Chapter Four

Interpreting utterances with modal particles and tracking referents in the online processing of referring expressions

4.1 Introduction

In Chapter 3, I presented a study focused on whether comprehenders predictively process utterances containing morphosyntactic cues to epistemic stance. More specifically, I investigated whether comprehenders draw predictive inferences about upcoming lexical information when processing negated polar questions in German and English with different morphosyntactic forms. Crucial for my analysis was the assumption that different forms of a negated polar question – such as, in English, *Didn't you buy a ticket?* vs. *Did you not buy a ticket?* – signal different epistemic meanings, either a commitment to the truth of the proposition expressed in the question in the case of the former or no commitment to that same state of affairs in the case of the latter. In the present chapter, I present a study focused on a different class of modal markers, namely lexical rather than morphosyntactic markers of stance.

In contrast to the morphosyntactic cues investigated in Chapter 3, here I investigate

German adverbs which serve a specialized function as modal particles, i.e., as markers of common ground relations and intersubjective stance. I start by reviewing previous work on the function and interpretation of discourse particles, with a particular focus on particles which serve modal functions like German *eigentlich* and *tatsächlich*, before turning to the empirical study which makes up the present chapter.

4.1.1 Discourse particles and the marking of modality

Discourse markers are pervasive communicative devices found across the languages of the world (Heine et al., 2021). As their name indicate, these are devices used to manage linguistic communication and interaction itself (Fedriani and Sansó, 2017; Jucker and Ziv, 1998; Maschler and Schiffrin, 2015), covering a multitude of functions and meanings, from more subjective, attitudinal ones to fundamentally intersubjective, interaction-oriented ones. Not only do discourse markers, as a class, cover a wide range of functions, individual discourse markers themselves also tend to be polyfunctional, their exact meaning depending on both the utterance they occur in as well as the larger discourse context (see refereces above), as in the examples below which show the flexibility of the English marker *really*.

A: Did you see how fast that bird flew by?

(1) B: I *really* don't think it flew by that fast.

(2) B: I don't *really* think it flew by that fast.

(3) B: I don't think it flew by that fast, *really*.

A: Did you see how fast that bird flew by?

(4) B: *Really*? I didn't even notice.

(5) B: I don't *really* think it flew by that fast.

(6) B: It *really* did fly by fast, you're right.

Really, while being formally an adverb, serves a variety of different discourse-marking functions in the examples above: in (1)-(3) it acts as a marker of (subjective) epistemic stance, indicating, in each case, B's degree of commitment to the truth of the state of affairs at issue in the sequence; in (4) it acts as a surprise/ mirative marker, prefacing B's actual response to A's statement; in (5), much like in (1)-(3), it acts as a marker of subjective stance, indicating B's disagreement to A's statement; finally, in (6), it acts as an intensifier, indicating, respectively, both B's subjective and intersubjective understanding of A's statement.

Despite different meanings being encapsulated in a single form in the case of *really*, discourse markers can have a much more restricted meaning profile depending, among other things, on the lexical category they stem from. Indeed, while *really* formally consists of an adverb, discourse markers can stem from all sorts of lexical categories, from individual words (e.g. *really*) and phrases (e.g. *in fact*) to entire sentence frames (e.g. *What surprises me*) and even non-lexical expressions (e.g. confirmation/ hesitation markers like *mhm* or *huh*). Notwithstanding the diversity of forms or even the exact linguistic status of certain discourse markers, many accounts distinguish discourse *markers* more generally from the more restricted class of discourse *particles* (see Abraham (2016) and Fischer (2006), *inter alia*, for discussions on the terminological issue of calling an item a discourse particle vs. a discourse marker). Unlike other linguistic markers, particles are usually understood as non-inflected elements which can occur in isolation and which impart meaning to other linguistic elements, from individual lexical items to entire sentences or utterances. Discourse particles can therefore be distinguished from discourse markers more generally on the basis of their form, such that uninflected lexical elements like *really* and *in fact* might be classified as particles while larger morphosyntactic constructions like *What surprises me* might not[1].

[1]Although non-lexical expressions like *mhm* or *huh* might traditionally not be considered particles or even linguistic devices in their own right, from a functional perspective, they form a continuum with other lexicalized and grammaticalized markers.

While this grants discourse particles, by definition, a high degree of meaning flexibility relative to non-lexical discourse markers, their function as discourse-managing devices is, opposed to their morphological status, their most defining feature.

In the present study, I focus my attention on discourse particles which are used to mark modal relations related to epistemicity. Given their function as markers of modality, such particles are also known in the literature as modal particles, although their exact relation to discourse particles, and by extension discourse markers, is the subject of much debate (see, e.g., Abraham (2016) and Schoonjans (2013)). Indeed, both discourse particles in general and modal particles in particular function as indexical linguistic devices, not modifying the content of an utterance but rather adding to its illocutionary dimension and relating it to other information in the discourse context. However, unlike discourse particles which primarily relate their host utterance to non-propositional information, such as in the English example in (7) below, modal particles serve a more specialized function of relating an utterance to a proposition or an speech-act alternative, as in the German example in (8) below.

(7) A: Did you see how fast that bird flew by?

B: *Right*, but that's not really surprising. It was a swallow, and, as you know, they always fly so fast.

(8) A: Hast du gesehen, wie schnell der Vogel vorbeigeflogen ist?
Have you seen how fast SING.DEF bird flew by is?
"Did you see how fast that bird flew by?"

B: Ja, aber verwunderlich ist das nicht. Das war *ja* eine Schwalbe, die
Yes, but surprising is that not. That was *ja* a swallow PL.DEF
fliegen immer so schnell.
fly always so fast.
"Right, but that's not really surprising. It was, *after all*, a swallow, and they always fly so fast."

In (7), the particle *right* is used to mark a general sense of agreement to the previous

turn while also signaling a transition of turns between A and B. While it might be said to confirm the state of affairs at issue in the sequence, i.e., the speed with which the bird flew by, its primary function is that of a response particle, much like its German counterpart *ja* at the beginning of B's response in (8). Both *right* and the turn-initial *ja* serve as overt confirmation of the at issue state of affairs expressed in the statement by A. The *ja* in the second utterance in (8), on the other hand, plays a similar yet still qualitatively different role: it signals that the utterance it is embedded in is uncontroversial, doing so by establishing that the proposition expressed in the utterance is part of the common ground between A and B, i.e., that they both know that the bird might have been a swallow, as suggested by the approximate English translation of the utterance-medial *ja*. While both *right* in the English example and the first *ja* in the German example thus function primarily as discourse-managing devices – more specifically response or expectation-managing devices – facilitating the flow of the conversation while indirectly also contributing to the update of shared epistemic information, the second *ja* in the German example, a so-called modal particle, is used explicitly to manage common ground between the interlocutors, not serving as a response particle in itself and only being felicitous because it's part of an utterance which has propositional content of its own.

Despite discourse particles being, cross-linguistically, a common means of marking relations of knowledge and truth and any uncertainty surrounding them (Heine et al., 2021), the two examples above illustrate the fact that languages differ in terms of the exact types of particles they employ in the expression of such relations. The empirical picture is such that while some languages, like English, draw mostly on manner and certainty adverbs (e.g. *really*, *right*, *apparently*), other languages, like German, rely more heavily on particles with a specialized modal function. Focusing on the modal-marking repertoire of German, there are a variety of particles dedicated primarily to the expression of epistemicity, for instance, *ja*, *doch*, and *eigentlich*, to name just a few. While all of these particles can be used to establish a relation of knowledge or belief between two propositions, they do so in different

ways. Let's look at the examples below.

A: Hast du gesehen, wie schnell der Vogel vorbeigeflogen ist?
Have you seen how fast SING.DEF bird flew by is?
"Did you see how fast that bird flew by?"

(9) B: Ja, aber verwunderlich ist das nicht. Das war *ja* eine Schwalbe, die
Yes, but surprising is that not. That was *ja* a swallow PL.DEF
fliegen immer so schnell.
fly always so fast.
"Right, but that's not really surprising. It was, *after all*, a swallow, and they always fly so fast."

(10) B: Der ist *doch* nicht so schnell vorbeigeflogen.
SING.DEF is *doch* not so fast flew by.
"It didn't *actually* fly by that fast."

In (9), the utterance-medial *ja* can be said to mark B's commitment to the truth of the proposition *Das war eine Schwalbe*, or to put it differently, it can be said to indicate their belief in the strength or validity of that particular state of affairs (Döring and Repp, 2019; Gast, 2008; Unger, 2016), which is pragmatically reinforced by the ensuing statement "and they always fly by so fast". More specifically, as discussed with regard to the example in (8), *ja* indexes that B believes that the proposition at hand is part of the common ground between them and their interlocutor, such that that particular piece of information is almost trivial or unremarkable when uttered (Döring and Repp, 2019; Gast, 2008; Unger, 2016). As such, *ja* can be said to mark not only B's *subjective* stance toward the relevant state of affairs (i.e. "I commit to the truth of that being a swallow") but also an *intersubjective* understanding of how their response relates to A's original statement (i.e. "I think we both agree that that's not surprising given that the bird was a swallow and they always fly so fast"). *Doch* functions in a very similar way to *ja*[2], except that it indexes a mismatch of

[2]Note, however, that *doch* is more flexible than *ja*, occurring both sentence-medially as well as sentence-initially as an intensifier/ focus particle (*Doch ist der schnell vorbeigeflogen, die fliegen immer so schnnell*; English: "*Of course* it flew by fast, they always fly really fast").

information in the common ground, indicating that something is in fact false while it was originally believed to be true (Döring and Repp, 2019; Gast, 2008; Unger, 2016). Thus, while A states that they think that the bird flew by fast, in (10) B challenges that position stating that the bird didn't actually fly by that fast. Let's now compare the usage of *ja* and *doch* to that of two other similar particles, *eigentlich* and *tatsächlich*.

A: Hast du gesehen, wie schnell der Vogel vorbeigeflogen ist?
Have you seen how fast SING.DEF bird flew by is?
"Did you see how fast that bird flew by?"

(11) B: Der ist *tatsächlich* sehr schnell vorbeigeflogen.
SING.DEF is *tatsächlich* very fast flew by.
"It did *actually* fly by very fast." OR "It *really* did fly by very fast."

(12) B: Der ist *eigentlich* nicht so schnell vorbeigeflogen.
SING.DEF is *eigentlich* not so fast flew by.
"It didn't *actually* fly by that fast."

Tatsächlich in (11) serves a function similar to that of *ja* in (8) and (9): it confirms some piece of information which is part of the common ground between the two interlocutors. Unlike *ja*, however, it confirms that the proposition expressed in the statement by A is true without indexing that that information is trivial or unremarkable. It therefore plays a similar role to that of a positive response particle like *yes* or *right*; just like *ja*, however, it is only felicitous because it's part of an utterance which has propositional content of its own. In keeping with the parallelisms between particles, *eigentlich* is to *doch* very much like *tatsächlich* is to *ja*, that is, it disconfirms some information which is part of the common ground between interlocutors. More specifically, in (12), *eigentlich* disconfirms the statement that the bird flew by really fast, contrasting that proposition with the proposition that the bird didn't actually fly by that fast. *Eigentlich* therefore introduces a new proposition which stands in direct contrast to the proposition originally expressed in the statement by A, playing a similar role to that of a negative response particle like *no* while, just like *tatsächlich*, being dependent on the utterance it modifies.

In summary, *ja*, *doch*, *tatsächlich*, and *eigentlich* are all instances of discourse particles in German which serve the specialized function of relating the utterance they are embedded in to a proposition or speech-act alternative which may be more or less salient in the discourse context. Despite their common function as epistemicity markers, these particles differ from one another in terms of the specific modal meaning they index: while *ja* and *tatsächlich* both mark agreement between the propositional content of their host utterance and some contextually-relevant propositional antecedent, *doch* and *eigentlich*, on the other hand, both mark contrast. Interestingly, what the examples above show is that these particles are polyfunctional by definition: even in the same context of utterance, they can be said to have both a more propositionally-oriented function, which makes them modal particles in the strict sense, as well as a discourse-managing function, which brings them together with other discourse-managing devices like response particles. Crucially, when it comes to how these particles are interpreted in context, interpretation seems to depend on a pragmatic calibration between their specialized modal meaning and their general discourse-marking meaning. As discussed in Chapter 1, conventional implicatures, when construed as a form of pragmatic inference (Feng, 2011, 2010), can be thought to capture exactly such process, that is, the derivation of a context-sensitive interpretation by means of association between contextual constraints and the conventionalized meaning(s) of an expression.

In the study reported below, I address the question of whether modal particles, as discussed so far, serve as predictive cues to interpretation in the incremental processing of language, as measured in a mouse tracking task. More specifically, I focus on the interpretation of referring expressions containing the German particles *tatsächlich* and *eigentlich*. Given the similar and yet complementary functions of *tatsächlich* and *eigentlich*, I'm interested in whether processing one or the other particle gives rise to differential effects on real-time interpretation and the resolution of reference. Before introducing the study and reporting my findings, I review previous findings on the processing of modal particles and related non-lexical cues of epistemicity.

4.1.2 Processing profile of modal particles

Previous psycholinguistic research has explored the role of different epistemicity markers in the incremental processing of language. Indeed, many studies concerned with the role of information structure on sentence processing have looked at the processing of cues which can be said to function as non-lexical markers of epistemicity. For example, Kurumada et al. (2014) looked at whether contrastive prosody in English allow comprehenders to predict the likely referent in referring expressions of the sort "It looks like a zebra" when either the verb (i.e. "It LOOKS like a zebra") or the noun (i.e. "It looks like a ZEBRA") is accented. In order to test that, they had participants listen to the expressions while looking at visual displays containing images of potential referents, their eye movements being tracked in the process. Participants were instructed to select the most likely referent by clicking on the relevant image. Crucially, the visual displays contained not only the mentioned referent (e.g. a zebra) but also a perceptually similar competitor (e.g. an okapi) and two distractors. Kurumada et al. (2014) found that a contrastive accent on the verb, as opposed to the noun, induced comprehenders to look predictively at the non-prototypical referent, i.e., something that looked like a zebra but wasn't one, prior to the onset of the noun.

In a similar study, Roettger and Franke (2019) looked at whether contrastive prosody in German allow comprehenders to draw predictive inferences about the likely referent in referring expressions of the sort "Der Wuggy hat dann die Birne aufgesammelt" (English: the wuggy has then the pear picked-up, "The wuggy then picked up the pear"). Crucially, they manipulated whether the auxiliary verb or the object of the sentence was accented. Much like in the study by Kurumada et al. (2014), they had participants listen to the expressions while looking at a visual displays containing images of potential referents. Participants were instructed to select the most likely referent by clicking on the relevant image, their mouse movements being tracked in the process. Roettger and Franke (2019) found that a pitch accent on the auxiliary (i.e. "Der Wuggy HAT DANN die Birne aufgesammelt"; English:

the wuggy HAS THEN the pear picked-up) induced comprehenders to shift their mouse towards the mentioned referent earlier compared to when the pitch accent fell on the object of the sentence (i.e. "Der Wuggy hat dann die BIRNE aufgesammelt"; English: the wuggy has then the PEAR picked-up).

Taken together, these studies suggest that an intonational feature like pitch accent can guide comprehenders in the online resolution of reference, serving as a cue to intended meaning in context. While prosody can thus function as a non-lexical means of marking epistemic stance in reference resolution, at least when used in combination with other linguistic markers like the expression "It looks like" as in Kurumada et al. (2014), little research to date has directly addressed the processing of lexical markers of epistemicity like discourse and modal particles. Indeed, only four studies in the literature have explicitly looked at the processing of modal particles. Dörre, Czypionka, et al. (2018) investigated whether particles in German which function both as modal and non-modal markers are processed differently depending on the relative frequency of each meaning and on whether the discourse context biases either one or the other interpretation. They tested 10 different particles, having participants read sentences of the sort "Wer hat *bloß* den Flur gewischt?" (English: who has *bloß* the corridor wiped; "Who *bloß* wiped the corridor?") one word at a time against contexts which disambiguated between the particles' modal and non-modal meanings, as in 13 and 14 below.

(13) **Modal usage**

Hier ist noch Schneematsch von draußen, wer hat *bloß* den Flur gewischt?
Here is still mud from outside who has *bloß* the corridor wiped?

"Here there's still some mud from outside, who *on earth* wiped the corridor?"

(14) **Non-modal usage**

Die anderen Zimmer sind ebenfalls schmutzig, wer hat *bloß* den Flur
The other rooms are also dirty who has *bloß* the corridor
gewischt?
wiped?

"The other rooms are also dirty, who wiped *only* the corridor?"

Crucially, Dörre, Czypionka, et al. (2018) measured the reading times at each word in the sentence, their analysis including a measure of the relative meaning frequency of each particle, as computed from a corpus of spoken German. They found that comprehenders read the particle itself at the same speed regardless of the type of context and the relative frequency of its different meanings. One word downstream from the particle (i.e. "Wer hat bloß *den* ..."), while the reading times did not differ depending on the context, they were reliably higher when the processed particle had a more frequent modal meaning. Two words downstream from the particle (i.e. "Wer hat bloß den *Flur* ..."), the reading times were reliably higher when the context biased a modal interpretation as opposed to a non-modal one. Finally, at the last word in the sentence (i.e. "Wer hat bloß den Flur *gewischt*"), three words downstream from the particle, while the reading times were generally shorter when the processed particle had a more frequent modal meaning as well as when the context biased that same meaning, they were longer when considering more frequent modal meanings in modal-biasing contexts. All in all, the results suggest that both context and the relative frequency of a particle's modal meaning impact the online processing of sentences containing modal particles, such that downstream from the particle processing is, by and large, more costly the more conventionalized its modal meaning. Interestingly, while a modal-biasing context and a high relative modal meaning frequency seem to facilitate processing at the main verb in the sentence when considered separately, their interaction actually seems to hinder processing, suggesting that, overall, the modal meaning of a polyfunctional particle is harder to process than its non-modal meaning.

In a subsequent study, Dörre and Trotzke (2019) further investigated the processing of two of the particles originally investigated in Dörre, Czypionka, et al. (2018), namely *bloß* and *nur*. In their modal usage, both *bloß* and *nur* index a (subjective) strengthening of a proposition[3], as in "Warum hat sie *bloß* das Kleid gekauft?" (English: why has she *bloß* the

[3]See Dörre and Trotzke (2019) for their account on how the modal meaning of *bloß* and *nur* might relate

dress bought), which can be interpreted as "Why *on earth* did she buy the dress?". *Bloß* and *nur* can also be used as focus particles, functioning like the English particle *only*, as in "Warum hat sie *bloß* das Kleid gekauft?" (English: why has she *bloß* the dress bought), meaning "Why did she *only* buy the dress?". As in Dörre, Czypionka, et al. (2018), Dörre and Trotzke (2019) had participants read questions of the sort "Warum hat sie *bloß* das Kleid gekauft?" one word at a time against contexts which disambiguated between the particles' modal and non-modal meanings. Despite measuring reading times at each individual word in the sentence, they analyzed the entire region starting at the particle and ending at the last word in the sentence (i.e. *particle* + das + Kleid + gekauft). They found that both in the case of *bloß* and *nur* comprehenders read the aggregated critical region at the same speed regardless of whether the context biased the particle's modal meaning or its meaning as a focus marker.

In a second experiment, Dörre and Trotzke (2019) looked at whether prosodic focus affects the processing of *bloß* and *nur*, having participants listen to the same questions as in their reading experiment and manipulating whether the noun (i.e. "Warum hat sie *bloß* das KLEID gekauft?") or the auxiliary verb (i.e. "Warum HAT sie *bloß* das Kleid gekauft?") was accented, just like in the studies by Kurumada et al. (2014) and Roettger and Franke (2019). Crucially, while the focus-marking function of *bloß* and *nur* is said to be realized prosodically via a pitch accent on the noun (i.e. the focused element), their modal-marking function does not strongly correlate with any particular intonational contour, either the question form, the auxiliary verb, or the noun being accented in such a usage. Participants were instructed to listen to the test sentences, presented in the absence of any discourse context. After listening to the sentences, they were presented with the same discourse contexts used the reading experiment, namely one biasing a modal interpretation (e.g. "Hier ist noch Schneematsch von draußen"; English: Here is still mud from outside, "Here there's still some mud from outside") and one biasing a non-modal interpretation

to the particles' meaning as focus particles.

(e.g. "Die anderen Zimmer sind ebenfalls schmutzig"; English: The other rooms are also dirty, "The other rooms are also dirty"). Participants were then asked to select the context that was most appropriate. Dörre and Trotzke (2019) found that when comprehenders were presented with sentences where the auxiliary was accented they selected the modal-biasing context at a reliably higher rate (87%) than the non-modal-biasing context (13%), whereas when they were presented with sentences where the noun was accented they selected the non-modal-biasing context at a reliably higher rate (63%) than the modal-biasing context (37%). The results suggest that contrastive prosody acts as a reliable cue in the processing of modal particles, guiding comprehenders in their interpretation of polyfunctional particles such as *bloß* and *nur*.

The two other studies in the literature which have looked at the processing of modal particles have dealt with the particles *inderdaad* and *eigenlijk* in Dutch. These two particles have the same meaning profiles as *tatsächlich* and *eigentlich* in German, respectively, meaning that *inderdaad* marks agreement between two discourse elements while *eigenlijk* marks contrast. Much like their counterparts in German, they also serve both a more specialized modal function and a general discourse-managing function. While in their discourse-managing function *inderdaad* and *eigenlijk* mark a general sense of (dis)agreement between two discourse elements, like a response particle would, in their modal function they mark (dis)agreement between the propositional content of their host utterance and some contextually relevant propositional antecedent.

van Bergen and Bosker (2018) investigated whether *inderdaad* and *eigenlijk* allow comprehenders to draw predictive inferences about the likely referent in referring expressions of the sort "Ik schrok *inderdaad* van de rondrennende *BEEP* aan het eind" (English: I startled *inderdaad* from the running *BEEP* at the end; "I was indeed scared by the running *BEEP* at the end"). The referring expressions were embedded in mini dialogues, as in (15) below.

(15) Ondanks haar angst voor dieren is Marie naar het circus geweest.
Despite her fear from animals is Marie to the circus gone.
"Despite her fear of animals, Marie went to the circus."

Je vond de dierenact zeker doodeng?
You found the animal act surely terrifying?
"You must have been terrified by the animal act?"

Ik schrok *inderdaad* van de rondrennende *BEEP* aan het eind.
I startled *inderdaad* from the running *BEEP* at the end.
"I was *indeed* scared by the running *BEEP* at the end."

Participants were instructed to listen to the dialogue while looking at visual displays containing images of potential referents, their task being to select the most likely referent by clicking on the relevant image. Their eye movements were tracked in the process. Crucially, the displays contained a referent which was biased by the discourse context as being the target of the referring expression (e.g. a lion in (15)) as well as a contextually-relevant competitor (e.g. a clown) and two distractors. While a modal interpretation of *inderdaad* should lead to a confirmation of the proposition expressed in "You must have been terrified by the animal act?", resulting in a preference for expecting the referent biased by the context (i.e. the lion), a modal interpretation of *eigenlijk* should lead to a disconfirmation of the proposition expressed in "You must have been terrified by the animal act?", resulting in a preference for expecting the unbiased referent (i.e. the clown.) By design, however, the actual referent was beeped in the playback of the target sentence, as indicated in (15). The omission of the critical noun was meant to avoid biasing a strictly modal interpretation of the particles, such that participants could interpret them not necessarily in propositional terms but perhaps in discourse-managing terms more generally. While it is unclear what referent in the visual display could be mapped onto a non-modal interpretation of *inderdaad*, a non-modal interpretation of *eigenlijk* could induce comprehenders to interpret it as marking a general sense of disagreement between its host utterance and the preceding context, thus not strongly supporting a referential contrast between "You must have been terrified by the

animal act" and "I was scared by the running *BEEP* at the end", which should lead to no particular preference for the unbiased referent.

van Bergen and Bosker (2018) found that processing *inderdaad* induced comprehenders to select the referent biased by the context (i.e. the lion) at a reliably higher rate (about 100%) than the unbiased competitor (i.e. the clown). They also found that it induced comprehenders to direct their gaze at the biased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 50% of the eye fixations were directed at the biased referent while about 20% of them were directed at the unbiased referent. When it comes to *eigenlijk*, van Bergen and Bosker (2018) found that processing the particle induced comprehenders to select the unbiased referent (i.e. the clown) at a reliably higher rate (about 60%) than the biased referent. They also found that, when this was the case, it induced comprehenders to direct their gaze at the unbiased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 40% of the eye fixations were directed at the unbiased referent while about 20% of them were directed at the biased referent. When comprehenders selected the biased referent after processing *eigenlijk*, they did not direct their gaze at the unbiased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 25% of the eye fixations were directed at the unbiased referent while about 35% of them were directed at the biased referent.

In a second experiment, van Bergen and Bosker (2018) extended their original design by introducing a new *eigenlijk* condition where the discourse context did not bias any particular referent, as in (16) below.

(16) Op zaterdagmiddag is Marie naar het circus geweest.
On Saturday afternoon is Marie to the circus gone.
"On Saturday afternoon, Marie went to the circus."

Wat is je het best bijgebleven?"
What is you the best remembered?
"What did you find most memorable?"

Ik schrok *eigenlijk* van de rondrennende *BEEP* aan het eind.
I startled *eigenlijk* from the running *BEEP* at the end.
"I was *actually* scared by the running *BEEP* at the end."

They predicted that this would result in no particular referential expectation compared to the *eigenlijk* condition where the context biased one of the referents. As for the remaining conditions, they did not predict any changes compared to the first experiment. Indeed, just like in their first experiment, van Bergen and Bosker (2018) found that processing *inderdaad* induced comprehenders to select the biased referent (i.e. the lion) at a reliably higher rate (about 100%) than the unbiased referent. They also found that it induced comprehenders to direct their gaze at the biased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 55% of the eye fixations were directed at the biased referent while about 15% of them were directed at the unbiased referent. When it comes to *eigenlijk* in the biasing condition, van Bergen and Bosker (2018) found that processing the particle did not induce comprehenders to select the unbiased referent (i.e. the clown) at a reliably higher rate than the biased referent. However, when comprehenders selected the unbiased referent, they did direct their gaze at the unbiased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 45% of the eye fixations were directed at the unbiased referent while about 25% of them were directed at the biased referent. On the other hand, when comprehenders selected the biased referent after processing *eigenlijk* in the biasing condition, they did not direct their gaze at the unbiased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 25% of the eye fixations were directed at the unbiased referent while about 40% of them were directed at the biased referent. Interestingly, when it comes to *eigenlijk* in the unbiasing condition, van Bergen and Bosker (2018) again found that processing the particle did not induce comprehenders to select the unbiased referent (i.e. the clown) at a reliably higher, much like in the biasing *eigenlijk* condition. They also found that, when comprehenders selected the biased referent, it did not induce them to direct their gaze at

the biased referent prior to the onset of the beep at a reliably higher rate, such that, at the beep onset, about 40% of the eye fixations were directed at the biased referent while about 25% of them were directed at the unbiased referent.

All in all, the results suggest that processing modal particles affects the interpretation of referring expressions and the online resolution of reference. While processing an agreement-marking particle like *inderdaad* induces comprehenders to expect an agreement between its host utterance and a contextually-relevant antecedent, allowing them to predict a referent biased by the context, processing a contrast-marking particle like *eigenlijk* induces comprehenders to expect a disagreement between its host utterance and a contextually-relevant antecedent, allowing them to predict a contextually-relevant referential alternative. Interestingly, *eigenlijk* does not seem to be biased in terms of either a modal or a non-modal interpretation, regardless of whether the discourse context biases a modal interpretation. However, when *eigenlijk* is interpreted as marking a modal relation, it affects online interpretation to a similar extent as *inderdaad*.

In a subsequent study, Rasenberg et al. (2019) investigated whether *inderdaad* and *eigenlijk* allow comprehenders to draw predictive inferences about upcoming lexical information in sentences of the sort "We zijn *inderdaad* elke dag naar een museum geweest" (English: We are *inderdaad* every day to a museum been; "We *indeed* went to a museum every day"), which were embedded in mini dialogues, as in (17) below.

(17) Diana is een weekend met haar klas van de kunstacademie naar Parijs
Diana is a weekend with her class from the art school to Paris
geweest.
been.
"Diana spent a weekend in Paris with her art academy class."

Haar vriendin vraagt: Jullie hebben vast een hoop kunst gezien?
Her friend asks: You have plenty of art seen?
"Her friend asks: You guys must have seen a lot of art?"

Diane zegt: We zijn *inderdaad* elke dag naar een **museum**/ **park** geweest.
Diana says: We are *inderdaad* every day to a **museum**/ **park** been.

"Diana says: We *indeed* went to a **museum**/ **park** every day."

Participants were instructed to read the dialogue for comprehension, their event-related potentials (ERPs) being measured relative to the words in the sentences. Crucially, Rasenberg et al. (2019) compared the ERP signatures relative to the noun following the modal particle (bolded in (17)), which was either predictable (e.g. museum) or unpredictable (e.g. park) given the discourse context. In order to draw inferences about differences in the amplitude of the relevant ERP components, they compared the signal for each particle to that of a control condition containing adverbs instead of discourse particles, as in "We zijn *daar* elke dag naar een museum geweest" (English: We are *there* every day to a museum been; "We went *there* to a museum every day"). They predicted qualitative changes to the amplitude of the N400 and P600 components, the former functioning as an index of semantic plausibility and the latter as an index of prediction disconfirmation. Whereas upon processing *inderdaad* the amplitude of the N400 component should be reduced for predictable nouns (e.g. museum) and increased for unpredictable nouns (e.g. park) relative to processing an adverb, upon processing *eigenlijk* the amplitude of the N400 component should be reduced for unpredictable nouns and increased for predictable nouns. Similarly, whereas upon processing *inderdaad* the amplitude of the P600 component should be decreased for predictable nouns and increased for unpredictable nouns relative to processing an adverb, upon processing *eigenlijk* the amplitude of the P600 component should be decreased for predictable nouns and increased for unpredictable nouns. Rasenberg et al. (2019) found that processing neither *inderdaad* nor *eigenlijk* affected the N400 component of the ERP signal. Similarly, they found that neither particle affected the P600 component.

Taken together, the studies in the literature paint a complex picture of the processing profile of modal particles. While the earlier results from van Bergen and Bosker (2018) suggest that processing modal particles does affect incremental interpretation, such that comprehenders are able to draw predictive inferences about subsequent material in a sentence when

that material is relevant for interpretation, the more recent results from Rasenberg et al. (2019) do not provide evidence that modal particles allow comprehenders to draw predictive inferences about linguistic material downstream in a sentence. The results from Dörre, Czypionka, et al. (2018) further suggest that processing modal particles is at least partially more costly the more conventionalized the modal meaning associated with a particular particle form. These findings raise a number of questions about the role of modal particles as cues to epistemicity in the incremental processing of language.

First and foremost, given the conflicting results, it remains unclear to what extent modal particles affect incremental interpretation. While they do affect the outcome of the interpretive process, as shown in van Bergen and Bosker (2018), interacting with the propositional meaning of the sentences they occur in in expected ways given their modal meaning, it is not clear what precise role the non-sentential discourse context plays in biasing any relevant interpretitve expectations. For instance, in the study by van Bergen and Bosker (2018), the same discourse context constrained interpretation qualitatively differently depending on the particle at hand, as *eigenlijk* was interpreted both as a modal marker and as a non-modal discourse-managing particle while *inderdaad* was not. While this was a feature of the experimental design, allowing the authors to tackle the polyfunctionality of *eigenlijk* and *inderdaad*, it obscures the inferences one may draw about the interaction between context and particle meaning. Even when a particular meaning of a particle might be said to be constrained in context, as in Dörre and Trotzke (2019), potentially overriding any concerns about non-trivial interactions between context bias and particle meaning, it remains unclear to what extent modal particles serve as reliable cues to interpretation in context, especially when considering their role as *predictive* cues to interpretation. In the present study[4], I tackle precisely that question, focusing on the incremental interpretation of sentences containing the German modal particles *eigentlich* and *tatsächlich*. Crucially, I investigate the

[4]This study has been conceived and carried out in collaboration with Michael Franke, Laine Stranahan, and Timo Roettger.

processing of referring expressions of the sort "Das ist *eigentlich/ tatsächlich* ein Bild von einem Wolf" (English: This is *eigentlich/ tatsächlich* a picture from a wolf; "That's *actually/ indeed* a picture of a wolf") against discourse contexts which bias not only concrete referential expectations but also clear mappings between those referential expectations and a modal interpretation of the respective particles. Thus, I manipulate the predictability of any given referent as a function of an interaction between the discourse context and the usage of the particles in context. I also manipulate whether the particles are used reliably according to their modal meaning. In other words, I look at whether pragmatically incoherent usages of *eigentlich* and *tatsächlich* lead to qualitative modulations of the processing signatures compared to the pragmatically coherent usages. In the section below, I describe the exact predictions of the study, as they relate to the test items, as well as the details and setup of the experimental task.

4.2 Experiment

Method

Participants

Fifty native speakers of German were recruited among the cognitive science student population of the Osnabrück University. Participants were given course credit for their participation in the experiment. Data collection was conducted entirely online and all collected data was stored at servers from the Osnabrück University.

Materials and design

Participants were asked to listen to 120 miniature dialogues mimicking an interaction between a parent and a child, each anchored to a visual world display containing two pictures, one depicting a referent explicitly mentioned in the dialogue and another depicting a semantically

and/ or perceptually-related competitor, e.g., a wolf and a dog, as shown in Figure 4.1. Each dialogue consisted of a question-answer pair where the question, posed by the child, was polar in nature and made direct reference to one of the pictures visible on screen (e.g. "Ist das ein Bild von einem Wolf?"; English: "Is this a picture of a wolf?"), while the answer, provided by the parent, either confirmed or disconfirmed the truth of the proposition expressed in the question. Crucially, the answer did not contain any response or negation particle but rather a discourse particle with a modal meaning, either *eigentlich* or *tatsächlich* (e.g. "Das ist *eigentlich*/ *tatsächlich* ein Bild von einem Wolf"; English: "That's indeed/ actually a picture of a wolf"). Participants' task was to listen to the question-answer sequence and to select the referent they thought was most representative of the topic of the conversation.

Each participant was assigned to one of two experimental groups with a different distribution of critical trials. In the **reliable** group, participants encountered 60 fillers and 60 dialogues in which the discourse particles were used according to their conventionalized modal meaning, i.e., where *eigentlich* marked a contrast and *tatsächlich* marked a confirmation, as shown below.

Question

Ist das ein Bild von einem **Wolf**?
Is this a picture from a **wolf**?
"Is this a picture of a **wolf**?"

Answer

Das ist *eigentlich* ein Bild von einem **Hund**.
This is *eigentlich* a picture from a **dog**.
"That's actually a picture of a **dog**."

Das ist *tatsächlich* ein Bild von einem **Wolf**.
This is *tatsächlich* a picture from a **wolf**.
"That's indeed a picture of a **wolf**."

In the **unreliable** group, participants encountered 20 fillers, 60 dialogues in which the particles were used according to their conventionalized modal meaning, and 40 dialogues in

which the particles were used unreliably, i.e., where *eigentlich* marked a confirmation and *tatsächlich* marked a contrast, as shown below.

Question

Ist das ein Bild von einem **Wolf**?
Is this a picture from a **wolf**?
"Is this a picture of a **wolf**?"

Answer

Das ist *eigentlich* ein Bild von einem **Wolf**.
This is *eigentlich* a picture from a **wolf**.
"That's actually a picture of a **wolf**."

Das ist *tatsächlich* ein Bild von einem **Hund**.
This is *tatsächlich* a picture from a **dog**.
"That's indeed a picture of a **dog**."

Filler trials contained various sorts of questions, such as polar questions similar to the ones from critical trials (e.g. "Ist das ein Fisch?"), polar questions with miscellaneous forms (e.g. "Findest du das Tier mit der dicken Nase auch lustig?", "Ist das, etwas was man zum Kämpfen benutzt?"), as well as wh-questions (e.g. "Wie heißt das orangene Tier?"; "Was ist das mit dem Schabel für ein Tier?"). The respective answers to these questions did not contain any discourse particles.

Procedure

Written instructions were provided prior to the actual task, followed by 10 practice trials which mimicked the exact procedure of the test trials. Participants were instructed to listen to the question-answer sequence and to select the picture that best matched the topic of the exchange. On each trial, they were presented with a "Start" button at the center bottom of the screen. Upon clicking on it, the button disappeared from the screen and playback of the question started. After the playback was over, a "Go" button appeared at the same location as the "Start" button together with the two visual stimuli, as shown in Fig. 4.1. Upon

Figure 4.1 Sample visual world display. The display contains two pictures, one mentioned in the dialogue and an unmentioned competitor, each presented in one of the top corners of the screen.

clicking on the "Go" button, playback of the answer started together with the tracking of participants' mouse cursor from the "Go" button at the bottom of the screen towards one of the pictures at the top of the screen. Upon moving the cursor over one of the pictures, a choice was automatically made and participants were thus directed to the next trial.

Hypotheses

As discussed in the introduction, *eigentlich* and *tatsächlich* are particles which, in their modal usage, index how the utterance they are embedded in relates to a contextually-relevant propositional antecedent. Importantly, this entails that the presence of either one particle in a sentence is expected to affect incremental interpretation differentially depending on the specific modal meaning the particle signals in context. This general effect is expected to be modulated by the reliability of the input, as discussed above. While *eigentlich* signals a propositional contrast, *tatsächlich* signals a propositional agreement. Given the setup of the present experimental task, I am therefore interested in how the measured mouse trajectories

are affected by the presence of each particle in the test sentences. More concretely, I am interested in the shape of the trajectory segments measured after the onset of the particle and before the onset of the disambiguating noun, i.e., "Das ist PARTICLE [ein Bild von einem] NOUN". Crucially, any changes in mouse position within this time window can reasonably be assumed to be an effect of processing the particle, such that movements either towards or away from the image representing the referent mentioned in the question indicate a sensitivity to the meaning of the processed particle.

While the general account entertained here does not generate precise predictions about the shape of the mouse trajectories within the critical time window, it generates predictions about qualitative changes to the trajectories after the onset of the particles. For the **reliable** group, it predicts that upon hearing *eigentlich* participants will shift their mouse trajectory towards the referent not mentioned in the question, as that referent constitutes a contextually-relevant propositional alternative to the referent mentioned in the question. It also predicts that upon hearing *tatsächlich* participants will continue their trajectory towards the referent mentioned in the question. Note that, in both cases, if participants initiate any mouse movement prior to the onset of the particle, their trajectories are assumed to be en route towards the mentioned referent at the onset of the particle, as the mention of the referent in the question biases them to entertain that particular referent before any relevant semantic information from the response is processed.

For the **unreliable** group, the present account predicts the same qualitative patterns as in the **reliable** group but it also predicts an attenuation of those patterns, such that initiation of the mouse movements towards the selected referent might start later in the trial, or the curvature of the trajectories might be larger due to the higher degree of uncertainty.

Results

The data and analysis scripts are available for inspection under https://osf.io/dt8r3/.

Data cleaning and pre-processing

Prior to any analysis, the data set was cleaned according to both participant-based and response-based exclusion criteria. I first excluded any participant who selected a picture before the onset of the particle in more than 30% of their trials (n=3). This is motivated by the fact that any such decision would not be based on the processing of the relevant linguistic material in the test sentence. This exclusion resulted in a data set which included data from 47 participants.

After excluding three individual participants as per the criterion above, I excluded any individual trial in which a participant took longer than 1000 ms to make a decision after the offset of the disambiguating noun (5% of the data) as well as any trial with fewer than 10 observations (7% of the data), an observation corresponding to a single change in horizontal mouse position as sampled during the recording of the mouse movements. I also excluded all trials in which participants selected the competitor instead of the target referent, i.e., *eigentlich* trials where participants selected the referent mentioned in the question and *tatsächlich* trials in which participants selected the unmentioned referent. Finally, all **unreliable** trials were discarded, such that the remaining data only contained **reliable** trials.

The cleaned data, plotted and analyzed below, consists of mouse trajectories which were linearly extrapolated from the actual recorded trajectories. This procedure was adopted in order to circumvent the loss of relative temporal information when aggregating over trajectories with different absolute durations. Data points were extrapolated at every 10 ms time interval for each individual participant-item pair.

Confirmatory analyses

First I report and inspect the mean mouse trajectories, as shown in Figure 4.2.

The figure is split in two columns, each showing the data from one of the two experimental

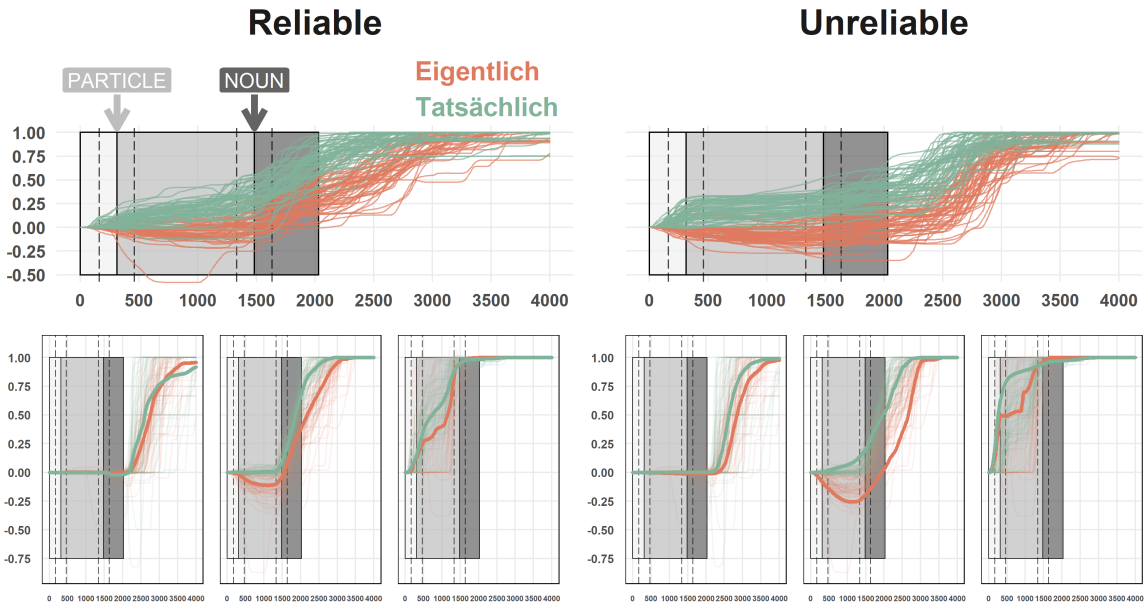

Figure 4.2 Mean mouse trajectories. **Top**: Mean item trajectories aggregated over participants. **Bottom**: Clusters of mean trajectories aggregated over items and participants, with mean item trajectories shown in the background. The shaded areas in the graphs show the (average) time windows between the onsets of the discourse particle and the disambiguating noun. The left column shows data from the **reliable** group, while the right column shows data from the **unreliable** group. *Tatsächlich* items are shown in green while *eigentlich* items are shown in orange.

groups, **reliable** on the left and **unreliable** on the right. Each graph at the top of the figure shows the individual item trajectories aggregated over participants. Trajectories unfold over time, mapped onto the x-axis, relative to the position of the target referent, mapped onto the top of y-axis. The shaded areas in the graphs show the (average) time windows between the onsets of the relevant linguistic material in the signal, namely the particle and the noun which disambiguates the referring expression. As the figure shows, there is considerable variation in terms of the shape of the mouse movements: while some trajectories show an early bias for the mentioned referent, either continuing en route towards it after the onset of the discourse particle or changing its course towards the unmentioned referent, others linger around the center of the screen until after the onset of the disambiguating noun. Given this variation in the data, instead of producing a single aggregate for each discourse particle I

first split the data into clusters of similarly shaped trajectories, as shown at the bottom of Fig. 4.2. The clustering of the data was done using an automated procedure implemented in the `mt_cluster()` function from the `mousetrap` package for *R* (Kieslich and Henninger, 2017). `mt_cluster()` generates a user-defined number of trajectory clusters by computing a matrix of distances between pairs of individual trajectories and then splitting it by the number of desired clusters. Before inspecting in more detail the aggregates shown within the clusters at the bottom of Fig. 4.2, which consist of mean trajectories aggregated over all experimental items and participants, I look at the relative proportion of trajectories per cluster, as shown in Figure 4.3.

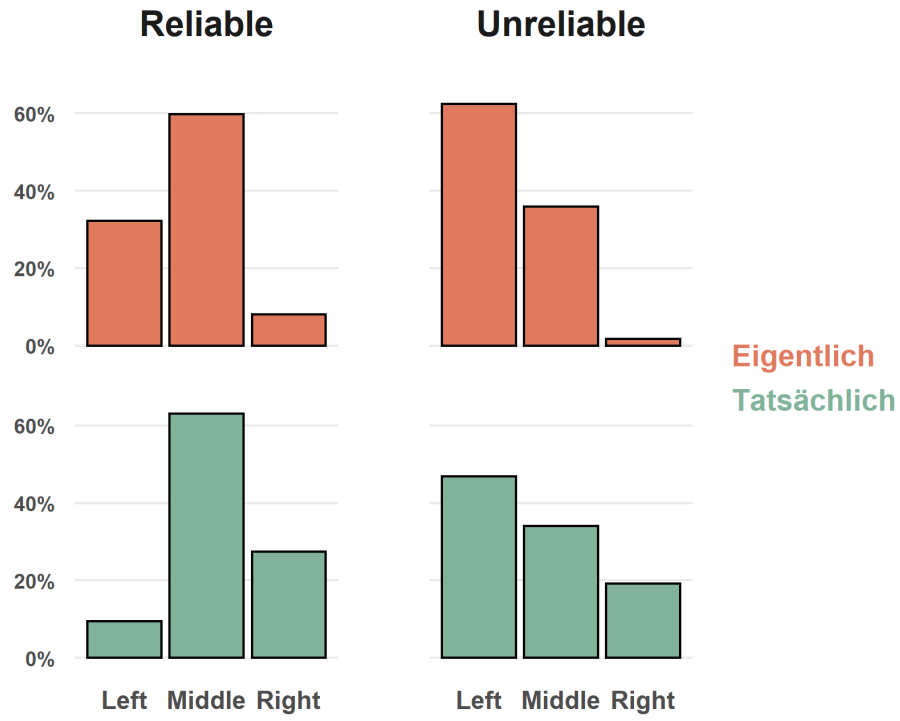

Figure 4.3 Proportion of trajectories per cluster. Proportion of data points per trajectory cluster in Figure 4.2. The left column shows data from the **reliable** group, while the right column shows data from the **unreliable** group. *Tatsächlich* items are shown in green while *eigentlich* items are shown in orange.

In rough terms, the clusters represent three types of trajectories: those in the left-hand cluster can be said to represent a strategy where one waits for disambiguation from the

noun; those in the middle cluster can be said to represent a hesitant interpretation of the particle; those in the right-hand cluster can be said to represent an early drift towards the referential target. As the figure shows, both in the **reliable** and **unreliable** groups the data is unevenly split across clusters. In the **reliable** group, 32% of the *eigentlich* trajectories are found in the cluster in the left-hand panel ($P(Y = left) = .32$, 95% CrI [0.19, 0.48]), 60% of the trajectories being found in the cluster in the middle panel ($P(Y = middle) = .60$, 95% CrI [0.42, 0.75]) and 8% in the cluster in the right-hand panel ($P(Y = right) = .08$, 95% CrI [0.03, 0.18]). Ten percent of the *tatsächlich* trajectories are found in the cluster in the left-hand panel ($P(Y = left) = .10$, 95% CrI [0.04, 0.19]), 63% of the trajectories being found in the cluster in the middle panel ($P(Y = middle) = .63$, 95% CrI [0.45, 0.78]) and 27% of them in the cluster in the right-hand panel ($P(Y = right) = .27$, 95% CrI [0.15, 0.44]). In the **unreliable** group, on the other hand, 62% of the *eigentlich* trajectories are found in the cluster in the left-hand panel ($P(Y = left) = .62$, 95% CrI [0.48, 0.76]), 36% of the trajectories being found in the cluster in the middle panel ($P(Y = middle) = .36$, 95% CrI [0.23, 0.50]) and only 2% in the cluster in the right-hand panel ($P(Y = right) = .02$, 95% CrI [0.00, 0.07]). Forty-seven percent of the *tatsächlich* trajectories are found in the cluster in the left-hand panel ($P(Y = left) = .47$, 95% CrI [0.32, 0.61]), 34% of the trajectories being found in the cluster in the middle panel ($P(Y = middle) = .34$, 95% CrI [0.19, 0.52]) and 19% of them in the cluster in the right-hand panel ($P(Y = right) = .19$, 95% CrI [0.12, 0.31]).

In order to quantify uncertainty over the results reported above, I fitted a Bayesian hierarchical model predicting cluster membership as a function of an interaction between the particles and the experimental groups, with *eigentlich* in the **reliable** group dummy-coded as the reference level. The model estimated the log odds of each observation being categorized as belonging to any given cluster, the estimates reported below being a conversion of (predicted) log odds to probabilities. It included the maximal random effect structure justified by the design, which in the present case is random slopes and intercepts for the

effect of particle for both participants and items. Since participants were assigned to only one experimental group, at the participant level I only estimate random slopes for the effect of group. The model, fitted using the *R* package `brms` (Bürkner, 2017) and described in detail in the scripts available in the supporting material, had the following form, shown in `brms` syntax:

```
cluster ~ particle * group
                (1 + particle | participant) +
                (1 + particle * group | item),
                family = categorical()
```

I now turn back to Fig. 4.2 to inspect the actual trajectories within each of the clusters. As visual inspection of the mean trajectories within each cluster in the **reliable** group shows, within two of the three clusters, the trajectories for the two particles are qualitatively different from one another. The three distinct types of *tatsächlich* trajectories, shown in green, consist of shifts towards the target after the onset of the disambiguating noun (left panel, 10% of the data), shifts towards the target within the particle window (middle panel, 63% of the data), as well as shifts towards the target before the onset of the particle without subsequent change of course (right panel, 27% of the data). The *eigentlich* trajectories, shown in orange, consist of shifts towards the target after the onset of the disambiguating noun (left panel, 32% of the data), change-of-minds shifts towards the target within the particle window (middle panel, 60% of the data), as well as shifts towards the target before the onset of the particle without subsequent change of course (right panel, 8% of the data). Taken together, these results suggest that both *eigentlich* and *tatsächlich* are used predictively in the interpretation of the referring expressions, such that their presence in an utterance induces participants to move their mouse towards the referential target before hearing information from the disambiguating noun in 90% of *tatsächlich* trials and 68% of *eigentlich* trials.

Visual inspection of the mean trajectories within each cluster in the **unreliable** group

shows that, as in the **reliable** group, in two of the three clusters, the trajectories for *eigentlich* and *tatsächlich* differ qualitatively from one another. Crucially, while the qualitative patterns within each cluster are similar to those in the **reliable** group, the exact shape of the trajectories is different when participants are exposed to unreliable usages of the particles. Interestingly, in the case of the cluster in the left-hand panel, the shifts towards the target are delayed compared to the same cluster in the **reliable** group. In the cluster in the middle panel, the shift towards the target for *tatsächlich* items occurs earlier compared to the same cluster in the **reliable** group, following a more curved trajectory. The change-of-mind shift towards the target for *eigentlich* items, on the other hand, is exacerbated compared to the cluster in the **reliable** group, meaning that due to a larger initial bias for the mentioned referent the curvature of the trajectory is also more pronounced. Finally, in the cluster in the right-hand panel, the initial bias for the target for both *eigentlich* and *tatsächlich* items is exacerbated compared to the same cluster in the **reliable** group. Following this exacerbated initial bias, the trajectory for *tatsächlich* items is flatter than in the same cluster in the **reliable** group, while the trajectory for *eigentlich* items first flattens before shifting abruptly towards the target. Taken together, these results suggest that exposure to unreliable trials generally leads to temporal modulations of the qualitative patterns found in the **reliable** group, together with an overall reduction in the number of trials in which the particles are used predictively, from 90% to 67% in the case of *tatsächlich* and from 68% to 57% in the case of *eigentlich*. Moreover, in the case of the target bias found in the predictive trajectories, there is a modulation of the magnitude of the bias, with an increase in the bias in the **unreliable** group.

In order to better assess the impact the particles have on the resolution of reference, I focus on the shapes of the mouse trajectories within a critical time window spanning 150 ms after the onset of the particle up until 150 ms after the onset of the disambiguating noun, as shown at the top of Figure 4.4.

Crucially, any changes in the relative horizontal mouse position within this time window

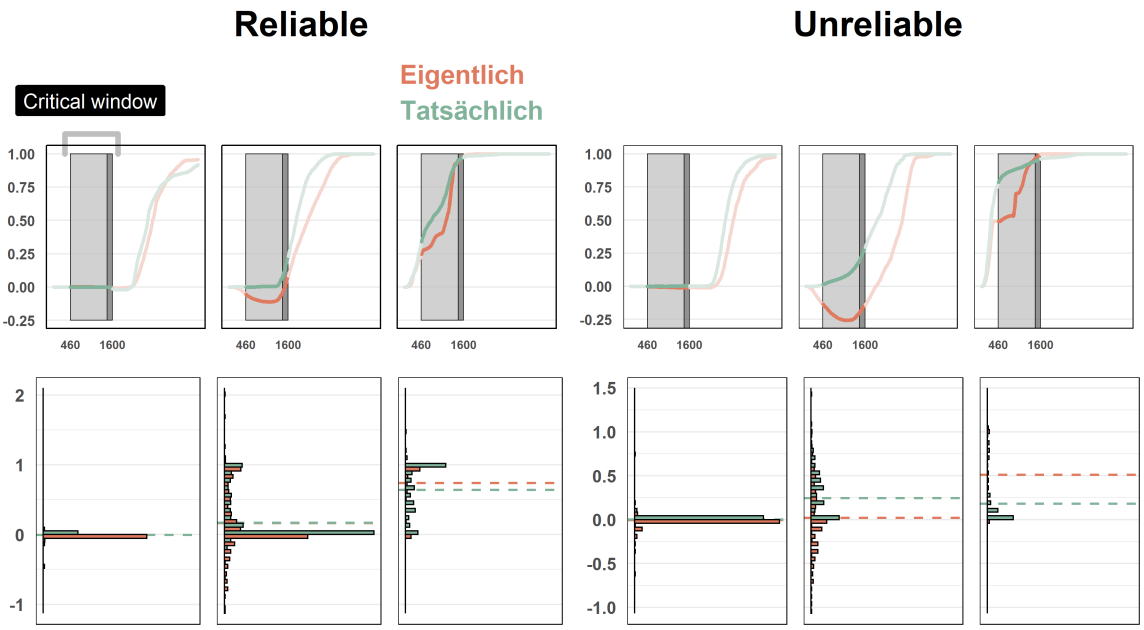

Figure 4.4 Horizontal mouse position within the critical time window. **Top**: Clusters of mean trajectories aggregated over items and participants, with the segments falling within the critical time window highlighted. The shaded areas in the graphs show the (average) critical time window, which starts at 150 ms after the onset of the particle and ends at 150 ms after the onset of the disambiguating noun. **Bottom**: Distributions of differences in the horizontal mouse position for each participant-item pair, shown in terms of each trajectory cluster. The dotted lines show the mean differences for each discourse particle. The left column shows data from the **reliable** group, while the right column shows data from the **unreliable** group. *Tatsächlich* items are shown in green while *eigentlich* items are shown in orange.

can reasonably be assumed to be an effect of processing the particle, such that movements either towards or away from the mentioned referent indicate a sensitivity to the meaning of the processed particle. In order to quantify such effects, I calculate for each participant-item pair the difference between the absolute horizontal mouse position at the start and end of the critical time window, which, on average, are at the 460 and 1600 ms marks, respectively. The computed differences are shown at the bottom of Fig. 4.4, with values around zero indicating no change in horizontal position, positive values indicating movements towards the target, and negative values indicating movements towards the competitor. Note that the target is different for each particle: while in the case of *tatsächlich* it is the referent

mentioned in the question, in the case of *eigentlich* it is the unmentioned referent. In order to quantify uncertainty over the results reported below, I fitted two Bayesian hierarchical models predicting differences in horizontal mouse position within each trajectory cluster as a function of an interaction between the particles and the experimental groups, with *eigentlich* in the **reliable** group dummy-coded as the reference level. Each model estimated changes in the magnitude of the horizontal position when making pairwise comparisons between clusters. Both models included the maximal random effect structure justified by the design, which in the present case is random slopes and intercepts for the effect of particle for both participants and items. Since participants were assigned to only one experimental group, only at the item level do I estimate random slopes for the effect of group. The models, fitted using the *R* package `brms` (Bürkner, 2017) and described in detail in the scripts available in the supporting material, had the following form, shown in `brms` syntax:

$$\begin{aligned} \text{horiz\_diff} \sim\ & \text{particle} * \text{group} \\ & (1 + \text{particle} \mid \text{participant}) + \\ & (1 + \text{particle} * \text{group} \mid \text{item}) \end{aligned}$$

First, I report the mean differences within each cluster in the **reliable** group, anchoring them to the actual trajectories shapes. As visual inspection of the figure shows, in the cluster in the left-hand panel, both the *tatsächlich* and the *eigentlich* trajectories are flat, resulting in differences which are practically zero. In the cluster in the middle panel, the *tatsächlich* trajectory is flat through most of the critical window, suddenly shifting towards the target shortly before the onset of the noun, resulting in a difference of 0.17 ($\beta = 0.18$, 95% CrI [0.12, 0.25], $P(\beta > 0) = 1$). The *eigentlich* trajectory, on the other hand, shows a light curvature, first following an initial bias for the mentioned referent before changing course towards the unmentioned referent, resulting in a difference of 0.17 ($\beta = 0.16$, 95% CrI [0.08, 0.25], $P(\beta > 0) = 1$). In the cluster in the right-hand panel, the *tatsächlich* trajectory follows an initial bias for the mentioned referent, being mostly linear and resulting in a difference

of 0.64 (β = 0.73, 95% CrI [0.63, 0.83], $P(\beta > 0) = 1$). The *eigentlich* trajectory, on the other hand, follows an initial bias for the unmentioned referent before suddenly accelerating towards it midway in the critical window and thus resulting in a difference of 0.74 (β = 0.77, 95% CrI [0.62, 0.93], $P(\beta > 0) = 1$).

As for the cluster in the left-hand panel in the **unreliable** group, visual inspection of the figure shows that, much like in the same cluster in the **reliable** group, both the *tatsächlich* and *eigentlich* trajectories are flat, resulting in differences which are practically zero. In the cluster in the middle panel, the *tatsächlich* trajectory follows an initial bias for the mentioned referent, being more curved than the same trajectory in the **reliable** group and thus resulting in a difference of 0.25 (β = 0.30, 95% CrI [0.20, 0.41], $P(\beta > 0) = 1$). The *eigentlich* trajectory, on the other hand, shows a strong curvature, first following an initial bias for the mentioned referent before changing course towards the unmentioned referent, resulting in a difference of practically zero. In the cluster in the right-hand panel, the *tatsächlich* trajectory follows an initial bias for the mentioned referent, being more curved than the same trajectory in the **reliable** group and thus resulting in a difference of 0.18 (β = 0.31, 95% CrI [0.15, 0.47], $P(\beta > 0) = 1$). The *eigentlich* trajectory, on the other hand, starts mostly flat before suddenly accelerating towards the unmentioned referent midway in the critical window, thus resulting in a difference of 0.51 (β = 0.65, 95% CrI [0.35, 0.95], $P(\beta > 0) = 1$).

4.3 Discussion

In the present study, I investigated the role modal-marking discourse particles play in incremental interpretation and the online resolution of reference. To that end, I used a mouse-tracking task where participants were asked to listen to mini dialogues which included referring expressions containing the German particles *eigentlich* and *tatsächlich*. For each dialogue, participants were presented with two images on a screen and instructed to select

the one that best portrayed the topic of the exchange. Crucially, I was interested in the potential differential effects of processing *eigentlich* and *tatsächlich* in context, as measured in terms of any relative changes to the position of the mouse after hearing the particle. I was also interested in the effects of varying the pragmatic reliability of the particles, as manipulated in a separate condition which included both reliable and unreliable particle usages.

The results show that, in the reliable group, both *eigentlich* and *tatsächlich* were used predictively in the interpretation of the referring expressions, such that their presence in an utterance induced participants to move their mouse towards the relevant referential target – which was different for each particle – before hearing information from the disambiguating noun. The results also show that exposure to unreliable particle usages not only led to temporal modulations of the qualitative patterns found in the reliable group but it also reduced the predictive potential of the particles, as evidenced by an overall reduction in the number of predictive trajectories.

What these results suggest is that, as found for the Dutch particles *eigenlijk* and *inderdaad*, discourse particles with a modal-marking function do affect incremental interpretation, at least in the context of reference resolution. Crucially, they affect not only the outcome of the interpretive process but also how it unfolds in real-time. While the contrast-marking particle *eigentlich* modulates interpretation insofar as it induces comprehenders to revise an assumption at issue in the discourse, e.g., that the discourse referent under consideration is actually a dog as opposed to a wolf, the agreement-marking particle *tatsächlich* modulates interpretation insofar as it induces comprehenders to confirm a prior assumption at issue in the discourse, e.g., that the discourse referent under consideration is indeed a dog as opposed to a relevant alternative like a wolf.

These findings raise a number of questions about the status of modal particles both in relation to discourse particles more generally as well as in relation to other linguistic markers of epistemicity. Regarding their role as particles with the specialized function of marking

modality, the present findings illustrate how modal particles can impact the interpretation of their host utterances, underlining their modal character and indeed warranting their special status among the more general class of discourse particles. At the same time, while in the present study the materials were designed so as to explicitly elicit a modal interpretation of *eigentlich* and *tatsächlich*, in van Bergen and Bosker (2018), for instance, the materials allowed both a modal as well as a non-modal interpretation of *eigenlijk* and *inderdaad*. As their results regarding *eigenlijk* suggest, a particle's modal meaning might not necessarily be the most salient meaning in context, at least not when the two meanings are relatively similar to one another, as in the case of *eigenlijk*/ *eigentlich* and *inderdaad*/ *tatsächlich*, which signal complex pragmatic meanings which are hard to distinguish from one another.

As discussed before, however, the difference between marking a general sense of (dis)agreement between two utterances and (dis)agreement between two propositions might be one of degree as opposed to type. Indeed, even in the case of particles which signal arguably quite different meanings depending on their particular usage – such as *bloß* and *nur*, which index either a subjective sense of strengthening or a restriction depending on what they relate their host utterance to – the difference between a particle's modalized and non-modalized meaning might not be as apparent as it seems. For instance, Dörre and Trotzke (2019) argue that there is a "common semantic denominator" between the two senses of *bloß* and *nur*, one which involves restriction of a particular relevance domain. In the case of, *eigenlijk*/ *eigentlich* and *inderdaad*/ *tatsächlich*, both meanings are related to the management of expectations in discourse, whether those are expectations to do with concrete propositions which can be modalized over or expectations related to the flow of conversation and the general maintenance of understanding, which can nevertheless also be signaled upon.

Regarding their role as markers of epistemicity, modal particles should be considered vis-à-vis other such markers, like other lexical modal markers as well as morphosyntactic and prosodic modulations of the linguistic signal, to highlight only canonical linguistic means of marking epistemicity. Indeed, prosodic as well as morphosyntactic cues can be used to

mark one's stance in discourse, both in the absence of and in conjunction with particles. For instance, let's look back at examples (4)-(6) from the introduction, reproduced here as (18)-(20).

A: Did you see how fast that bird flew by?

(18) B: *Really*? I didn't even notice.

(19) B: I don't *really* think it flew by that fast.

(20) B: It *really* did fly by fast.

In the examples above, B uses many different strategies in addition to their usage of *really* to mark their stance with regard to the content of A's utterance: in (18), B employs the particle *even* to emphasize an aspect of their own response to A's statement; in (19), B employs the intensifier *that* to highlight information about the degree of the event at issue; in (20), B employs a non-canonical syntactic structure to emphasize the validity of A's statement; finally, both in (19) and (20) a pitch accent could be used to signal different degrees of epistemic belief in the strength of the respective propositions, i.e., "I don't REALLY think it flew by that fast" vs. "I don't really think it flew by THAT fast" (19) and "It REALLY did fly by fast" vs. "It really did a bird" in (20). This illustrates not only the variety of different linguistic strategies which can be used to mark epistemicity in the case of English but also the expressive potential of combining different strategies.

As shown by Kurumada et al. (2014), for instance, combining different cues can help guide comprehension in context, as when prosody modulates the interpretation of a lexical expression like "it looks like". Similarly, as shown by Dörre and Trotzke (2019), a non-lexical marker can be used to constrain the meaning of a lexical marker even in the absence of otherwise overt contextual cues, as when pitch accent disambiguates the meaning of a modal particle. All in all, what this suggests is that modal particles should be seen as yet another tool in the toolkit of modality markers, one which shows functional behaviors similar to those of other modal tools despite, from a formal perspective, belonging to a special

class of markers. In that vein, future research could address the role of competing modal cues in incremental interpretation, such as when pitch accent and a modal particle provide conflicting information in context. Is interpretation affected by such a pragmatic mismatch between different linguistic sources of information? Does one cue take precedence over the other, such that maybe specialized markers like particles are interpreted more or less reliably depending on whether non-lexical cues like pitch accent or word order alternation are used unreliably? The findings of the present study provide evidence that a pragmatic mismatch in the form of distributional unreliability in the usage of the particle itself does affect the predictive protential of a particle in context. Future research should investigate the role played by other forms of pragmatic mismatch in the interpretation of modal particles and lexical markers of stance more generally.

In terms of the cognitive nature of the meanings associated with a polyfunctional particle, one could ask whether their modal meaning is represented differently than their non-modal counterpart. Since in the case of some particles the two meanings seem to be intertwined not only diachronically but also synchronically, with subtle differences in terms of their exact synchronic functions, that raises questions about the nature of not only the representations themselves but also the interpretive processes underlying the understanding of modal particles in context. While earlier work suggested that the more conventionalized the modal meaning of a particle, the harder it is to process it (Dörre, Czypionka, et al., 2018), the present findings suggest a role of pragmatics in guiding the interpretation of particles in context. Future research could thus more directly address these sorts of questions, dissecting the exact dynamics of modal meaning representation.

All in all, the findings of the present study provide evidence that the particles *eigentlich* and *tatsächlich* serve as predictive cues in the incremental processing of language, allowing comprehenders to anticipate lexical material in a sentence before actually processing that material. These particles thus seem to function as pragmatic cues to interpretation, modulating the illocutionary force of the utterance they occur in. However, the nature of the

modulation is one which depends on the modal meaning associated with either particle, suggesting an interpretive process akin to the derivation of conventional implicatures, that is, one which builds on both conventionalized (non-truth-conditional) meaning as well as pragmatic expectations.

REFERENCES

Abraham, Werner (2016). "Discourse marker = discourse particle = thetical = modal particle? A futile comparison". In: *Discourse Particles: Formal Approaches to their Syntax and Semantics*. Ed. by Josef Bayer and Volker Struckmeier. De Gruyter, pp. 241–280. DOI: doi:10.1515/9783110497151-010. URL: https://doi.org/10.1515/9783110497151-010.

Bürkner, Paul-Christian (2017). "brms: An R package for Bayesian multilevel models using Stan". In: *Journal of statistical software* 80.1, pp. 1–28.

Döring, Sophia and Sophie Repp (2019). "The modal particles ja and doch and their interaction with discourse structure: Corpus and experimental evidence". In: *Experiments in Focus*.

Dörre, Laura, Anna Czypionka, et al. (2018). "The processing of German modal particles and their counterparts". In: *Linguistische Berichte* 255, pp. 58–91.

Dörre, Laura and Andreas Trotzke (2019). "Chapter 6 The Processing of Secondary Meaning: An Experimental Comparison of Focus and Modal Particles in Wh-questions". In: *Secondary Content: The Semantics and Pragmatics of Side Issues*. Leiden, The Netherlands: Brill, pp. 143–167. ISBN: 9789004393127. DOI: https://doi.org/10.1163/9789004393127_007. URL: https://brill.com/view/book/9789004393127/BP000006.xml.

Fedriani, Chiara and Andrea Sansó (2017). "Pragmatic Markers, Discourse Markers and Modal Particles. What do we know and where do we go from here?" In:

Feng, Guangwu (2011). "A neo-Gricean pragmatic analysis of Chinese pragmatic markers". In: *Language Sciences* 33.3, pp. 417–434. DOI: https://doi.org/10.1016/j.langsci.2010.10.005.

— (2010). *A Theory of Conventional Implicature and Pragmatic Markers in Chinese*. Leiden, The Netherlands: Brill. DOI: https://doi.org/10.1163/9789004253179.

Fischer, Kerstin (2006). "Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume". In: *Approaches to discourse particles*. Brill, pp. 1–20.

Gast, Volker (2008). "Modal particles and context updating – the functions of German ja , doch , wohl and etwa 1". In:

Heine, Bernd et al. (2021). *The Rise of Discourse Markers*. Cambridge University Press. DOI: 10.1017/9781108982856.

Jucker, Andreas H. and Yael Ziv (1998). "Discourse markers : Introduction". In: *Pragmatics and beyond. New series* 57, pp. 1–12.

Kieslich, Pascal J. and Felix Henninger (2017). "Mousetrap: An integrated, open-source mouse-tracking package". In: *Behavior Research Methods* 49, pp. 1652–1667.

Kurumada, Chigusa et al. (2014). "Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings". In: *Cognition* 133.2, pp. 335–342.

Maschler, Yael and Deborah Schiffrin (2015). "Discourse markers: Language, meaning, and context". In: *The handbook of discourse analysis* 1, pp. 189–221.

Rasenberg, Marlou, Joost Rommers, and Geertje van Bergen (2019). "Anticipating predictability: an ERP investigation of expectation-managing discourse markers in dialogue comprehension". In: *Language, Cognition and Neuroscience* 35, pp. 1–16.

Roettger, Timo B and Michael Franke (2019). "Evidential strength of intonational cues and rational adaptation to (un-) reliable intonation". In: *Cognitive science* 43.7, e12745.

Schoonjans, Steven (2013). "Modal particles: Problems in defining a category". In: *Discourse markers and modal particles: Categorization and description*, pp. 133–161.

Unger, Christoph (2016). "DEGREES OF PROCEDURE ACTIVATION AND THE GERMAN MODAL PARTICLES JA AND DOCH - PART 3". In: *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 2016, pp. 31–45.

Bergen, Geertje van and Hans Rutger Bosker (2018). "Linguistic expectation management in online discourse processing: An investigation of Dutch inderdaad 'indeed' and eigenlijk 'actually'". In: *Journal of Memory and Language*.