Research Article
Oliver Ehmer*, Malte Rosemeyer

Inferences in Interaction and Language Change

https://doi.org/10.1515/opli-2018-0026
Received September 21, 2018; accepted October 5, 2018

Keywords: inference, implicature, interaction, language change

1 Why inference matters

Utterances usually convey more meaning than is expressed. This ‘surplus’ of meaning can be explained by the process of inferencing. A typical definition is given, for example, by Huang, who defines inference as the “process of accepting a statement or proposition (called the conclusion) on the basis of the (possibly provisional) acceptance of one or more other statements or propositions (called the premises)” (Huang 2011: 397). This definition rests on the basic distinction that there is an encoded meaning for linguistic signs from which further meaning may be arrived at by inferences. Two types of inference can be distinguished: semantic inference, i.e. logical entailment, and pragmatic inference. Entailment reflects logical connections between sentences; for instance the sentence All of my friends like reading inescapably entails Some of my friends like reading. In contrast, pragmatic inference is based on default logic, i.e. “reasoning on the basis of stereotypes and prototypes” (Eckardt 2006: 86). For instance, in the correct context and with the correct intonation the sentence ALL of my friends like reading might lead to the inference on part of the hearer that she is either not considered a friend or should pick up reading as a hobby. Given that pragmatic inferences are based on non-monotonic, i.e. probabilistic, logic, they can be canceled, whereas entailments cannot.

From a usage-based perspective on language, it is pragmatic inference that is particularly important to the study of interaction and language change. Whereas entailments are unlikely to be discussed in discourse (e.g., upon hearing All of my friends are reading I am unlikely to react by asking Are some of your friends reading?), pragmatic inferences are frequently dealt with in interaction and may, for example, become the topic of conversation (e.g., I might react to ALL of my friends are reading by saying So does this mean I am not your friend?). In addition, it is a commonplace in historical linguistics that meaning change is often derived from pragmatic inferences. For example, historical and typological studies have observed that future tense constructions frequently derive from modal constructions expressing obligation, on the basis of the inference that speakers inferred intentional from obligation readings, and, in a second step, prediction from intention readings.

Many linguists thus take pragmatic inference to be an important part of pragmatics, and in fact inference has even been used as the basis for defining pragmatics as opposed to other domains, mainly semantics (Ariel 2010). However, pragmatic inferences have been considered to a far lesser extent in approaches like conversation analysis and interactional linguistics, which are concerned with the organization and formation

Article note: This paper belongs to the special issue on Inferences in Interaction and Language Change, ed. by Oliver Ehmer & Malte Rosemeyer.

*Corresponding author: Oliver Ehmer, Albert-Ludwigs-Universität Freiburg, oliver.ehmer@romanistik.uni-freiburg.de
Malte Rosemeyer, KU Leuven
of social action. But for interaction, too, pragmatic inferences play a crucial role; they have a profound impact on how social meaning is construed. As an example of the intricacies of pragmatic inference processes in interaction, consider the text passage in (1), taken from a novel by U.S. American writer T.C. Boyle. Alfred Kinsey, the character referred to in third person, has just invited the protagonist John Milk over for dinner, stating at the same time that Milk probably does not get “all that much home cooking”. After a brief discussion of this point, Kinsey returns to his unanswered invitation, which is where the excerpt starts.


01 “Well,” he said, pulling up short and swinging round on me, “so how about it? Shall we
02 say Saturday, six p.m.?”
03 I must have hesitated, because he added, in a different tone altogether, almost as if he
04 were preparing to build a case, “If you’re not doing anything, of course. You don’t have
05 any plans, do you? For Saturday?”
06 I looked off down the deserted street, then came back to him. “No,” I said. “Not really.”

In lines 1 and 2, Kinsey re-opens the question under discussion, namely his invitation to dinner. Apparently, Milk does not react immediately to Kinsey’s two questions (at least this is what the narrator Milk formulates in line 3). This hesitation leads to an inference on Kinsey’s part. Specifically, Kinsey’s continuation in lines 4–5 suggests that he inferred from Milk’s hesitation that Milk might not accept the invitation. This is however not the case, as Milk claims that he is not busy on Saturday (line 6) and thereby accepts the invitation, as also evidenced in the progress of the novel.

Kinsey’s inference that motivates his utterance in lines 4–5 is a prototypical example of an inference in the sense that it is an abduction. Upon noting Milk’s failure to take the turn immediately, Kinsey infers that he might decline the invitation. We may additionally say that a further inferencing leads to Kinsey’s assumption that the most likely reason for Milk to decline the invitation would be to already have an appointment on Saturday. The fact that Kinsey makes this inference explicit — in the form of a question in lines 4–5 — should not be seen as facilitating Milk’s ability to refuse the invitation. Rather, Kinsey effectively pressures Milk into accepting the invitation by excluding the most likely reason for this refusal. Our interpretation is supported by the fact that Kinsey adopts “a different tone altogether”, which Milk interprets as him “preparing to build a case”. Note also that Kinsey’s question You don’t have any plans, do you? is biased in the sense that it presupposes that Milk does not have any plans. By introducing the proposition as a presupposition, Kinsey makes it more difficult for Milk to decline the invitation. Note that for this rhetorical effect it is not actually necessary for Kinsey to believe in the proposition. One page later in the text, Milk narrates that he indeed canceled a date with his future wife on that Saturday. This example makes apparent some of the inferential processes of the characters (as constructed by the author of the novel) that can be captured by definition of inference given above.

However, our example also demonstrates possible problems with the definition of inference given above, as well as additional important characteristics of inferences in conversation. First, Kinsey’s inference is not based on a verbally expressed proposition, but rather nonverbal behavior, i.e., Milk’s assumed pause. This goes to show that as domain-general mental processes, inferences can be based on quite different kinds of indexical cues and are not limited to verbal stimuli. Second, the example shows that a given ‘premise’ can lead to different inferences. Kinsey’s inference (Milk might have hesitated because he might decline the invitation) is a likely inference due to the social convention of hesitating before realizing a non-preferred action such as refusing an invitation. Other inferences would have been possible, but are less likely. For instance, Milk might have simply not been following the conversation. This means that there will usually be not only one but a set of possible inferences from one ‘premise’. This observation leads us to formulate a third problem. Interactants have to know which one of these possible inferences is most likely given a specific situation. This process of choosing between possible inferences — and actually the raising of any inference itself too — therefore presupposes different kinds of knowledge, such as social conventions. Consequently, inference is not a process that runs automatically and unequivocally from premises to conclusions.
The definition given above implies that inference processes occur in the mind of one person. However, our example demonstrates that inference processes are frequently attributed to other interactants. In line 3, Milk argues that Kinsey’s questions with regard to Milk’s plans on Saturday might have resulted from Milk’s hesitation. In other words, Milk construes a mental model of Kinsey in which he simulates the inferential process that might have led to Kinsey’s question (cf., e.g., Deppermann 2014). We are thus dealing with what one could call a “second-order inference”: inference over a possible inference process realized by the interlocutor. As evident in Grice’s (1975) original approach to the topic, such second-order inferences are what enables implicatures, i.e. “any meaning implied or expressed by, and inferred or understood from, the utterance of a sentence which is meant without being part of what is strictly said” (Huang 2011: 407). In order to implicate a certain interpretation a speaker needs to be able to assume that the hearer will draw exactly the inference that leads to the implicated meaning. At a more basic level, the study of inferences thus touches upon the same issues dealt with in research on Theory of Mind (Malle and Hodges 2005; Call and Tomasello 2008; Apperly 2011), Common Ground (Clark and Brennan 1991; Clark 1996: 92-121; Stalnaker 2002) and intersubjectivity (Zlatev et al. 2008; Sidnell, Enfield and Kockelman 2014). These considerations lead us to conclude that the study of inferences requires a principled account of the role of speaker and hearer in interaction. Ultimately, the example also demonstrates that inference in interaction is subservient to joint action. Kinsey and Milk are “caught in a web of inferences” (Levinson 1983: 321) not because inferring is an end in itself, but because they are negotiating a specific social action, i.e., an invitation.

Although the study of inferences is traditionally viewed as falling within the domain of linguistic pragmatics, inference has also come to be considered an important explanatory factor in related aspects of language description such as interaction and meaning change. However, these disciplines tackle the concept of inference from very different angles, using different methodologies. Whereas studies on language change focus on the potential of inferences to bring about meaning change, they have only recently begun modeling the role of the relationship between speaker and hearer in these processes. In contrast, studies on conversation have long focused on the dialogical emergence of meaning and understanding, but usually have given preference to observable actions, without considering cognitive processes such as inferencing in discourse. To quote Maynard (2012: 28):

CA [conversation analysis] crucially eschews the rule-based approach of speech act theory and is agnostic about cognitive maxims, intentions and other psychological features that are said to explain the meaning of utterances. Rather, the focus is on participants’ observable attributions and displays as these occur through visible, hearable ways in everyday talk.

Not all scholars in interaction and conversation studies, however, share this stance but rather include inferences in different ways into their analysis (cf. for example the contributions in Molder and Potter 2005). The present special issue brings together scholars working on language change and interaction who attach importance to inferences. The papers are based on the results of a colloquium held in November 2016 at the University of Freiburg. Specifically, the central aims of the special issue are threefold:

1. To assess the importance of inferential processes in different domains of language use (e.g., morphology, syntax, pragmatic markers) both in interaction and language change.
2. To contribute to the development of a contextualized model of the roles of speaker and hearer in the synchronic and diachronic emergence of meaning.
3. To bring together studies that illustrate similarities and differences in currently used methods in the analysis of inferences from the perspective of conversation analysis, interactional linguistics, and historical linguistics.

1 We would like to thank all participants in the colloquium for the lively discussions. Our special thanks go to the invited discussants: Bert Cornillie, Daniel Jacob, Ekkehard König, Stefan Pflünder and Esme Winter-Froemel. We would also like to express our gratitude to the reviewers, who put a lot of effort in commenting on the papers collected in this issue: Galina Bolden, Hendrik De Smet, Andreas Dufter, Michael Haugh, Bettina Kluge, Benjamin Meinsitzer, Álvaro Octavio de Toledo y Huerta, Florence Oloff, Jacqueline Viscotti and several anonymous reviewers. The colloquium was generously funded by the Deutsche Forschungsgemeinschaft (DFG) in the context of the Graduate School “Frequency effects in language”, as well as the Freiburg Institute of Advanced Studies (FRIAS).
2 What can studies on inferences in interaction and language change learn from each other?

In this section, we summarize the relevance of the study of inferences in interaction (2.1) and language change (2.2) and in each case highlight points of contact between the two disciplines. We do not, however, aim to provide a comprehensive overview of the study of inferences in both fields.

2.1 Interaction

In the study of conversation and interaction several approaches are concerned with inferences. The main point of those studies is that both speaker and interlocutor(s) are involved in the local management of inferences in context as part of the organization of social interaction.

The importance of inferences for everyday reasoning has already been pointed out in Garfinkel’s (1967) studies on ethnomethodology. His famous ‘breaching experiments’ were designed such that an experimenter purposefully deviated from an implicit norm, e.g., by insisting on a clarification of the sense of a commonplace remark by an interlocutor. One important observation from these breaching experiments was that the deviation usually did not lead to a failure of comprehension or interactional disorder. Rather, the participants in such an experiment tended to interpret the experimenter’s actions as motivated deviations from the norm, and gave rise to inferences about the motivation of the experimenter for this deviation, who most often was judged to be offensive and was treated as such (Heritage 1984: 97-101). The inferences raised thus often contributed to the maintenance of the norms they were based on.

Building on such observations, conversation analysis has shown that a large part of the mechanisms whereby participants organize social interaction rests on standard assumptions maintained by the participants and the inferences they allow for. In turn, any deviation from what is expected will give rise to further inferences, for which the speaker will be held accountable (Levinson 1983; Heritage 1984: 97-101). For instance, sequence organization and turn-taking rests on expectations about how discourse continues (Sacks, Schegloff and Jefferson 1974; Schegloff 2007). Even minuscule deviations from standard routines, such as a hesitation, may result in a noticeable departure (Schegloff 2007), which leads to inferences on behalf of their partners (cf. Levinson 1983: 321), as illustrated in our introductory example.

That situational and institutional factors play a central role in inferencing has been shown, for example, by Drew and Heritage (1992). Talk in institutions may be highly dependent on specific ‘inferential frameworks’, in the authors’ terms. Participants may also use those frameworks strategically to invite certain inferences (Drew and Atkinson 1979; Drew 1985; 1992). Drew and Atkinson (1979) show, for example, that cross-examinations in present-day Anglo-Saxon court interactions are characterized by the fact that once a witness has finished his/ her turn, the next turn is automatically allocated to the counsel. The counsel may now use this pre-allocation of the next turn strategically by producing a ‘significant pause’ before starting his next turn. Such a pause is intended to lead the jury to certain inferences, e.g., that the answer should be doubted.

Inferences and standard assumptions, however, do not play a role only at the sequential level of action organization, but also, for example, in the constitution of situations. In his lecture ‘The inference-making machine’, Sacks (1989) proposes the membership categorization device as one central mechanism – or, in his words, ‘machinery’ – for the constitution of situations and social identity. This device rests on the existence of ‘inference-rich’ social categories, which allow to make default assumptions about any representatives of the respective category (cf. Sacks 1972). That culture-specific background-knowledge may play a central role in interaction has been emphasized in Interactional Sociolinguistics (Gumperz 1982; 1992; 1993; 2000). Inferences are here seen as the “mental operations we engage in to retrieve such knowledge and integrate it into the interpretative process” (Gumperz 2000: 131), at several levels of granularity, for example the level of the clause, sequence organization, and more global levels of the activity (Gumperz 1993: 200).
Such inferences usually remain ‘embedded’ in the course of the actions that the participants perform (Haugh 2017). They may, however, also be ‘exposed’, i.e. be elevated to a topic that is to some extent explicitly discussed by the participants. Members of a speech community have different resources for dealing with such exposed inferences in interaction. For example, Schegloff (1996) analyses a practice that participants employ to deal with a contribution in which a speaker has made an allusion. This allusion is subsequently explicitly formulated by a partner. In a third sequential step the first speaker produces a verbatim repetition of the partner’s formulation, and thereby confirms that the partner has understood his allusion correctly, or, in Schegloff’s words, he confirms “both the allusion and that it had been an allusion” (1996: 210).

Besides such specific sequential patterns inferences have been shown to play a role in on-line syntax, e.g., in projection (Auer 2005; 2009; 2015) and ellipsis (Imo 2011; 2014). For example Imo (2014) analyses certain kind of syntactic breakoffs, by which the speaker leaves the continuation and a possible meaning of the contribution open to be inferred by his/her interlocutors. In example (2) speaker A produces a break off after aber ‘but’. With her continuation, speaker B signals that she understood the gist of A’s contribution, i.e., that the dishcloth they are talking about is really old, so old it already looks alive.

Example (2). Lappenphobie (adapted from Imo 2014: 145)

01 A: <ich hab keine LAPpenphobie aber- <lachend>>
   ‘I have no dishcloth phobia but <laughing>>’

02 B: he he he weil dEr hier so lebEndig AUSSieht.
   ‘he he he because this one looks as if it is alive’

As Imo shows, speakers systematically use such syntactic breakoffs as an interactive device to induce other participants to draw and explicate certain inferences. It has not only been shown, that fragmented syntactic structures may be used as a resource for negotiating alignment (Pfänder 2016), but also that their systematic use may lead to the conventionalization of such patterns, as for example in the case of but as a the grammaticalized turn-final particle (Mulder and Thompson 2008; cf. also Thompson and Suzuki 2011; Barth-Weingarten 2014). Such studies show that the inference that the speaker wants his/her partners to draw (the implicature) may be spelled out or become manifest in interaction to different degrees.

Moreover, studies in Conversation Analysis and Interactional Linguistics have identified explicit linguistic devices for managing inferences. For example, speakers have resources to introduce utterances as the formulation/explication of an inference. Bolden (2010) has shown for English that a speaker may use the connective and to introduce a turn in which he/she articulates an element that is ‘missing’ in the addressee’s prior talk, but which is claimably inferable. For German, Deppermann and Helmer (2013) have shown (i) that also ‘so/ then’ and dann ‘so/ then’ both function as grammaticalized resources to explicate a meaning that is implicit in a prior turn of the partner, but (ii) that the two connectives attribute different types or degrees of epistemic authority to the speaker and the hearer. More specifically, dann may be used to introduce the formulation of a unilateral inference, which is not necessarily shared by the interlocutors. The following excerpt, which is taken from a medical discussion about abortion, provides a case in point.

Example (3). medizin (adapted from Deppermann and Helmer 2013: 20)

01 HT: <<all,f>dann hAm sie keine Ahnung von der mediZIN==
   so/ then you have no notion of medicine at all

02 =entSCHULdigen sie bitte;=
   please excuse me

03 =wenn ich ihnen das VORwerfen darf.>
   if I am allowed to criticize you about this
The formulation of a unilateral inference introduced with *dann* (01) is subsequently explicitly framed as a reproach (02-03). In contrast, utterances that are introduced with *also* are never competitive or used to express opposition (Deppermann and Helmer 2013: 32).

Besides connectives, there are additional devices to manage inferences. For example, Diewald and Fischer (1998) and Fischer (2007) have analyzed modal particles as resources that are used by speakers to allow partners to infer a relevant argumentative discourse for a current utterance. Furthermore, sentential negation in interaction has been analyzed as a resource to cancel unintended or unwanted inferences that are or might be drawn by interlocutors (Deppermann and Blühdorn 2013; Deppermann 2014). These authors also take into account different degrees of manifestness, ranging from the explicit formulation of an assumption to mere inferability. In addition to lexico-syntactic devices, bodily resources, have been shown to be instrumental in managing context-derived inferences. For example, Enfield has shown that “information about head orientation allowed for inference of direction of eye gaze, a deictic signal of obvious importance in pointing” (2009: 93).

The central advantage of synchronic studies is that they permit the investigation of the actual mechanisms through which inferences are dealt with in the actions of the participants. However, scholars in the study of conversation and interaction also face several problems that have been worked on extensively in historical linguistics. We mention just three of these problems. First, the inferences under investigation are located on different but interdependent levels (action, syntax, semantics, etc.), with no coherent model yet available. Second, variation in the data concerning the interactional function of linguistic structures may be due to layering, i.e., polyfunctionality resulting from diachronic processes. Third, longer, more monological contributions to discourse, but also ‘deviant cases’ in which actual conversational moves of a pattern are ‘missing’, are notoriously difficult to handle, although they may rely on the same inferential processes (cf. Ehmer 2016). Interactional linguistics can also profit from historical linguistics in that diachronic changes may offer evidence for interactional processes not easily observable in synchrony.

### 2.2 Language change

Studies on language change have long assumed inferences to play a prominent role in meaning change and grammaticalization/ constructionalization. Not long after Grice’s (1975) seminal paper on logic and conversation, historical linguists began exploiting the notion of implicature in the description of meaning change. Many of these studies assume that meaning change results from the conventionalization of conversational implicatures (cf., e.g., Dahl 1985: 11; Traugott and König 1991; Bybee, Perkins and Pagliuca 1994; Traugott and Dasher 2002; Hopper and Traugott 2003: 81-84). This notion eventually replaced the assumption, dominant at that time, that meaning change involved metaphorical mapping (Sweetser 1990).

Consider, for instance, the rise of the periphrastic future in Romance languages. Due to the metonymic link between obligation and future, speakers of Vulgar Latin may have used the deontic *cantare habeo* ‘I have to sing’ construction in order to implicate the future meaning ‘I will sing’ (Fleischman 1982; Pinkster 1987). Consequently, “intention is the crucial bridge to prediction” (Bybee, Perkins and Pagliuca 1994: 279-280): first, an expression is taken to imply speaker intention. This implicature becomes part of the meaning of the expression, which means that it can be used to express the intention of the agent of the main verb. A second inference can then lead to the meaning of prediction; since the speaker can only guess the intention of the agent of the main verb, she is effectively predicting the action. Consider, for instance, examples (4–5) taken from Pinkster’s study:

(4) *multos ferro, multos veneno (occidit); habeo etiam dicere quem ... de pote in Tiberim deiecerit*

‘Many he killed by the dagger, many by poison. I can even give you an example of one man whom he threw from the bridge into the Tiber’ (Cic. S. Rosc. 100)
(5) **repute enim non esse dignas passiones huius temporis ad futuram gloriam, quae in nos **habet revelari

'I reckon that the sufferings of this time are not worthy of the future glory that will be revealed to us'  

Early examples of the periphrasis such as (4) typically express a modal meaning. Although Pinkster glosses the example with 'can', an obligation reading ('I have to tell you') seems licit, as well. Both a dynamic and an obligation reading can be taken to imply Cicero's intention to give the reader an example in the following context. This inference in turn licenses another inference, namely that Cicero will indeed give the example (and this is, of course, what happens in the example). While this second inference is backgrounded in example (4), it is foregrounded in example (5), which Pinkster classifies as an example with a future reading. In (5) the intention is attributed not to the speaker but to the implicit agent of the passive construction, thus suggesting that intention has part of the construction's encoded meaning. Given that the writer does not have direct evidence of this intention, he is making a prediction which then leads to the future reading. This future reading, in turn, then came to be conventionalized and promoted to the coded meaning of the construction. Note that Pinkster argues that such early future uses of *habere* + infinitive retain a modal nuance that can be paraphrased with the deontic ‘is bound to’ (Pinkster 1987: 206).

As to the reasons for speakers to use a modal construction in order to implicate future, we might argue with Detges (1999: 43) that the communicative benefit of such indirect strategies lies in the fact that they make a stronger prediction than already grammaticalized future tenses (such as the synthetic Vulgar Latin future *cantabo* ‘I will sing’). This is due to the fact that intention is typically interpreted as entailing a high probability that the action will indeed be realized. Periphrastic, indirect future expressions are therefore better suited to convince the hearer that the action will occur than ‘neutral’ future tenses. Such mechanisms thus motivate the observation that the initial stages of such grammaticalization processes are characterized by ‘expressivity’, ‘emphasis,’ or ‘relevance’, claims which date back at least to Meillet (1912) (cf. the discussion in Hopper and Traugott 2003: 24).

One problem with this approach, noted especially in formal descriptions of language change, is that not all inference processes lead to change. To quote Eckardt (2006: 10),

> The urge to reanalyze cannot be stimulated by the mere occurrence of pragmatic inferences alone. Pragmatic inferencing happens all the time. Practically all investigations in discourse semantics show that virtually no sentence is ever understood on the basis of the literal contribution of its words alone. [...] We may therefore conclude that sentences that do give rise to reanalysis need something in addition to a conventionalized pragmatic inference.

Eckardt argues that a much better understanding and formalization of the usage contexts is necessary in order to explain why some inferences lead to meaning change and others do not (cf. also Detges this issue). For instance, she demonstrates that while a construction such as *Ich gehe mal ein Bier holen* ‘I am going to fetch a beer’ has exactly the same inferential potential in German as in English, it never developed a future tense reading in German because it “is never used as a conventional means to express one’s intentions for the near future” (Eckardt 2006: 103, italics in the original). Similarly, Deo argues that formal pragmatics and specifically game-theoretic pragmatics “offers the formal tool of choice for a precise modelling of the complex processes of reasoning and learning that must underlie such developments” (Deo 2014: 404).

What both the functionally- and formally-oriented studies cited until now have in common is that they assume interaction to be the locus and in some cases the trigger of language change (see Traugott this volume for a summary). Such ‘channel-bias’ theories of language change (Moreton 2008) thus differ from ‘analytic bias’ theories such as Universal Grammar that focus on cognitive biases that aid learning of some

---

2 Detges (1999) actually analyzes the development of the periphrastic future *aller* + infinitive in French, which has replaced the synthetic *chanter-ai* ‘sing-fut.1sg’ future in many contexts. Given that the *chanterai* future is nothing but the contracted form of the previously periphrastic *cantare habeo* future instantiated in (4), one might argue that the renewal of the future tenses in French (and other Romance languages) is an instance of an onomasiological semantic/pragmatic cycle as defined by Hansen (this issue). See Section 3 for a brief description of this notion.
patterns and hinder that of others (e.g., Lightfoot 1991) in that they focus on the role of the precursors of the investigated constructions. If Detges’ (1999) functional account of the grammaticalization of the periphrastic future is correct, it is necessarily tied to speaker motivations in concrete interactions, which create innovations that can then be diffused and conventionalized in communities of practice (Jucker and Kopaczyk 2013). However, given that data reflecting real interactions in earlier time periods are hard to come by, historical studies frequently have to stipulate patterns in interaction. Given the type of documents that serve as the data for reconstructions of language change, it is probably no coincidence that historical studies typically place emphasis on the role of the speaker in interaction. In a relatively recent development, a number of studies have argued that meaning change may not only occur as the result of the conventionalization of (speaker-based) conversational implicatures, but also as hearer-based reanalyses (Detges and Waltereit 2002; Eckardt 2009; Schwenter and Waltereit 2010; Grossman and Polis 2014; Rosemeyer and Grossman 2017). For example, the notion of “Avoid Pragmatic Overload” (Eckardt 2009; Schwenter and Waltereit 2010) assumes that there are instances of meaning change that start with the use of a construction involving a certain presupposition in contexts in which this presupposition does not hold. The hearer of the sentence can either accommodate this presupposition or infer a new meaning for the construction, potentially triggering meaning change. For instance, using the particle too in John had dinner in New York, too presupposes that someone other than John had dinner in New York. In contexts such as example (6) where this presupposition is implausible, the hearer may reanalyze the meaning of too, for instance as an adversative marker (Schwenter and Waltereit 2010: 83).

(6) A: You didn’t do your homework!
B: I did too!

The notion of hearer-based reanalysis has important consequences for models of meaning change for at least three reasons. First, it explains the apparent discrepancy between the motivation (‘expressivity’ etc.) and the outcome of the change, i.e., the mechanism of the ‘invisible hand’ (Keller 2014) in that it locates correctly the locus of intention (local communicative goals in interaction) and removes the need for both teleology and assuming that speakers want to change language. The author of example (5) did of course not want to ‘invent’ a new future tense in Vulgar Latin. Rather, he made creative use of the inferences connected to notions such as obligation and future in order to express a certain discourse function (certain future) that did not have a grammatical expression. The eventual result of this rhetorical strategy, i.e., the creation of a new, unmarked future tense, was not intended by the speaker and is the result of hearer reanalysis.

Second, hearer-based reanalysis can be creative as well in that it allows for the possibility of meaning change based on unintended or ‘uninvited’ inferences. This idea is reminiscent of current models of sound change and specifically the work by Ohala (1981; 1993). As summarized in Grossman and Noveck (2015: 145-146), Ohala proposed that sound change is a result of the way hearers perceive the speech signal. In the perception mechanism, hearers typically filter contextual variation out from the speech signal. However, they sometimes fail to do so, analyze a part of the contextual variation as the articulatory goal and even filter out a part of the signal that was part of the original articulatory goal. Thus, errors in speech perception can in the long run lead to sound change. The presumed mechanism of interpreting a contextual mechanism as the coded one in meaning construction might look very similar, in that the hearer-based reanalysis might sometimes be based on such ‘mistakes’ in the retrieval of the intended meaning. This would predict that we have to identify the typical contexts in which hearers make these mistakes in order to establish the potential for a construction to undergo meaning change.

Third, hearer-based reanalysis is a very helpful concept for understanding the notion of “bridging contexts” (Diewald 2002, 2006; Heine 2002), which Heine (2002: 84) defines as contexts that “trigger an
inferential mechanism to the effect that, rather than the source meaning, there is another meaning, the target meaning, that offers a more plausible interpretation of the utterance concerned”. Example (5) can thus be described as a bridging context in the sense that both the modal and the temporal reading are possible, but the temporal reading is preferred. Crucially, the inference that yields the temporal reading is time-constant, in that both a contemporary and a Present-Day reader such as Pinkster have to realize the same inference leading to the establishment of the new meaning.

In summary, recent studies of meaning change argue that a full understanding of meaning change requires a context-sensitive model that includes both the speaker and the hearer (Grossman and Polis 2014 call this the ‘Two to Tango Principle’). Research in Conversation Analysis and Interactional Linguistics has been working heavily towards such a model, a fact from which diachronic studies could profit.

3 Contents of this special issue

The first three papers in the special issue are concerned with the problem of whether inferences have to be considered explanatory parameters for interaction and language change. Elizabeth Traugott’s paper “Rethinking the role of invited inferencing in change from the perspective of interactional texts” revisits the concept of invited inferences from a perspective that explicitly combines assumptions from historical and interactional linguistics. Traugott analyzes the historical process by which the imperative look came to be used as a discourse-structuring marker marking a shift of the topic at hand. She distinguishes between three types of inference, local inferences associated with specific expressions, discourse-structuring inferences pertaining to coherence, backgrounding, and foregrounding, and turn-taking inferences, which are inferences about relevant transitions at particular points in an interaction. Traugott argues that the starting point of the change was the spread of the use of look to complement clauses (as in Bonne loca dú here hu þu scealt þin gear rihtlice gedafian ‘then consider how to organize your year correctly’), inviting the discourse-structuring inference that what needs attention is not a concrete object but a linguistic contribution. Contrary to previous studies on this topic, she therefore assumes that the turn-taking inferences that have been postulated for Italian guarda ‘look’ cannot be projected onto English and are but a by-product of this change.

Arnulf Deppermann’s paper is entitled “Inferential practices in social interaction: A Conversation Analytic account”. Adopting an interactional perspective, the author argues that inferences play a central role in the organization of social interaction and matter for the interactants, an issue that has up to now been largely neglected in conversation analysis studies. Deppermann analyzes a broad range of practices in German that are used to signal or display implicatures or inferences of a certain meaning or communicative intention. He distinguishes three kinds of such practices for inference management. First, inferences may be made explicit (e.g., by introducing them with connectors such as dann ‘then’ and also ‘so’). Second, community members may use conventionalized indices (such as the German particle eben ‘exactly’) that inferentially relate the current turn to a preceding one. Third, the author demonstrates that inferences may also be necessary for understanding even in cases in which they are not explicated or indexed and thus remain implicit (e.g., in the case of analepsis and ellipsis). These considerations lead to the establishment of a catalogue of criteria by which inferential practices can be described systematically.

In his paper on “Inferences and indirectness in interaction” Paul Drew explores the interconnections between inferences that participants draw about each others’ actions, possible implications the participants attribute to each other and indirectness as one way of dealing with other participants’ actions. Inferences, the author points out, are not restricted to certain kinds of utterances, but are involved in the understanding of any turn at talk. Inferences are thus “ubiquitous in naturally occurring interactions”. Crucially, the author proposes to view implications not as an outcome of a speaker’s intention (as suggested by, for example, Grice) but as an attribution a recipient makes to the turn of a prior speaker. The main argument for this view is that the assumed meaning of a turn may differ between speaker and (different) hearer(s). The author analyzes cases in which a speaker explicitly marks her or his turn as formulation of an inference from a prior contribution, thereby attributing the responsibility for this meaning to some prior speaker(s). This strategy
is typically used in conflictual contexts. The disaffiliative character of such inference usage in interaction is even more apparent in indirect responses to enquiries. In such indirect responses, inferences function as an interactional device to challenge (push back or correct) a prior enquiry, offering an alternative to realizing an explicit correction or rejection.

The next three papers in the special issue are concerned with the relevance of inferences for the use of interrogatives and exclamatives in interaction and changes in their distribution. Richard Waltereit’s paper “Inferencing, reanalysis, and the history of the French est-ce que question” analyzes the historical rise of est-ce que, which changed from expressing a complex sentence with a matrix clause, in which ce cataphorically refers to the subordinate clause, towards a monoclausal construction in which est-ce que has turned into an interrogative particle. Waltereit demonstrates that est-ce que is first used in low answerability contexts, i.e., contexts with “a strong rhetorical flavor” in which the question utterer already knows the answer to his or her question. Over time it conventionalizes an information question reading, i.e., starts appearing in high answerability contexts. This change coincides with a change in the interpretation of the pronoun ce, which no longer refers to an antecedent. The author thus argues that the functional change from low to high answerability goes hand in hand with a reanalysis process, in which the compositional interrogative construction receives a holistic interpretation. The author argues that two types of inferences play a crucial role in this change: hearers need to have inferred (a) that the construction is used in a non-literal way and (b) that this use has spread to the community-level.

Oliver Ehmer and Malte Rosemeyer’s contribution “When ‘questions’ are not questions. Inferences and conventionalization in Spanish but-prefaced partial interrogatives” analyzes the use and development of pero ‘but’-prefaced partial interrogatives in Spanish, such as ¿pero qué dices? ‘but what are you saying?’? The authors combine methods from interactional and variationist diachronic corpus analysis and demonstrate that especially with situational verbs such as hacer ‘do’, pasar ‘happen’ and decir ‘say’, pero serves as an explicit marker of an interpretation of the interrogative as an interactional challenge. The interpretation of a partial interrogative as a challenge arises via inference; when such an interrogative is used in a low-answerability context, the hearer infers that the interrogative utterer does not request information but rather criticizes a previous utterance or action. The use of pero bolsters this interpretation because — due to its concessive semantics — pero typically implies an acknowledgement of the previous utterance or action by the speaker. However, the challenge function of pero-prefaced partial interrogatives is not only achieved compositionally, but has been conventionalized, leading to a conventionalization of entrenched patterns of the construction both in spoken language and historical texts.

Uwe-A. Küttner’s paper is entitled “Investigating inferences in sequences of action: The case of claiming “just-now” recollection with oh that’s right”. Within the framework of interactional linguistics, the study analyses the use of oh that’s right in English conversations. The main function of this lexicosyntactic format is to signal, that the speaker ‘just now’ recollects information that she had previously known but temporarily forgotten or presently not taken into account as relevant. The author shows that oh that’s right systematically occurs in a tripartite sequential pattern. In the first step, participant A realizes an action that conveys a certain presumption. In the second step, participant B challenges this presumption and attributes to A that she actually knows better. B thus produces a reminder of sorts, actually inviting A’s subsequent claim of a momentary forgetfulness or confusion. In the third step, A uses oh that’s right to accept B’s epistemic attribution and the inappropriateness of the initially realized action. Inferential processes at work in this sequential pattern are firstly that B infers that A’s action is motivated by some sort of forgetfulness, and secondly that B infers that A’s contribution points to an assumed forgetfulness. It is thus the specific sequential pattern and the involved inference processes that give rise to the meaning of oh that’s right as embodying ‘just now’ recollection.

In the third part of the special issue, two papers analyze the role of inferences for the synchronic and diachronic description of pragmatic markers. In “The role of inferencing in semantic/pragmatic cyclicity: the case of Latin NUNC and French or/ maintenant”, May-Britt Mosegaard Hansen proposes that these pragmatic markers, all of which have a content-level meaning similar to English now, evolved in a cyclic fashion. In line with results from previous studies, the author demonstrates that the meaning of these markers developed from content-level to context-level use. This change was motivated by various types of inferences that lead
to a reinterpretation of temporal as textual deixis (discourse-marking uses), as a speaker’s assessment of the relationship between parts of the text (context-level connecting uses) and as cause-result relationships between parts of the text (uses in which ‘now’ is used to mark the transition from a subjective assessment to a directive speech act). Given that all three of these etymologically unrelated markers undergo similar processes of meaning change, the author argues that these diachronic changes constitute an “onomasiological” semantic/pragmatic cycle. In other words, both in the development from Latin to French and within the development of French, speakers again and again started using new linguistic expressions (nunc, or, maintenant) for similar context-level meanings, thus renewing these meanings or functions.

The paper by Kerstin Fischer and Maiken Heide is entitled “Inferential processes in English and the question whether English has modal particles”. While modal particles play an important role in some languages like German, others, such as English, seem to not have an established equivalent grammatical category. Taking this observation as their point of departure, the authors pursue the question of whether the inferential processes that are involved in the use of German modal particles, can also be evoked by the use of pragmatic markers in English. The authors define modal particles by their main function to link an utterance to some contextually given proposition, e.g., to anchor the utterance in the Common Ground of the interlocutors. Fischer and Heide show that although English alright/ all right, already, then and other markers do indeed fulfill this main function, they also have other and more important pragmatic functions. Based on this result and further formal considerations, the authors conclude that from a Construction Grammar perspective no schematic modal particle construction needs to be posited for English. However, the inferential processes by which English pragmatic markers anchor an utterance in the Common Ground are comparable to those associated with the use of modal particles in German.

In the fourth section of this special issue, two papers analyze the importance of inferencing for processes of language use and change in morpho-syntactic phenomena. In “Te lo tengo dicho muchas veces. Resultatives between coercion, relevance and reanalysis”, Ulrich Detges investigates uses of the Spanish resultative tener ‘have’ + PP construction, especially cases where the past participle is formed from the non-transitional verb decir ‘to say’ (te lo tengo dicho muchas veces ‘I have told you many times now’). Detges argues that these expressions are marked by a mismatch between the resultative semantics of the construction and the non-transitional meaning of the lexical verb decir ‘to tell’. Nevertheless, this mismatch is licensed by coercion, which the author describes as an inferential repair mechanism giving rise to special meaning effects. But why do speakers create semantically defective expressions in the first place? Detges demonstrates that tener dicho ‘have told’ is typically used in contexts where it expresses a strong reproach in support of a directive speech act. Thus, the mismatch between the semantics of the construction and its lexical filler is motivated by the inference that the ‘result’ expressed, i.e. the illocutionary effect of the past speech act is still valid at the moment of speech, thereby lending the reproach a particularly strong illocutionary force. While this usage pattern can probably be considered a prime bridging context for the past speech act is still valid at the moment of speech, thereby lending the reproach a particularly strong illocutionary force. While this usage pattern can probably be considered a prime bridging context for

The paper by Peter Auer and Anja Stukenbrock is entitled “When ‘you’ means ‘I’: the German 2nd Ps.Sg. pronoun du between genericity and subjectivity”. While the second person singular pronoun du ‘you’ in German is generally used to directly address an interlocutor, there are also non-addressee deictic uses. The authors identify four different kinds of such uses that differ with regard to whether the addressee and/or speaker is part of the referent group. On the one hand there are uses in which a general rule or social category is being referred to or invoked, which encompasses both speaker and hearer, only the speaker or none of them. On the other hand there is a ‘subjective’ use of du ‘you’, in which no category is made relevant and the pronoun exclusively refers to the speaker alone. The authors identify the inferential processes on which these different uses rest and which deictic shifts/mental displacements need to be performed by the interlocutors. While there has been a clear increase of non-addressee deictic uses of du in the past 70 years, the authors provide evidence that such uses date back at least 200 years and that thou-monologues – as already described by Grimm in 1856 – can be seen a predecessor of those uses.
4 Summary and avenues for future research

The findings from this special issue can be summarized in at least three points.

4.1 Inferences are needed to explain social action and linguistic behavior

First, the papers united in the special issue argue that the notion of inference is crucial for advancing current models of the description of interaction. Inferences are not only a feature of certain kinds of utterances, but are involved in the understanding and reaction to any turn at talk (cf. Drew this issue). Inferring is an integral part of any social activity. The study of inferences is thus a central object for the study of conversation and interaction, since it allows describing certain actions more appropriately. The importance of inferences, however, needs to be disentangled from the notion of implicature. It is definitely the case that speakers sometimes use inferences intentionally and strategically, anticipating possible inferences on behalf of their partners and designing their turns respectively to invite certain inferences. Such uses are only one way in which possible inferences matter in interaction, so to say prospectively from the speaker’s perspective. There are numerous other ways in which inferences matter in interaction and different degrees in which they surface or become manifest in the actions of the participants. As Deppermann (this issue) puts it, there are at least three ways in which inferences matter to social interaction: they (i) may remain tacit, (ii) may be conventionally indexed or (iii) may be formulated explicitly. Interactants furthermore employ different devices to deal with inferences, ranging from sequential patterns, syntactic procedures (like break-offs, ellipsis), over lexico-syntactic means (like connectives, particles) to morphology (like person marking) and others. Those devices furthermore exhibit different degrees of routinization/sedimentation/conventionalization, regarding the function of managing certain inferences and contributing conventionally to the organization of social interaction.

4.2 An explicit speaker-hearer model is crucial for explaining meaning change

Historical linguistics has assigned the notion of inference a central place in explanations of meaning change, as evinced by the fact that all of the papers in this special issue that work with diachronic data assign a central explanatory value to inferences. However, the results from this special issue clearly demonstrate the importance of employing a contextualized model of the roles of speaker and hearer in the synchronic and diachronic emergence of meaning. Thus, meaning change commonly appears to arise in situations in which the hearer draws an inference on the basis of the use of a linguistic construction in a context in which it use is unexpected (such as the use of the imperative look in complement clauses, interrogatives in low-answerability contexts, particles with the meaning ‘now’ in contexts in which temporal deixis is not at stake, second person singular pronouns in contexts in which the addressee is no longer part of the category of people about which the statement is made, and tener + participle constructions with auxiliated verbs that do express a resultant state). Crucially, the resulting historical change is unmotivated from the perspective of the speaker. Although the speakers exploit the semantic potential of using a linguistic construction in novel contexts and anticipate the inference by the hearers that a divergent reading is intended, they do not necessarily expect the conventionalization of this inference (which, in turn, might be modeled using the Principle of Reference or Avoid Pragmatic Overload, as outlined in Waltereit’s paper).

The systematic exploration of the use of inferences in interaction in the synchronic studies in this special issue is therefore of crucial interest to studies on meaning change, as it clarifies exactly how such inferences arise. For instance, the more explicit speaker-hearer model proposed by these studies can help to explain the notion of gradualness in language change. Recent historical studies such as De Smet (2012) propose that actualization, i.e. “the process following syntactic reanalysis whereby an item’s syntactic status manifests itself in new syntactic behavior” (601), is sneaky, in the sense that the spread to new syntactic contexts will first affect those contexts that most resemble the original usage contexts of the construction. While De Smet
explains this observation in terms of salience and analogy, the results from this special issue suggest that it can also be explained in terms of the degree of expectedness of hearer-based inferences. We could thus expect scenarios such as the ones described above in which the original reanalysis is highly unexpected (and consequently, salient) in discourse. However, once reanalysis has taken place, the same inference becomes much less unexpected in those usage contexts that most resemble the original reanalysis context. These contexts are favored in the actualization process because of cognitive ease; the hearers can use an already established reanalysis pattern based on a more or less conventionalized inference to deal with this new utterance type. This means that (a) in order to be able to predict whether meaning change will occur, historical studies need to assess the degree to which a given hearer-based inference can be expected in a given discourse situation and (b) historical studies need to start paying more attention to the how inferences are used in ‘intermediate’ positions in grammaticalization processes, as also noted in Winter-Froemel (2014: 516-521).

4.3 Assuming degrees of conventionalization of inferences is useful for analyzing variation in language and interaction

In line with the description of the findings from the diachronic papers from the previous section, a distinction between ad-hoc inferences and conventionalized inferences has to be assumed. While the use of a construction in a novel contexts leads to an ad-hoc inference by the hearer (corresponding to a particularized implicature on the speaker side), repeated exposure to the same novel usage will lead to the conventionalization of this inference. The degree of conventionalization of an inference has an important influence on the perception and management of inferences, as conventionalized inferences are arguably drawn on a less conscious level and are more robust. This may impact the usage contexts of the constructions that the inferences are associated with. Take, for instance, the two studies on interrogatives in this special issue. In Waltereit’s paper, the conventionalization of the inference that a French cleft-wh-interrogative is indeed used to ask for something (as opposed to a rhetorical use) leads to its use in contexts in which the complementizer ce is no longer anaphorical. Likewise, in Ehmer and Rosemeyer’s paper it is claimed that the conventionalization of the challenge interpretation of Spanish pero-preposed interrogatives leads to a higher usage frequency in reported speech. This means that assuming degrees of the conventionalization of an inference and observing the reflexes of this process in interaction can be useful in determining at which point an inference has become part of the encoded meaning of a construction.

The distinction between ad-hoc inferences and conventionalized inferences can thus have important implications for the analysis of interaction. Interactional linguistics and Conversation Analysis often distinguish between “prototypical” and “deviant” cases in the analysis. A historical approach to the analysis of inferences could be instrumental to the process of distinguishing such cases.

References


