Explaining Particle Typology in Minimal Bayesian Pragmatics

Henk Zeevat,
ILLC, Amsterdam University (retired)

If one witnesses an event one tries to explain it using the information that one has available by the most probable explanation that one can find within these limits. If the event is somebody's action, the explanation is only proper if it attributes an intention to the agent. If the event is an utterance, proper most probable explanations attribute an intention to the speaker and derive the utterance from the intention using grammatical knowledge in the sense of tradtional grammar: knowledge about how to express content in a particular language. Interpreting events, actions and utterances is finding their most probable explanation given the information that one has.

Minimal Bayesian Pragmatics (MBP) is the view that this almost gives all of natral pragmatics. One needs to add the standard account of information structure (for implicatures) and disallow very low priors or likelihoods (these prompt non-literal interpretations) but that is all. The explanations supply —unlike the existing accounts of non-literal language use and pragmatic enrichment causal and identity inferences (implicatures that are not derivable in Gricean pragmatics and its successors) and thereby allow crucial improvements to presupposition projection (Karttunen's filtering is only correct if logical entailment is replaced by actual causation), the integration of argumentation pragmatics (Anscombre and Ducrot) and the full reduction of discourse relations to general pragmatics. For more details, see Zeevat & Winterstein (under review).

Explanations can become more probable if they get larger. For this, unreduced elements in the explanation must systematically be assigned probabilities based on experienced frequencies or default priors, instead of assigning them probabilities by calculating in the effects of causality as perceived by the subject. Identifying an unreduced x with y can then improve the probability to the product of the probability of y and the probability that $x = y$ given what is known about x and y. Assigning y as its cause to x can similarly give a better probability: the product of the probability of y and the proability that y causes x given what we know about x, y and their relation. Making causal and identity assumptions is then a way of obtaining more probable larger explanations.

Stochastic disambiguation is the first thing that pragmatics needs to do. Minimal Bayesian Pragmatics (MBP) is the view that pragmatic enrichment happens already at this point, but that the object that has to be maximally probable in the context is not its semantic representation or the speaker

intention, but the causal explanation of the utterance: how did it come to happen?

That it should be explanation and not semantic representation follows from the notion of enrichment: extra information can only decrease the probability of representations. But extending an explanation by identifying causes for unreduced elements in the context or by finding these elements in the context itself increases its probability. According to MBP, all enrichment is a side-effect of causal and identity inferences that extend explanations in order to make them more probable.

To guide such reductions, stochastic information is needed: identities and actual causations need priors that can increase and decrease as more information comes in about the objects to be identified or causally linked. It seems also clear that the computation of the explanations must be based on both the context and the utterance. Explanations are initialised by the utterance (a structured event) and a context and would use simulated language production to causally derive utterances from postulated speaker intentions.

An explanation is complete if it cannot be further reduced. It is improper if it does not contain a speaker intention (utterances are intentional actions). It is also improper if it is not substantially better than its competitors (complete explanations resulting from different reductions, the explanation is not recognised as correct if there are close competitors). It is also not proper if it has the best probability that is still very low. MBP needs an inclusive concept of causality.

Notice that if there is a conflict between the result of selecting the best literal meaning, speaker intention, conversational contribution and what the best explanation fixes them to be, the explanation always wins. Notice too that explanations work also for the interpretation of non-verbal utterances (here the explanation fixes the meaning) and even for the interpretation of non-communicative behaviour of others.

The strongest linguistic argument for MBP is the light that it throws on the typology of particles:why certain kinds of particle use occur in so many languages.

It is well known that a number of particle categories are instantiated in many languages (if they are not universal) without serving a clearly expressive purpose. These categories emerge from comparative and historical linguistics and are attempts at getting a handle on particle meanings across languages and stages of languages. The additive particles are a case in point. *Bill sleeps too* does not say more than that Bill sleeps, the particle *too* merely indicates that somebody different from Bill is also given in the context as sleeping.

But the same point can be made with contrastive, concessive, approximative, adversative, mirative, antiexplanative, downtoning and antiinferential particles: the host says what it says, the particle does not add to that meaning, but says somehing about the relation of the context and the host or the way in which the host needs to be integrated in the context.

In each case, the question arises why it is so important to mark the particle meaning that in many languages the meaning has been the target of a grammaticalisation process that led to the particle. Why is a language with addtive particles better than a language without them? And so on on for all particles without a clear expressive function.

Other particles have transparently useful functions. Causal markers like *so* mark that a causal relation from the pivot to the host needs to be adopted, *therefore* marks a conclusion from the pivot, *in particular* introduces a further specification of the pivot, *then* a temporal relation, the negation (in German and Dutch plausibly a particle that inverts the polarity of the host).

MBP does a good job in predicting causal and inference particles and in predicting idenfication marking strategies (in English, this is mainly personal pronouns and the definite article, but also old marking on sentences by markers such as *indeed*, *doch* (no accent) and *ja* in German). But it also predicts that it is important to prevent identifications and causal reductions where it is not correct to have them. In addtion, it predicts that non-standard and non-literal interpretations should be prevented when the interpretation of the host has a low prior but is still correct.

To start with the last category, one way to achieve this is to mark that the host's intended interpretation has a low prior, thus inhibiting the search for non-standard interpretations. These are the mirative markers, which include also the adversative and concessive particles. The other mirative markers tend to be more specific as to what is surprising: the high cardinality of the plural referent (*only, just, merely*), the earlyness of the occurrence (*already*). the lateness of the termination (ıstill), the referent itself (*even*) among the other possible referents. The adversative and concessive markers add the identification of the reason why the prior is low and correct it: at this occasion the reason for disbelief does not apply. The extreme case is here correction: the host 's prior is zero and update is not possible without adapting the context.

The additive and contrastive markers prevent identifications (contrastive intonation also has this function) against the Bayesian preference for identification if this leads ot better explanations (known as DOAP and *NEW in the optimality theory). It is to be expected that causal reductions cannot

be prevented as directly: the main problem are possible sources for such anticausal particles and directing them at the non.cause. Explicit denials of causal relation are found in anticausal readings of concessives: *although A, B*, which deny *A actually caused not B*. It follws that the closest anticausal particle is the proconcessive. The interpeter must find X in the connect that would cause the negation of the host.

The other form a strange collection.

A. antiexplanative particles like german *halt* and *eben* or dutch *nu eenmaal* , or *just* (eng) marking that the interpreter should not attempt to explian the host.

B. anti indferential markers like german eigentlich, (uses of) eng. *really* indicating that normal consequences do not apply (*My name is really Hendrik* can indicate that people do not call me that way, though this normally follows).

C. approximatives (almost, nearly, barely) prevent the assumption of normal host consequences by replacing them with the consequences assoicated with whatever the host comes close to (*he barely made it*: consequences as for *he did not make it*).

D. Downtoners (kind of, wel (dutch), schon (german)). Mark the absence of the cause that the speaker knows the host is true and thereby that the speaker really knows that the host is true.

But one finds confirmation of the typological predictions and crucially there appear to be no particle classes that need a functional explanation that do not get it from MBP.

Reference

Henk Zeevat & Gregoire Winterstein (under review). Minimal Bayesian Foundations for Pragmatics.