiring pure contradiction.

Some existing analyses attempt to extend the discourse-effect of the contradiction hypothesis to the unstressed *doch* as well. But doing this requires a more flexible interpretation of what kind of content can be targeted for contradiction. In an exchange like (206), while Tom can contradict Sophie with *doch* in (206a), he appears to contradict one of the *presuppositions* of Sophie's statement in (206b), namely, that he has a dog:

(206) **Sophie:** Dein Hund hat gestern Abend die ganze Nacht gebellt!

Your dog barked all night last night!

a. **Tom:** *Hat er* ***doch*** *nicht.*
has he *doch* not
No he didn't.

b. **Tom:** *Wir haben* ***doch*** *keinen Hund.*
we have *doch* no dog
We don't have a dog!

If *doch* were simply a marker of contradiction, we might expect that only (206a) would be felicitous with *doch*. We would miss the other pragmatic information that it can target in an utterance. As well as providing alternative information to the context, *doch* adds commentary about Tom's view of the current conversational context: he is surprised that Sophie thinks he has a dog. The otherwise bare utterance with neutral intonation does not allow this pragmatic interpretation, and neither does the stressed MP, which is contextually infelicitous in this case. Though there is much similarity, the unstressed *doch* has yet a subtler function in a discourse than its stressed counterpart.

Setting aside for the moment the issue of contradiction, we can turn to examples like (207), which demonstrate again that speaker mood plays a large part in the interpretation of sentences containing the MP *doch*:

(207) **Hannah:** Kommst du mit in die Oper?

Are you joining us for the Opera?

a. **Sophie:** *Nein, ich habe abgesagt.*
no I have declined.

No, I declined.

b. **Sophie:** *Nein, ich habe* ***doch*** *abgesagt.*
no I have *doch* declined.
No, *you should know that* I declined. (Rojas-Esponda 2013)

The discrepancy here between Sophie's responses is subtle in terms of truth conditional meaning; in fact, one would likely evaluate the two responses as truth conditionally equivalent. Yet whereas (207a) is relatively neutral, the response in (207b) has a clear undertone of expectation violation. *Doch* in this case implies that there had been some previous discussion of Sophie's attendance at the opera, and given that previous discussion, Hannah's question should not be relevant—Sophie's use of *doch* indicates that she's surprised that she has to answer this question, since the answer already lies in the common ground.

In part, it is this feeling that some participant in the discourse *should have known better* that drives the violated expectation interpretation of *doch*. A speaker may have a certain view or expectational causal model of the world that they are approaching a discourse with, and given new information and facts that are introduced, they can update their outlook accordingly. In cases where they have to revise a certain assumption that they have made, *doch* can be appropriate. Formulating the contribution of the particle in this way is reminiscent of the treatment of the English discourse particles in Chapters §3 and 4: as markers of mirativity. Mirative markers are a grammatical device that languages use to cue a listener in to the speaker's belief state at the time of utterance. For a speaker, they help to place recent receipt of information into the larger communicative context. The German particle *doch* shows all the hallmarks of a mirative strategy: it is a particle that comments on how a speaker's expectations have been violated, given her knowledge of the discourse context.

In addition to contexts that set up an obvious expectational contrast between a speaker and an addressee (such as question-answer scenarios (207), or situations where a speaker is correcting a previous utterance (206), *doch* can be used more in more obviously emotive cases where a speaker really wants to foreground the fact that expectations have been violated:

(208) *Tom walks in to find Jeff watching football.*

Super Bowl?! Das ist ***doch*** *total langweilig.*
Super bowl that is *doch* totally boring
The Super Bowl? Come on, that's totally boring.

(209) *Jeff announces that he'll be running a marathon next weekend. Tom responds:*

Du kannst ***doch*** *nicht ohne Training einen Marathon laufen.*
you can *doch* not without training a marathon run
(Oh please.) You can't run a marathon without training.[7]

Not only are the examples above an expression of a speaker's violated expectations, they also implicitly ask the addressee for a response. In (208), Tom notes that Jeff's actions do not accord with his view of world, and they also seem to seek a response; Jeff needs to somehow defend his actions. In (208), perhaps he will assert that he likes football, or that the team from his hometown is playing. In (209), he can submit that he has been training, or that he doesn't need to train, he's fit enough as it is. But in both cases, if Tom is met with silence, the conversation does not seem felicitous. *Doch* warrants a response.

The use of both stressed DOCH and unstressed *doch* in these declarative sentences carries with it three important observations. First, in all cases of the stressed MP and the unstressed MP, expectation violation plays an important role in navigating the particles' precise conventional discourse effect. Second, *doch* can indicate to a addressee that a certain aspect of the conversation *should have been known*. Third, this tendency of *doch* to mark statements as evident to (at least) the speaker gives rise to an implicit pressure on the part of the hearer to respond to the *doch*(p) utterance.

The differences between these three instances of *doch* can be summarized in Table 5.3. Checkmarks in the table represent *necessary* conditions for felicitous interpretation of the particle in question; the unstressed *doch* may, for example, occur with the conversational background that sets up an explicit contradiction, but it *need not*:

[7]Examples taken from "German is Easy!" blog, https://yourdailygerman.com/meaning-of-doch/, published 27 February, 2012.

		stressable	sets up a contradiction	requires negation	sentence initial	positive response to negation	"should have known"
MP	*doch*						✓
	DOCH	✓	✓	✓			✓
PP	**DOCH**	✓	✓	✓	✓	✓	

Table 5.3: Broad strokes differences between *doch*, DOCH and **DOCH**

While the stressed MP DOCH can occur in many of the places that the polarity particle can, their uses do not completely overlap. The unstressed particle, at least relative to the conditions in this table, seems to have none of its necessary conditions overlap with the polarity particle. And in terms of stressed and unstressed *doch*, the stressed variant requires many more conditions to be fulfilled. What can be said is that a variant of *doch* is used in situations where the speaker and the addressee have come to a point in the conversation where their expectations on how to proceed conflict in some way. The particle is a signal that some resolution needs to happen in order to come to a consensus on a given conversational topic.

The remainder of this section focuses primarily on cases of the unstressed MP *doch* in various other sentence types, and attempts to pin down what exactly the core meaning component this particle contributes to a discourse.

5.2.2.1 *doch* in Imperatives

The use of *doch* in imperatives has been said to have a "strengthening" effect on the particular type of imperative (Thurmair 1989). In requests, the effect is to soften the speech act, making it closer to an invitation (210). In commands, it places the speaker's expectations in full view, implying an urgency with which the speaker requires the imperative to be fulfilled (211):

(210) *Sophie and Tom are leaving work and run into Jeff:*

Wir gehen ein Bier trinken. Komm ***doch*** *mit!*
we go a beer drink come with

We're going to have a beer. Come with us, if you'd like!

(211) *Jeff is whistling while Sophie is trying to work. Sophie::*

Sei ***doch*** *endlich still!*
be *doch* finally quiet

Shut up now, will you!

In (210), the request is interpreted as a suggestion, and does not obligate the addressee to comply with the imperative, but rather, simply suggest that they do. With the command in (211), Sophie's utterance leaves no room for Jeff not to comply–it makes the command more forceful.

The behavior of *doch* in imperatives can shed some light on just how the particle navigates violated expectations in other sentence types. With declaratives, *doch* modifies an utterance from an neutral statement to one that expresses the speaker's mood, either in a way that shows a speaker's surprise at a situation, or in a way that shows surprise at the actions of another discourse participant. With imperatives, this modification happens in much the same way. *Doch* directly reflects the speaker's attitude toward the utterance. Speaking theory neutrally, whereas a bare imperative can be interpreted as having the effect of removing the speaker from the utterance ("Don't talk in class!" = *I don't want you to talk in class* OR *The rules advise not to talk in class*), adding *doch* seems to force an interpretation that orients the imperative toward the speaker's expectations. This can have the effect of downgrading a request to a suggestion, and strengthening the effect of a command.

There are further cases where *doch* in an imperative seems to rely on a speaker's view of how others should formulate their expectations. Take, for example, the context in (212):

(212) *A mother and her young child are walking in a parking lot. Normally, she allows her child to walk without holding her hand, but this afternoon, the parking lot is very busy, and she is worried about her child's safety. She utters:*

Nimm ***doch*** *meine Hand.*
take *doch* my hand
Just hold my hand.

In this case, it seems that a speaker may use *doch* to flag unexpected events. Normally, the child would not be expected to have to hold the mother's hand, yet in this outlier case where the parking lot is exceptionally busy, expectations shift. Such an environment warrants a mother's wanting to have her child be safe from danger, and using *doch* is an attempt to bring this exceptional situation to the child's attention. She does not expect a struggle–given the contextually defined situation, she expects that the child should hold her hand, and the child should be able to infer why. *Doch* forms a sort of explanatory link between the command and the reason for it.

5.2.2.2 *doch* in Questions

Doch's use is limited to questions that express speaker bias, which is perhaps unsurprising under a view that *doch* references violated discourse expectations. The function of a pure information seeking questions should set up the speaker as being ignorant about the likelihood of one answer over another, and the question itself should not introduce bias into the equation. If we talk about *doch* as referring to someone's violated expectations, this inherently encodes some sort of speaker bias. As might be expected from this, *doch* is allowed in biased questions like the tag question in (213a), the rising declarative in (213b), but ill-formed in the pure polar question in (213c):

(213) a. *Du hast* ***doch*** *Sekt mitgebracht, oder?*
you have *doch* sparkling-wine brought-with or
You of course brought sparkling wine, right?

b. *Du hast* ***doch*** *Sekt mitgebracht...?*
you have *doch* sparkling-wine brought-with
(I'm assuming) You've brought sparking wine...?

c. *# Hast du* ***doch*** *Sekt mitgebracht?*
have you *doch* sparkling-wine brought-with

If we assume that questions semantically denote sets of alternatives, and that for polar questions, these sets are the prejacent of the question and its negation, then $^{?}$p semantically denotes $\{p, \neg p\}$. Speaking in very general terms, a polar question is inherently a move that requests information about the highlighted alternative. If *doch* introduces a speaker's expectations about the prejacent, as well as commenting on how others in the context should view the proposition, this is no longer a pure request for information. A similar pragmatic restriction is found with *doch* in Wh-questions, as in (214):

(214) **Sophie:** Do you remember when Germany beat Brazil 7-1?

Tom: *Ja... was war das* ***doch*** *für ein Fußballspiel?*
Yes what was that doch for a soccer-game
Yes, what kind of soccer game was that again?

Wh-questions with *doch* are interpreted as reminders, rhetorical questions, or momentary lapses in a speaker's memory. (214), for example, can only be interpreted in a context where Tom is asking to be reminded about the specifics of the match. He knew at one time, but now it escapes him—*Was it in the World Cup? Was it the semi-finals?* An answer that is relevant for these two questions is also a relevant answer to the question in (214). In other apparent Wh-question uses of the particle, the speaker uses this structure to ask a rhetorical question:

(215) *Wie bemerkt Goethe* ***doch*** *so treffend?*
how remarked Goethe *doch* so aptly?
What was it again that Goethe said so aptly?

The use of *doch* in questions provides another key insight into the conventional discourse effect of *doch*. Questions provide perhaps the clearest case of distinguishing just what kinds of things *doch* can anchor a speaker's violated expectations to. With tag questions like (213a), the speaker seems to be referencing some pre-established fact, and asking for confirmation or clarification. From Wh-question forms, it is clear that a speaker's own lapse in memory can be the anchor of surprise—the violated expectation that *doch* makes reference to is the fact that this question needed to be asked in the first place.

5.2.2.3 *doch* in Exclamatives

In addition to questions, imperatives and declaratives, *doch* is also possible in exclamatives. This is perhaps the corner of the scholarly work on *doch* that is the most untouched. Thurmair (1989) notes that *doch* can be used only in WH-Exclamative contexts as in (216a-b):

(216) a. *Was hast du* ***doch*** *für einen schönen Pelzmantel!*
what have you *doch* for a nice fur-coat
'What a (very) nice fur coat you have!'

b. *Was hast du* ***doch*** *für schöne Beine!*
what have you *doch* for pretty legs
'What (very) pretty legs you have!' *(Thurmair, 1989, 115)*

It is difficult to capture just what extra information *doch* brings to an exclamative utterance, as exclamatives seem to be the closest approximation of what *doch* actually contributes. Rett (2009) characterizes exclamatives as content that comments on the fact that the degree to which some property holds exceeds a relevant contextual standard. For an exclamative to be expressively correct, there must also be a salient context for the utterance, and the speaker must find the content of the utterance surprising. In other words, a speaker uses an exclamative to comment on her own violated expectations relevant to a contextually established degree. The expectations that *doch* introduces into a context must then be understood with respect to their interactions with exclamatives.

Rett makes a further distinction between types of exclamative contexts. *Propositional exclamations* differ from pure *exclamatives* in virtue of the fact that propositional exclamations do not require a contextually salient degree. The relevance here is that the contexts in which *doch* occurs do require a degree: they are *exclamatives*. Assuming that *doch* is an attempt for the speaker to make her expectations clear, one possible explanation is that the particle comments on the speaker's expectations about where this degree threshold lies. In exclamatives, what *doch* contributes is something like "though I expected your coat to be pretty, the degree to which it is is higher than I thought." Exclamatives with *doch* require a contextually salient degree.

The degree threshold with *doch* in this sentence type is very highly contextually relevant. Take, for example, the utterance in (217) below:

(217) *A five year old comes home from Kindergarten with a painting in her hand, and proudly presents it to her father. The painting is an obvious improvement, and showcases the child's developing motor skills very nicely. Her father exclaims*:

Das ist ***doch*** *so schön!*
that is *doch* so pretty
'That is just so pretty!'

A five year old child's Kindergarten painting will not be the epitome of high quality art. But in context, the overall detail of the painting may surprise the father enough that degree to which this painting is pretty exceeded the previous paintings that she may have done for him. What *doch* in (217) does is reference just this. The painting being pretty is not surprising, but it is surprisingly good for a piece of children's art.

In WH-exclamatives with *doch*, the speaker is expressing their amazement at the degree to which something holds. Most WH-exclamatives express some degree of amazement in terms of a gradable property. In (216a), this property is *niceness*. The coat could be nice, it could be pretty nice, it could be at the upper end of the extreme, commenting on the extreme niceness of something. The presence of *doch* in (216a) tends toward this extreme interpretation, implying that the item in question is on an extreme end of the gradable niceness scale.

Grosz (2012) mentions that in addition to WH-exclamatives, polar exclamatives also allow *doch*, but are restricted by the fact that they must occur with what he calls 'overt particles'—*glatt*, 'outrightly' or *tatsächlich*, 'indeed'. In fact, as he notes, all polar exclamatives must obligatorily occur with some marker of acknowledgment that the surprise that is being conveyed is to an extreme degree. That is, not only are the sentences below ungrammatical with only *doch*, but they are also no good without an overt degree marker:

(218) a. *Kennt der* ***doch*** **(**glatt**) den Kaiser von China!*
know he DOCH/*doch* outrightly the emperor of China
'[I'm shocked that] he knows the emperor of China!'

b. *Hätte der dem* **DOCH** **(**tatsächlich**) das Buch gegeben!*
had$_{subj}$ he him DOCH indeed the book given
'[I'm shocked that] he would have indeed given him the book!' *Grosz, 2012*

Grosz (2012) notes that a polar exclamative "expresses a polar opposition between what is the case and what was to be expected." In a sense, a speaker's expectations have been completely surpassed to such a degree that this amount must be verbalized. One must also observe that the *doch* that is licensed here is not always the unstressed MP, but can also be the stressed variant. In fact, though the polar exclamative in (218a) can be either stressed or unstressed, the variant in (218b) may only be the stressed particle. The distribution of the stressed and unstressed particle in these polar exclamatives is, contrary to other environments where this particle is licensed, highly irregular. This lends credence to the interpretation of the *doch* utterances above, where a speaker's expectations have been violated relative to some proposition. In cases where the speaker has some idea of a degree or gradable property, an exclamative context is allowable, whether wh- or polar.

What does seem clear is that *doch* in an exclamative context signals surprise to a very high degree, which is strikingly evident in interpreting WH-exclamatives, and coercable in polar exclamatives that force a degree reading. Again, we can boil down the proposed meaning of this MP in terms of a speaker's violated expectations, which will be key when formulating the core meaning of *doch*.

5.3 Deriving a meaning for *doch*: modal particles, mirativity, and expectation

5.3.1 Modal Particles and the table model

Modal particles occur as just one part of the larger discourse strategy, and as such, many proposals have sought to capture their functions by embedding their contributions in particular models of discourse. There have been relatively few attempts to analyze MPs in the Farkas & Bruce (2010) table model of discourse, which seeks to make clear the discourse commitments of conversational actors in a scoreboard-type system. Such a model is a convenient way of representing these MP contributions, as the model itself makes the relationships between propositions in the common ground, issues on the table, and the particular discourse commitments of participants very clear. For *doch*, this is especially convenient, given the particle's tendency to reference both its prejacent, as well as a separate salient proposition or event. Below, I outline two recent approaches to understanding MPs in a table model framework. I then introduce my own update to the system and to the function of the particles, which relies on the notions of expectation and expectation violation.

5.3.2 A brief review of the Table Model

Recall from Chapter §2 the five main components of the Farkas & Bruce (2010) Table model, which grows in part out of the works of Stalnaker (1978) and Gunlogson (2001, 2008):

(219) a. **Table:** s stack of issues modeling salience in the current context

b. **Common Ground (*cg*):** the set of all propositions that all members of the conversation are publicly committed to during the course of the conversation

c. **Discourse Commitments (DCs):** for every participant *x*, the set of propositions that *x* has publicly committed to, but which are not yet part of the CG

d. **Projected Set (*ps*):** the set of propositions currently under consideration for addition into the *cg*

e. **Context Set:** the set of all worlds that are compatible with the propositions in the Common Ground

Table model is a system that decompose the act of assertion into specific effects that it has on a discourse that do not become part of the common ground. This grows out of the desire to unify the way in which declaratives and interrogatives are formally represented.

Conversational actors negotiate a conversation by way of a commitment to shrinking the Context Set by adding content to the Common Ground. When speakers raise issues (which are sets of propositions), they place them of the Table for consideration, Issues are resolved (either by way of agreeing to move a proposition into the common ground, or rejecting the proposal) by clearing them from the top of the stack of relevant issues on the Table. For a more detailed look at this, see Chapter §2.3.

5.4 Two theories of Modal particles on the Table

Müller (2014) takes on the particles *doch* and *ja* in her table model approach to the interpretational differences between the MP combinations *ja doch* and *doch ja*. While her main goal is to understand how the apparent ban on the particle combination *doch ja* arises in most situations, she must first explain the contribution of the particles in isolation. She does this by appealing to a speaker's commitments about propositions in the discourse.

Using *doch* in this system puts a speaker's commitment to the prejacent p in their personal discourse commitments list, but it additionally places the issue of $^{?}p$ on the table, and suspends the resolution of a salient q, which is calculated from presuppositions introduced by p. By doing this, the speaker indicates to the addressee that these two propositions are in some way incompatible, and one or the other can be accepted, but not both. In an exchange like (220), the issue of q stands in opposition to the issue of p, the prejacent of *doch*:

(220) **Sophie**: Jeff's not home. $= q$

Tom: *Aber sein Auto ist* ***doch*** *da.*
but his car is doch there
But his car is there! $= p$

Under this view, all the pragmatic heavy lifting that *doch* and *ja* do is done on the table. In this system, a *ja*-assertion by a single speaker puts p into the discourse commitments of both the speaker and the hearer and in one fell swoop, places p onto the table and then into the conversational common ground. This seems to model the pragmatic contribution of *ja*, which is to treat the content of p as known to all members of the conversation.

This also seems to fare well for *doch* in many contexts. In immediate response scenarios like (220), this is a reasonable tactic: the information that a speaker is responding to with the *doch* statement is simply the preceding utterance. Putting this information on the table allows the speaker and the hearer to access this in order to come to a joint resolution of the issue. But in out of the blue cases, or cases where there is no clear contradiction (as in non-contrastive cases like (207)), dealing with these issues by suspending the acceptance of a previous proposition on the table to introduce a contradiction does not work.

Relying on the table for the discourse contribution of this particle is problematic for another reason: it assumes that the contribution of the particle is at-issue. This is contra many of the properties of German MPs outlined in Section §5.1. By introducing an alternative that *doch* contrasts with, the analysis gets at the heart of the particle's ability to reference two contrastive propositions, but allowing the alternative to be negotiated on the table suggests that the contribution of *doch* can be subject to negotiation, which it cannot.[8]

Csipak & Zoebel (2016) also use the table model to show the discourse function of conditional *denn*, another common discourse particle in German. The function of *denn* has many similarities with *doch*, especially in its effect of emphasizing a speaker's disbelief of a particular proposition. But unlike *doch*, this reading is apparently environment-specific: it may only arise with this function in the antecedent of a conditional.

The meaning component of *denn* is added into the discourse commitments of the speaker and at the same time, are introduced into the projected set. Csipak & Zoebel (2016) argue that *denn* contributes a relative probability measure which is contextually determined. Using *denn* indicates that the speaker viewed the probability of that utterance as lower than than some contextually supported threshold.

The core of Csipak & Zoebel (2016)'s proposal also treats *denn* within the table model of discourse. They assume that a *denn*-utterance comments on tacit proposals, which should be entered into the common ground as presupposed content when an issue p is placed on the table. These can be any content that is not explicitly conveyed, and includes presuppositions and defeasible inferences:

(221) **Csipak & Zoebel (2016)'s Tacit Proposals:**

The proposition p is tacitly proposed or can reasonably be inferred to be tacitly proposed by a participant α, where p is a necessary precondition for the validity of the content of a previous utterance by α (or a part of that utterance).

The pragmatic contribution of *denn* is to make a speaker's uncertainty with respect to the prejacent clear. The core of their proposal is in the way tacit propositions are introduced. If there is evidence in the discourse context, either explicitly or tacitly introduced, *denn* can make reference to this, indicating that the speaker is uncertain as to whether she can commit. *Denn* can either target the prejacent, or an issue that has been tacitly raised previously.

[8] See example (198) in Section §5.1 for discussion of this.

In more familiar terms, tacit proposals are those propositions that a speaker assumes an addressee will accommodate with any utterance. In order to allow tacitly proposed content directly into the common ground, and able to bypass the table, Csipak & Zoebel (2016) propose that there is an ordering relation that dictates which contributions are added to the table at which time. Tacit proposals that arise from an utterance p are added as not-at-issue content in a speaker's discourse commitments before p is placed as an issue on the table. When an issue is resolved, tacit proposals are accepted into the common ground alongside the proposition. Introducing tacit proposals is essentially a way for the discourse model to be able to keep track of proposals that have not explicitly be introduced into the discourse. This is important for a treatment of *denn* which allows the particle to make reference to its prejacent, as well as other salient propositions that may or may not be explicitly on the record.

I argue for a simpler view of modal particles in a table model system. I take the position that MPs are inherently a marker of a speaker's expectations in a particular context. Being able to reference the expectations of particular discourse participants allows for a wide range of ways for a speaker to mark how an utterance either accorded or did not accord with their view of reality. I assume that these expectations are not-at-issue commentary that a speaker enters into her discourse commitments at the time of utterance. Doing this allows for the speaker to express her view of the situation and make it known to all other discourse participants, while at the same time making it off-limits for acceptance or rejection by other participants. The following section goes deeper into this relation between expectation, MPs, the meaning of *doch*, and how it fits into a table model of discourse.

5.4.1 A brief review of expectation and mirativity

Many models of discourse rely heavily on the notion of a speaker's commitment to issues that they raise, or that have otherwise been introduced. This commitment can be variable. When speakers use assertive strategies (falling intonation, and in some languages, a default word order), they are assumed to be committed to the propositional truth of their utterance. When speakers use interrogative strategies, they defer commitment to another conversational agent. While the notion of commitment can capture much of the back and forth of a conversation, tracking a speaker's pragmatic contributions requires a more nuanced view of a conversational context. In particular, conversations track not only how speakers work together to grow the common ground between them, they also are sensitive to the expectations of the speakers themselves, and how their views of the actual world may be shaped or changed with the exchange of information.

In previous chapters, I have argued that some English discourse particles and prosodic contours are strategies that make a speaker's expectations known to their addressee. This is a useful grammatical tool, as tracking a speaker's basic expectations in a discourse can alert conversational participants to when a speaker views a proposition as contextually neutral, exciting, or surprising. While many languages encode this speaker affect in the prosodic component, many more encode this grammatically as particles, verb forms, or morphological affixes. Mirative strategies in particular rely on how a speaker's expectations can be *violated*. If actors in a discourse are sensitive to pragmatic notions of violated expectations, they must also be passively aware of the general expectations of their interlocutors as well.

The following is a brief summary of Chapter §2's comprehensive discussion of mirativity. Miratives are strategies that speakers use to mark a speaker's violated expectations in a discourse. This information is integrated into their view of the current discourse context, and helps their addressee understand how is this information is integrated as well. Miratives, like intonation, contribute illocutionary not-at-issue content, which in a tabletop model of discourse navigation, can be registered in the speaker's Discourse Commitment list. Placing this here makes the contributions of miratives known in the context, but not as public or projected content.

Miratives in English can take the form of prosodic contours and discourse particles. Chapter §3 of this work outlined two different kinds of mirative markers in English: those that make reference to the violated expectations of the speaker alone, and those that mark a speaker's violated expectations in view of her interlocutors. While the basic effect is similar, the pragmatic contributions are slightly different. Simple mirative strategies like English *oh* and *huh* can be used to indicate surprise without requiring an overt explanation for it, or to anchor it to a salient addressee. But more complex mirative strategies require some sort of explanation or follow up from another member of the conversation. English requires this of the Surprise-Redundancy Contour, and in what follows, I argue that German *doch* is also a complex marker of mirativity. In fact, *doch* and the SRC play extremely similar roles in their respective languages, both indicating a speaker's expectation for the truth of one proposition in view of another participant's commitment to a contrasting one in the discourse.

The previous section was in part an overview of the various environments that *doch* can be found it, but moreover, it was also an exploration of the role that a speaker's expectations play in calculating licit uses of the particle. From context to context, and across sentence types and speech act types, the constant factor was *doch*'s ability to make a speaker's expectations known to discourse participants. Other effects, such as contradiction, unexpectedness, or a speaker's insistence that some relevant fact *should be known* falls out from an expectation-based account of this particle. Coupled with environmental factors (such as conventional effects contributed by prosodic contours, or the semantic licensing conditions of particular speech act types), *doch* uniformly projects expectation violation. The next section is dedicated to formalizing the contribution of this mirative marker.

5.4.2 The contribution of *doch*

Modal particles are one strategy that German uses in order for a speaker's expectations to be realized in a conversation. Previous sections have laid out a variety of empirical facts that frame how *doch* places the expectations of a speaker into public view. For Csipak & Zoebel (2016)'s *denn*, instead of raising tacit proposals, the particle could be construed as talking about how a speaker's expectations have *not* accorded with their internal model of the world. With *doch* utterances, the expectation calculation is particularly nuanced: a speaker expected the addressee to share similar beliefs, yet something in the discourse context has lead her to believe the opposite. *Doch* is a mirative strategy: it inherently expresses a speaker's violated expectations in view of a particular context. I propose the following as an expectational-account of the contribution of *doch*:

(222) *doch*(q) is anaphoric to a salient proposition or event p in a discourse context C and for

discourse-salient participants x, s.t.:

a. *doch*(q) adds the following to the speaker's *DC* list:

$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

The formulation of *doch* in (222) has flavors of many of the previous analyses of the particle. There is an element of contrast between two propositions p and q, which signals that the speaker expects that the two propositions are incompatible. It also signals that the speaker's expectations factor in an assumption of her addressee's expectations as well. The real difference in this formulation of *doch*'s pragmatic contribution lies in how the contribution integrates into the discourse. Like English discourse particles, German modal particles contribute not-at-issue commentary about a speaker's expectations. In the table model, the contribution is logged into a speaker's discourse commitments list, where it is presented as information committed to by the speaker, but not raised as an issue available for public scrutiny.

In contrast to Csipak & Zoebel (2016), who assume that updating with *denn* enters elements directly into the common ground, I argue that *doch* does not make any common ground updates: there are negotiations that still must go on between participants in a discourse. Making an assertion with *doch* does not close an issue. In an exchange like (223), when Tom makes his *doch* utterance, it is strange if there is no overt accepting move by Sophie:

(223) **Sophie:** There's no cheese left in the fridge.

Tom: *Ich habe* ***doch*** *gestern Käse gekauft!*
I have doch yesterday cheese bought
I bought cheese yesterday!

Tom expects an answer or an elaboration from Sophie when he uses *doch*. Before Sophie's utterance, he has certain expectations about what food is and is not available. Given his knowledge that he had just purchased cheese, it is reasonable for him to be surprised when Sophie mentions that there isn't any. *Doch* is a particle that asks the interlocutor for a reply—it pragmatically lays out the speaker's expectations about the world, and introduces a contrast that must somehow be resolved. The conversation takes a strange turn if Sophie does not respond to Tom's statement. For this reason, I place the contribution of *doch* solely in the discourse commitments of the speaker.

5.4.3 Extending the analysis: *doch*, expectation, and sentence types

One main reason to favor an expectation-based account of *doch* is that it can easily be used to explain the contribution of the particle in non-declarative sentence types. As opposed to QUD models, or accounts that assume contradiction of a salient proposition in context, an expectation account of *doch* bypasses those that limit the use of the particle to only particular sentence types. In base cases like those in (223), the particle can respond to the propositional content of an utterance. In question-answer pairs like (224), repeated from (240), *doch* targets the fact that the question was asked at all:

(224) **Jeff**: Are you coming with us to the Opera?

Tom: *Nein, ich habe* ***doch*** *abgesagt.*
No I have doch cancelled

No, I cancelled.

In (224), Tom uses *doch* as a response to Jeff's act of questioning, given the fact that he expects Jeff to know that he had declined the trip to the opera. Jeff's speech act is the salient *p* here: the fact that he asked whether Tom was coming with. (225) shows the breakdown in semantic and pragmatic contributions that go along with Tom's utterance:

(225) a. Jeff asks whether Tom is coming to the opera. (Speech act = *p*, contextually salient event)

b. **Tom**: Ich habe doch zugesagt (=q)

c. Update the context with q and *doch*:

DC$_{Jeff}$	**Table**	**DC**$_{Tom}$
	{q}	$\text{Exp}_T(q) \approx 1 \wedge$
		$[\text{Exp}_{J,T}(p\|q) \approx 0]$

$cg : p,\ ps = \{\{p \cup q\}\}$

Here, the contextually salient p is already in the common ground, as it is Jeff's speech act itself. What *doch* indicates here is Tom's surprise that he must answer this question he assumes is known to all parties.

The use of *doch* in interrogatives is limited to questions that express some sort of speaker bias toward a particular resolution of the question. This is a straightforward prediction of an analysis that places a speaker's expectations into view for other discourse participants. Since *doch* comments on the expectations a speaker has on both the content of the utterance as well as another salient proposition, it would be strange if *doch* questions were *not* biased.

In a biased interrogative question like (226), though Tom uses a question form, he pragmatically indicates that he has an expectation for the highlighted alternative. In these cases, it becomes clear why pure contradiction to some salient p cannot be a contribution that *doch* makes to a discourse. Though Tom has a fairly high expectation for the highlighted alternative of his question, he leaves room for Steven to respond:

(226) **Steven:** I can't remember what Ross says about gapping.
Tom: *Du hast* ***doch*** *den Artikel gelesen, oder?*
you have doch the article read or
You read the article, didn't you?

In *doch* questions like (226), Tom has good reason to believe that Steven has in fact read the article in question. The question form without *doch* indicates this. Using *doch* pragmatically indicates that Tom's belief in the highlighted alternative is higher than with a normal tag question. This is presented in contrast to Steven's utterance, which Tom takes to mean that Steven has not retained the information necessary from Ross' paper. (227) shows the contributions Steven and Tom make in this situation:

(227) a. Steven utters p:

$\mathbf{DC}_{Steven}$	**Table**	$\mathbf{DC}_{Tom}$
q, $\text{Exp}_S(p) \approx 1$	{p}	

$cg : s_0,\ ps = \{\{s_0 \cup p\}\}$

b. **Tom:** Du hast doch den Artikel gelesen, oder? (highlighted alternative = q)

c. Update the context with q and *doch*:

$\mathbf{DC}_{Steven}$	**Table**	$\mathbf{DC}_{Tom}$
p, $\text{Exp}_S(p) \approx 1$	$\frac{\{q \vee \neg q\}}{\{p\}}$	$\text{Exp}_T(\neg\text{q}) < \text{Exp}_T(\text{q}),\ \text{Exp}_T(q) \approx 1$
		$\wedge\ [\text{Exp}_{S,T}(p \vert q) \approx 0]$

$cg: s_0 = s_1,\ ps = \{\{s_0 \cup p\}, \{s_0 \cup p \cup q\}, \{s_0 \cup q\}\}$

I make no major theoretical claims on the analysis of tag interrogatives like *oder* here; instead I blindly assume that they bias the speaker to expect the highlighted alternative ($Exp_{spkr}(\neg q) < Exp_{spkr}(q)$). What I can say is that the tag question and the use of the modal particle have a cumulative effect. Tag questions bias a speaker toward the highlighted proposition, and *doch* increases this bias. This, coupled with Steven's utterance alluding to the fact that he has not read the article in question, gives us the intended interpretation. The question tag positions the speaker as seeking confirmation of q, while the pragmatic effect of the modal particle betrays their actual intention.

Though I make no explicit claims on the contribution of exclamatives or imperatives in a table model of discourse, a descriptive explanation of *doch* in these sentence types lends insight into how to understand the contribution of the particle in these contexts. Imperatives with *doch* have a softening effect, as (228) does:

(228) **Sophie:** Should I call Jeff? I don't want to bother him. Do you think I should call?

Tom: *Ruf ihn* ***doch*** *an!*
call him doch on!
(Just) call him!

Assuming a standard theory of imperatives (see Portner 2004, or Condoravdi & Lauer 2012, for example), we can assume that an imperative commits the speaker to an effective preference for a proposition of the form *addressee performs VP*. In (228), this amounts to Tom committing to the preference for Sophie to call Jeff. *Doch*'s contribution is to pick out a behavior salient in the context such that it somehow conflicts with this preference expressed by the imperative. In the scenario above, Tom expects Sophie to call Jeff, yet Sophie is acting in such a way that indicates that she won't. This doesn't accord with Tom's expectations, and he indicates as such with *doch*. In an exclamative, Rett (2009) assumes that there must be salient content that the speaker finds surprising *to a contextually relevant degree*. Like *doch*, exclamatives comment on violated speaker expectations, and such contexts seem like the perfect environment for mirative particles to occur. Yet as noticed in Section §5.2.2.3, *doch* actually has some interesting restrictions. Exclamative contexts with *doch* appear to limit the particle to modifying the degree to which the speaker's expectations have been violated. An illustrative example takes the form of (229):

(229) *Steven walks out in shorts for the first time since winter*:

Tom*: Was hast du* **doch** *für schöne Beine!*
what have you doch for nice legs
What nice legs you have!

Picking out a salient *p* for *doch* to be anaphoric to in (229) is tricky. Tom has already committed himself to the content of his exclamative utterance. But the content of the utterance is in itself the surprising element. The salient *p* here is actually the degree of niceness that the speaker expected. *Doch* here indicates that Tom's expectations about the *niceness* was violated in such a way so as to exceed his expectations.

Recent work by AnderBois (2017) has identified mirative elements in Yucatec Maya (*bakáan*), as well as Tagalog (*pala*), that also appear to act as miratives whose contribution is to content above the speech act. The particle *bakáan* in particular is relevant here, as it functions as an *illocutionary mirative*, which can reference the sudden revelation of a speaker, which leads them to make the *bakáan*-utterance in the first place. Like *doch*, the particle can be used in a variety of sentence types, and its use is dictated more by contextually relevant pragmatic conditioning factors, rather than on pointing out a clear logical inconsistency, contradiction, or equivalence. One key insight taken from AnderBois (2017) is that even given its flexibility in speech act types (*bakáan* is found in declaratives, imperatives, and interrogatives), the clearest generalization about the particle's use is in the speaker attitude that it imparts on an utterance, and is an indication of an actor's internal model should be updated to reflect this. At its core, *bakáan* is a speech act modifier, revealing sudden realization or revelation on the part of the speaker. This is strikingly similar to the core function of *doch*.

From his investigation, AnderBois (2017) predicts that there are three main classes of miratives that we might find in natural language use, two of which he has clear evidence for, and the third which is proposed but is under-researched in the mirative literature. Class 1 miratives encode revelation or realization, and are uniformly *illocutionary updates*. These are particles like the Yucatec Maya *bakáan*, and looking back to the data in previous chapters, English *oh, huh* and the SRC. For German, *doch* appears to also fit into this class. These miratives help model participant attitudes in a discourse. Class 2 miratives at their core encode new information, and crucially, are *propositional updates*. They help navigate discourse content, indicating source of information. The members of this class are hypothesized to be those that function on the level of dual evidential/mirative markers, such as Turkish *-miş* or Cheyenne *neho* (see Chapter §2 for discussion). Class 3 miratives are not as clear cut. AnderBois (2017) hypothesizes that members of this class are mirative markers that encode only surprise, but is agnostic on what their mirative argument is (illocutionary, propositional, other?). He notes that these might include English surprise intonation, or the Tagalog particle *a*, but ultimately leaves this up to future research.

So far, I have argued for an interpretation of English and German mirative particles that is built off of speaker expectation, which seems to pattern in some respects quite closely to AnderBois (2017)'s Class 1 miratives requiring sudden revelation, and in other respects closer to his Class 3 miratives, which encode surprise. While I hesitate to classify these particles squarely into one bin or another based off of the mirative interpretation they contribute to an utterance, his observation that miratives can operate on the illocutionary level is exactly on track with what I propose here. The mirative elements that I engage with in this work share this core insight, which is a contribution unto itself.

Part of a broader investigation of mirativity and mirative strategies cross-linguistically is a sensitivity to other pragmatic indications of speaker expectations or mood. In exclamatives in particular, as well as in other sentence types, intonation plays a key role in how a speaker intends an utterance to be interpreted, and how an addressee is able to then parse the various semantic and pragmatic contributions. Exclamative utterances with *doch*, for example, rely very highly on the performance of the utterance in addition to the semantic content that the exclamative form provides, assertions with *doch* can also be manipulated by pitch variation, voice quality, focus marking, and other compounding factors. One avenue for further investigation into mirative classifications of Surprise (in the style of AnderBois (2017)'s Mirative Classes) would be with a sensitivity to these paralinguistic factors as well. While Chapters §3-4 of this work are a start to this research program, they are just a start. Much more investigation needs to be done.

5.4.4 The difference between *doch* and DOCH

One small puzzle remains in this view taken of the German modal particle *doch*. What is it exactly that distinguishes the contributions of the stressed and unstressed particles? As opposed to the treatment of unstressed *doch*, scholars almost uniformly agree that stressed MP DOCH indicates direct contradiction between a proposition and its opposite ($p, \neg p$). A unified treatment of these particles is clearly preferable, as their discourse functions are so similar. Under this expectational account, the conventional discourse effect of the stressed DOCH is simply the case that happens when the contrastive, discourse salient p turns out to be the opposite of the prejacent.

Instead of setting up a speaker's expectations for a proposition p and placing them in contrast to an uttered proposition q, focusing the particle sets up a contrast between the truth and the falsity of the proposition itself. Just as with the unstressed version, stressed DOCH commits the speaker to a high expectation for the particle's prejacent. But extra prosodic marking has an effect as well. Focus marking in the style of Rooth (1985, 1992), von Stechow (1996) relates an utterance to a set of alternatives to that utterance. Focus on the modal particle, whose contribution is above the level of the proposition, has the effect of relating the the proposition with its opposite—the only member in its alternative set. Contradiction is then not an inherent property of this modal particle; it is what arises when a speaker's expectations for the prejacent are no longer approximate, and when they wish for the truth of a proposition to be put into contrast with its logical opposite.

Consider again the contexts where DOCH is deemed appropriate. They are situations like (230), where a speaker holds a belief about q which is then proven to be verifiably false. Or they can take the form of situations like (231), in which the speaker and the addressee hold opposing views, and one of them is proven to be unequivocally right:

(230) *Tom tells Sophie that he's going to dye his hair. Later, she sees him with a fresh haircut,*

but his hair color is the same. She says:

Du hast die Haare DOCH *nicht gefärbt!*
you have the hair DOCH not colored
You DIDN'T dye your hair!

(231) *Sophie thinks that Tom can dance, and Steven does not. Later they observe him doing a perfect tango. Sophie, pointing out that she was right to Steven:*

Er kann DOCH *tanzen!*
he can DOCH dance
(You see?) He CAN dance!

The accented DOCH here is masquerading as a separate particle. What we actually see is its interaction with focus. In each case above, the speaker indicates that expectations have been violated in a very particular way. Note that the unstressed *doch* in reponses to these scenarios does not have the same effect: in (230), the negation must receive focus in order to be pragmatically licensed. In (231), a speaker may choose to place focus on the modal, but this emphasizes his ability, not the fact that Steven holds a belief that contradicts reality. What sets DOCH utterances apart from their unstressed counterpart is the focus placement, which forces an actual contradiction. Recall the definition from (232), repeated in (222):

(232) *doch*(q) is anaphoric to a salient proposition or event p in a discourse context C and for discourse-salient participants x, s.t.:

a. *doch*(q) adds the following to the speaker's *DC* list:
$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

When focus marking on the particle intervenes, the salient p is simply the negation of the prejacent. The formulation of *doch*'s conventional discourse effect may stay the same for both stressed and unstressed versions of the particle, but substituting in $\neg q$ in the place of p sets up an actual contradiction. Working through a derivation for (230) looks like (233):

(233) a. Sophie assumes a proposition p is true:
p = *Tom$_i$ dyed his$_i$ hair*

b. Sophie sees Tom, and has visual evidence that her prior assumption was false.

c. She utters DOCH($\neg q$):

$\neg q$ = *You$_i$ didn't dye your$_i$ hair.*

d. $p = q$

e. Add the following to the speaker's discourse commitments:

$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ \text{Exp}_x(\neg q|q) \approx 0$

While the discourse effect of stressed DOCH does not change from this analysis to any other current theory, one beneficial effect of formulating the particles' contribution in terms of expectations is that both the stressed and unstressed particles can be collapsed into a single description. Their difference lies in what each of them consider to be the relevant q in the context. *Doch* is much more permissive, while the stressed DOCH (composed of Focus + *doch*), presents only one possible alternative in the form of the prejacent's logical opposite.

5.5 Previous approaches

There have been many attempts in the past that propose to capture the meaning and functions of the stressed and unstressed *doch*. Of the many proposals, common themes include contradiction to a salient proposition in the context, a marker of known material (to either the speaker or the addressee), re-resolution of a previously closed issue, or as a marker of some unknown material relative to a salient participant. And while many of these analyses are adequate to describe a good deal of the usage cases for the particles, many of the analyses only extend to the uses in declarative utterances. What's more, the performance of the utterances are also left out of many of the current analyses. Just as English discourse particle meaning is influenced by intonational meaning, German modal particles interact with the pragmatic effects introduced by utterance-level prosody. The following sections will outline a few of the previous approaches to *doch* and its contribution to a discourse.

Many previous approaches to pinpointing the meaning of *doch* are tied in to the meaning of another MP as well: *ja*. This particle is thought to be a kind of foil to *doch*, indicating uncontroversiality where *doch* introduces controversy, which marks a proposition as mutually known, wit an extra effect of indicating that this information *should be obvious* (leading to an inference that someone in the discourse has acted in a way that requires this to be brought to attention). In some of the approaches that I will outline below, the meaning of *doch* is taken to be built off of the meaning of *ja*.

5.5.1 Grosz (2014): *doch* marks contradiction, *ja* marks uncontroversiality

A very common theme in much of the literature on *doch* is that the particle acts to signal some contradiction to some other salient proposition in the discourse (Thurmair 1989, Rinas 2007, Diewald 2007, Müller 2014). One such analysis stems from Grosz (2014), who argues that the contradiction element of *doch* arises as a presupposition, and that the context holds a salient focus alternative that contradicts the prejacent of *doch*. This is in addition to the particle also assuming that the prejacent is uncontroversial in the common ground:

(234) For any sentence p, $[\![doch(p)]\!]^c$ (where c is the utterance context) is only defined if:

a. The speaker in c takes p to be firmly established in w_c and it is safe to disregard $\neg p$ as a possible answer to the question of whether p or $\neg p$ holds in w_c.

b. There is a contextually salient proposition q such that:

i. q is a focus alternative of p

ii. The current context c entails $\neg[p \textit{ and } q]$

If defined, $[\![doch(p)]\!]^c$ equals the denotation of $[\![p\ ]\!]^c$.

Like the expectation analysis presented above, Grosz (2014) assumes that there must be a salient proposition that stands in contrast to the prejacent of *doch*. This argument hinges on the fact that *doch* lexically associates with focus. In an utterance context like (235), focus on *Sophie* activates alternatives, which are the corrective and contrastive element of *doch*. In addition, the prejacent of *doch* introduces a contradiction to some other proposition, in this case, the fact that Steven might have baked the cookies under discussion:

(235) Jeff: Steven wasn't home this afternoon.

Tom: *Dann hat* ***doch*** *[Sophie]$_F$ diese Kekse gebacken!*
Then has doch Sophie these cookies baked
Then [Sophie]$_F$ must have baked these cookies!

a. p = Sophie baked these cookies.

b. q = Steven baked these cookies.

c. Assuming that Sophie and Steven didn't bake the cookies together, it must be the case that $\neg[p$ and $q]$.

Grosz (2014) assumes that formulating *doch* in this way bakes in the entire meaning of *ja* into (234a). *Ja* is just a less restrictive version of *doch*, which only carries a presupposition that the speaker feels it is safe to disregard $\neg p$ in the context. Whereas *doch* assumes uncontroversiality as well as contradiction as core meaning components, *ja* only assumes uncontroversiality. An utterance like (236a) with *ja* seeks to establish p as firmly established in the common ground, whereas the corresponding *doch* utterance establishes p as uncontroversial as well as contradictory to some salient focus alternative:

(236) a. *Oprah hat* ***ja*** *keine Kinder.*
Oprah has *ja* no children
(As you know,) Oprah has no children.

b. *Oprah hat* ***doch*** *keine Kinder.*
Oprah has *doch* no children
Oprah obviously has no children.

By marking *ja* as established, participants should be able to conclude that it is safe to discard $\neg p$ as a possible way to resolve the issue raised. In effect, this way of analyzing *ja* has a pragmatic discourse effect of treating a proposition as if it is already established in the common ground, even if the information that proposition contributes is hearer-new. It is also attractive for its treatment of the two particles together, as it situates the two as sharing a single core component, with the more pragmatically rich particle requiring one more specification.

But analyses that require *doch* to contradict a salient proposition in the discourse falter in the face of exchanges like (237):

(237) Steven: These flowers are beautiful!

Tom: *Diese Blumen sind* ***doch*** *traumhaft schön!*
these flowers are doch dreamlike beautiful
These flowers are gorgeous!

In (237), though Tom appears to be agreeing with Steven's utterance, he uses *doch* in order to upgrade the degree to which he views the flowers to be beautiful. Under the analysis in (234), we do not predict this behavior. Though Tom's contribution is a salient focus alternative of Steven's utterance, the two alternatives are not contradictory. Both *p* and *q* can be true at the same time. *Doch* can instead be used to pick out a contrast between the perceived degree of the gradable predicate. This use is predicted from an expectation-based account of *doch*, where the contrast lies in speaker expectations about other discourse participants behaviors and expectations, not between the truth and falsity of individual propositions.

Rojas-Esponda (2013) notes that the same contrastive-but-not-contradictory distinction is relevant for question-answer pairs as well. In particular, the formulation of *doch* in (234) does not hold in responses to polar questions. In such cases, the highlighted alternative should be unavailable as a focus alternative, but can indeed appear with *doch* in an answer to a question:

(238) **Sophie:** Is Jeff going to climb Mont Blanc?
Tom: *[Den Berg]$_F$ hat er* ***doch*** *bestiegen.*
the mountain has he *doch* climbed
He has climbed that one (already).

Again, as with the flower example, though, as Grosz predicts, the *doch* utterance is a focus alternative of the highlighted alternative, it is not one that contradicts it, or that would necessarily entail ¬[*p and q*]—for all we know, Jeff is going to climb Mont Blanc again. Analyses that assume that *doch* must require contradiction at the propositional level cannot easily be extended to non-contrastive cases like those in (238).

5.5.2 A Question Under Discussion account of *doch*: Rojas-Esponda 2013

In response to analyses that presuppose that *doch* requires a contradiction introduced by focus alternatives of a salient *p*, Rojas-Esponda (2013) introduces a QUD account of the MP. This is in part motivated by non-contrastive examples of *doch* like in (237), and more clearly, by minimal pair examples like (239) and (240), which introduce particular question-answer patterns with *doch* that a contradiction account would not predict. This contrastive-but-not-contradiction distinction holds here as well. When *doch* responds to a question, such accounts predict that the contrastiveness arises between the highlighted alternative of a question and the prejacent of *doch*. This works well with a question-answer pair like in (239). However, the *doch* response in (240), though, is perfectly well formed, is not predicted by the contradiction accounts:

(239) Jeff: Are you coming with us to the Opera?

Tom: *Nein, ich habe* ***doch*** *abgesagt.*
No I have doch cancelled

No, I cancelled.

(240) Jeff: Are you coming with us to the Opera?

Tom: *Ja, ich habe* ***doch*** *zugesagt.*
yes I have doch confirmed

Yes, I said I was coming.

In these examples, contradiction to a salient *q* appears to falter as a core meaning component of *doch*. Rojas-Esponda (2013) instead argues that the interpretation of *doch* arises from the specific ways that the speaker chooses to answer the current QUD. Unstressed uses of *doch* can appear as contrastive or non-contrastive as in the cases above, but the core function of the particle is to indicate something about how the answer relates to the more general conversational context. Unstressed instances of *doch* are predicted to mark an utterance as previously answered somewhere in the discourse. Stressed uses of the MP *doch* are purely contradiction of a previous issue on the discourse record. Rojas-Esponda (2013) formalizes the contributions of the particles in the following way:

(241) ***doch*** marks the current QUD Q as previously closed in one of the following ways:

a. signals Q was previously resolved (C is contained within a single cell of Q)

$$\forall v, w \in W, (v \in C \wedge w \in C) \rightarrow \langle \text{v,w} \rangle \in Q$$

b. signals Q was previously shown invalid (the presupposition of Q does not hold)

$$C \cap \hat{Q} = 0$$

Additionally, the stressed and unstressed particles are licit in the following situations, which make reference to cells in a set of worlds that has been partitioned into p and $\neg p$ worlds:

(242) a. Unfocused *doch*(p) is used when the cell containing C properly contains p:

$$\forall w \in C\ [C \subseteq c_Q(w) \wedge p \subsetneq c_Q(w)]$$

b. Focused DOCH(p) is used when C and p pick out different cells:

$$p = \bigcup_{w \in S \subset \hat{Q}} \quad \text{and} \quad C \cap p = 0$$

When a speaker uses the unstressed *doch*, she is indicating that the current issue under discussion has already been resolved, and re-answering it with *doch* indicates that its resolution should be the same as before. This is the pragmatic information that the addressee computes in (239) and (240): Tom is in effect reminding Jeff of how he has previously resolved the issue of whether he is joining for the Opera. With stressed DOCH, the speaker indicates that there must be some revision to the discourse record; he has had a change of plans and cannot join:

(243) Jeff: You're coming to the Opera, right?

Tom: *Tja, Ich kann* DOCH *nicht mitkommen.*
Well I can doch not come-with

Well, I can't come after all.

Perhaps the biggest take away from this analysis is the fact that *doch* signals common knowledge between all discourse participants. Using this MP indicates that an issue that is currently under discussion is already logged in some accessible conversational scoreboard. The differences between the two particles is straightforward: one signals an identical re-resolution of an issue, the other signals a revision.

However despite the clear merits of this approach, the QUD analysis does bake in the fact that *doch* issues are re-raised issues. This fact requires there to already be an answer to the QUD entered into the common ground. But one of the peculiarities of *doch* is that is can be an out-of-the-blue reaction by a speaker to an event or observation:

(244) *You are sitting in the grass with a friend, and all of the sudden see something out of the ordinary. You say:*

*Das ist **doch** ein vierblättriges Kleeblatt!*
that is doch a four-leaf clover

That's a four-leaf clover!

In such situations, the unstressed particle is licit, yet the issue is not a previously closed QUD (and arguably only a QUD in virtue of the issue raised with the *doch* utterance). On an interpretation that the re-raised QUD could be something like a general assumption about the world (e.g. *one does not generally find four-leaf clovers*), this would predict that the stressed particle, rather than the unstressed particle should be acceptable here. This is not what we find.

What's more, for both the stressed and the unstressed particle, there need not be an explicit answer to the QUD accessible in the common ground. For the scenario in (245), it seems that *doch* reacts to the likelihoods of particular outcomes given the evidence available, not to an explicitly answered QUD:

(245) *Two doctors are discussing the possible diagnoses for a patient. They are listing the symptoms that the patient is experiencing in order to narrow down the possibilities*:

Doctor A: The patient says she has been experiencing stomach pain all day.

Doctor B: According to her charts, she has low blood pressure and nausea.

Doctor A: I've run some tests for infectious diseases and they all came back negative.

Doctor B: The chart here says that she also has a low white blood cell count.

Doctor A: *Es könnte* DOCH***/doch*** *Krebs sein.*
it could DOCH/*doch* cancer be
It could indeed be cancer. / It could (just) be cancer.

It seems that the function of *doch* in these cases is to activate possibilities that each doctor privately has access to, but may not be taking into account in the current context. A QUD account is slightly too restrictive to capture such examples.

A more concerning aspect of a QUD account of *doch* is the correct treatment of the particle when it appears in questions and in imperatives. As it stands, the QUD analysis would predict that using *doch* in a question would indicate that an issue raised with the question were already answered, and further, answered in a particular way. In reality, questions with *doch* do indicate a bias for a particular way to answer the QUD, but they need not indicate that this issue has been previously settled. Consider (246), repeated from (213a):

(246) *Du hast* ***doch*** *Sekt mitgebracht, oder?*
you have *doch* sparkling-wine brought-with or
You of course brought sparkling wine, right?

Though the speaker indicates that she has a bias for the highlighted alternative, her tag *oder* indicates that this is only a bias—it has not been fully resolved. This conflicts with the contextual requirements for *doch* as outlined in (242a), which places the prejacent of the the unfocused particle squarely in one cell or another of a partitioned world space, co-occuring with the common ground. This seems to predict that *doch* may only be used in contexts where the speaker is sure of her contribution. But this does not leave room for those situations where the speaker only *expects* that her contribution can be taken as common ground.

5.5.3 Kaufmann & Kaufmann 2012

Kaufmann & Kaufmann (2012) try yet another approach to the problem of capturing the meaning component of *doch* and *ja*. They, like Thurmair (1989), note that the two particles are extremely similar in their functions, but that *doch* is the less restrictive of the two, occurring in more possible sentence types than its counterpart *ja*. Even with this restriction, the two particles share the observation that both commit the speaker to the belief that *p* is in some sense given, obvious or uncontroversial.

They propose that the difference in meaning between *doch* and *ja* lies in the presuppositional component of the particles. Under this view, *ja* presupposes normality, while *doch* has a presupposition of abnormality. Their proposal is below (p. 212):

(247) **Level of Assertion:** both *ja*(p) and *doch*(p) are equivalent to p

(248) **Level of Presupposition:** In addition, the speaker is committed to the belief that the following is in the common ground:

a. both *ja*(p) and *doch*(p): normally in a situation like c, any rational agent whose goal is to find out whether p, does find out whether p (from information already available or in the immediate surroundings). ***uncontroversiality***

b. i. *ja*(p): c is *normal* in the sense of (248a). ***normality***

   ii. *doch*(p): c is *not normal* in the sense of (248a). ***abnormality***

Formally, they define **uncontroversiality**, **normality** and **abnormality** in the following terms, invoking a circumstantial modal base $f^c(w)$ and an ordering source g^n:

(249) p-**uncontroversiality**

$$\Box^{CG_{s,h}}[\Box^{[f^c,g^n]}\forall\alpha\ [\Box^{T_\alpha}[\Box^{B_\alpha}p \leftrightarrow p] \rightarrow [\Box^{B_\alpha}p \leftrightarrow p]]$$

(250) h/p-**normality**

$$\Box^{CG_{s,h}}[\Box^{[f^c,g^n]}\varphi \rightarrow \varphi \text{ where } \varphi = \forall\alpha\ [\Box^{T_\alpha}[\Box^{B_\alpha}p \leftrightarrow p] \rightarrow [\Box^{B_\alpha}p \leftrightarrow p]]$$
$$\therefore \Box^{T_h}[\Box^{B_h}p \leftrightarrow p] \rightarrow [\Box^{B_h}p \leftrightarrow p]$$

(251) h/p-**abnormality**

$$\Diamond^{CG_{s,h}}[\Box^{[f^c,g^n]}\varphi \wedge \neg\varphi] \text{ where } \varphi = \forall\alpha\ [\Box^{T_\alpha}[\Box^{B_\alpha}p \leftrightarrow p] \rightarrow [\Box^{B_\alpha}p \leftrightarrow p]]$$
$$\therefore \Diamond^{CG_{s,h}}\ [\Box^{T_h}[\Box^{B_h}p \leftrightarrow p] \wedge\neg\ [\Box^{B_h}p \leftrightarrow p]\ ]$$

Doing this effectively treats *ja*-marked utterances as if they are common ground, while treating *doch*-marked utterances as information the hearer is evidently not aware of (or not taking into account). This formulation is especially important in their account of the distribution of *doch* and *ja* imperatives; *ja* is not allowable because its presuppositions (uncontroversiality, normality) clash with the general conditions required of imperatives. Under the analysis of Kaufmann (2012), the prejacent of imperatives must meet an *Epistemic Uncertainty Condition* (EUC), requiring the speaker to keep open some future paths that lead to p, and some which lead to $\neg p$. For an imperative, the speaker must always leave room for the addressee not to comply with their request. With these two pieces, interpreting an imperative with *ja* would lead to contradiction: the hearer, under the normality condition required of *ja* (250), would inherently clash with the EUC: one cannot ensure that p will be brought about while also leaving room for the EUC. On the other hand, *doch* would be licensed under the abnormality condition in (251), and would not clash with the inherent possibility left open by the EUC. The inherent possibility in both conditions on imperatives and on *doch* utterances predicts that the particle will be good in imperatives.

But there are a few oddities that arise from some of the assumptions that are made under this view. Assuming that *ja* marks elements that are common ground and verifiable from the uncontroversiality and normality conditions predicts that exchanges like (252) should be good, when in fact they are quite odd:

(252) Jeff: My arm is broken. I'm going to the doctor later.

Sophie: *#Ja, du sollst* ***ja*** *zum Arzt gehen.*
Yes you should ja to-the doctor go

Yes, (as we know) you should go the doctor.

Jeff has raised an issue, and Sophie, with the polarity particle *Ja*, has agreed to make it common ground. Yet repeating this proposition with the MP *ja*, though it has just been made common ground, is pragmatically odd. There has to be some distance associated with *ja*'s being in the common ground, and the time that the speaker has uttered the *ja* utterance. This is part of *ja*'s conventional discourse effect: reactivation of a salient q. Since the q here needs no reactivation, it is anomalous. [9]

Setting this temporal anomaly aside, the analysis also makes predictions about the particles in other sentence types. For *ja*, it makes the correct prediction that the particle is not allowed in questions, and for *doch*, though some more explanatory work would need to be worked out about the how these conditions would apply to sets of alternatives, the abnormality condition could be argued to account for the particle's distribution in question environments. As shown in (5.2.2.2), *doch* is only allowed in biased questions, and wh-questions with rhetorical effect. Assuming that the highlighted alternative could serve as the proposition a speaker favors to be a correct resolution of the question, one could imagine an analysis that committed the speaker to believing the abnormality condition, just in those cases where she suspects that there

[9] Thanks to Ramona Wallner for pointing this out, and to Arno Goebel and Gisela Grohne for discussion.

is a possibility that the addressee holds an inconsistent belief state that conflicts with what she assumes is true about the world.

While the system proposed by Kaufmann & Kaufmann (2012) seems to work for most of the cases here, the pragmatic effect of these particles is hard-coded into a speaker's beliefs. Though the system correctly predicts the distribution of both *doch* and *ja*, it does so using some unnecessary machinery. The account is predicated on the particles' particular behavior in imperatives, and requires the particular account of imperatives and modality built in Kaufmann (2012). Of course, one benefit of using this theoretical framework is the logical accuracy that it provides. Yet one could make an argument based off the Law of Parsimony: the solution here is far from the simplest explanation.

The mechanism for the interpretation of *doch* argued for in Section §5.4.2 is an argument for simplicity. I propose that the driving force behind the interpretation of *doch* is *speaker expectation*. And as I will show in the following chapter, the difference between *doch* and *ja* lies not in inferences calculated from an uncontroversiality belief, but in whether the speaker's expectations about the context have been violated. This is a system that is framed in a theory of conversation which assumes a causal model underlying structure, but which could easily be adapted to other more traditional theories of discourse representation. The basic intuition of the particles is captured in the following way: when a speaker uses *doch*, she is inherently signaling something about her own expectations as well as the expectations she has assumed play a role in an addressee's behavior. *Doch* says "given q, we should not expect p," and places these expectations into view for everyone. Doing this pressures an addressee to respond: she feels obligated to help navigate the situation toward a resolution. As will be presented more thoroughly in the chapter to follow, *ja* simply makes clear that a speaker commits to the fact that her expectations and those of her addressee will align with respect to a salient proposition. Though this account and the Kaufmann & Kaufmann (2012) proposal explain the same data, I argue that an expectational account is the more intuitive solution.

Forming the effects of these particles in terms of expectations also allows the observations about mirativity to fall out cleanly from the account. As markers of surprise, miratives should be anchored to the expectations of a speaker. Under this system, *doch* sets up a clear juxtaposition between a speaker's expectations, the actual state of the conversation, and common ground. It is clear why *doch* is used in many cases where surprise is a clear inference to draw: the particle inherently marks violated expectations.

5.6 Summary: a step back

This chapter has sought to do three things. Its primary goal was to provide a comprehensive analysis of the German modal particle *doch*. Part of the difficulty in this task lies in keeping the descriptive insights of previous attempts intact while also reformulating the way in which modal particles update a conversational context. This chapter proposed a meaning for *doch* which sought to do just that, while also providing a definition of the particle that could account for its behavior in all sentence types. In addition, the pragmatic discourse effect of the unstressed particle, coupled with simple focus marking, can unify the apparent differences shown by unstressed *doch* and stressed DOCH; DOCH is simply an instance of the unstressed particle where focus forces a contrast in sentence polarity rather than with another salient proposition or event.

This characterization of *doch* as a marker of expectation violation brings us to the secondary goal of this chapter, which was to place this particle into the larger category of mirative strategies. Like the English discourse particles *oh* and *huh*, *doch* grammatically signals a complex emotional state that a speaker finds herself in. As actors move through a conversation, they gather knowledge and form beliefs based on this knowledge that can drive their expectations for particular conversational outcomes forward. These expectations are malleable, and can change and be reevaluated when a speaker encounters information that may seem to challenge their beliefs. Reevaluating expectations and beliefs is a core component of mirative strategies, and it is an integral part of the discourse contribution of *doch*. This mirative element distinguishes *doch* from its oft cited counterpart, *ja*. Yet being able to formulate both *doch* and *ja* in terms of a speaker's expectations keeps the intuitions of previous work, which calculate the meaning of *doch* as an extra inference built off of the core contribution of *ja*. Both particles express a speaker's chain of reasoning at the time of utterance, the difference being that one communicates violated expectations, the other mutual belief.

The final goal of this chapter has been kept relatively covert, but becomes clear upon further reflection. One of the main goals of this dissertation is to look at a language's different ways of expressing mirativity. In English, both utterance-initial discourse particles as well as intonational contours were shown to be markers of mirativity. So far for German, the sentence-medial particle *doch* has also been shown to be a member of this category. In a striking way, the English SRC and this unstressed MP *doch* appear to perform the very same function in their respective languages. Both mark speaker surprise, and both rely on the same type of clash in expectations between the speaker and the addressee. Both function at the level of illocution, and both pragmatically prompt the addressee to respond. The difference here is that English prefers to communicate this information on the intonational level, while German's preferred tactic is with a semantically bleached but pragmatically rich particle. Compare the conventional discourse effects of the two strategies below:

(253) *doch*(q) is anaphoric to a salient proposition or event p in a discourse context C and for discourse-salient participants x, s.t.:

a. *doch*(q) adds the following to the speaker's *DC* list:
$\mathrm{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\mathrm{Exp}_x(p|q) \approx 0]$

(254) The English Surprise Redundancy Contour (H) L* H*-L% is anaphoric to a salient proposition or event p in a discourse context C, and is admissible for discourse-salient participants x when

a. q is the proposition expressed by the speaker (uttered content or the presuppositions introduced by a question),

b. add the following to the speaker's *DC*s:
$\mathrm{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\mathrm{Exp}_x(p|q) \approx 0]$

Admittedly, part of the motivation for an expectational view of *doch*, and modal particles more generally, is out of a desire for coherence between similar strategies cross-linguistically. Intonation is thought to contribute information about a speaker's commitments to propositions in a discourse. In Chapter §3, I reframe these commitments to be more precise, and talk about how they commit a speaker to some *expectation* for a proposition's truth in a discourse. Intonational idioms like the SRC further modify a speaker's commitments, commenting on what the speaker believes of herself, and what she expects of her addressee. German *doch* has this same effect of modifying a speaker commitment to a proposition through their discourse expectations. This chapter was an exercise in extending this way of thinking to a related realm with many disparate analyses in an attempt to bring together the core generalizations observed by previous researchers.

Chapter 6

Bits and Pieces: German *ja*, prosody, and a question of salience

This chapter is a bit of a mixed bag, and combines a few themes that address some open issues from the previous chapters. First, it provides an analysis of *ja* in terms of speaker expectation, in parallel with the analysis of *doch* presented in Chapter §5. Since both particles have been given a dual treatment in much of the literature to date, giving an accurate account of the conventional discourse effects of this particle is a test of this combined speaker expectation and Table model system. The second portion of this chapter looks at how the interpretation of German modal particles can interact with utterance-level prosody, in an attempt to tease apart the pragmatic contributions that may intervene from both paradigms. This is just the beginning of an investigation on the scale of that in Chapters §3-4, but touches on how a thorough investigation of these facts might be done in a language with clause-internal particles. Lastly, this chapter addresses the notion of salience, which has been a theme throughout this work, and how we might explain what it means to be a proposition or event that is "salient in the discourse."

6.1 The modal particle *ja*

Many analyses of *ja* come hand in hand with an analysis of *doch*. This is not without good reason: in many respects, *doch* seems to be the foil of *ja*. Whereas *doch* marks an imbalance in speaker and hearer expectations, *ja* appears to mark equilibrium toward a proposition in the discourse: all participants agree that a proposition is uncontroversial. To capture this, many previous account have built the discourse effect of the two particles off of each other. Grosz (2014) assumes that the meaning of *ja* is the basis for the meaning of *doch*: *ja* marks that a proposition has already been established in the common ground, while *doch* triggers a further presupposition of contradiction. This account is similar to that of Kaufmann & Kaufmann (2012), in which *ja* and *doch* differ only in terms of presuppositions of relative *normality* and *abnormality* (see Chapter §5.5.3 for a detailed account). In the following sections, I attempt to frame the differences between *ja* and *doch* in terms of the *expectations* that they place in the discourse commitments of a speaker. *Ja* signals that a speaker's expectations surrounding the discourse have followed a pre-assumed course, while *doch* indicates that expectations have somehow been violated.

6.1.1 The use of *ja*

German *ja* is traditionally taken to mark information that is known and uncontroversial to both the hearer and the speaker. In the feature-driven system developed by Thurmair (1989), *ja* is given only one of the many attributes: <KNOWN>. When Watson uses *ja* in (255) below, he is signaling just this to Holmes: the content of his utterance should be taken as uncontroversial.

(255) *The detective and the doctor hear loud banging on the wall next door.*

Holmes: *Hören Sie das auch?*
hear you that also
Do you hear that, too?

Watson: *Ich bin* ***ja*** *nicht taub.*
I am *ja* not deaf
I'm certainly not deaf.

When used as an answer to a question, the speaker appears to signal that this is the only valid answer to consider—in other words, it is a signal that it is safe for the addressee to reject any alternative propositions as answers to the question (Grosz 2012). But this MP need not only occur as an answer to a question. It can also be used in situations where the speaker is pointing out information that she thinks should be clear to anyone in the conversation. Imagine that Tom and Sophie have been asked to feed their neighbor's cat. It's Friday afternoon, and Tom calls Sophie with a suggestion:

(256) **Tom:** Let's go get dinner straight from work tonight.

Sophie: *Wir müssen* ***ja*** *zuerst die Katze füttern.*
we must *ja* first the cat feed
We have to first feed the cat, you know.

The use of *ja* serves as a sort of reminder to Tom, as well as a signal that Sophie is not committing to, but rather negotiating his contribution to the common ground. The particle in Sophie's response marks her utterance as uncontroversial, and indicates that she views her contribution as uncontroversial. As this contribution conflicts with Tom's suggestion, Tom takes this as an indication that Sophie rejects his proposal in view of the issue she has just raised. Kratzer & Matthewson (2009) propose that utterances like this one form the core meaning of *ja*, and suggest that *ja* should be assigned the following semantics:

(257) For any sentence α, $[\![\, ja\ \alpha\,]\!]^{c,g}$ is only defined if:

a. the speaker in c takes $[\![\alpha']\!]^c$ to be firmly established and therefore

b. doesn't consider the question $\lambda w[\![\alpha']\!]^c(w) = [\![\alpha']\!]^c(w_c)$ to be an issue for inquiry in c or after c.

If defined, $[\![\, ja\ \alpha\,]\!]^{c,g} = [\![\, \alpha\,]\!]^{c,g}$

This formulation of *ja* can capture its behavior in contexts where the prejacent is inferable from other information as well. Sophie's answer to Tom could have been less direct, yet still recoverable from the fact that they share the responsibility of feeding the cat:

(258) **Sophie:** *Wir müssen* ***ja*** *zuerst nach Hause.*
we must *ja* first to house
We have to go home first, you know.

The use of *ja* here makes it clear that the particle can place the speaker as committing firmly to the prejacent, by way of some other relevant pre-established fact. Using *ja* here is a way for Sophie to 'reactivate' this knowledge that she and Tom share. It follows from *We have to feed the cat* that in order for this task to be done, they must be in the same physical location as the cat. What Sophie is indicating to Tom is something like *since we both know that we have to feed the cat, it must also be true that we have to go home in order to accomplish that. Ja* seems to have the added ability to mark a proposition other than the prejacent as uncontroversial.

Under the analysis in (257), Sophie's utterance is conditioned off of the fact that she as the speaker considers her contribution to be uncontroversial in the discourse context. Though this proposition has not been explicitly added to the common ground, based off of other relevant details, she can take this fact as "firmly established"[1] at least as far as she is concerned. But formulating this only in terms of the speaker seems a bit too restrictive—this would predict that *any* proposition a speaker might consider firmly established could be indicated as such with *ja*. As a cultural anthropologist, Sophie could take linguistic determinism as firmly established, while Tom might have exactly the opposite stance. Yet though Sophie has this firm belief, she cannot indicate this with *ja* in a conversation with Tom: in doing so, she seems to imply that he also has this belief:

(259) *Sophie, an anthropologist, to Tom, a linguist:*

Die Sapir-Whorf-Hypothese ist ***ja*** *völlig richtig.*
the Sapir-Whorf-hypothesis is *ja* completely right

(As we both know), the Sapir-Whorf hypothesis is completely right.

Yet the information that *ja* marks as uncontroversial need not have been explicitly negotiated by all participants in every context. What (259) shows is that there must be some acknowledgment of an addressee's beliefs as well. In contexts where the topic of discussion is common or world knowledge, *ja* is completely felicitous, even when the speaker and the addressee have not explicitly discussed the topic. In a context where one can assume that interlocutors are familiar with the events that lead up to World War II, the speaker might use

[1]Kratzer & Matthewson (2009) use the term "firmly established" but specify that this need not mean common ground.

ja as in (260) to signal that the fact that Hitler rose to power in 1931 is not new information to anyone in the current context. But this same utterance would be infelicitous if were made by a teacher to a class just learning about the pre World War II-era.

(260) *1931 war Hitler* ***ja*** *noch nicht an die Macht.*
1931 was Hitler *ja* yet not on the power
(As we know,) In 1931, Hitler had not yet come to power. *Thurmair (1989), 105*

In contrast, not all historical or world events can be marked by *ja*. While the timeline of World War II can be taken as common knowledge in many circles, the course of events of smaller conflicts or other cultural developments are infelicitous with *ja* outside clear contextualization. If you are not a scholar or history buff of the Bosnian War, (261) is anomalous; if you are not intimately familiar with Eurovision Song Contest winners, the speaker of (262) falsely assumes shared speaker and addressee common ground:

(261) *(#) Goražde war* ***ja*** *während des Krieges eine UN-Sicherheitsrat erklärte*
Goražde was *ja* during the war a UN-Security-Council designated
Schutzzone.
protected-zone
Goražde was declared a protected zone by the UN Security Council during the war.

(262) *(#) ABBA hat* ***ja*** *mit "Waterloo" den Eurovision-Wettbewerb gewonnen.*
ABBA has *ja* with Waterloo the Eurovision-Contest won
ABBA won the Eurovision Song Contest with "Waterloo."

Thus, *ja* need not point out an explicitly common ground proposition; it need only mark it as generally known and agreed upon by all participants in the relevant discourse context.

The particle can also be used to introduce a topic that the speaker is not sure the hearer has information on already, but would like to mark as possibly known:

(263) **Sophie:** *Du kannst dich noch an Jeff erinnern, ne?*

You remember Jeff, right?

Tom: *Ja, so 'nen großen, blonden? War eine Weile in Japan?*

Yeah, he's like, tall and blonde? Was in Japan for a while?

Sophie: *Genau. Er hat* ***ja*** *nach Japan ein Kätzchen adoptiert.*

Exactly. You know, after Japan he adopted a kitten.

Tom: *Ach, echt? Das wusste ich nicht.*

Oh really? I didn't know that.

Sophie: *Ach so. Ja! Eine kleine graue. Aber, genau— er kommt Morgen zu Besuch.*

Oh. Yeah! A little grey one. But right— he's coming for a visit tomorrow.

In a conversation like (263), *ja* is an attempt by the speaker to get the hearer on the same page without having to discuss the issue at hand. Sophie's main point with her *ja* statement is not to tell Tom that Jeff has adopted a kitten—it's information that she uses to get Tom up to speed as a sort of 'by the way' marker. Hentschel (1986) argues that uses of *ja* that seem to introduce hearer-new information are part of the inherent function of this MP, specifically, to emphasize (friendly) familiarity or closeness between the speaker and the hearer. In (263), *ja* effectively treats the prejacent as material that the speaker would like to introduce to the hearer as important in the conversation, but not something that is either the main point of the conversation turn, or a point that they are necessarily keen to dwell on. She acts as though the information was already known in order to move on to a new topic.[2]

But rather than simply politeness markers, uses of *ja* in situations like (263) seem to actually function as a way for a speaker to exploit the use of the particle as a marker of shared knowledge. In a case like (263) where the speaker takes the liberty of assuming information that is not widely shared is common ground, listeners are able to negotiate the particle's contribution, by either explicitly or implicitly calling it into question (Thurmair 1989). Notice that this is what Tom does with his response to Sophie's *ja* utterance above. In this case, Sophie notes this discrepancy and elaborates a bit further before moving on to her main point. But notice, too, that when a speaker is sure that her information is shared knowledge, it is anomalous for her addressee to call *ja*'s contribution into question:

(264) *Tom, Sophie, and Jeff carpool to work in the mornings. Tom drove this morning, and*

[2]Note that the English paraphrase of the relevant portion of (263) achieves a similar thing: the speaker does not know if the hearer has access to the information that she does, yet marks this with 'you know', a string whose semantics seems to indicate specifically that the hearer *does* know what is about to come next, though the speaker might *not* know for sure about the hearer's knowledge.

at the end of the day, the three are walking out to the parking lot together. Jeff starts walking to the right, but Sophie motions to the left:

Sophie: *Er hat* ***ja*** *da hinten geparkt.*
he has *ja* there back parked

He parked back there.

Jeff: *# Ach echt? Das wusste ich nicht.*
oh really that knew I not

\# Oh really? I didn't know that.

It is also important to note the environments in which *ja* cannot occur. Unlike *doch*, which can appear in a variety of sentence types, unaccented *ja* is used exclusively in declarative sentence types. Taking the impressionistic view of *ja* into account, this is perhaps unsurprising. If the conventional discourse effect of an interrogative is to raise the issue of whether a proposition is true, then a particle whose purpose is to signal information known to all parties should be ill-formed in an environment that raises unresolved issues. This is exactly what we find: marking a question with *ja* mixes signals. It indicates uncontroversiality with a MP, but does so in a construction that is meant to introduce uncertainty. Such a use is anomalous. It is also similarly barred in rising declaratives (265b):

(265) a. *# Hast du* ***ja*** *Pflaumen mitgebracht?*
you have *ja* plums brought
Intended: Did you bring plums?

b. *# Du hast* ***ja*** *Pflaumen mitgebracht?*
you have *ja* plums brought
Intended: You brought plums?

Unstressed *ja* is also infelicitous in imperatives and exclamatives, though a phonologically similar, stressed particle JA is not. Thurmair (1989) categorizes this as a "strengthening" particle—one which simply intensifies the force of the utterance. JA is limited to imperative sentences, and sentences modalized with deontic uses of *sollen, wollen* or *dürfen*, English 'should', 'want' and 'be allowed to.'[3] This particle is only anchored to the speaker's desires, and as opposed to unstressed *ja*, does not as clearly mark an utterance as uncontroversial to discourse participants. As in (266), it can signal that the imperative is a warning or a threat, but as in (267) it can also simply reaffirm necessity as it relates to a speaker:

(266) *Mother to child:* *from Thurmair (1989), p.109*

Komm JA nicht zu spät heim!
come JA not too late home

Don't (you dare) come home too late!

(267) *Ich darf JA meine Autoschlüssel nicht vergessen!*
I permit JA my car-keys not forget

(Of all things) I can't forget my car keys!

We set these uses of the stressed particle JA aside for now, but return to them in Section §6.1.3.

The use of the unstressed *ja* so strongly signals shared knowledge between the speaker and the hearer that some argue that sentences that contain the particle **must** be interpreted as shared information. Take for instance, the pair of sentences below:

(268) a. *Ich habe neulich die Inge getroffen. Sie hat ihre Dissertation fertig!*
I have recently the Inge met she has her dissertation ready

I just ran into Inge. She's done with her dissertation!

b. *Ich habe neulich die Inge getroffen. Sie hat* ***ja*** *ihre Dissertation fertig!*
I have recently the Inge met she has ja her dissertation ready

I just ran into Inge. (As we know,) she's done with her dissertation! *Thurmair (1989)*, p. 108

If the speaker hasn't seen Inge in a while, and has just run into her on the street and heard about her progress, she will be inclined to use (268b) when relaying this message to accommodate the possibility that her interlocutor has already been informed of the good news. But if the speaker has just run into Inge in the copy room, with the finished dissertation coming out page by page from the printer, it would be very strange for her to utter (268b) to

[3] Note that these modals impart a deontic flavor to their utterances, and can be interpreted in a very similar way to imperatives. Both sentence types tend to impose the desires of the speaker onto their interlocutors.

an addressee that she meets directly afterward in the hallway. There is no way that addressee could have known that Inge's dissertation is already finished, and no reason for the speaker to see the need to either accommodate this for the addressee, or to ask the addressee to accept this information as anything but new. (268b) in this context is pragmatically odd. One question to raise here is why a speaker would feel the need to share already publicly known material with their interlocutor. As it stands, *ja* seems to be an overt marker that signals a violation of Grice (1975)'s Maxim of Quantity. Marking something as known or uncontroversial to all conversational participants is giving more information than is strictly necessary. But having an overt marker of Gricean conversational maxim violation could be helpful in managing a discourse. Being able to pragmatically signal agreed upon material allows the hearer to not have to move to accept or reject the proposition at hand. The information is simply being used to activate particular themes, and can be used as a basis for building common ground on other issues.

In some cases, a speaker can use this overt signposting of a Maxim of Quantity violation to her advantage. Because the pragmatic effect of this particle so strongly signals that information is shared, *not* using the particle amounts to raising an issue that is new information. But since the language has a dedicated particle that marks known material, by using this in situations where the speaker is not certain of her addressee's knowledge state, the particle becomes something of an active backgrounding strategy. Like in (263), *ja* can introduce material that is new, but that the speaker wants to treat as if it were old, or in some way less informationally crucial to the issue at hand. The addressee is faced with a choice: either take this in stride and accommodate the speaker's utterance as if it were known, or make a move that corrects the implied contribution of *ja*. Like *oh* or *huh* in English, *ja* is linked to two propositions in a discourse: the utterance in which it occurs, and another utterance that it is a reaction to. If there is no utterance that it is reacting to, the hearer has a right to negotiate this. As Kratzer & Matthewson (2009) point out, for *ja*, "there must always be some suitable connection to some salient fact."

Lindner (1991) notes that a speaker using *ja* assumes that she will not be contradicted upon a thorough assessment of the context. Assuming this asserts something about *ja*'s effect on the discourse as a whole: that the particle is used to modify illocutionary types in a particular way. Normally when a speaker makes an assertion, the content of that utterance is raised as an issue that can be freely accepted or rejected by others in the conversation. With *ja*, as we have seen, the speaker both asserts the content of her utterance, while effectively marking the content as a pre-settled issue. The same is true for the contrast in (269):

(269) **Sophie**: Today is Tuesday and Jeff still hasn't replied. I wonder if he got my letter? I sent it off last Wednesday.

a. ***Tom**: Dann hat er ihn am Freitag gehabt.*
then has he it on-the Friday had
Then he'll have gotten it on Friday.

b. ***Tom**: Dann hat er ihn **ja** am Freitag gehabt.*
then has he it *ja* on-the Friday had
Well, in that case, he got it on Friday. *Modified from Lindner, p.170*

If Tom uses the variant in (269a), he assumes that Sophie will take him at his word, but that there is still a chance that he might have to back up his claims. But if he uses (269b), and Tom and Sophie both know that the post takes a maximum of two days, there will be no felicitous way for Sophie to contradict Tom. She can agree with him, but she cannot question or contradict this[4]:

(270) **Sophie**:

a. *Ja, richtig.*
Yes, that's right.

b. # *Ach so, echt?*
Oh, really?

c. # *Nein, das glaube ich nicht.*
No, I don't believe that.

[4] Of course, if Sophie doubts the reliability of the mail carriers, or if the two have conflicting views on postal service, (b) may not be used. But on the assumption that the two share common background on this issue, (b-c) are infelicitous.

In fact, introducing explicit doubt as to whether an addressee agrees with the speaker renders the contribution of *ja* somewhat ungrammatical. Notice the contrast in felicity between (271a) and (271b)[5]:

(271) *A professor addresses the class at the start of lecture:*

a. *Letztes mal habe ich* ***ja*** *erzählt...*
last time have I *ja* explained
Of course, last time I explained...

b. $^{?}$*# Ich weiss nicht, ob Sie sich daran erinnern, aber letztes mal habe ich* ***ja*** *erzählt...*
I know not if you self on-it remember but last time have I *ja* explained
$^{?}$# I don't know if you remember, but last time I of course explained...

While it is common for a professor to assume and review some common ground information at the beginning of a lecture as in (271a), explicitly calling into question whether the students remember this information before going on to mark it as common ground is odd. It must be the case that using *ja* builds in a speaker assumption that the addressee has access to the information that is currently the topic of discussion. Kaufmann & Kaufmann (2012) notice a similar pattern based off of examples using *ja* with ostensibly addressee-new information, but which is directly perceptible to anyone in the discourse:

(272) *Speaker notices that the addressee has spilled coffee on her shirt:*

Du hast ***ja*** *gekleckert.*
you have *ja* spilled.
You spilled, you know.

(273) *Addressee walks in with a gash in her right leg, and it's bleeding:*

Du sollst ***ja*** *zum Arzt gehen!*
you should *ja* to-the doctor go
You really should go to the doctor!

From this, the speaker uses *ja* to modify her own attitude toward the proposition, placing the speaker in a position of knowledge given what she thinks is reasonable evidence to

[5]Thanks to Ramona Wallner for pointing this out.

all the participants in a discourse. Thus, *ja* can be used as a marker of uncontroversiality in the context, but this uncontroversiality need only be licensed from the point of view of the speaker.

In sum, while *ja* generally marks uncontroversiality, it is further sensitive to a few things. First, *ja* need not comment on the uncontroversiality of its prejacent, but can use the *content* of the prejacent to point to another common ground element that is established for all discourse participants (see examples like (258)). Second, this referenced proposition does not have to be explicitly on the discourse record, but rather, simply clear or inferable from context. Further, it seems that the contribution of *ja* can be generalized in terms of a speaker's expectations in a particular discourse context. The following section attempts to build off of previous analyses to come to a comprehensive account of the contribution of *ja*.

6.1.2 The contribution of *ja*

In order to formalize the meaning component of *ja*, one must be sensitive to multiple things. First, it must operate at the level of illocution, as its modal particle status places it on the pragmatic level of interpretation. In addition, propositions that *ja* takes as its prejacent must be taken by the speaker to be discourse-old. Kratzer (1999) offers the following approximation for *ja*, inspired by Lindner (1991):

(274) **Ja**(α) is appropriate in a context *c* if the proposition expressed by α in *c* is a fact of w_c which – for all the speaker knows – might already be known to the addressee

Refining this original notion, Viesel (2015) proposes the following definition for *ja*, modified from a proposal by Hinterhölzl & Krifka (2013) and Kratzer & Matthewson (2009). She takes as given the general idea that *ja* introduces some proposition as *uncontroversial*, giving the following definition of this uncontroversiality of *ja*, taken from Gutzmamnn (2009):[6]

(275) $[\![\lambda p.\mathbf{ja}(p)\,]\!]^c$ = that function $f \in \{f | f : \mathrm{D}_{<s,t>} \longmapsto \mathrm{D}_{<s,u>}\}$ such that $f(p)(w) = \checkmark$ (is felicitous) if the speaker believes that p is common knowledge of the speaker and the hearer in w, or it is verifiable on the spot that p, else $f(p)(w) = \lightning$ (is infelicitous).

[6]The definition in (275) has been modified from the original presentation, but only in terms of readability.

Viesel introduces another dimension of meaning that *ja* contributes as well. Along with marking the content *p* as common knowledge between the speaker and the hearer, she notes that *ja* must really take two semantic arguments, and *ja*(*p*) as outlined above only does justice to *ja*'s relationship to the first. The second argument must say something about the explanatory relation between *p* and some other discourse move *q*, outlined in (276):

(276) $[\![ja]\!](p)(q)$ iff *p* is true, *p* is not used to answer to the QUD, and *p* explains *q*, where *q* **is the proposition that the speaker asks a question or makes an assertion or a request.**

In these terms, the presence of *ja* has been somehow motivated by another discourse move, and stands in an Explanation relation to it, similar to the Explanation relation proposed by Asher & Lascarides (2003). This view of *ja* casts the particle as a causal update, which explains the relevance of a proposition by relating it to others in the common ground. This proposal focuses in on how *ja* navigates a discourse by what effect it has on the speech act type; by being able to comment on the fact that the speaker is making the very speech act, *ja* can be interpreted as an explanation for why the speaker is asserting the content of *p*, raising the question of whether *p*, or requesting more information regarding *p*.

In cases like (256), repeated here as (277), Viesel (2015)'s thoery is understood as contributing the content of the proposition *p*, which is in this case given, as well as providing an additional contextual update. In this case, the update is the act of asserting the content of *p*, which is a reaction to the issue of whether Sophie commits to Tom's proposal.

(277) **Tom:** Let's go get dinner straight from work tonight.

Sophie: *Wir müssen* ***ja*** *zuerst die Katze füttern.*
we must *ja* first the cat feed

We have to first feed the cat, you know.

Accounts that treat the prejacent of *ja* as uncontroversial, given, or known to the speaker and the hearer in the context make the prediction that the particle should be infelicitous when used as an answer to a pure information-seeking question. The accounts outlined above, as well as many others, assume a QUD-based model of discourse, a hallmark of which assumes that questions explicitly raise a current question under discussion. Assuming that *ja* cannot be a direct answer to the QUD, as Kratzer (1999) and Viesel (2015) do, is a nice way of capturing this fact: information-seeking questions mark the speaker as not being able to commit to $^{?}$p. This prediction turns out to be true. *Ja* is in fact disallowed as a response to unbiased information seeking questions (278).

(278) *Tom and Sophie are researching European capital cities.*

Sophie: What is the capital of Estonia?

Tom looks at a map, finds the capital, and responds:

Tom: *# Das ist* ***ja*** *Tallinn.*
that is *ja* Tallinn

It's obviously Tallinn. **Response to Info-seeking Q**

But a side effect of assuming a QUD model is that questions, no matter their pragmatic intent, will always introduce a QUD. In cases where question-askers are assumed to lack the answer to the question that they pose, *ja* is correctly predicted to be infelicitous: anything that raises an issue or is part of the QUD should not be targetable by *ja*. However, questions whose pragmatics assume a different question-answer relation between the speaker and the addressee respond differently to answers with *ja*, an outcome not predicted by QUD-based models. With quiz master questions in particular, a QUD analysis does not directly calculate the pragmatics of these questions, and indeed, answers to these questions may indicate that the speaker as well as the addressee know the answer. Quiz master questions, for example, pose a problem for Kratzer (1999)'s and Viesel (2015)'s theories, which assume that *ja* may not be used to answer a current QUD, regardless of its pragmatic intent:

(279) *Tom and Sophie are studying for a test on European capitals. Sophie is quizzing Tom on his knowledge, and has the answers in front of her.*

Sophie: What is the capital of Estonia?

Tom: *Das ist* ***ja*** *Tallinn.*
that is *ja* Tallinn
It's obviously Tallinn. **Response to Quiz Master Q**

Even though (279) is perhaps a cheeky answer to the question, it is not infelicitous. It simply acknowledges that as far as the speaker is concerned, the speaker as well as the quiz master find the answer to be known/uncontroversial/common ground. What's more, in these situations, it can be used to directly answer the current QUD that the quiz master unquestionably raises, a prediction ruled out by QUD-based accounts of the particle. A similar issue arises with QUD accounts of rhetorical questions, which would similarly assume that they are not licit as an answer to the issue raised. An excerpt from the German version of *Harry Potter and the Sorcerer's Stone* shows that *ja* can in fact be used in response to such questions (the English below shows the original author intent):

(280) **German Version**

Harry stand auf.

"Sir, Professor Dumbledore? ***Darf ich Sie etwas fragen?*** *"* Rhetorical Q

"Nun hast du ***ja*** *eine Frage schon gestellt," sagte Dumbledore lächelnd.*

English Version

Harry stood up.

"Sir, Professor Dumbledore? **Can I ask you something?**" Rhetorical Q

"***Obviously***, you've just done so," Dumbledore smiled.

I propose that capturing the conventional discourse effect of *ja* in terms of a speaker's *expectations* bypasses this issue while keeping the generalization that the particle comments on the common ground or uncontroversial status of the proposition it introduces (Lindner 1991, Kratzer 1999, Kratzer & Matthewson 2009, Viesel 2015). It can also retain the observation made by Viesel (2015), citing *ja*'s explanatory relation to another salient proposition in the discourse. By formulating *ja* in terms of a speaker's expectations in a context as in (281), a speaker's chain of reasoning is made explicit: using *ja* makes it clear that she is committed to believing that all participants in the conversation know some salient fact q, and based off this q, they should be able to expect p to be true as well. (281) is a proposal for the conventional discourse effect of *ja*:

(281) *ja*(q) is anaphoric to a salient proposition or event p in a discourse context C and for discourse-salient participants x, s.t.:

a. *ja*(q) adds the following to the speaker's *DC*s:

$\forall x \in C$, $\text{Exp}_x(p) \approx 1 \wedge \text{Exp}_x(q|p) \approx 1$

When a speaker uses *ja*, she signals to the other discourse participants that she believes that everyone's expectations about a particular proposition in the discourse are approximately equal, based on information that is accessible through p. In other words, she uses discourse-salient information (p) to assume that everyone views her contribution (q) as uncontroversial. Doing this allows the speaker to act as though the proposition is in the common ground already, even though it might not be. In some cases, indicating this p reactivates it in memory for the addressee; p may be a proposition already in the common ground, and q could simply be act of uttering the content of p, identical to q. In others, the speaker uses *ja* as an accommodation technique when she does not know for certain that a topic is common ground, but there is reason to suspect her addressee might already know the information in q.

In an example like (277), the update to a speaker's discourse commitments is trivial. The prejacent q of *ja* is *We have to feed the cat first*, which references the salient common ground p, which is *Tom and Sophie are responsible for feeding the neighbor's cat*. In a table model, this update might look something like the following:

(282) a. *Common ground:* Tom and Sophie both know that they are responsible for feeding the neighbor's cat. (p)

b. Tom suggests a proposition r for consideration:

Tom: Let's go get dinner straight from work tonight. (=r)

DC$_{Sophie}$	**Table**	**DC**$_{Tom}$
	{r}	$\text{Exp}_T(r) \approx 1$

$cg : p,\ ps = \{\{p \cup r\}\}$

c. **Sophie**: Wir müssen **ja** zuerst die Katze füttern. (=$\boldsymbol{ja}(q)$)

d. Update the context with q and *ja*:

$\mathbf{DC}_{Sophie}$	**Table**	$\mathbf{DC}_{Tom}$
p, $\text{Exp}_{S,T}(p) \approx 1 \wedge$ $\text{Exp}_{S,T}(q\|p) \approx 1$	$\frac{\{q\}}{\{r\}}$	$\text{Exp}_{T}(r) \approx 1$

$cg: p,\ ps = \{\{p \cup q\}\}$

In this case, Sophie's discourse move with *ja* is a rejection of Tom's suggestion. With it, she effectively dismisses the issue Tom has raised with his utterance in the previous move, while placing *q* on the table, and using *ja* to bolster her evidence for why her expectations are the way they are. But she does not outright reject Tom's proposal without good cause— she uses *ja* to point to contextually salient information as a cue for her initial refusal. It's not that she doesn't want to agree to Toms proposal, but other obligations that the two of them have prevent this from being a valid outcome.

An expectation-account of *ja* also correctly predicts that *ja* will be felicitous in response to quizmaster or rhetorical questions. In these situations, *ja*'s felicity is determined only with respect to a speaker's expectations. Such questions raise an issue, but in opposition to information-seeking questions, their pragmatics suggest that the question-asker has access to a correct answer, and is testing whether the addressee also knows this information. If the question-answerer is certain about their response, they may use *ja* to show that their answer is uncontroversial. Since the context has already established certain expectations about each participant's knowledge with respect to the answer, *ja* answers are felicitous, as they can reference the fact that this is known to everyone.

Formulating *ja*'s contribution in terms of expectations is able to capture the apparent corner cases of the particle as well. Recall the conversation between Tom and Sophie, repeated in (283) below:

(283) **Sophie:** *Du kannst dich noch an Jeff erinnern, ne?*

You remember Jeff, right?

Tom: *Ja, so 'nen großen, blonden? War eine Weile in Japan?*

Yeah, he's like, tall and blonde? Was in Japan for a while?

Sophie: *Genau. Er hat* ***ja*** *nach Japan ein Kätzchen adoptiert.*

Exactly. You know, after Japan he adopted a kitten.

Tom: *Ach, echt? Das wusste ich nicht.*

Oh really? I didn't know that.

Sophie: *Ach so. Ja! Eine kleine graue. Aber, genau— er kommt Morgen zu Besuch.*

Oh. Yeah! A little grey one. But right— he's coming for a visit tomorrow.

In these cases, the speaker uses her expectations about information that *could* be common ground in order to introduce a statement with *ja*. The speaker here assumes too much, but she does this as a sort of in-group marking—she acts as though her addressee could have this information as common ground in order to signal that it is not the most critically informative contribution that she is making. Using the particle in this way is taking advantage of the *ja*'s role as an overt Maxim of Quantity violation. By treating *After Japan, he adopted a kitten* as known material, the speaker signals that she wants to be able to raise this issue without having to do much to negotiate it. Note that while the addressee can request confirmation and can also accept the contribution with no problem (as in (284a) below), he cannot outright reject her *ja* statement in this context (284b):

(284) a. **Tom':** *Ja, das habe ich schon gehört.*
yes that have I already heard
Yes, I heard that already.

b. **Tom'':** # *Nein, hat er nicht.*
no has he not
No he didn't.

What Tom *can* do is negotiate the fact that the contribution is on the discourse record, but not the truth of the proposition itself. In this sense, a speaker using *ja* in this way is still placing the same expectations about the common ground and the proffered content of the proposition on the table, but she is doing so without knowing for sure that what she takes as established in the common ground is in fact established for all participants.

Formulating the contribution of *ja* in terms of a speaker's expectations is able to keep the core meaning of the particle constant from previous proposals, while also being able to navigate between the prejacent and another contextually relevant proposition in the common ground. By placing these contributions directly into the discourse commitments of the speaker, these expectations are in view to all conversational participants, yet not negotiable. And as we have seen, when they do become negotiable is when a speaker has been too confident in assumptions: if no common ground exists between two speakers, it is infelicitous to use *ja*.

6.1.3 The interaction between *ja* and JA

Just like *doch*, *ja* has a stressed variant that occurs in a restricted set of utterance contexts and sentence types. Stressed JA is only allowed in imperative utterances and deontically modalized declaratives, as in (285) and (286), repeated from (266-267):

(285) *Mother to child:* *from Thurmair (1989), p.109*

Komm JA nicht zu spät heim!
come JA not too late home
Don't (you dare) come home too late!

(286) *Ich darf JA meine Autoschlüssel nicht vergessen!*
I permit JA my car-keys not forget
(Of all things) I can't forget my car keys!

Thurmair (1989) observes that stressed JA in imperatives is acting as a particle that strengthens the illocutionary act. She notes that the particle seems to only occur in imperatives that act as threats, warnings, or expressions of necessity or obligation. In (285) above, the mother can only say this to the child if there is a certain expectation held by both of them. If the child has a curfew at 11pm, and has been known for ignoring it, using (285) is a way for the mother to use knowledge of these previous transgressions as a warning to her child: *Don't come home too late, or there will be consequences.* In (286), we interpret a necessity reading. The speaker is asserting with authority that she cannot forget her car keys.

I argue that just like stressed DOCH cases, stressed instances of JA arise in conjunction with focus marking on the particle, coupled with the discourse effect of the unstressed particle. Assuming a theory of imperatives along the lines of Condoravdi & Lauer (2012) or Portner (2004), imperatives express an effective speaker preference for the addressee to comply with the order, something like *addressee does VP*. In doing this, the speaker must leave room for the fact that the addressee *might not comply* with the imperative. Using the unstressed *ja* in an imperative signals a pragmatic clash: it is infelicitous for the speaker to both express a preference for a proposition (contributed by the force of the imperative), as well as use *ja* to signal that the speaker expects that proposition to be known or uncontroversial to all participants in the discourse.

Stressed JA, like stressed DOCH, is a way to highlight the intrinsic effect of the particle at a level above the semantic content of an utterance. Focus marking in a Roothian-based theory assumes that a focus-marked utterance sets up a set of salient alternatives to that utterance. When *ja* is focused, the set of salient alternatives is calculated at the level of the speech act itself. The focus semantic value includes only two alternatives: the prejacent and its opposite. Knowing this allows the hearer to make the following broad pragmatic calculations:

(287) a. The speaker utters an imperative q, which has a basic effect of committing her for an effective preference of the form *addressee does* q.

b. The speaker uses JA, committing her to a high expectation for all discourse participants to expect the prejacent of *ja* to be true based off of expectations about another salient proposition p, and introduces a set of alternatives.

c. The only available focus alternative is $\neg$(*addressee does* q). There must have been a reason for the speaker to highlight this: the addressee has previously *not* complied with the imperative.

d. The salient p must be the fact that *the addressee already knows that the speaker has a preference for the addressee to perform* q.

e. Previous non-compliance with a speaker's effective preferences has forced the speaker to double down: an imperative with JA is a way for the speaker to express her preferences, as well as reference the fact that these preferences were already raised and not complied with.

The "strengthening" use of JA in imperatives here is then an additive effect of the particle, focus marking, and the conventional discourse effect of an imperative. Its limited use in cases where a speaker is making a threat, warning, or expressing necessity is a confluence of factors. The imperative form contributes a speaker's effective preferences with respect to a discourse situation. Focus marking on the modal particle highlights these preferences even more, indicating that there must be a reason for the speaker to make this kind of contribution. The expectations that *ja* places on discourse participants makes this focus marking even more salient: the speaker indicates that she believes the addressee is already aware of her effective preferences. Highlighting this, along with the fact that the addressee has already not acted in accordance with the speaker's wishes comes off particularly strong. The addressee concludes that the speaker intended this as a threat, with implied repercussions for non-compliance.

6.1.4 Summary: wrapping up the effects of *ja* and *doch*

Both *ja* and *doch* have a similar discourse function, wherein they make explicit the expectations of all conversational participants. But where they differ is in how these expectations are expressed. For *ja*, the goal is holistic: the particle attempts to place the speaker and all other interlocutors on the same page about a particular proposition, given contextually salient information. Formulating this in terms of expectations also has the added benefit of deriving the fact that *ja* is disallowed with other speech act types that assume the speaker and the hearer have asymmetrical expectations. In all cases where a contrast in participants' expectations arises, *ja* is simply infelicitous.

Doch's contribution is less straightforward. Instead of indicating that everyone's expectations are aligned, it is a marker that the speaker's expectations have somehow been violated. As the goal with any conversation is to grow a consistent common ground between interlocutors, *doch* is a signal to an addressee that some tension exists relative to this goal. In particular, the speaker's expectations with respect to to some proposition indicate one path forward, while some salient event or proposition indicates another. In other words, the speaker's expectations about the progression of the discourse has been violated.

6.2 German Intonation and its interaction with *doch*

German intonation is very similar to that of English. The language is generally described in typical ToBI fashion, as a system of tones and tonal movements throughout an utterance. These tones can fall on various different stressed syllables in a phrase, but default accenting places the most prominent sentential tone on the stressed syllable of a verb's complement. As with English, the prosodic structure of an utterance is not rigidly fixed, and speakers can manipulate the pragmatic message of an utterance by varying the prominence relations between constituents. In particular, a speaker's pitch is responsible for conveying a healthy chunk of the pragmatic information that she intends to express. Much of the systems of German and English overlap, leading to a nice parallelism in the two languages' prosodic realizations. The following sections outline a few prosodic tunes in German and compare them to their English counterparts.

6.2.1 German Neutral Falling Intonation

German's default, neutral intonation is much the same as English. It is realized as an utterance-prominent H* tone, which falls to a low tone at the end of the intonational phrase. While this seems identical to the English contour, Féry (1993) proposes that German intonational phrases are realized slightly differently, arguing that they do not need boundary tones. Instead, they undergo a final lowering process, where the height of the nuclear accent at the end of a phrase is realized lower than it would be were it not phrase-final. This is a reaction to the multiplicity of final tones in English, which she notes "has a much richer inventory of post-nuclear tones than German" (78). For this work, boundary tones (%) will still be indicated here when discussing utterance-level tunes in German to indicate that the contour in question is assumed to be at the sentence level.

The most natural tone for declarative utterances in German is the neutral final falling tone, as described for English. This is used for normal assertions, as in (288).

(288) *Tom will Käsespätzle.*
H* L %
Tom wants cheese noodles.

This is also the default tone for wh-questions, with the accent falling either on the wh-phrase (289a) or on a narrow-focused element (289b):

(289) a. ***Mit wem*** *warst du gestern Abend?*
H* L %
Who were you with last night?

b. *Wo hast du Sophie kennengelernt?*
H* L %
Where did you meet Sophie?

Rising tones have a similar distribution in German and English as well. Both languages use this tonal pattern in canonical polar question formation and with rising declaratives. Féry (1993) also notes that is can be used by a speaker to indicate that an answer to a question should be very obvious in particular contexts:[7]

(290) A: *Was ist das?*
What's that?

B: Das ist ein Pandabär?
L* H %

B: (Uh,) that's a panda bear?

For the German neutral final fall, I assume the same basic pragmatic effect as the final fall in English:

(291) The neutral falling H*L% in German adds the following to the speaker's *DC*s:
$\text{Exp}_{spkr}(\text{p}) \approx 1$

Just like modal particles, a speaker's expectations about a proposition can be expressed by their intonation. When *doch* or *ja* are paired with an utterance with an overall neutral

[7] I do not take up the issue of rising utterance-level tunes in German. For more information on the pragmatic effects of similar tunes in English, I defer to Jeong (2017), Jeong & Potts (2016), and Rudin (2017a).

contour, nothing unexpected happens. The speaker already indicates with both particles that she expects the content of her utterance to be true. This small bit of redundancy further strengthens the fact that the speaker is confident in her commitment to the truth of her contribution.

6.2.2 German intonational idioms

German also makes use of intonational idioms, whose component parts cannot be further broken down into specific meaningful units. Chapter §3 introduced the English SRC as one of these intonational idioms, and investigations in Chapter §4 showed that the interpretation of this contour was dependent on factors like relative pitch height and f0 excursion. German makes use of a similar contour expressing a similar sentiment. Like English, this contour has a high variant (292a) and a lower overall pitch variant (292b):[8]

---

[8]The SRC in German contrasts with the so-called 'hat pattern' which is similar in form, but not in function. Whereas the SRC comments on a participant's specific expectations in a discourse, the hat pattern has a less clearly pragmatic function. Though it is also composed of a low pitch accent that rises to a high sentential accent and falls phrase-finally, the hat pattern is more of a topic-focus pattern which is not anchored to prominent positions in a clause. Rather, it is the default for a number of (non-canonical) syntactic constructions in the language (examples taken from Féry (1993)):

i. A: Did you sleep well?

B: *(Na,)* GESCHLAFEN *hat (doch)* KEINER *von uns*
L* H H* L

SLEPT, none of us has.
intended: NONE OF US were able to sleep. **VP Topicalization**

ii. *Jedes mal wenn es klingelt, bellt der Hund.*
L*H H*L

Every time the doorbell rings, the dog barks. **Temporal clauses**

iii. *Leo ist nicht gekommen, um Maria zu ärgern.*
L* H H* L

Leo didn't come in order to annoy Maria. **Negative scope disambiguation**

iv *Maria ist nach Berlin gefahren, und Martin nach Hannover.*
L* H H* L L* H H* L

Maria went to Berlin, and Martin to Hannover. **Gapping/Contrastive**

The difference between the hat pattern and the SRC lies in what Gussenhoven (1984) calls *tonal linking*. This is a distinction between bitonal and monotonal pitch accents, wherein particular phrasal tones are composed either as linked to a single position, or through a movement of the tone in the utterance. The SRC is a (partially) linked contour, with the L* tone of the secondary accented syllable rising to the H* of the last stressed syllable of the utterance with no other significant tonal movement between. That is, there is a single position that represents the

(292) a. *Speaker is walking from stall to stall at a farmer's market:*

Eine blaue Melone!
L* H* L%

A blue melon!

b. **Sophie**: Looks like we're going to need more sauerkraut after all.

Hannah: *Das habe ich doch gesagt!*
L* H* L%
I SAID that!

The figure in (6.1) shows the pitch tracks of the same utterance spoken in a surprised context and in a redundancy context. Here, they have been time-aligned to show the similarities in their contours, and also to show the difference in relative f0 excursion for each production. Each were elicited in particular contexts–the surprise context is shown in (293a) below, while the redundancy context is in (293b). This shows a very similar pattern to the investigations and experiments in Chapter §4.4: differences between surprise and redundancy readings of the contour are distinguished at least in part by the difference in pitch height between L* and H* movements, and H* and the L% boundary tone.

(293) a. **Surprise context:** Sophie leaves her house a mess in the morning when she goes to work. When she gets home, she walks into her house to find that her living room has been completely cleaned from top to bottom.

b. **Redundancy context:** Jeff is marveling at how clean the living room is, since it's usually quite a mess. Jeff asks Sophie how the living room got that way. Sophie responds as if to imply that the answer is obvious: someone cleaned the living room.

Jemand hat das Wohnzimmer geputzt.
L* H* L L%

---

high pitch accent of the contour. In contrast, the hat contour is not linked: the L* pitch accent rises to a high tone, but this high tone is a distinct segment from the following H* accent, which then falls to a low tone. While the shape of the two contours is very similar, the hat pattern is a composition of two distinct tonal movements: one that rises to a peak (L* H), and another that falls from a different peak (H* L%).

For the SRC, the L* and H* tones are anchored to primary and secondary phrasal stresses in the utterance, and the movement from the L* to the H* are linked–there are no intonational pauses between them and the low tone is carried through stressed and unstressed syllables until the high tone receives prominence. This is not the case for the hat pattern, where a sequence of two tonal movements–one low to high, the second high to low–is broken by stresses, pauses or intervening tones. The hat pattern is two smaller, phrase level tunes, and not the subject of investigation here.

Someone cleaned the living room.

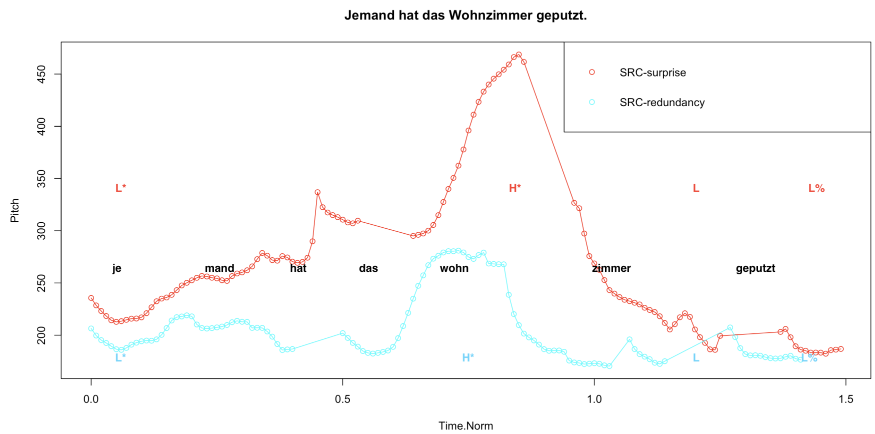


Figure 6.1: Relative f0 height differences between a 'surprised' and a 'redundancy' reading of the SRC in German

As with English, the pragmatic meaning component of the German SRC is a commentary on violated expectations in the current discourse context. But whereas the English SRC is a strategy that expresses the expectations of the speaker while pragmatically signaling a clash between other participants' expectations, the German SRC is less pragmatically complex. The English SRC functions similarly to *doch*, indicating violated expectations, along with implicitly requiring a response from the addressee. The German SRC does not appear as pragmatically complex. While it does comment on a speaker's surprise, there is no pressure for the addressee to respond. In surprise cases, the speaker's expectations are interpreted as violated in some way. But in redundancy cases, there does not seem to be an unresolvable juxtaposition of speaker and addressee expectations that arises. Instead, the speaker simply indicates that while she viewed some propsition as uncontroversial, there is someone else in the discourse who was surprised by this. Instead of putting the addressee on the hook for a successful resolution of the issue, it simply points out that the proposition should be obvious. Given this, a tentative description of the meaning component of the German SRC is below:

(294) For a proposition p in a discourse context C and for discourse salient participants x, the

German SRC adds the following to a speaker's *DC* list:

$\exists x \in C, \text{Exp}_x(p) \approx 0$

Though simple, the interaction between the German SRC and a speaker's overall f0 height lead to almost the same effects as *doch* or the English SRC: high f0 with the SRC indicates a speaker's own surprise at the prejacent, and a hearer interprets a 'surprise' reading. Neutral or low f0 commits the speaker to p, which leads to the inference that it must be someone else in the discourse whose expectations have been violated. Though I do not go so far as to formalize exact expectation calculations for high f0 or low f0 here, the basic effect is schematized in (295) and (296):

(295) *Jemand hat das Wohnzimmer geputzt.* **+ *High f0***
L* H* L L%
Someone cleaned the living room.

a. **Speaker uses SRC:** L* H* L% contributes $\text{Exp}_x(p) \approx 0$ for some participant x

b. **High f0:** Speaker indicates newness and a positive reaction to p (similar to the English Excited Contour, Chapter §3.3.2)

c. **Hearer computes that surprise is anchored to the speaker:**
Speaker-oriented newness toward $p \rightarrow$ Speaker's expectations for p were low.

(296) *Jemand hat das Wohnzimmer geputzt.* **+ *Low f0***
L* H* L L%
Someone cleaned the living room.

a. **Speaker uses SRC:** L* H* L% contributes $\text{Exp}_x(p) \approx 0$ for some participant x

b. **Low f0:** Speaker indicates clear commitment to the proposition p

c. **Hearer computes that surprise is anchored to the someone else:**
Speaker commits to $p \rightarrow$ Addressee's expectations for p are low.

One thing that is predicted by a compositional account of intonation and modal particles is that the two German mirative strategies can combine: *doch* should be able to be used with the German SRC. And they do. When *doch* is present with the SRC, the effect is very similar to the English SRC strategy. Though acoustic analysis does not find any measurable pitch or tone difference between SRC environments with and without *doch*, the pragmatic effect is different once the particle is thrown into the mix.[9] Adding *doch* to (293b) with lower pitch at once changes the interpretation from simply signaling that the speaker and the addressee are not on the same page about a proposition: the speaker also insinuates that something in the discourse context should have lead the addressee to believe something other than what they have committed to. The pragmatic clash between speaker and addressee expectations is foregrounded.

(297) **Redundancy context:** Jeff is marveling at how clean the living room is, since it's usually quite a mess. Jeff asks Sophie how the living room got so clean. Sophie responds as if to imply that the answer is obvious: someone cleaned the living room (*duh!*).

Sophie: *Jemand hat* ***doch*** *das Wohnzimmer geputzt.*
L* H* L L%

Someone cleaned the living room!↘

The contribution of *doch* here is clear, as the MP-utterance, as opposed to the bare utterance, seems to require a response from the addressee to reconcile the differences in participants' expectations. Ultimately, while both strategies can be used at the same time, the pragmatic force of *doch* seems to overshadow the SRC. While their contributions are additive, violated expectations contributed by the SRC entailed by the discourse effect of *doch*. What surfaces is a more expressive version of *doch* which a speaker may or may not choose to use.

Though their formalizations depend on language-specific requirements related to their positions in a clause, the conventional discourse effects of *doch* and the English SRC are identical. The effects of overall pitch height throughout an utterance also translate well to disambiguating "surprise" and "redundancy" readings in German. Just as an English, an utterance with higher average overall pitch is indicative of surprise readings when *doch* is present, while lower overall pitch induces a pragmatic effect of redundancy. Though this has not been experimentally validated as is the case for the English SRC (see Chapter §4), impressionistic judgments from multiple native speakers of German have tentatively confirmed these claims.

[9] A very small corpus study performed in which identical utterances from four speakers (3 females and 1 male, n = 15/speaker, 60 total) were compared. Measures of relative pitch minima and maxima were tagged and word-aligned in SRC contexts, and then compared based on the presence or absence of *doch* in the utterance. There were no significant differences, or differences even approaching significance.

6.3 The question of salience

Much of the effect of both *ja* and *doch* depends on identifying a salient proposition or event *p* that a speaker uses as evidence or a rationale for their MP-utterance. At times, this *p* can be the propositional content of the utterance, but it can also be the speaker's *act of utterance*, or a proposition or event that provides an explanation for the utterance in the first place. But given this, there must be reasonable constraints on the discourse relations between modal particle (*doch* or *ja*), the prejacent *q*, and this salient proposition *p*.

Formulating expectation in terms of causal models of belief and knowledge structure provides a way forward. For both *doch* and *ja*, this external, salient proposition provides an explanatory link between the propositional content of the utterance itself, and its relation to the current discourse context. This returns back to Halpern & Pearl (2005)'s stated goal: "Whenever we undertake to explain a set of events that unfold in a specific scenario, the explanation produced must acknowledge the actual cause of events"(p.194). *Doch* and *ja* do just that. With an underlying casual structure, *doch* and *ja* indicate different causal paths that speakers assumed to get to the current world of conversation, and how they may or may not differ from the paths that they assume their addressee has taken as well. A salient *p* in these cases indicates a general, causal link between the state that the speaker finds herself in, and some intermediate state along the chain of reasoning. In a very general, schematic way, *doch* represents a state of affairs in which a speaker expects that her chain of reasoning has diverged from that of her addressee (6.2), while *ja* represents a causal network where the speaker expects that she and her addressee have made the same inferential leap (6.3):

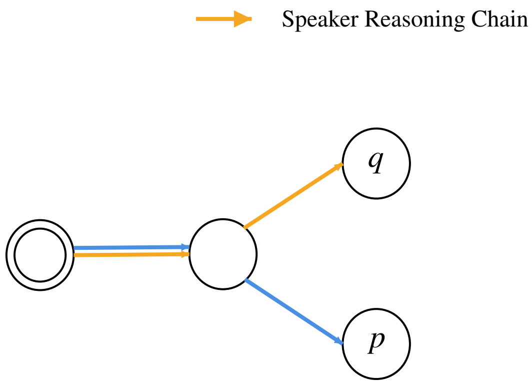


Figure 6.2: Generic causal model for *doch*(q)

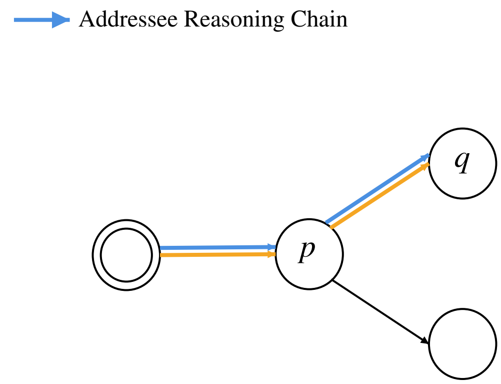


Figure 6.3: Generic causal model for *ja*(q)

While these illustrations are simplifications of the expectational structure that *doch* and *ja*, they do clearly represent salience relations between the particles, their prejacent, and their antecedent. For *ja*, p serves as an explanatory link: it is evidence off of which a speaker can base her claim. For *doch*, p is a node on an alternative path that a participant could have taken. Using the particle indicates that the speaker's and the addressee's expectations have diverged at a particular point: the speaker's commitment to q is used to suggest that the addressee reevaluate their apparent commitment to p.

6.4 Summary

This section aimed to present three things: an extension of the *doch* analysis from the previous chapter to the MP *ja*, similarly modeling it off of speaker expectations, a look at modal particles and intonation, and finally, a small discussion on the idea of contextual salience. This chapter argued in particular for a treatment of *ja* that differs from *doch* in that it indicates shared speaker and addressee expectations, rather than conflicting expectations. In this way, while *doch* was a marker of mirativity and violated expectations, *ja* is **not** mirative. Instead, the particle is a way for a speaker to comment on collective knowledge as a way to communicate a larger message. This could surface in the form of reminders, of generally known information, or could be used as way to signal the uncontroversiality of one proposition given another, related one. While it does not comment on violated expectations, it does allow a speaker to comment on the common ground that she believes she shares with her interlocutors. Like *doch*, *ja* is a discourse navigating technique speakers employ to make their views about a discourse clear.

A broader goal of this chapter is to provide an explanation of how to interpret modal particles as part of the larger performance of an utterance. Much of the semantic work that has been done on particles ignores the way in which intonation is used to contribute pragmatic information about the utterance. The contribution of intonation is also intricately tied to the notion of expectation, and is similarly calculated as not-at-issue speaker commentary about the uttered proposition in context. Disentangling the pragmatic contributions of modal particles from the contributions of prosodic contours is no small problem; the way in which German MPs are integrated into an utterance makes it difficult to fully separate out their pragmatic meaning from the meaning a speaker intends to convey with the presentation of her utterance. Yet it is clear that the effects are at least cumulative: pitch differences in an utterance's performance have an effect on the pragmatic interpretation. This places some confidence in the hypothesis that claims modal particles and intonation are both separate contributors of pragmatic information to a discourse.

Chapter 7

Conclusion

This dissertation seeks to answer a few central questions, all of which are built off of a central theme: can we use prosodic meaning as a metric to inform the meaning of other semantically underspecified elements? To this end, I have proposed an updated view of what kind of information is handled in a Table Model of discourse representation (Farkas & Bruce 2010), which is underlain by a causal model framework of discourse structure (Halpern & Pearl 2005), and a probabilistic notion of expectation (Chapter §2). Once these tools have been established, I show how they can be extended to capture various exponents of mirativity, which exist in both English (Chapter 3), and German (Chapters §5 and 6). In order to put some teeth behind the theoretical claims that I make in Chapter §3, I introduce a series of experiments designed to test the predictions that this new view on discourse management and coherence makes (Chapter §4). Overall, a picture emerges that casts intonation and discourse particles as related, but different strategies that help to oversee and communicate not-at-issue speaker expectations to all participants in a discourse. Doing this is a crucial part of the pragmatic channel, functioning at a communicative level above the utterance itself. In all, I argue that what these elements have in common is their ability to enrich a discourse with pragmatic material that contributes speaker-oriented commentary on a participant's beliefs and expectations in a conversation.

Chapter §2 is the most theoretically rich portion of this dissertation. It introduces some the central concepts that serve as the core of the analysis, and of the investigated phenomenon more generally. This chapter covers the way that expectations are structured for an individual, and provides a definition of the technical term *Expectation* that is then used throughout the rest of the work. This is built in part off of the idea from **?** that beliefs are conditioned on a speaker's relative credence level toward a proposition. Here, I define expectation as a function of the joint probability of an at-issue proposition relative to other propositions in the context set. This allows us to have a grounded view of what it means when we say that a speaker *expects* some proposition, and precisely what it means when these expectations have been violated.

This chapter also introduces an underlying model of discourse that is easily extended to situations that involve violated expectations. It uses a causal model framework to condition a speaker's expectational state at certain points in a discourse. Doing this is a way to make the causal connections between propositions explicit, and also a way to narrow down the set of expectational pathways that can serve as the catalyst for a speaker's surprise. Doing this makes the relations between a speaker's mental states clear in order to be able to tell a coherent story about mirativity, a realm that deals very intimately with the notion of violated expectations.

Chapters §3 and 4 function as a unit. Within this unit, they serve as first a theoretical grounding for an additive approach to discourse particle and intonational meaning, and second, a proof of concept. Chapter 3 looks at the discourse particles *oh, huh*, and *what* and proposes formulations for their pragmatic discourse effects, based in part off of their interactions with intonational contours, which also contribute pragmatic information. In a pairwise fashion, the meanings of these particles and contours are compared, and discourse effects of these elements are proposed, based on this core notion of expectation (and expectation violation). As the particles under discussion in this chapter, as well as one of the contours, are mirative strategies, we see complex discourse navigatory functions emerge from the additive relations between the two. From this, the experiments in Chapter 4 are born, which test this claim of additivity, as well as perceptions of emotion that these particle and contour pairs convey in various neutral and biased contexts. The results here are promising, and at the very least suggest that an additive approach to multiple pragmatic discourse navigation strategies in a single utterance is an analysis that is on the right track.

Chapters §5 and 6 also form a pair of related topics, as the focus switches from mirative strategies in English to mirative strategies available in German. As opposed to the literature on the relevant discourse particles in English, the literature on German modal particles is vast and controversial. I look at two particles here that are assumed by many to be each other's foil: *doch* and *ja*. Here, I use the same approach to analyze these particles, focusing on how they contribute to the expectations of a speaker in a discourse, and their navigatory power. I diverge from previous analyses of these particles, and provide an account of these that takes the prosodic environment and utterance type differences into account, providing a way of analyzing these particles that at once captures their discourse function, while also fitting them into the existing typology of mirativity. In particular, *doch* functions in almost the exact same way as the English Surprise Redundancy Contour, while *ja* functions as an anti-mirative, marking content that is known and expected in the context by all speakers. In all, the path through this dissertation weaves itself through the fields of expectation, intonation, discourse particles, mirativity, and ultimately, right back to expectation. In all, these themes collectively serve as ways for a speaker to reveal her nuanced beliefs and commitments in a discourse.

But investigation into this realm is far from complete. What follows in the rest of this chapter is a handful of open questions, avenues for inquiry, and an outlook on the future of exploration and research in this domain.

7.1 Variation in expressing (violated) expectation

The conclusions and generalizations drawn about the particles and contours in play throughout this work also bring up questions about variation. Specifically, what kind of variation do we find with respect to mirative strategies in English, in German, and more broadly? What kind of variation do we see in the kinds of expectations that can be referenced in a discourse? And if we expect these claims about expectation to be a feature of not just these two languages, but of languages more broadly, how do we capture the variation here? These questions, you will notice, have been saved for the end, as they are not questions that I purport to have the answers to. Instead, I will leave them as open questions, but not without first entertaining them each individually.

7.1.1 Variation within English

One of the main takeaways from this work is that languages use a variety of methods to indicate speaker expectation violation. This includes, but is not limited to discourse particles, modal particles, and intonation. As we have seen for English, each of these strategies has its own place in the arsenal of discourse navigation techniques. The particles *oh*, *huh*, and *what* are all particles that are in some way forward looking: they occur as a response to another proposition, event or fact and provide an explanatory link between that proposition and a following utterance. The SRC is a bit more complicated, linking a proposition that it occurs on with another salient proposition in the discourse. In addition, all of these strategies differ in their function, and all play crucial roles in helping the speaker externalize their own mental representation of the discourse and compare it to the actual world. The variation in meaning from one particle or intonational contour to another is meaningful, and their different roles in navigating the discourse represents a complex awareness of the way speakers' internal states differ from the actual world, and from the states of the other actors that they are participating with.

One way in which all of these elements are the same is with respect to anaphoricity. In all the cases that we discuss here, in both English and German, these strategies crucially provide a discourse-connective link to other propositions that have already been entered into the discourse context, whether by means of explicit acceptance of propositions into the common ground, or acknowledgment or receipt of facts in the larger global discourse structure. But there is an added complexity in this anaphoric aspect. Some particles, like *oh, huh* or *what* are anaphoric to a single proposition, while other strategies, like the SRC or German *doch* are inherently more complex due to their dependent nature; they must make explicit reference to a host utterance as well as an anaphoric utterance. But we can make a generalization: anaphoricity is a constant. What varies is the discourse relations that each strategy sets up in a specific context.

There is also the open question of the various interpretations of the SRC in English, and the similarities as well as differences that it poses as a contour in the German inventory.

Within English, the contour has a single pragmatic function, but can be used to indicate two quite different mental states that the speaker encounters. What the contour seems to indicate is that this particular kind of surprise can occur with both a positive and a negative valence, while still expressing the same conventional discourse effect at its core. In this work, I have identified this as two variations of the same contour, one the SRC-High, whose overall higher pitch skews the conversation toward a 'surprise' interpretation, and an SRC-Low, with an overall lower f0 and bias for a 'redundancy' reading. I have been willing to break these elements down into their component parts, but have only staked a claim on the discourse effect of one of these parts. I have defined the SRC as in (298) below, but have been a bit handwavey about the exact effect that is contributed by the SRC and an overall high f0, and the SRC with an overall lower relative f0.

(298) The English Surprise Redundancy Contour (H) L* H*-L% is anaphoric to a salient proposition or event p in a discourse context C, and is admissible for discourse-salient participants x when

a. q is the proposition expressed by the speaker (uttered content or the presuppositions introduced by a question),

b. add the following to the speaker's *DC*s:
$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

In part, I have done this so as to not have to assign a meaning to parts of the speech stream like 'high overall pitch' or 'low overall pitch', as doing this seems much too restrictive. But it is worth looking into what kinds of canonical affective meanings are associated with this type of delivery. It is true that an utterance with high overall pitch for a speaker's range is not very likely to be interpreted as anger, frustration, or any other more negatively skewed emotion. Even semantically biasing the utterance is little help in remedying this effect. Imagine the sentence in (299) being said with high overall f0, and it is very difficult to interpret a negative reading. Similarly, though the sentence in (300) is semantically skewed in the positive direction, saying it with a lowered overall f0 seems to indicate that the speaker has a negative bias toward that proposition.

(299) Tom was annoyed by the constant banging on the wall. + *High overall f0*

(300) Tom was bursting with pure joy as he entered the bowling alley. + *Low overall f0*

What seems to be generally the case is that lower f0 involves the speaker's emotional state, as well as the state of some addressee, while higher f0 need only reference the emotional (or expectational) state of the speaker. I do not have anything more to say on this matter, other than we can hope that further study of these effects will yield a cleaner picture of the complex interactions between fluctuations in pitch (among other variations in voice quality) and the intonational idioms that they can combine with.

Step[ping back for a moment, we can look at the variation in the pragmatic effect of the SRC between two closely related languages and see what we can draw from this. While English's use of the SRC has a very complex pragmatic meaning, the SRC in German, is more restrictive. Setting aside for the moment some of the differences in the two languages' overall prosodic mappings, we can ask, why would we predict a difference in the interpretation of the SRC, given all the similarities between German and English? One reason is that German has another way of expressing this same sentiment: through the use of the modal particle *doch*:

(301) *doch*(q) is anaphoric to a salient proposition or event p in a discourse context C and for discourse-salient participants x, s.t.:

a. *doch*(q) adds the following to the speaker's *DC* list:
$\text{Exp}_{spkr}(q) \approx 1 \wedge \forall x \in C\ [\text{Exp}_x(p|q) \approx 0]$

Given this information, it is not hard to imagine why the SRC in German does not have the same affective power as the SRC in English; there is already a strategy for that which has been grammaticalized in the language, namely, *doch*. I speculate more about why languages might use one communicative channel or another to express similar emotional or not-at-issue content in the following section.

7.1.2 Variation Crosslinguistically

One question that we might ask is why there is so much variation in the expression of violated expectations both within English, and cross linguistically. Intuitively, we might expect there to be a clear division between channels of information, corresponding roughly to the prosodic realm and the segmental channel. We might expect that information that is entered into the discourse record, placed on the table for discussion, agreed upon and then entered into the common ground is a part of the segmental channel. Many languages tend toward the generalization that it is better for discrete pieces of language to show up in this channel. Words have the property of having clear, unbiased informational consistency, and in a system that seeks to establish mutual consistency between participants, the best way to do this is through material that is at-issue, and can be easily discussed, accepted, and rejected.

Intonation, on the other hand, is part of the prosodic channel, and it is much less discrete. There are many uses for this channel that do not correspond to meaningful pieces of language, such as affective expressions, and indexical or speaker-specific verbalizations, which can be either voluntary or involuntary. This channel is potentially loaded with other material that is also non discrete and not at-issue. Violated expectations is, in this respect, like the information that can be conveyed in the prosodic channel. It is a highly speaker-oriented notion, and expresses emotive content that is not-at-issue to others in the discourse. But there are properties of intonational mirative markers that are not like other non-discrete content. For one, it is anaphoric, and has the quality of discourse connectiveness. Expressing violated expectations does not just express information about a speaker's emotion, it is also information that can be relativized to other participants in the discourse. It shares in common with words and other forms of discrete language the property that it seeks to be informationally concordant. Intonational markers of expectation have the goal of establishing mutual information, but not in a way that establishes this goal as at-issue. By marking a proposition as surprising, they are not conveying that the conversation itself is inconsistent, but that the speaker's own internal model is. Expressing expectation violation through intonation has the advantage of being able to express public consensus-building information that discrete language is well suited for, but it does so on a channel that is itself non-discrete.

If we believe this, then we are immediately faced with the question of why a language would use the segmental channel as its main method of communicating surprise. If intonation is such a good way to carry this info, then why would we ever expect not-at-issue content to show up in the discreet channels? But there are many mirative strategies that only show up as word-like morphemes in their respective languages. This sketch above cannot be the whole story. We must remember that expressions of expectation are not the only not-at-issue content that can show up in the discreet channels: other emotive or speaker-oriented content shows up here too. If we think of internal states as important information for guiding conversational moves, it is not a far leap to say that these elements might benefit from discretization for many of the same reasons that at-issue content benefits from discretization. Making them part of the channel that normally signals at-issue information can have the effect of elevating them in terms of their conversational importance, while still allowing their contributions to go unnegotiated in the larger discourse. Under this view, it is really not surprising at all that languages have developed different strategies for expressing violated expectations.

Another point of variation that we want to address is whether all mirative markers all share some common core property. AnderBois (2017) touches this topic briefly, and sketches an outline of a typology for mirative markers which draws divisions between markers that comment on a core predicative relation of revelation, receipt of new information and surprise. While this may be an informative division, a tentative proposal that I put forth, which may prove useful to further categorizing miratives into the relevant subtypes, is that these markers build in the notions of *anaphoricity* and *not-at-issueness* into their meaning. Doing this builds in the notion that these particles are extremely context sensitive, and could perhaps lend some insight into their discourse-navigatory power. Other lines that could be useful in testing which of the typological categories certain mirative strategies fall into come from two concepts that Rett & Murray (2013), Rett (2017) allude to: *recency* and *speaker-orientedness*. While these two properties are not inherent to all miratives that I am aware of (*doch*, for example, seems to be a counter example to both points), they are recurring themes throughout the mirative literature. I leave this area unsettled, but with aspirations toward future endeavors in these topics.

7.2 Mirativity and Anti-mirativity

One of the major points of investigation in this work concerns how miratives strategies in different languages are used to guide the receipt and acknowledgement of information by the speaker. A lot has been said about the particular strategies that are used in English and German, but a topic that has not been broached throughout this discussion is why languages would have mirative strategies at all. And further, for these languages and these systems, why is there such an abundance of mirative particles and a lack of anti-mirative particles?

One simple explanation for the variety and variation in mirative markers comes from the nature of violating expectation in itself. In a complex network of expectation and belief structures, there are many different ways that something can violate the expectations that we have built for ourselves. In many cases, this is out of our control, and subject to the context that we find ourselves in. We are constantly in the process of adding unknown variables to our models, in the forms of our conversational partners or new situations. If Expectation can indeed be defined in terms of a minimum function quantifying over all propositions in the context set, if even one of these has a low conditional probability, we conclude that expectations were violated. Recall the definition from Chapter 2:

(302) Expectation for q by a speaker α

$$\mathrm{Exp}_\alpha(q) = \min_{p \in CS} P_\alpha(p,q)$$

From this, we say that a speaker believes or expects q if $\mathrm{Exp}_\alpha(q) \approx 1$. Under this view, violated expectations are sensitive to even one low conditional probability between it and another contextually relevant proposition. It follows then that from the way that we have set up the notion of expectation that there are many ways in which a proposition or fact about the world can be expectation-violating to an actor in a discourse. From this we might conclude that the variation that we see both within and across languages for expressing mirative meanings is not an accident.

The question then, is how can we derive the apparent restriction on so-called anti-miratives, or strategies that speakers use in order to mark a proposition as known, given or expected? In terms of the definition for expectation, this too, seems to fall out. In order to avoid violated expectations, it is necessary for all propositions in the context set (p) to have high conditional probability relative to the at-issue proposition (q). What this means is that our established expectations (ps in the context set) are expected with respect to the at-issue proposition (q)—very simply, there is nothing more to comment on when all is as we expect.

If we think of anti-miratives as representing the default—the way we believe the world should work given our prior information—it seems odd or even redundant to overtly signal this state of the discourse. We can think of this as the unmarked case. Miratives are a way to flag the marked cases. But recall German *ja*, which is perhaps the closest strategy in this work to be considered an anti-mirative. We might expect that this kind of marker to mark redundant information, as a sort of Gricean Maxim of Quantity violation. Why should this information be marked when not marking it would have the very same effect of treating it as expected? It turns out that marking anti-miratives has a clear discourse function as well, in that they can serve as a marker of propositional not-at-issueness. In the case that we have looked at in this work, using an anti-mirative is a way for a speaker to structure a discourse in a particular way. By making explicit reference to the common ground status of all propositions being considered, a speaker can indicate to others that the content of the utterance is less important to the discourse than the way the proposition fits into the discourse.

7.3 What can we gain from an experimental approach?

Developing a methodology for research on discourse particles and intonation is difficult for a variety of reasons. One very clear issue that arises is in the nature of the comparisons. While it is not impossible, doing work with auditory stimuli presents various issues, both when using stimuli taken a single speaker and when using multiple. Production od audio stimuli is much more messy than our theoretical analyses of it make it out to be. Not only do we have to make sure that speakers are clear, that their productions are correct, and that their performances are comparable, we also have to reconcile with various quirks or irregularities of the speech stream that might be unique to our particular speaker. These issues are particularly relevant when meaning and intonation are being investigated, as it is sometimes up to the experimenter's best judgment whether a given token's effect can be attributed to a hypothesized piece of the prosody, or whether it is a confluence of other paralinguistic factors, or both.

One concrete difficulty in the data presented here comes from particle and contour pairs. Even particles and utterances that listeners might perceive as having identical contours can have significant pitch height differences, vowel durations or overall time-course measures. Discrepancies in any of these could lead a listener to assume informationally relevant material where none was intended. Of course, there are ways that we can go about preventing this noise in the signal, but none of them seem to be able to capture just what it is that we are looking for without losing other potentially meaningful information. One way to prevent discrepancies is to elicit neutral or "flat" utterances and artificially manipulate more prosodically complex contours on top. This is a methodology used by Jeong & Potts (2016) in order to normalize rising declarative utterance-final pitch height. Such a method leads to similar recordings that can be very precisely measured against one another, but perhaps at the cost of sounding like natural speech. As a level of utterance-level contributions, this seems like a small price to pay. But at a smaller level of analysis, as it is when we look at discourse particles and the interactions that they have with intonation, sacrificing naturalness in the presentation of the utterances seems to be a large price to pay.

One way to potentially minimize this effect in naturally produced speech is to rigorously train voice talents on the features of speech that will be important in an experimental investigation. Hayes (1994) shows (anecdotally) that even non-actors who have been trained to hear the difference between particular prosodic shapes perform reliably better than naive production of the same material (p. 73). But again, we run into the issue of naturalness: if actors have been trained to produce these utterances in a particular way for a particular purpose, does this really track with what we find in natural speech?

Once we are able to solve the issue of comparability across contours, the question shifts to how to probe a listener's intuitions about the naturalness of the utterances. In the pilot experiment in Chapter $ 4, participants heard utterances out of context, which allowed for very generous acceptability and naturalness ratings. Tokens that the researchers assumed would elicit negative naturalness ratings were easily accommodated with no context provided. Placing these utterances into specific contexts would account for this, and presumably force interpretations intended by the researchers. This is what is shown in Experiment 3, and results here are extremely clean, but this of course might be due to too much experimenter manipulation. There is something to be said about introducing naive listeners to utterances out of context; because many of the meanings being probed in such studies are limited only contextually, and their effects are in many cases diffusible, it's easy for a listener to accommodate something that a researcher might not actually be investigating. But too much lack of contextualization leads to listeners interpreting utterances in ways that were not intended, or in some cases, not anticipated. Future research seeks to find the balance between forced contexts and no context situations.

A final question is one that all experimental studies must face: what is the best way to represent the data? Especially on impressionistic, semantico-pragmatic judgments, asking the right questions that probe the intended intuitions is a difficult problem in itself. Quantifying these after the fact in a statistically valid way is an added challenge.

Phonetically and prosodically grounded work in intonational semantics puts armchair theoretic judgments to the test: can intuitions about felicity and theoretical conceptualizations of the contributions of a particles and contours be verified experimentally? If they can, then our theories can (for now) rest easy. If not, this is also positive for our pragmatic theories of language understanding, as it shows that we must approach the problem from yet another angle.

7.4 Looming in the distance: Syntax-Prosody interactions

One of the key issues that has been entirely absent from the discussion of discourse particles, modal particles and intonation is the how these elements fit in to a larger structure. Part of the reason for this is because it is not a trivial question. Restricting the view to German for the moment, the question becomes a bit more manageable. Modal particles like *doch* and *ja* in the language are only licensed in the Mittlefeld of a clause, and as such, must somehow slot in to an overarching syntactic structure. Most agree that these elements are adverbial in nature, and as such, there has been a good number of pointed investigations into where exactly in the clausal spine they sit. Viesel (2015) shows convincingly that these elements can show up in a position between vP and CP, but that these positions are only reliable for matrix clauses. Modal particles in embedded clauses and nominal phrases have another distribution. Work by Coniglio (2007) shows how a handful of these particles slot in into a Cinque (1999)-style cartographic framework of the clausal spine, arguing that there is reason to believe that a structured subclass of these particles are base-generation in the specifier of MoodP, above VP and below TP. But while these inquiries give us a detailed look at the positioning of these elements within the Mittelfeld, it leads us to ultimately wonder, why should we expect the *Mittelfeld* in German to be the host of utterance-level modifiers?

One way to approach this is to think about what it is exactly that the different parts of a German clause contribute in terms of information structure. The Vorfeld, or pre-verb region, is very informational structurally prominent. The specifier of CP is generally thought to be a topic position, both for A and A' moved phrases, which is a host for information structural as well as prosodic prominence. The Nachfeld, or post-verb region in a clause, is perhaps less clearly a host for prominent phrases or heads in a clause, but is nevertheless a marked position. This is the position of right-dislocated phrases, and of afterthoughts, as seen in the examples below, taken from (Truckenbrodt 2013):

(303) *Ich habe sie gesehen [die Maria]$_{NF}$.*
I have her seen, the Maria
I saw her, Maria. *Right Dislocation*

(304) *Er hat ein Buch gelesen [von Chomsky]$_{NF}$.*
he has a book read from Chomsky
He read a book, of Chomsky's. *Afterthought*

What we see for these two positions is that they are syntactically easy to pinpoint. Attachment sites are unambiguous, and information can be packaged in these sites in a structured way. In contrast, the Mittelfeld is much less syntactically and information structurally marked. In general, it is a position where one expects to find scrambled phrases and deaccented pronouns, and where constituent order is thought to be relatively free. It is the part of the German clause that is maximally ambiguous—in a sense, a syntax-free zone. When thought about in this way, we can form a hypothesis as to why the Mittelfeld is the most likely position to host scope-taking modifiers, and especially those that contribute not-at-issue content. As utterance-level modifiers, we expect these elements to arise in a position where deaccented content is generally found, where attachment is the most syntactically ambiguous, and where the configuration has the least propensity for information structurally-relevant positioning. Perhaps there is some deeper correlation here between the anaphoric nature of these particles, and their ability to reference facts outside of the immediate spoken content that leads it to want to appear in the least marked position possible. Since they does not pick out a single constituent or phrase in order to realize their discourse effect, appearing in the most neutral position possible would be the preferred placement. In the case of German, the Mittelfeld is the perfect catch-all: a place that is the lowest informationally marked, where deaccented and backgrounded material is found, and where attachment is maximally ambiguous.

So for German, this answers part of our question. Yes, they are present in the syntax in a generically specified position, and in contrast to the English particles, don't co-occur with prominent positions in the utterance. But we cannot tell the same story with the English particles. Here, we see that not only are *oh, huh*, and *what* able to host entire prosodic idioms, they are also detached from the clause in a way that we do not see for the German particles. English discourse particles appear utterance-initially, and are independently prosodically phrased.

In one sense, this might be the English way of dealing with the attachment ambiguity issue that we hypothesize is solved in with the Mittlefeld solution for German. If we take prosodic phrasing as being fed by informationally-relevant pieces of the syntax, then we might conclude that the discourse particle's phrasing outside the prosodic phrasing of the following utterance is a way of bypassing the question of attachment as well as sidestepping the issue of placement in informationally relevant positions (English, like German, attaches topicalized elements in high positions in the clause). If the particle were phrased separately, this could take care of these issues, as well as signal that the particle takes utterance-level scope. If we believed this, we could also use prosodic phrasing to skirt around the issue of where this element attaches to the larger syntactic structure. Since it is in its own prosodic phrase, we can ostensibly not include it as a part of the structure of the following utterance. But accepting this reasoning leads us right back to where we started, and no closer to an answer as to what kind of syntactic element these things are, or even how they integrate into the syntax at all. In fact, if we assume, as mentioned above, that syntax informs prosodic phrasing, what can we make of these particles, which seem to be syntax-independent, yet prosodically complete?

But perhaps their independent prosodic phrasing does in fact lend some insight into their integration into a larger clause. As a prosodic adjunct, these elements would be assumed to be adjoined to the highest intonational phrase level of the subsequent clause. This could perhaps help to explain their pragmatic scope, as it would be to the highest node in the adjoined structure. But again, if syntactic positions are taken to inform prosody, how is it that this intonational idiom is prosodified on this particle?

Even if we could come to an agreement on the syntactic category or attachment site of these English discourse particles, one final wrench in the gears for any syntactic analysis is the difference between intonationally independent particles, which we have dealt with exclusively here, and particles that have been prosodically integrated into the clause. Recall the difference in interpretation that arises in using (305) and (306) in different contexts:

(305) Oh! I put my keys on the hook last night. Independent

(306) Oh I put my keys on the hook last night. Integrated

In (305), the particle is phrased independently from the following utterance, while in (102), *oh* is integrated into the overall utterance as a sort of cliticized element of the overall prosodic phrase. Their cases of use overlap in most cases, but there are instances where the integrated and the independent particles pattern differently. If Jeff isn't already looking for his keys, and walks around the corner and sees them hanging on a hook, he can utter (305), but (306) is odd. Discourse initially, we see that only the independent particle is felicitous. However, if someone asks Jeff where his keys are, he can respond with either of the responses above. Given a lack of understanding about the overall syntax of these particles, this piece of variation in their performance and prosodic phrasing is troubling. We must add this, too to the growing list of questions regarding their syntax: how can we handle the difference in the syntactic variation between integrated and independent particles? And also, how can we tell if there is any?

7.5 Outlook and ways forward

This paper is the start of a larger research program focusing on English and German discourse particles and the varying interpretations they are given when paired with prosodic contours. It has argued that any analysis of discourse particles must be sensitive to their prosodic environment. Similarly, it has lobbied for a narrower view of the semantic interpretation of prosodic contours, asserting that the meanings of complex contours cannot be determined by analysis of their component parts.

In addition to these broader claims, this work has identified three discourse particles and one prosodic contour that serve as mirative strategies in English. This is a significant finding, both in terms of better understanding the contribution of the discourse particles *oh, huh* and *what*, and also in terms of the general realm of Mirativity. This work fits English into the existing typology of other diverse languages with mirative markers and strategies.

Another contribution this paper has attempted to make is to differentiate the contributions of the discourse particles *oh, huh* and *what* in different prosodic and discourse contexts. Though the particles do have significant areas where they overlap, they are not the same. This has been shown with impressionistic judgments, and have been preliminarily validated through experimentation utilizing a novel setup. The work here is ongoing; though claims about which discourse contexts these particles and contour combinations are sensitive to have not been validated, it is clear that even when contexts are not strictly controlled for, *oh, huh* and *what* have nonidentical conditions of use. Future research seeks to verify the broader claims made for English, as well as extend the reasoning and design to the realm of German modal particles. There is much work still left to do.

Appendix A

The following twelve sentences were elicited by a single male speaker of Standard American English. Particles were then cut and saved into their own files, and then cross-spliced with all sentences in order to produce 144 audio files fulfilling 36 experimental conditions. Final audio files can be found here:
`https://www.dropbox.com/sh/hqnsmz3505ifwcs/AAAdBE2YHbrSeCgTB5xkJvvBa?dl=0`.

Note that the experiment originally included an extra particle condition, *god*, which was ultimately excluded from the final analysis. The researchers originally assumed that this particle functioned in a similar way *huh, what* and *oh*, but closer inspection of free-response questions by participants in the experiment showed that this particle introduced an overlooked confound. The similarity of the discourse particle *god* to the phonologically identical deity name enabled many participants to assume a vocative reading of the particle, resulting in skewed ratings for particle + sentence pairings which were expected to be otherwise judged infelicitous.

<u>Sentences and particles with neutral prosody:</u>

(1) Huh. We're out of flour.

(2) Oh. I put the toolbox in the shed.

(3) (#) What. I'm on the phone.

(4) God. Another pile of papers to grade.

<u>Sentences and particles with surprise-redundancy prosody:</u>

(5) Huh? You don't know how to tango.

(6) Oh? I thought apples were in the sunflower family.

(7) What? She's never done that before.

(8) (#) God? I've never heard that before.

Sentences and particles with exclamative/surprised prosody:

(9) Huh! He really likes it!

(10) Oh! A package arrived for you!

(11) What! You can't be serious right now!

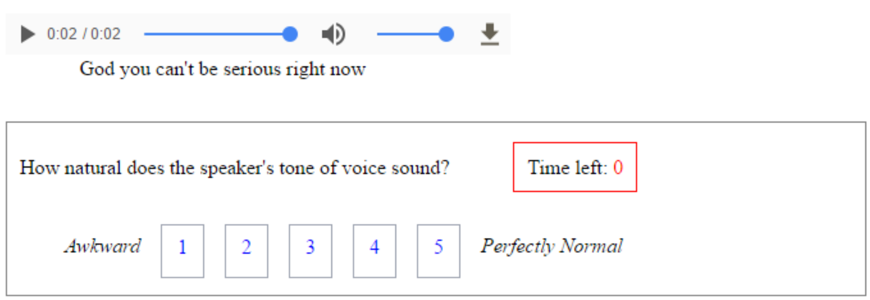
(12) God! There are a lot of cakes to choose from!

Participants were given the following instructions prior to receiving two audio files:

After you listen to the following audio clip, you will have 5 seconds to judge the utterance. Use the mouse to select your choice. You will then be asked three additional questions. After the initial question, you will be able to listen to the audio again as many times as you like. To continue, press any key.

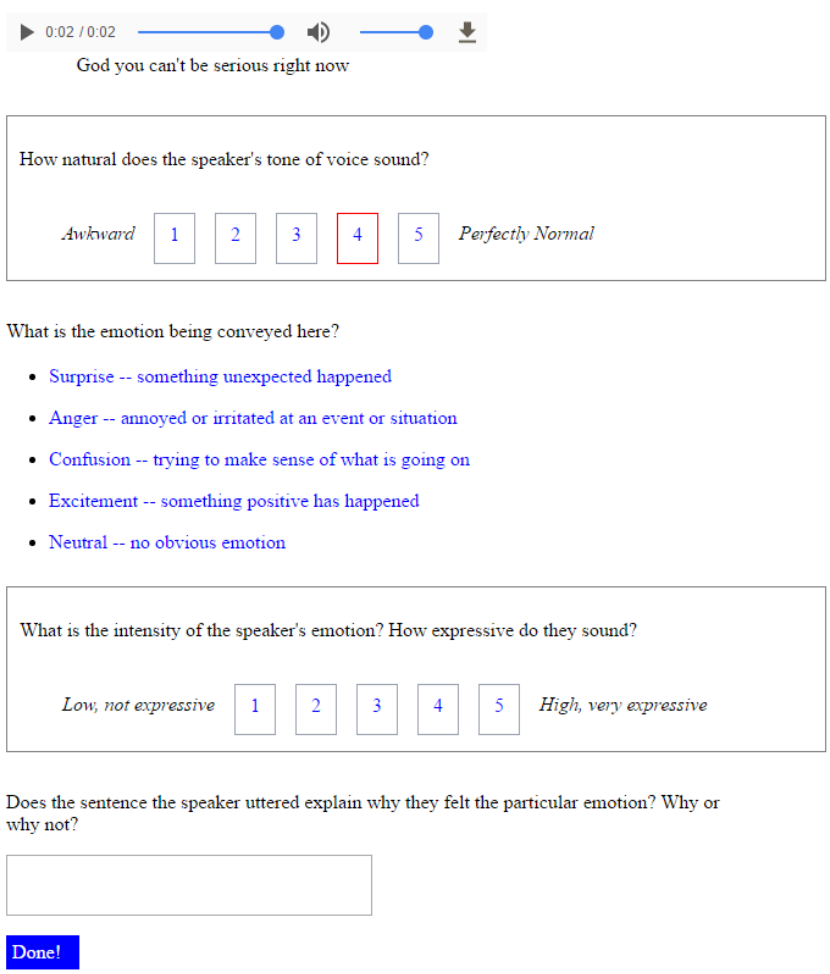
After a key press, participants were directed to a new screen with an audio file. They were prompted to click play, and after listening to the audio clip, the text of the clip (punctuation removed) and an initial question were presented. Participants were given five seconds to respond to this first question, as in Figure .1. Participants were not allowed to change this response.

Figure .1: Initial presentation of audio and Likert response scale



Upon answering the first question, three more questions were revealed, and the participants had the option to replay the clip as many times as they felt necessary. Each play was recorded in the data file. The questions first ask the participant to judge the overall emotion of the sentence out of five potential responses. Participants could change their selected response. Each response change was recorded. They were also asked to judge the intensity of the emotion they heard, on a scale from “Low, not expressive” (1) to “High, very expressive” (5). These answers, too could be changed, with the changes recorded. The last question was a free response, which asked for impressionistic evaluations regarding the sentence the speaker uttered and the emotion that they logged. This is shown in Figure .2.

Figure .2: Audio clip and all subsequent questions



Appendix B

List of items used in Experiment 2 and Experiment 3:

(1) They spread butter on the sourdough.

(2) She poured coffee grounds on the flower beds.

(3) He made pizza for the potluck.

(4) They set up parasols in the garden.

(5) She grew sunflowers in the greenhouse.

(6) He sent an email to the band members.

(7) They served cocktails at the ballgame.

(8) She used a towel as a bath mat.

(9) He painted turtles on the moonroof.

(10) They used school busses for the ski trip.

(11) She climbed a mountain with a guide dog.

(12) He took chicken feed from the rooster cage.

(13) They wore beanies to the concert.

(14) She slipped case files under the news reports.

(15) He picked wildflowers for the centerpiece.

(16) They kicked branches off the building.

(17) She traded pennies for a lollipop.

(18) He taped messages to the blackboard.

(19) They built shelters in the bushes.

(20) She drew portraits in the waiting room.

(21) He mailed parcels to the senators.

(22) They gave raises to the secretaries.

(23) She stamped visas for the tourists.

(24) He wore khakis to the wedding.

These audio files can be found at the following link:

`https://people.ucsc.edu/~knkraus/ABX/audio/`.

Experiment 2 can be previewed here:

`https://people.ucsc.edu/~knkraus/ABX/main.html`

Experiment 3 can be previewed here:

`https://people.ucsc.edu/~knkraus/ABX2/main.html`

Bibliography

Abraham, Werner. 1991. Discourse particles in German: How does their illocutionary force come about? In Werner Abraham (ed.), *Discourse particles*, 203–252. John Benjamins.

Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. John Benjamins.

Aikenvald, Alexandra. 2004. *Evidentiality*. Oxford University Press.

Aikenvald, Alexandra. 2012. The essence of mirativity. *Linguistic Typology* 16. 435–485.

AnderBois, Scott. 2016. Illocutionary mirativity: the case of Yucatec Maya *bakáan*, SULA 9.

AnderBois, Scott. 2017. Illocutionary revelations: Yucatec maya bakáan and the typology of miratives. *Journal of Semantics* .

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

Austin, John. 1962. *How to do things with words*. Harvard University Press.

Baker, Rachel & Ann Bradlow. 2009. Variability in Word Duration as a Function of Probability, speech style, and prosody. *Language and Speech* 52. 391–413.

Baldi, Pierre & Laurent Itti. 2010. Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks* 23. 649–666.

Bartels, Christine. 1999. *The intonation of English statements and questions*. Garland Publishing.

Beaver, David. 2001. Presupposition and assertion in dynamic semantics. *Stanford, CA: Center for the Study of Language and Information (CSLI), StanfordUniversity* .

Beaver, David & Brady Clark. 2008. *Sense and Sensitivity: How focus determines meaning*. Oxford: Wiley Blackwell.

Beckman, Mary & Gayle Ayres Elam. 1997. Guidelines for ToBI Labeling, version 3.

Beddor, Bob & Simon David Goldstein. 2017. Believing epistemic contradictions. *The Review of Symbolic Logic* .

Bell, Alan, Jason Brenier, Michelle Gregory, Cynthia Girand & Dan Jurafsky. 2009. Predictability Effects on Durations of Content and Function Words in Conversational English. *Journal of Memory and Language* 60. 92–111.

Beltrama, Andrea & Emily Hanink. 2017. From marking imprecision to expressing surprise: *like* between hedging and mirativity. *Ms., Universität Konstanz and University of Chicago* .

Brown, Penelope & Steven Levinson. 1987. Politeness .

Chen, Ajou. 2005. *Universal and language-specific perception of paralinguistic meaning*: dissertation.

Cinque, Guglielmo. 1999. *Adverbs and functional heads. a cross-linguistic perspective*. New York, Oxford University Press.

Condoravdi, Cleo & Sven Lauer. 2012. Imperatives: Meaning and illocutionary force. *Empirical Issues in Syntax and Semantics* (9). 37–58.

Coniglio, Marco. 2007. German modal particles in the ip domain. *Rivista di Grammatica Generativa* 3–37.

Cooper, William E., Stephen J. Eady & Pamela R. Mueller. 1985. Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America* 77(6). 2142–2156.

Csipak, Eva & Sarah Zoebel. 2016. Discourse particle denn in the Antecedent of Conditionals. *Empirical Issues in Syntax and Semantics 11* 31–60.

Cuenca, Maria Josep. 2013. The fuzzy boundaries between discourse marking and modal marking. In Liesbeth Degand, Bert Cornille & Paola Pietrandrea (eds.), *Discourse markers and modal particles: Categorization and description*, 191–216. John Benjamins.

Davidson-Nielson, Nils. 1996. Discourse particles in Danish. *Content, Expression and structure: Studies in Danish Functional Grammar* 283–314.

DeLancey, Scott. 1998. Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1. 33–52.

Diewald, Gabrielle. 2007. Abtönungspartikeln. In Ludger Hoffmann (ed.), *Handbuch der deutschen wortarten*, 117–142. de Gruyter.

Doherty, Monkia. 1982. Doch.

Egg, Markus & Malte Zimmermann. 2016. Stressed out! accented discourse particles: the case of 'doch'. *Proceedings of SuB 16* .

Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*, vol. Cambridge Studies in Linguistics 84. Cambridge; New York: Cambridge University Press.

Faller, Martina. 2006. Evidentiality and Epistemic Modality at the Semantics/Pragmatics Interface. *Journal of Phonetics* 30(4).

Farkas, Donka. 2010. The grammar of polarity particles in Romanian. *Edges, Heads, and Projections: Interface properties* 87–124.

Farkas, Donka & Kay Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27. 81–118.

Farkas, Donka & Floris Roelofsen. 2012. Polar initiatives and polarity particles in an inquisitive discourse model (. *UCSC* .

Farkas, Donka & Floris Roelofsen. 2017. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics* 1–53.

Féry, Caroline. 1993. *German intonational patterns*. Tübingen: Max Niemeyer Verlag.

Féry, Caroline. 2011. German sentence accents and embedded prosodic phrases. *Lingua* 121. 1906–1922.

Fox Tree, Jean E. & Josef C. Schrock. 1999. Discourse markers in spontaneous speech: Oh what a difference an oh makes. *Journal of Memory and Language* (40). 280–295.

Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14. 383–95.

Ghezzi, Chiara & Piera Molinelli. 2014. Deverbal pragmatic markers from Latin to Italian (Lat. quaeso and It. prego): The cyclic nature of functional developments. In Chiara Ghezzi & Piera Molinelli (eds.), *Discourse and pragmatic markers from latin to the romance languages*, Oxford University Press.

Ginzburg, Jonathon & Ivan Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, Calif.: CSLI Publications.

Goodhue, Daniel, Lyana Harrison, Yuen tung Clémentine Siu & Michael Wagner. 2015. Toward a bestiary of English intonational contours. *Proceedings of the 46th Conference of the North Eastern Linguistic Society* 311–320.

Grice, Paul. 1975. Logic and Conversation. *The Logic of Grammar* .

Grosz, Patrick. 2012. On the grammar of optative constructions. *Linguistik Aktuell* 193.

Grosz, Patrick. 2014. German doch: an element that triggers a contrast presupposition. 46, .

Grunow-Hårsta, Karen. 2007. Evidentiality and mirativity in Magar. *Linguistics of the Tibeto-Burman Area* 30(2). 151–194.

Gunlogson, Christine. 2001. *True to form: rising and falling declaratives as questions in English*: UC Santa Cruz dissertation.

Gunlogson, Christine. 2008. A question of commitment. *Commitment: Belgian Journal of Linguistics* 22. 101–136.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, Carlos. 2002. Intonation and Interpretation: Phonetics and Phonology. *In Proceedings of Speech Prosody 2002* .

Gutzmamnn, Daniel. 2009. Hybrid semantics for modal particles. *Sprache und Datenverarbeitung* 33:1-2. 45–59.

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London: Routledge.

Halpern, Joseph Y. 2005. *Reasoning about uncertainty*. MIT Press.

Halpern, Joseph Y. & Judea Pearl. 2005. Causes and explanations: A structural-model approach. part i: Causes. *British Journal for the Philosophy of Science* 56(4). 843–887.

Halpern, Joseph Y. & Riccardo Pucella. 2007. Causes and explanations: A structural-model approach. part I: Causes. *Journal of the ACM* 56(4).

Hayes, Bruce. 1994. "gesture" in prosody: comments on the paper by Ladd. *Phonological Structure and Phonetic Form: Papers in Labrotory Phonology II* 64–75.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and case studies*. Chicago: The University of Chicago Press.

Hentschel, Elke. 1986. *Funktion und Geschichte deutscher partikeln 'ja', 'doch', 'halt', und 'eben'*. Max Niemeyer Verlag.

Heritage, John. 1984. *A change-of-state token and aspects of its sequential placement* 299–345.

Hinterhölzl, Roland & Manfred Krifka. 2013. Modal particles in adverbial and adnominal clauses. *Ms., University Ca'Foscari Venezia, Humboldt Universität zu Berlin* .

Hintikka, Jaako. 1969. Semantics for propositional attitudes. *Philosophical Logic* 21–45.

Hirschberg, Julia. 1993. Pitch accent in context: predicting intonational prominence from text. *Artificial Intelligence* 63. 305–340.

Hirschberg, Julia. 2002. The Pragmatics of Intonational Meaning, Aix-en-Provence.

Horn, Larry. 1969. A presuppositional approach to only and even, 98–107. Chicago: Chicago Linguistic Society.

Hölker, Klaus. 1990. Französisch: Partikelforschung, 77–88. Tübingen.

Jeong, Sunwoo. 2017. Intonation and Sentence Type Conventions: Two types of rising declaratives. *Presented at the UCSC/Stanford Workshop on Clause Types, 29 January 2017* .

Jeong, Sunwoo & Christopher Potts. 2016. Intonational sentence-type conventions for perlocutionary effects: An experimental investigation. *SALT 26* 1–22.

Kaiser, Elsi. 2015. Perspective-shifting and free indirect discourse: Experimental investigations, vol. 25, 346–372.

Kamp, Hans. 1981. A theory of truth and semantic representation. *Formal Methods in the Study of Language.* .

Karagjosova, Elena. 2001. Towards a comprehensive meaning of German doch, 131–142. ESSLI 2001.

Katz, Jonah & Elizabeth Selkirk. 2011. Contrastive Focus vs. Discourse-New: Evidence from Phonetic Prominence in English. *Language* 87:4. 771–816.

Kaufmann, Magdalena. 2012. *Interpreting imperatives*. Springer.

Kaufmann, Magdalena & Stefan Kaufmann. 2012. Epistemic particles and performativity. *Proceedings of SALT 22* 208–225.

Krahmer, Emiel & Marc Swerts. 2001. On the alleged existence of contrastive accents. *Speech Communication* 34:4. 391–405.

Kratzer, Angelica. 1999. Beyond Ouch and Oops: How descriptive and expressive meaning interact, UMass.

Kratzer, Angelica & Lisa Matthewson. 2009. Anatomy of two discourse particles. *Handout, SULA 5 at Harvard/MIT* .

Kratzer, Angelika. 1991. Modality. *Semantics: An International Handbook of Contemporary Research* 639–50.

Krifka, Manfred. 2013. Response particles as propositional anaphors. In *Proceedings of SALT 23*, 1–18.

Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. In *Proceedings of SALT 25*, 328–245.

König, Eberhard. 1991. Identical values in conflicting roles: the use of German ausgerechnet, eben, genau and gerade as focus particles. In *Discourse particles*, 11–36. John Benjamins.

Ladd, Robert. 1980. The Structure of Intonational Meaning: Evidence from English .

Ladd, Robert. 2008. *Intonational phonology*. Cambridge Studies in Linguistics.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75(3). 522–551.

Lassiter, Daniel. 2011. *Measurement and modality: the scalar basis of modal semantics*: New York University dissertation.

Lassiter, Daniel. 2012. Presuppositions, provisos, and probability. *Semantics & Pragmatics* 5. 1–37.

Lassiter, Daniel. 2017. Bayes nets and the dynamics of probabilistic language. *Proceedings of Sinn und Bedeutung 21* `http://web.stanford.edu/~danlass/Lassiter-SuB-21.pdf`.

Lauer, Sven & Prerna Nadathur. 2017. Causal necessity, causal sufficiency, and the implications of causative verbs. *Manuscript* .

Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8. 249–336.

Lindner, Katrin. 1991. "Wir sind ja doch alte Bekannte." the use of German ja and doch as modal particles. In Werner Abraham (ed.), *Discourse particles*, John Benjamins.

Livescu, Michaela. 2014. Ma rog: A pragmatic marker in Romanian. In Chiara Ghezzi & Piera Molinelli (eds.), *Discourse and Pragmatic Markers from Latin to the Romance Languages*, Oxford University Press.

MacMillan, Neil A. & C. Douglas Creelman. 2005. *Detection theory: A user's guide*, vol. 2. Lawerence Erlbaum Associates, Inc.

Malamud, Sophia & Tamina Stephenson. 2015. Three Ways to Avoid Commitments: Declarative Force Modifiers in the Conversational Scoreboard. *Journal of Semantics* .

Matthewson, Lisa, Hotze Rullmann & Henry Davis. 2007. Evidentials as epistemic modals: Evidence from St'at'imcets. *The Linguistic Variation Yearbook* 7.

Meriçli, Benjamin. 2016. *Modeling indirect evidence*: University of California Santa Cruz MA thesis.

Mithun, Marianne. 2001. *The Languages of Native North America*. Cambridge University Press.

Murray, Sarah. 2017. Evidentiality and Illocutionary mood in Cheyenne. *International Journal of American Linguistics* 83:4.

Müller, Sonja. 2014. Zur Anordnung der Modalpartikeln ja und doch. *Linguistische Berichte* 238.

Northrup, Oliver. 2014. *Grounds for Commitment*: UC Santa Cruz dissertation.

Ohala, John J. 1983. Cross-language use of pitch: an ethological view. *Phonetica* 40:1. 1–18.

Ouyang, Iris & Elsi Kaiser. 2015. Individual differences in the prosodic encoding of informativity, .

Ouyang, Iris, Sasha Spala & Elsi Kaiser. 2017. Speakers' rapidly-updated expectations influence prosodic realization of information structure, .

Papafragou, Anna, Peggy Li, Youngon Choi & Chung hye Han. 2007. Evidentiality in language and cognition. *Cognition* 103:2. 253–299.

Peirce, Charles S. 1955. Abduction and induction. *Philosophical Writings of Pierce* .

Peterson, Tyler. 2015. Problematizing mirativity. In *The linguistic expression of mirativity*, John Benjamins.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation.*: MIT dissertation.

Pierrehumbert, Janet & Julia Hirschberg. 1990. The Meaning of Intonational Contours in the Interpretation of Discourse. In P. Cohen, J. Morgan & M. Pollack (eds.), *Intentions in communication*, 271–311. Bradford Books, MIT Press, Cambridge MA.

Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. *Proceedings of SALT 14* 235–252.

Prieto, Pilar. 2015. Intonational meaning. *WIREs Cogn Sci* 2. 1–11.

Reitveld, Toni & Carlos Gussenhoven. 1985. On the relation between pitch excursion size and pitch prominence. *Journal of Phonetics* 15. 273–285.

Rett, Jessica. 2009. A degree account of exclamatives. *Proceedings of SALT 18* 608–681.

Rett, Jessica. 2017. The semantics of attitude markers and other illocutionary content. *Handout for the 6th Cornell Workshop on Linguistics and Philosophy* .

Rett, Jessica & Sarah Murray. 2013. A semantic account of mirative evidentials. *Proceedings of SALT 23* 453–472.

Rinas, Karsten. 2007. Abtönungspartikeln-Kombinationen und Skopus. *Sprachwissenschaft* 32(4). 407–452.

Roelofsen, Floris & Donka Farkas. 2015. Polarity particle responses as a window onto the interpretation onto the interpretation of questions and assertions. *Language* 91(2). 359–414.

Rojas-Esponda, Tanja. 2013. A QUD account of German doch. *Proceedings of Sinn und Bedeutung 18* .

Rooth, Mats. 1985. *Association with Focus*: UMass Amherst dissertation.

Rooth, Mats. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 75–116.

Rudin, Deniz. 2017a. Intonation as a Speech Act Modifier: Rising Declaratives and Imperatives. *Handout for Questioning Speech Acts 17 Conference, Konstanz* .

Rudin, Deniz. 2017b. Uncertainty and persistence: a bayesian updatesemantics for probabilistic expressions. *Journal of Philosophical Logic* 1–41. https://link.springer.com/article/10.1007/s10992-017-9431-4.

Sag, Ivan & Mark Lieberman. 1975. The intonational disambiguation of indirect speech acts. *Chicago Linguistics Society* 487–497.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schultz, Katrin. 2011. "if you'd wiggled a, then b would've changed": Causality and counterfactual conditionals. *Synthese* 179. 239–251.

Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Silverman, Kim, Mary Beckman, John Pitrelli, Marie Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert & Julia Hirschberg. 1992. ToBI: A Standard for Labeling English Prosody. In *Proceedings of the International Conference on Spoken Language Processing, Banff, Alberta*, 867–870. University of Alberta.

Stalnaker, Robert. 1978. Assertion. *Syntax and semantics* 9. 315–332.

von Stechow, Arnim. 1996. Focusing and Backgrounding operators. *Arbeitspapier 6 der Fachgruppe Sprachwissenschaft* 1.

Swerts, Marc, Emiel Krahmer & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: a comparative analysis. *Journal of Phonetics* 30(4).

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Max Niemeyer Verlag.

Toosarvandani, Maziar. 2013. Corrective *but* coordinates clauses not always but sometimes. *Natural Language and Linguistic Theory* 31. 827–863.

Truckenbrodt, Hubert. 2013. Some distinctions in the German Nachfeld.

Turco, Giuseppina, Christine Dimroth & Bettina Braun. 2012. Intonational means to mark verum focus in german and french. *Language and Speech* 56:4. 461–491.

Uldall, Elizabeth. 1960. Attitudinal meanings conveyed by intonation contours. *Language and Speech* 3. 223–234.

Viesel, Yvonne. 2015. Discourse Structure and Syntactic Embedding: The German Discourse Particle *ja*, 418–427.

Waltereit, Richard. 2001. Modal particles and their functional equivalents: A speech-act theoretic approach. *Journal of Pragmatics* 33.

Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The Pragmatics of Fall-Rise Intonation. *Language* 61(4). 747–776.

Wennerstrom, Ann. 2001. *The Music of Everyday Speech: Prosody and Discourse Analysis*. Oxford University Press.

Winans, Lauren. 2016. *Inferences of "will"*: UCLA dissertation.

Mei Lee Wong, Song. 1993. Face-support: Chinese particles as mitigators: a study of ba a/ya and ne. *Pragmatics* 8(3). 387–404.

Xu, Yi & Ching X Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33. 159–197.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916â937.