

What Time Is It? – It Depends On Where You're Going.  
Relevance and Precision in Telling the Time.

Diplomarbeit  
von  
Johann Jacoby

Angefertigt unter den Bedingungen der Diplomprüfungsordnung für den Studiengang  
Psychologie an der Universität Göttingen, Biologische Fakultät, vom 2. November 1998.

Datum der Themenausgabe: 08. 04. 2003  
Datum der Abgabe: 01. 07. 2003

Erstprüferin: Prof. Dr. Margarete Boos  
Zweitprüfer: Prof. Dr. Michael Waldmann

Betreuerin: Prof. Dr. Margarete Boos  
Anschrift des Verfassers: Burgstraße 4, 37073 Göttingen, [johann@johannjacoby.de](mailto:johann@johannjacoby.de)

This paper (my first one) is dedicated to my parents (R. I. P.).

Much gratefulness is due to Prof. Dr. Margarete Boos, Prof. Dr. Michael Waldmann, and Prof. Dr. Willi Hager, my academic teachers.

Thanks for different reasons are due to Barbara Schauenburg, Schmitti (diese sinnlose Prüfung), Gesine Schneider, and John R. Cash.

And everybody else whose listing is omitted due to the author's wish to keep at least this credit section short and simple.

My name is Johnny Knoxville, welcome to Jackass.

P. J. Clap

What Time is it? – It depends on where you're going.

Relevance and Precision in Telling The Time.

Johann C. Jacoby

Georg-August-Universität Göttingen

Running Head: RELEVANCE AND PRECISION

Please address correspondence to:

Johann Jacoby  
Institut für Psychologie  
Goßlerstraße 14  
37073 Göttingen  
Germany  
[johann.jacoby@psych.uni-goettingen.de](mailto:johann.jacoby@psych.uni-goettingen.de)

## Abstract

Relevance Theory (RT, Sperber & Wilson, 1986) aims to explain how comprehension in human communication succeeds when there is no unambiguous code as posited by the classical theories of communication (e. g. Shannon & Weaver, 1949). In recognizing and emphasizing inferential processing by the listener in understanding what a speaker means by an utterance, it psychologizes pragmatics to an extent where description of the cognitive basis of interpersonal communication becomes possible.

Drawing on van der Henst, Carles and Sperber's (2002) study on answers to a question for the current time, the present work looks at the consequences of RT for the choice and construction of utterances on the speaker's side. van der Henst, Carles and Sperber (2002) tested the claim that speakers would tune their messages to the requirements of their listeners in accord with the second principle of relevance as posited by RT. This principle asserts that every act of intended communication comes with a guarantee of relevance. Securing relevance should then lead to more precise answers given to a question for the time if higher precision is relevant. This prediction was confirmed.

In the present study, standard starting times of lectures were set as an appointment that an askee may presume the asker to want to make. RT predicts that even with such an implicit cue to the differential relevance of precise answers earlier vs. later before the assumed appointment, participants would give less precise answers longer before these times than shorter before the starting times.

The predictions were confirmed. However, the characteristics of the present study allow for a different, possibly simpler explanation of the results which is discussed in turn. Conclusions for future studies in the paradigm are drawn in order to bring more clear-cut evidence for the validity of the RT-explanation of choice and construction of utterances.

Relevance Theory (RT, Sperber & Wilson, 1986) aims at explaining how communication succeeds, especially if no firmly established, unambiguous and mutually known code exists which allows a mapping of the speaker's mental representation onto an acoustic signal onto a mental representation in the listener that closely corresponds to the original representation.

In doing so, RT proposes a set of axioms and definitions and two principles of relevance which together make predictions as to what a listener will understand from an utterance in a certain social and cognitive context and what a speaker will utter in order to achieve comprehension in the listener. The following outline of RT will essentially follow Sperber and Wilson (1986, 1987, 1995).

Basic to RT is the assumption that both speaker and listener each have a cognitive environment which consists of assumptions that are manifest to them at a given time. In the present context, for an assumption to be manifest to an individual means that "the individual is capable [...] of representing it [the assumption, J. J.] mentally and accepting its representation as true or probably true." (Sperber & Wilson, 1987). Manifestness is a continuous concept in that an assumption may be more or less manifest. Also, manifestness of an assumption to an individual may vary between times and situations. In this sense manifestness is conceptually somewhat similar to the notion of accessibility in social cognition research (cf. Higgins, 1996).

On an individual level, these manifest assumptions provide a background or a context in which stimuli are cognitively processed and interpreted. Of course not every stimulus will be processed or even attended to. The criterion governing which stimuli will be attended to and processed further, according to RT, is relevance.

In the current context, utterances will be treated as a special case of acoustic stimuli and, since human communication is at the focus here, discussion will be confined to utterances.

### Relevance

RT specifies relevance as a function of contextual effects to be gained from the processing and the cognitive effort spent in interpreting it. Thus, a stimulus or an assumption is relevant to the extent that – *ceteribus paribus* –, it yields contextual effects and is easy to process.

Contextual effects are inferences drawn from the stimulus and the cognitive environment that were not contained in the stimulus or the cognitive environment before. They add, modify or even delete assumptions to, in or from the cognitive environment. These effects contribute to a larger, more differentiated, hopefully more accurate and hence more useful representation of the world. The process of inference may be understood as a contextualization of the stimulus in the cognitive environment: the effects are concluded from a stimulus in the

context of the cognitive environment. Contextualization both yields a certain amount of contextual effects and determines the processor's cognitive environment as the "context". This process may also be considered as the interpretation of a stimulus in a context.

An example may illustrate this. Consider the stimulus information "Barbara is nice" and a context of assumptions such as "Nice people are trustworthy.", "Nice people are a pleasure to have around.", "If a person, that is a pleasure to be around, pays me a visit, then I will stop my current work and have a drink with her.", "Barbara came to my house yesterday night" and "If a person is unintelligent, she can't be nice". The contextual effects derivable from this information in the context are, for example, "Barbara is trustworthy.", "I had a drink last night.", "I stopped working for a while last night.", or "Barbara cannot be unintelligent". All of these informations are new assumptions and were not contained in the original stimulus or cognitive environment, but they were inferred from both together. Of course, some of these inferences, considered individually, may be more useful for a representation of the world and therefore individually yield more contextual effect than others, depending on the situation.

As can be seen from the above example, the context and the stimulus are used as something similar to premises of a logical deduction, and the contextual effects figure as the conclusion from this deduction. Of course this process is in many ways dissimilar to logical reasoning. For example it provides for fuzziness: the assumptions contained in the context can be more or less manifest and hence contribute unequally to the inference etc. Furthermore, valid yet uninformative inferences are discarded (such as  $p \supset \neg \neg p$ ). The inference process is a non-demonstrative one: it may draw on premises differing in strength, it does not yield each and every possible logically valid conclusion, and it may fail.

The cognitive environment is not fixed in most cases: as indicated above, it contains assumptions of different degrees of manifestness. The assumptions are considered as a context for contextualization of a stimulus in order of accessibility. It is hence possible in the inference process that the context continually changes. Then the context, or the set of assumptions presently most manifest, is expanded or narrowed so that more contextual effects can be inferred or processing becomes easier (see below). Individual assumptions may lose or gain manifestness even to the degree that assumptions that were not manifest at all are added or assumptions that were highly manifest before are completely dropped. This activity of changing the context points to the second determinant of relevance: cognitive effort.

The modification of the context and the contextualization of the stimulus herein draw on cognitive resources. The cognitive effort expended in interpreting an utterance is the degree to

which central and peripheral cognitive processes (reasoning, memory, perception etc.) are taxed when one engages in the interpretation of a stimulus.

Consequently, the notion of relevance is one of cost-benefit economy: the larger the contextual effects achieved by interpreting an utterance and the smaller the effort needed for this interpretation the larger the stimulus' relevance and the more attention it will be given. A stimulus that only yields very small effects in a context is considered less relevant than a stimulus yielding more contextual effects, given the same effort. In turn, a stimulus yielding a certain amount of contextual effect only if much cognitive effort is invested is less relevant than a stimulus yielding the same amount of effect at a lower cost of cognitive resources.

The notion of relevance in RT is not developed to a degree where absolute quantitative assessments of relevance can be made, but it allows for rough assessments of how useful processing a specific stimulus is to an individual.

#### First Principle of Relevance

Hence relevance provides a guide for attention and cognitive resources to be allocated. The foregoing usefulness of a stimulus in terms of relevance is crucial to understanding the world and acting in it. It is useful for individuals who want to choose from an enormous number of stimuli that could potentially be attended to – a situation typical for our everyday live and possibly even more for the phylogenetic past of humankind.

The first principle of relevance posits the natural tendency of human cognition to maximize the relevance of stimuli it attends to. It serves as a guideline in allocating cognitive resources, such as – and temporarily first – attention, and, to a certain degree, regulates the adjustment of the context if an immediate interpretation does not offer itself: Sometimes the most manifest assumptions from the cognitive environment together with the stimulus do not yield contextual effect, or too low of an effect. Then the context is successively adjusted – within certain limits –, until it affords for the extraction of contextual effects at acceptable cost. However, if the expected relevance of a stimulus drops too low, it will be abandoned and processing will be refrained from altogether. This may occur when the effects achievable from processing the stimulus in context of the present cognitive environment fall below a minimal degree, the effort – for whatever contextual effect is achievable –, is considered too large or the combination of both (the degree of relevance) is too low compared with other stimuli.

Note that the first principle of relevance (also labeled "cognitive principle of relevance") is a concept on an individual psychological level. Its role in communication as a social interaction will be outlined in the next section.

## Second Principle of Relevance

A special kind of stimuli processed by humans are utterances by other humans: language. They are not only special in that, in general, they have a potential of yielding richer contextual effects if processed in an adequate context or they lend themselves to conveying more complex and sophisticated messages than non human-made stimuli, but the stimuli are also intentionally produced by other humans in order to communicate something. To give an example of a worn-out kind, the sound of an approaching predator may attract more attention than a bird singing in a tree and more processing effort may be allocated to its processing. But a hungry tiger's grunting and the bird's song were not "made" for a human so that she "understand something". These stimuli are not intended to communicate something to the processor, in fact, if one wanted to impute any intention to the predator, one would actually have to assume that it intends not to communicate its approaching.

If a human, on the other hand, intends to change another's cognitive environment in a certain direction (i. e. communicate something) she will have to make sure that the communicative partner have some expectation of relevance concerning the acoustic stimuli she produces. She has an interest in the other human's paying attention to her utterances and interpret them.

RT therefore proposes the second principle of relevance or the "communicative principle of relevance". It asserts that any communicative act between humans carries a presumption of optimal relevance, that is a prima facie assurance that the speaker will secure contextual effect at the lowest possible cognitive effort or, given the effort, the highest contextual effect possible, given her own preferences and abilities. The speaker's preferences and abilities constrain what level of relevance she will provide: Of course she will not communicate something that she does not want to communicate. The guarantee of relevance posited by the second principle does not amount to a guarantee of total and maximal relevance, regardless of whether the speaker wants certain contextual effects in the listener or not. Also, the speaker will be bound by her, for example, linguistic abilities in achieving relevance laden processing of her utterances by the listener.

This principle has interesting implications for both speaker and listener: The speaker will choose her utterance according to the constraints mentioned (effect and effort for listener and her own preferences and abilities) and the listener will use these constraints in bounding the search for a context in which the utterance may achieve relevance and ideally finds a single context in which the utterance achieves optimal relevance by yielding sufficient contextual effect at reasonable cost.

For this to work, a crucial subset of assumptions in the cognitive environments of communicators is the set of mutually manifest assumptions. Mutually manifest assumptions are previously manifest assumptions for which there is another manifest assumption that the original assumption is manifest to the conversational partner as well. Such mutually manifest assumptions include utterances made in a conversation up to the present point, inferences presumably drawn by both communicators so far and assumptions assumed to be shared. Comprehension will draw on these mutually manifest assumptions in the inference process described next and if assumptions are not actually mutually manifest, communication will be prone to fail. The coordinated nature of the cognitive environments, the partial overlapping, is necessary for the inference process to work out effectively. If there were no mutual manifest assumptions, comprehension and understanding would merely be a result of chance in the absence of a safe and unambiguous code.

That such a principle governs utterance choice and construction by the speaker is a consequence of the assumption that in communication the speaker wants to secure the listener's attention, and she will utterly fail at this if she provides utterances yielding too small an effect or claiming too much effort by the listener – shortly: utterances that are not relevant enough to bind attention and cognitive processing. An extreme example may illustrate and make this argument more convincing: Suppose that you are telling me something you absolutely want me to know about the technology of digital movie encoding (a field in which you have vast knowledge) but rely on concepts, ideas and knowledge (or more generally: assumptions) that are not manifest to me: I have never heard of and thought about them and I will not be able to recover contextual effect from what you say unless I engage in year long studies of, say, optical engineering. I will consider this effort exceeding the effect I could retrieve from your utterances at the moment, hence your utterances will not seem or be relevant to me and I will stop listening to you. So you will see to it that I be able to understand what you are saying without this effort and simplify things to the extent that you are able and willing to. However this simplification must not demand too much of your own effort: you will tell me about the topic in a manner that is of course not too demanding to yourself, but increases the chances that I will not stop listening and do extract contextual effect from what you say and what I already (presumably) know together.

Since every listener is also (in other situations) a speaker, a listener will know that the principle will probably steer utterance choice and construction by the speaker and use the same constraints in gradually constructing the context in which the utterance may yield the expected effect.

Of course, the principle does not assert that communication cannot go awry: for example, the speaker's preferences as perceived by the listener may not match those actually held by the speaker in which case the choice of context may fail to allow the extraction of the expected effect.

Both speaker and listener of an utterance will be guided by the second principle of relevance in a conversation: the listener will use it to constrain the search for assumptions needed as a context for optimally relevant processing and the speaker will use it in the choice and construction of her utterance.

According to RT, comprehension by inference as described here is an extremely deeply entrenched mechanism, automatic for the largest part and possibly even based on a domain-specific language comprehension sub-module in an evolutionary psychological sense, as discussed by Sperber and Wilson (2002). Given the overwhelming ubiquity and importance of communication in human life it seems very plausible, from an evolutionary (or at least from a developmental) perspective that a task that basic and crucial for human social life should develop to be highly effective and efficient to a degree where it is largely triggered spontaneously and need not necessarily be conscious. This perspective on the functioning of human comprehension will be important in hypothesizing a similar spontaneous complementary tendency in speakers, as argued below.

### Empirical Evidence

The body of empirical evidence supporting RT is growing and has been collected in quite diverse areas. Three exemplary findings in two areas – developmental-clinical psychology and reasoning –, shall be briefly reviewed here.

#### Happé (1993) – Comprehension in autistic children

Happé (1993) found evidence that autistic children who lack basic so-called mind reading skills (as posited in the theory of mind theory, Baron-Cohen, Leslie & Frith, 1985) could not understand tropes such as metaphors or irony.

RT rests on the contention that listeners will have to assume or infer the speaker's assumptions and intentions ("read their mind"), specifically but not exclusively the intention to communicate something over and above what can be extracted from the mere code of an utterance, that is, what is explicitly and literally said<sup>1</sup>. This is of prominent importance if what is said is under no circumstances to be taken literally, such as in comprehending metaphors, irony or other "loose uses" of language (Wilson & Sperber, 2002). Autistic children, who lack mind reading skills and fail in false-belief tests (Baron-Cohen et al. 1985), should then not be

able to understand such tropes. Indeed, the participants in Happé's (1993) studies who had failed the false-belief test had difficulties choosing completing words for unfinished sentences from a set of items that only allowed for meaningful completion of the sentences as metaphors (experiment 1) or drawing inferences about the meaning of a speaker's using irony or metaphors in short vignettes (experiment 2). Thus, weak mind reading abilities as detected by the false belief test had the effect on language comprehension predicted from RT.

van der Henst, Sperber and Politzer (2002) – determinate answers to indeterminate problems

van der Henst, Sperber and Politzer (2002) analyzed determinate answers to indeterminate relational problems. Indeterminate relational problems require from participants to state conclusions derivable from sets of premises such as [Melanie is taller than Barbara; Melanie is taller than Margarete] (see also Byrd & Johnson-Laird, 1989). They are called indeterminate because the relationship between two of the three persons or objects in the premises (in the example: Barbara and Margarete) cannot be inferred. However, there are possible determinate answers to such problems, e. g. (1) and (2):

- (1) Melanie is the tallest.
- (2) Barbara and Margarete are the shortest.

These answers differ in the cognitive effort and hence in relevance, the recoverable effect being the same: (1), containing only one subject, is the easier to process than (2) which involves two subjects. They do not differ in effect, since they all convey the same relational information and hence the same contextual effect given the universe of the three persons. Determinate conclusions of type (1) – so called single-subject conclusions (SSC)–, should therefore be preferred to conclusions of type (2), which are double-subject conclusions (DSC).

Also, SSC drawn from same-subject premises (SSP), such as the ones given above, save less effort to an audience or to oneself in the future than SSC drawn from different-subject premises (DSP). The SSC [Melanie is the tallest] drawn from the DSPs [Barbara is shorter than Melanie] and [Margarete is shorter than Melanie] "condenses" the information given and saves effort in future inference to a larger degree than the same SSC drawn from the SSPs [Melanie is taller than Barbara] and [Melanie is taller than Margarete]. This SSC can almost be immediately recovered in a future context, hence their combination saves less effort than the combination of DSP into a SSC. SSCs are therefore more relevant given DSP than in the context of SSP and should hence be given more often in the former than in the latter context. While their contextual effects are equal, they differ in effort.

The detailed predictions for frequencies of SSC vs. DSCs drawn from SSP vs. DSP were confirmed. Relevance apparently determined the type of conclusion drawn by inference from the premises given.

Sperber, Cara and Girotto (1995) – the Wason selection task

Finally, Sperber et al. (1995) reanalyzed different well established empirical results from research on the Wason selection task (Wason, 1966) in light of RT. This task presents participants with four allegedly double-sided cards showing only one side each. A classic version presents two of the cards showing a feature "p" and "¬p" and the remaining ones the features "q" and "¬q". Participants' task is to indicate which of the cards are necessary and sufficient to turn over in order to determine whether a rule such as: "If a card has p on one side, then it has q on the other side" holds. Participants utterly fail in many versions of this task, selecting only the p- or the p- and q-cards instead of the p- and ¬q-cards as prescribed by the modus tollens argument form in informal standard logic; but they make the formally right selections in other versions.

RT provides an explanation for this disparity and Sperber et al. (1995) can show that the degree of relevance that turning over the respective cards provides predicts participants' choices. Based on this analysis they give a "recipe" for constructing card selection tasks eliciting the right responses. Following this recipe should make relevance of the respective cards coincide with the correctness of their choice according to modus tollens. Sperber et al.'s (1995) argument is of considerable greater complexity than the sketch presented here, but due to space restrictions, the reader is recommended to consult the original article for the reanalysis and the "recipe".

Four selection tasks constructed after Sperber et al.'s (1995) recipe yielded logically correct choices by participants in the laboratory (see also Girotto, Kimmelman, Sperber & van der Henst, 2001, for further examples). This goes contrary to the traditional theoretical accounts of performance in the task invoking general human weakness to reason in accordance to formal logic unless a domain-specific evolved mechanism such as a "cheater-detection module" is activated (Cosmides, 1989) or the task is sufficiently concrete (Johnson-Laird, Legrenzi & Legrenzi, 1972). Additionally, Sperber et al.'s (1995) account converges with the reanalysis of the selection task by Oaksford and Chater (1994) who could show, even though from a quite different (Bayesian) perspective, that the cards chosen by participants in a classic selection task judged to be "wrong choices" according to formal logic actually do afford larger information gain than the "right choices". In a fashion similar to Sperber et al. (1995) they show that valid intuitions – of what they refer to as informativeness rather than as

relevance –, predict participants' selections regardless of concreteness or domain characteristics of the task (see also Oaksford & Chater, 1995).

Empirical evidence thus supports the predictions that RT makes for how participants in experiments understand, receive and process information given to them. However, one important feature of communication is that it is essentially a coordinated enterprise (see for example McCann & Higgins, 1992). In order to understand the delicate interplay of speaker and listener exchanging communicative acts (which under many circumstances involves switching roles in the course of a two-sided conversation), it is necessary to look at the speaker side as well. If the second, communicative principle of relevance guides the listener in comprehending utterances, and interpretation of utterances by and large and systematically succeeds, then there must be a complementary activity on the speaker's side: that of doing the best in securing optimal relevance. If speakers did not try to secure optimal relevance and hence would only by chance cash in the guarantee embraced by the listener in line with the second principle of relevance, listeners would probably never have come to accept the guarantee of optimal relevance in the first place. Hence the sensitivity for relevance in speakers might well be just as entrenched and spontaneous as in listeners.

For a more comprehensive picture of communication according to RT, it is thus important to look at whether speakers really do strive for a solid, reliable degree of optimal relevance in their utterances. If they do so, they should, according to the definition of optimal relevance:

- not go beyond their preferences and abilities,
- save the listener unnecessary effort, given that
- they provide potential for extraction of adequate contextual effect.

These characteristics partially contradict the Cooperative Principle and the maxims derived from it (see below) posited by Grice (1975) as governing speech production (and reception) invariably.

#### Grice's Truth Maxim

Application of RT notions to the choice and construction of utterances by speakers has interesting implications for a maxim or convention often invoked by philosophical pragmatic theories: truthfulness. The following discussion is based on, but not copied from the discussion by van der Henst, Carles and Sperber (2002).

For example Grice's (1975) account of implicatures fundamentally rests on the adherence to the Cooperative Principle (CP) which asserts that in general communicators obey the rule that they "... make [their] conversational contribution such as is required, at the stage at which

it occurs, by the accepted purpose or direction of the talk exchange in which [they] are engaged" (p. 45).

The CP is spelled out into four maxims (and respective more specific sub-maxims, see Table 1).

-----  
 insert Table 1 about here  
 -----

The most prominent (according to Grice (1975) himself) is the first, the maxim of quality or truthfulness: A speaker is expected to say things that she believes to be true. An apparent non-adherence to this maxim (and also to the other maxims for that matter – appropriate specifications in the cases below notwithstanding) may be classified into one of these four cases:

1. Deception: The speaker says something she does not believe to be true herself in order to mislead the listener.
2. Opting out: The speaker gives or has given evidence that she is unwilling to cooperate and, hence, to obey the maxim of truthfulness.
3. Clash: The adherence to the maxim of truthfulness conflicts with adherence to another maxim.
4. Flouting: The speaker violates a maxim to achieve certain linguistic effects (such as in the case of irony), but the violation of the maxim is so obvious to the listener that she tries to recover unuttered assumptions (so called conversational implicatures) which bring the utterance in line of the principle of cooperation again.

A fifth case might be added which Grice (1975) did not explicitly discuss, but it is listed here for completeness of the discussion:

5. Ignorance of the Maxims: The speaker does not know about the maxims and hence does not adhere to them.

RT however has it that all of the Gricean maxims collapse into his third maxim, that of relevance. Grice (1975) did not explicate his concept of relevance thoroughly: In fact, judging from Grice's (1975) explanation of the maxim of relation, he probably means a different concept of relevance than the one laid down in RT, one that is more closely related to the everyday use of the word: "I expect a partner's contribution to be appropriate to immediate needs" (p. 47). The explication of the notion of relevance in RT on the other hand renders Grice's (1975) remaining maxims redundant: Adherence to them becomes arbitrary once

relevance is secured. As long as an utterance is relevant (i. e. allows the recovery of contextual effects at low cost) Grice's remaining "maxims" could be violated at libitum. Namely the truth maxim, of which Grice (1975) writes that "other maxims come into operation only on the assumption that this maxim of Quality is satisfied." (pp. 46/47) loses its imperative character.

Furthermore, RT posits that its principles of relevance, contrary to Grice's (1975) maxims, need not be learned. Relevance, as conceived of in RT, guides communication (utterance choice and construction as well as comprehension) through requirements of human cognition and the mutual confidence in this guidance in others as argued above in the explication of the second principle of relevance. Adherence to the RT principles securing successful communication will hence require no representation of these as rules or conventions or the learning of rules prescribing how an utterance should be constructed, but merely the normal functioning of the human cognitive apparatus and the general and basic assumption that other humans' apparatus functions very similarly. The processing of information and communicative acts in line with the principles of relevance arises not out of conventions or socially shared norms, but out of the evolved nature and functioning of human cognition.

Finally, according to RT, adherence to the principles will occur spontaneously, without extra effort experienced and the inferences on the base of communications will be triggered largely unconsciously – a notion that has been hinted at by Grice (1975), but not developed further: "The presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature..." (p. 50, emphasis in the original). Sperber and Wilson (1987), on the other hand, are very clear on the issue whether the inference needs to be able to be consciously worked out, and whether the principles and rules governing the inference need to be known: "Communicators and audience need no more know the principle of relevance to communicate than they need to know the principles of genetics to reproduce." (p. 704).

Thus, while Grice (1975) essentially postulates a two process model for communication (one route for the comprehension of "normal", strictly code-based linguistically coded utterances – the "standard case" –, and one for comprehension in the case of detected flouting – a less common case, cf. also Grice, 1978), RT manages with one process: that of relevance extraction on the listener's side and relevance securing on the speaker's side in any communication (for an overview on dual-process models in social psychology see Chaiken & Trope, 1999).

This contradiction between the Gricean and the RT account of how communication succeeds was subjected to empirical investigation by van der Henst, Carles & Sperber (2002). Only study three from this paper will be discussed next, as the present work is concerned with an extension of this study.

#### Experiment Three in van der Henst, Carles and Sperber (2002)

If strict truthfulness of an utterance is not necessary for optimal relevance and thus for comprehension, it may well be violated in the interest of cognitive economy viz. relevance (be it for the listener or the speaker). This is the argument put forth by van der Henst, Carles and Sperber (2002). The answer to the question "What time is it?", posed to a passer-by on the street, is taken to be a case in which, under normal circumstances, after having a look at one's watch, a rounded answer serves the interest of the asker well enough. No effort is made to read the watch more closely. Here, rounded answers, for example 9:05am, are those that state a time with a minute portion dividable by five, even though the minute portion of the time showing on the watch is actually not such a time, but rather, in the example, 9:07am.

Such a rounded answer is strictly speaking false. However, if the asker gives a hint of the requirement that the time information asked for be more precise, more accurate and strictly speaking "truer" than a rounded answer, askers should spontaneously tune their answer to secure the minimal contextual effect of their utterance to the effect that they will tend to be more precise in the minute portion (van der Henst, Carles & Sperber, 2002). The paradigm in which participants are asked for the time and answers are analyzed according to their accurateness will presently be referred to as the "telling the time" paradigm (TTT).

In terms of RT, giving the hint of an appointment that one wants to make amounts to making an assumption more mutually manifest to the askee. In experiment three of van der Henst, Carles and Sperber's (2002) study this is accomplished via the mentioning of an appointment (either the top of the hour or the half hour). By thus making more mutually manifest to the askee that there is a specific time at which the asker needs or wants to be in some specific place, the askee should spontaneously be less inclined to round her answer the closer the time of the question is to time of the appointment. If the appointment is another 30 minutes ahead two or three minutes of accuracy that are lost by rounding make less of a difference to the asker than if it the appointment only 5 minutes away: in the first case she may not particularly hurry yet, but in the latter case she might want to start running.

Indeed, van der Henst, Carles and Sperber (2002) found that a larger proportion of passers-by asked 30 to 16 minutes before an appointment casually mentioned by the asker

("earlier" group) gave a rounded answer (97.5%) to the time question than of the "later" group asked 14 to 0 minutes before the appointment (79.7%)<sup>2</sup>.

This result is presently taken as evidence for the sensitivity of humans to the needs and goals of the listeners when they construct utterances, a spontaneous tendency to secure optimal relevance by tuning their utterances to assumptions made mutually manifest by verbal presentation. In a case where the speakers assume that a high degree of accuracy is not relevant (longer before the appointment, "early" group) they save themselves the effort of giving an exact and strictly true answer, being more precise would not increase the relevance of their utterance. In fact, in the "early" group case a more precise answer could well be seen as lowering relevance, since extra effort is added for the speaker (observing their watch more closely) and the listener (a more precise answer may be longer and more difficult to calculate the time remaining until the appointment with<sup>3</sup> and thus might require more effort to be processed), but no extra effect arises from it. Contrarily, shorter before the appointment ("later" group), where a more precise and "truer" answer could make a difference in effect (and, assuming that the added effect is not outweighed by the added effort, a difference also in relevance) as sketched above, they give more accurate answers.

Note that only one of the cases of Grice's (1975) maxim violation classes mentioned above applies here. The participants asked for the time may safely be assumed not to want to deceive the experimenters, they do not opt out of a convention to cooperate (in fact they clearly do cooperate by answering), they probably do not want to achieve linguistic effects by means of a conversational implicature<sup>4</sup> and if there is such a thing as a cooperative principle to be known, they surely do know it. The only explanation for the results found by van der Henst, Carles and Sperber (2002) in the light of Grice's (1975) maxims is that fulfillment of the maxim of quality/truth clashes with or contradicts securing relevance. This analysis together with the argument for collapsing all Gricean maxims into two (cognitively and socially grounded) principles of relevance suggests that relevance (and not truth) is the lynchpin of communication: as long as it is secured, communication can succeed, no matter if what is said is strictly speaking true, obscure, lengthy, ambiguous etc.

### Hypothesis

The assumption made mutually manifest in van der Henst, Carles and Sperber's (2002) experiment three is already quite subtle. The present study aims at going further in making the hint at the required degree of precision even subtler.

Comprehension and utterance choice and construction draw upon any assumptions that are to some degree (mutually) manifest to the communicators. In experiment three of van der

Henst, Carles and Sperber's (2002) study, the crucial assumption that the asker wants to make a specific appointment is made manifest to the askee by an utterance of the asker herself. There are however circumstances in which such an assumption may already be manifest enough to elicit tuning of the answer, such as certain prominent appointments, that both asker and askee know and the asker may be reasonably assumed by the askee to want to make. One such case are the starting times of lectures and other courses at universities that largely start at certain standard hours usually known to every student and teacher. To persons on a university campus who are asked for the time by an experimenter who could easily be a student, the assumption that the person asking wants to make it in time to a course is considered more manifest than to a person asked somewhere downtown, where they could have any appointment or no appointment at all. The times around prominent starting times of lectures and seminars thus provide already mutually manifest assumptions, that the askee can use to tune her answer to the question what time it is when approached on campus. If a speaker's tendency to spontaneously adjust optimal relevance really is as pervasive, spontaneous and highly developed as the above argument suggests, than a difference in percentage of rounded answers similar to the one found by van der Henst, Carles and Sperber (2002) should also be found on a campus between two different time intervals: one temporarily relatively remote to standard starting times of course sessions and one closer to these times such that persons asked longer before the standard starting times will tend to be less accurate than those asked shorter before starting times. The difference may be smaller than van der Henst, Carles and Sperber's (2002), since the manifestness of the assumption of the appointment is not made manifest as strongly, but assumed to already be mutually manifest (even though possibly to a lesser degree than when it is mentioned by the asker). It should nevertheless be sufficient to yield observable differences.

The campus group will be compared to a group of participants approached downtown, remote from the main campus. The difference expected in the campus group between time-intervals should not arise in this group, since no prominent appointment is assumed to be mutually manifest to an asker and askee there.

## Study

### Method

#### Participants

Two experimenters in the mid-twenties approached 743 passers-by on the main campus of the University of Göttingen or in downtown Göttingen and asked them for the time in a

natural, casual way either longer or shorter before standard lecture starting times. Participants were not aware of their participating in a study. Only persons seeming to be between 20 and 30 years old were asked, so that the German informal form of you ("Du") could be used in asking the question. The participants were not asked for anything else but the time.

### Design

The study used a 2 (location: campus vs. downtown)  $\times$  2 (time interval: earlier vs. later) between-subjects design with the cells earlier/campus (EC), later/campus (LC), earlier/downtown (ED) and later/downtown (LD). Note however that none of the factors were manipulated experimentally, both location as well as time of the observation were given.

### Procedure

A survey of the schedule of classes at the University of Göttingen confirmed that most courses in the halls around the main campus on Mondays and Tuesdays held between 8am and 2:15pm start at 9:15 am, 11:15 am and 2:15 pm, some on the full hour and very few at other times (cf. Figure 1).

-----  
 insert Figure 1 about here  
 -----

Data collection was thus set to the Mondays and Tuesdays of two consecutive weeks at the beginning of the summer semester 2003 and to the times 8:15 – 8:45 am, 10:15 – 10:45 am, and 1:15 – 1:45 pm for the "earlier" time intervals, and 8:45 – 9:15 am, 10:45 – 11:15 am, and 1:45 – 2:15 pm for the "later" intervals.

The "campus" group participants were approached within a set area of the main campus ("Platz der Göttinger Sieben"), the "downtown" group were passers-by around a block in the pedestrian zone in the city of Göttingen (Weender Straße, Barfüßerstraße, Judenstraße and Prinzen-/ Theaterstraße). These two locations are withing about ten minutes walking distance of each other.

All participants were approached by a male or a female experimenter (who were balanced across weekdays and places) and asked "Excuse me, would you give me the time?" ("Entschuldigung, kannst Du mir sagen, wieviel Uhr es ist?", a common wording for this question in German). After they had responded, their gender, if possible the type of watch (analogue, digital or mobile phone), type of answer given (see below for types) and the response time, taken by a common digital stop watch to the hundredth of a second were recorded without them noticing it. The response times were recorded for simple exploratory

purposes. One might expect that if participants make some effort to ensure optimal relevance in giving the time, it may take a little longer for them to read their watch more accurately in the LC condition than in the other conditions (time interval  $\times$  location interaction). The measurement, however, was very insensitive: concealing the stop watch in their pockets, experimenters approximately started taking the time when they were done asking the question and stopped when they said "Thank you" to the answer given.

Persons either talking on a mobile phone, wearing headphones, carrying many things or otherwise involved in personal activities were not asked to ensure that the question would be minimally intrusive. For the same reason, only singly walking persons were asked.

For exploratory purposes, it was also recorded on the third and fourth day whether people were apparently just arriving and/or going to a lecture (in the campus conditions) or walking in the direction towards the campus (in the downtown conditions) or were evidently leaving from campus or/and a course session (for the campus conditions) or walking in the opposite direction from campus (in the downtown condition). These recordings were made as "A" (for arriving or walking towards campus) and "L" (leaving or walking away from campus) and the variable will be referred to as "A/L" below.

## Results

### Answers

Of the total of  $N=743$  observations made, seven were excluded because due to an error data collection had been started before the intended time. One observation was excluded since the participant's watch was evidently both digital and analogue.

Of the remaining 735 cases, 154 cases (21%) in which the person asked

- read the time off a public watch,
- did not answer at all,
- did answer very unclearly so that the experimenters could not understand,
- had obviously looked at her watch just before she would have been asked by the experimenter,
- guessed without looking at a watch, or
- were obviously asked a second time because of failure of experimenters to realize that right away

were also excluded from analysis. The answers thus excluded were approximately evenly distributed across conditions ( $\chi^2=5.043$ ,  $p = .169$ ).

Of the then remaining 581 observations, 30 (5.2%) of the recorded watch types were digital and 80 (13.8%) were mobile telephones. For 12 cases (2.1%) the watch type could not be determined. The proportions of digital watches, mobile telephones and 'undetermined watch type' cases did not differ across conditions ( $\chi^2 = 1.882$ ,  $p = .597$ ;  $\chi^2 = 3.117$ ,  $p = .374$ ;  $\chi^2 = 5.227$ ,  $p = .156$ , respectively). The cases in which participants read the time off digital watches or mobile telephones were excluded since van der Henst, Carles and Sperber (2002) found a significant difference in the basic propensity of participants' giving rounded answers between those with analogue vs. digital watches in experiment one. This left a sample of  $N = 471$  observations.

Not surprisingly, since efforts were made to these effects, observations were largely equally distributed across conditions with respect to the following possible confounding variables: experimenter ( $\chi^2 = 0.95$ ,  $p = .813$ ), day of data collection ( $\chi^2 = 6.731$ ,  $p = .665$ ), and gender of participants ( $\chi^2 = 2.430$ ,  $p = .488$ ).

The answers given by participants were classified into three categories:

- R-responses: answers that indicate a time with a minute information that is a multiple of five. This category also comprises such answers not indicating the hour (which is common if the hour part is assumed to be known by the addressed person) and answers indicating the full hour, the half hour and the quarter hour. Examples are "Five of nine", "Half past ten", "Eleven o'clock", "A quarter of two".
- X-responses: answers that indicated a time with a minute information that is not a multiple of five (including such answers only indicating the minutes and answers qualified by disclaimers such as "...but my watch is going wrong" or "approximately") – e. g. "Eight fifty six", "Approximately eight sixteen" "Two 0 seven, but that might be wrong".
- K-responses: responses similar to type R, but qualified by remarks such as "almost", "approximately" or "shortly after" – e. g. "Shortly after half past", "Soon nine", "Approximately two".

The numbers of answers of each type are presented in Table 2.

-----  
 insert Table 2 about here  
 -----

In van der Henst, Carles and Sperber (2002), the number of R-responses were adjusted based on the following reasoning: cases in which a multiple of five was given in the minute

part, the time showing on the watch of the participants might actually have been exactly that time. These cases however do not speak to the tendency of askees to give accurate vs. inaccurate responses since one does not know whether they had answered with an R-type response as well had their watch shown a time with a minute portion not dividable by five with no rest. These cases should hence be excluded from analysis. Since the actual time on participants' watches was not recorded, it can not be estimated how many of the R-type cases were actually "accurate". 20% of all answers in van der Henst, Carles and Sperber's (2002) experiment three can be expected to be such cases assuming that the times at which participants were asked and possible deviations of times showing on their watches from the "true" time were random. The percentage of rounded answers was therefore computed by means of the formula:

$$\text{percentage of rounders} = \frac{(\text{percentage of } 5x - \text{responses}) - 20}{80} \quad (\text{Formula 1})$$

where "5x-responses" are conceptually identical to responses presently classified as R-responses. This leads to an adjusted proportion of rounders, in the present classification and in the case of only R- and X-type responses being recorded, equivalent to:

$$\text{proportion of rounders} = \frac{R - .2(R + X)}{.8(R + X)} \quad (\text{Formula 2})$$

Hence if only one fifth of all answers are R-type answers (as can be expected by chance even if all participants accurately read off the exact time showing on their watch), the adjusted proportion of rounders will be zero, if, on the other hand, all answers are R-responses, it will amount to 1.

But van der Henst, Carles and Sperber (2002) only recorded answers involving minute parts at all. Since in this present study, answers of type K, (unequivocally "inaccurate" and/or indicating no definite minute portion at all), were recorded, a slightly different procedure, but based on the same argument, was followed:

First, all of the answers of type K will be counted as inaccurate, since in all of these cases, an answer was given that did not strictly correspond to the time (precise to the minute) actually showing on the watch. Of the remaining X and R answers, an expected 20% may be assumed to have been R-answers that could have been X answers under different circumstances (i. e., if the time on the watch had not shown a time with a multiple of five as a

minute portion). Therefore, in line with formula 2, 20% of X- and R- responses will be excluded from analysis. The sum of the remaining R-type responses and all K-type responses will then be the number S of "inaccurate answers":

$$S = K + R - .2(R + X) \tag{Formula 3}$$

S divided by the corrected total number per cell  $N_{corr}$  (which has also been adjusted for the 20% of X- and R-type responses dropped before) will constitute the measure I for "adjusted proportion of inaccurate answers":

$$I = \frac{S}{N_{corr}} = \frac{K + R - .2(R + X)}{K + .8(R + X)} \tag{Formula 4}$$

This term has the same characteristics as the formula from van der Henst, Carles and Sperber (2002), but takes K-type answers into account: if the percentage of type-R answers is one fifth (as expected by chance even if everybody who gives an answer with a minute portion and without a disclaimer reads off the exact time their watch shows), the adjusted proportion of inaccurate answers will be the proportion of K-type answers, if there are only K- and R-, but no X-type answers, the expression will yield  $I = 1$ .

Note that through these computations, non-integer numbers may result for S and  $N_{corr}$ . In order to be able to apply a  $\chi^2$ -test, these numbers will be rounded into S' and N' respectively, such that numbers with a decimal portion  $<.5$  will be rounded to the smaller integer and numbers with a decimal portion  $>.5$  to the next larger integer.

The resulting values of I within each cell of the design together with S, corrected total, S', and N' are shown in Table 3. Subsequent analyses are based on S' and N' unless otherwise indicated.

-----  
 insert Table 3 about here  
 -----

It was predicted that the proportion of inaccurate answers will be higher in the EC than in the LC group. This difference was found ( $\phi = .1296$ ) and is reliable ( $\chi^2 = 3.548, p < .03$ ). As hypothesized further, no difference emerged from the comparison of the ED and LD groups ( $\phi = -.0169, \chi^2 = 0.05, p = .41$ ). The difference in the campus group differs from the one in the downtown group in the predicted direction (interaction pattern, difference in  $\phi = .1466, \chi^2 = 4.167, p < .03$ ).

## Response Times

Since response times were not available for a few cases due to technical difficulties, the following analyses of response times is based on a subsample of the sample used for the analysis of answer types ( $N = 433$ ). Number of cases lacking response times recorded did not differ between earlier and later time intervals, but there were more missing cases in the campus than in the downtown condition (11% vs. 5 % respectively).

No differences in response times were found in a  $2$  (campus vs. downtown)  $\times$   $2$  (earlier vs later time interval) ANOVA (all  $p$ s  $> .28$ ).

Since it was suspected that experimenters were differently trained in taking reaction times with a stop watch of the sort used here, the analysis was repeated with experimenter as a third factor. Except for a marginal two way interaction of location and experimenter ( $F = 2.94$ ,  $p = .087$ ,  $R^2 < .01$ ), no effects emerged. In particular, no three-way interaction of location, interval and experimenter was found ( $F < 1$ ,  $p > .35$ ) which could be hinting at a difference in accordance with the predictions for at least one of the experimenters. No evidence for a more time consuming elevated effort for participants in the LC condition was thus found.

The presently reported analyses on response times have to be interpreted with great caution. The measurement technique was highly unsophisticated and experimenters did not train taking response times together. Efforts should be made in an improved future replication of this study to ensure that accurate and reliable response times be taken.

## Discussion

### Results

The predicted difference in the campus group was found, and virtually no difference emerged in the downtown group.

In order to insure that the artificial measure I introduced here has not created an artifact, for exploratory purposes, the predictions made were tested with measures that slightly differed from I, but should be correlated with them. Even though this does not provide any completely independent evidence, if the same pattern emerges, it would create confidence in the possibility to replicate the trend with I found here. It additionally speaks to the validity of the newly used measure.

Analysis based on unadjusted numbers (X-type or "accurate" answers vs. R-/K-type or "inaccurate" answers) yielded a significant difference in the predicted direction for the campus condition ( $\phi = .12$ ,  $\chi^2 = 3.7$ ,  $p < .03$ ), but none for the downtown condition ( $\phi = -.01$ ,  $\chi^2 = 0.03$ ,  $p > .4$ ). The difference in  $\phi$ s is significant ( $\phi = .13$ ,  $\chi^2 = 4.1$ ,  $p < .03$ ).

Using the adjustment procedure described by van der Henst, Carles and Sperber (2002) – thus dropping the answers here referred to as K-answers and adjusting the remaining answers like van der Henst, Carles and Sperber (2002)–, yields the same pattern ( $\varphi = .16$ ,  $\chi^2 = 4.7$ ,  $p < .02$  for the campus condition and  $\varphi = 0$ ,  $\chi^2 = .003$ ,  $p > 0.4$  for the downtown condition, difference in  $\varphi = .16$ ,  $\chi^2 = 4.4$ ,  $p < .02$ )<sup>5</sup>.

The convergence of these three analyses (even though they are not independent) weakens concerns regarding an artifact due to the newly introduced measure I.

Comparing the LC condition to the remaining three conditions, in each of the modes of calculation, answers in the LC condition are more accurate than those in the remaining conditions combined (all  $\chi^2$ s  $> 7.6$ ,  $ps < .01$ ). The remaining three conditions however do not differ (all  $\chi^2$ s  $< .2$ ,  $ps > .48$ , again, under all modes of calculation of the percentage of inaccurate answers). In particular, the EC condition does not differ from the two downtown conditions (ED and LD), all  $\chi^2$ s  $< .15$ ,  $ps > .35$ .

This allows for the interpretation that there was something special about the LC group that the other conditions lacked and that especially the EC group lacked. The specific situation for these participants in which they were asked by someone for whom an accurate answer might be relevant seems to have had an effect.

### An Alternative Explanation

The results do not contradict the prediction made by RT, however, a plausible alternative explanation can be brought to bear: The prediction tested here was derived from the argument that there should be a spontaneous tendency of humans to provide optimal relevance in communication to others, because otherwise the spontaneous tendency for humans to expect optimal relevance in communication from the another communicator (Sperber and Wilson, 2002) could not be very adaptive.

The situation in which the prediction was put to the present test however, allows for a different and more parsimonious explanation, namely that the participants of the study did not provide optimal relevance to the experimenter asking them for the time, but to themselves and the more accurate answers obtained by the experimenters were a mere by-product of this provision. The interpretational ambiguity is mainly due to the correlational character of the collected data. Both "independent variables" (location and time interval) were not manipulated; rather, they were given: participants were asked when and where they were encountered by the experimenters. It is reasonable that those encountered on campus in the "later" time interval were there with a specific intention, namely the same intention that they

were to impute to the experimenters as these asked them for the time: to go to a lecture or seminar. In this case, more accurate knowledge of the time should also be relevant to themselves. If then asked for the time in that situation they might as well provide that relevant information to themselves by retrieving a more accurate reading off their watch and, while they are at it, pass that information on to the experimenter asking without having to infer any relevance for the asker. This problem also arises in the interpretation of van der Henst, Carles and Sperber's (2002) study three: It is, in principle, conceivable that those asked shortly before the top of the hour or the half hour were themselves on their way to an appointment. Van der Henst, Carles and Sperber (2002) also point to this possibility themselves stating that "Most people's activities are commonly scheduled (if they are scheduled at all) to start on the the hour or half hour..." (p. 459). The people that they have encountered shortly before the appointments mentioned by the experimenters could have been on their way to an appointment at the same time themselves and then might have an interest in a more accurate time themselves (since it would be more relevant) which they could then pass on to somebody else asking them for the time. Such an explanation clearly does not need the postulate of a social-cognitive mechanism that is sensible to the needs and goals of a person talked to, let alone mutually manifest assumptions (and thus relevance), as it is invoked in the theoretical discussion above. It would suffice for a person to be interested in relevance to herself (cf. the first, cognitive principle of relevance).

A manipulation and random assignment of the time at which a person is asked by an experimenter would, of course, allow for a test of the competing explanations. Also, random assignment to the place would be important to strengthen the case, since even though efforts were made to only ask participants who could by their apparent age be students, it is not at all clear whether the populations from which the samples in the campus and downtown groups stemmed were identical. Of course, in order for this to raise serious doubts, one would have to argue that students have a generally different propensity to give a precise answer in a later condition vs. an earlier condition compared to participants from the general age group in the TTT – a matter to be decided by empirical investigation rather than of a priori plausibility.

The present observational context, however, would not allow for such manipulations. By collecting additional variables, one might nevertheless be able to further discard any competing explanation.<sup>6</sup> The variable A/L recorded in the present study may be such a variable.

As was indicated above, on the last two out of the four days of data collection, it was attempted to record whether participants were walking towards the campus halls or away from

it when they were asked by the experimenters on campus (variable A/L). In the downtown condition the same distinction was tried, however it was only possible for the experimenters to decide whether participants were walking in the general direction of the campus halls or not<sup>7</sup>. While this decision was very difficult and probably not a very reliable and valid measurement on campus already, the problems are exacerbated downtown. Nevertheless, analysis of these informations should give a first idea of the plausibility of such an alternative explanation.

Additionally, on these two days, an effort was made to prevent grossly disparate proportions of A- and L-cases in the samples. This might well have affected the results of the main analyses reported here, it is however impossible to retrospectively assess this impact. This point should, once again, be considered in future studies in the TTT.

These problems notwithstanding, if it is true that the above mentioned alternative sufficiently accounts for the difference found in accuracy in the two campus conditions, it should be mainly driven by answers by participants arriving on campus and presumably intending to go to a lecture or seminar (A). In particular, X-type responses should mainly come from those participants.

In the earlier interval the distribution of answer types was similar for A- and L-coded observations ( $\chi^2 = 0.535$ ,  $p > .7$ ) in the campus group<sup>8</sup>. In the LC group the distribution did not differ significantly either ( $\chi^2 = 2.229$ ,  $p > .3$ ). The same holds in the downtown groups (both  $\chi^2$ s  $< 1.5$ ,  $ps > .8$ ).

The issue should however again be systematically addressed in a future study in the TTT, possibly by controlling rather than measuring the A/L-distinction (see also below).

### Grice's Maxims and RT

Taking answers to the time question of the presently classified R- and K-types being inaccurate and thereby, strictly speaking, wrong answers, a difficulty for Grice's (1975) account of communication arises. Three of the four cases, in which violations of the maxim of truth and thereby the CP fall according to his account, do not apply in the present situation: Few, if any of the participants sought to deceive the experimenters, they did not declare or otherwise make clear that they "opted out" of any convention governing communication and they probably weren't aiming for any particular linguistic affects such as those ensuing in cases of flouting. Also, the additional case mentioned earlier, that one of the conversants does not know the convention known by the other, seems to be far-fetched. The only way the Gricean (1975) model could account for the neglect of accuracy and truth would be to invoke that their securing clashes with another maxim. However, the only maxim worth considering

at all would be the relevance maxim, but Grice does not specify this notion to a degree where it would be obvious that the different answers to the time question in the present study could be discerned as more or less relevant. Also, Grice (1975) asserts that the truth maxim has somewhat of a pre-emptive status in that if it is violated, considerations of violation of other maxims is suspended.

Of course, collapsing R- and K-type answers into one category of "inaccurate" answers calls for a remark: While R-type answers are strictly speaking "untrue", K-type answers are only imprecise. However, if the RT account of what answers askees will give to the question for the time is feasible, then strict truth as well as high precision should be adhered to in securing relevance, both features of an answer contribute to higher relevance in the LC, but not the other conditions. One might object that the answers obtained in the present study and classified as K-type responses are not untrue (neither strictly nor loosely speaking). But they are surely imprecise. They thus contradict the maxim of manner and namely the sub-maxim "Avoid ambiguity." (Grice, 1975). Then, the argument that Grice's account cannot explain why participants in the LC condition would give less imprecise answers goes through in an fashion similar to the one based on the truth maxim: only one of the maxim violation cases sketched by Grice (1975) could apply: the clashing of the postulate of unambiguity with the postulate of relevance. Again, securing relevance pre-empts the adherence to another maxim.

Clearly the present results (and those of van der Henst, Carles & Sperber(2002) for that matter) leave the Gricean account struggling with a phenomenon that can straight forwardly and readily be captured and explained by RT, the plausibility of yet a different competing explanation notwithstanding.

### Recommendations for Future Research in the TTT

The data presented here are purely correlational. This poses a major problem: a definite causal interpretation of the effects achieved must be refrained from (especially in the light of the competing hypothesis explaining the effect) since the distribution of possibly confounding variables was either not controllable due to resulting very small cell frequencies (such as for type of watch) or failure to record them in a reliable, valid way and for all cases of the study (such as the A/L-variable).

The problems of a field study as reported here of course poses some other possibly genuine problems such as that it cannot be guaranteed that a few participants were not asked twice on different days and times and/or on different locations. In some cases, it even happened to the experimenters that they asked a person which later they recognized as someone they had already asked at a different time and place.<sup>9</sup>

A future study in the presently used paradigm should secure the following characteristics in order to avoid the ambiguities that arose from the present data:

Response time taking should be trained and reliability increased. Differences in processing time consistent with RT should be very small. The measurement technique in the present study was very unsophisticated and possibly not fit to capture these differences if they exist.

There should be control for the level of relevance that an accurate or inaccurate answer could represent for the participants themselves. This might not be realized easily in a study on campus where people might pursue different plans (going to a lecture, going to have lunch, going to the library etc.), but it might be possible near a train station where no public clock can be seen and where people leaving the station building can be more strongly (than in the campus condition of the present study) assumed not to have an appointment themselves in compared to those entering the station. Of course this might only be a reasonable assumption on Saturdays and Sundays in a small University town like Göttingen, when there are very few people arriving to the city by train to make a specific meeting time (e. g. business or academic meetings).

In the present study, it was to a large extent possible to actually have a peak at the participants' watches. Explicitly making the effort to judge whether X-, and R-type answers do correspond to the time really showing on the watch from which the answer is read off could deliver more precise data since the adjustment procedures for proportions of inaccurate answers would become redundant.

Another possibility of conducting more clearcut tests for the prediction that speakers will adjust the accuracy and precision of their answers to the imputed requirements of the listener could be the creation of situations in the laboratory where participants may give more or less precise answers to the innocent question for participants' age or other information that can be given to different degrees of truth or accuracy. Pretendedly incidental information on the (possibly bogus) research topic to the effect that precise vs. not so precise information given by the participants makes a difference in relevance to the experimenter could serve as a hint similar to the starting times in the present study.

### Conclusion

The predictions made by RT were confirmed. The present study replicated the results of van der Henst, Carles & Sperber's (2002) experiment three in a different and more fine-grained test of the contention that speakers spontaneously secure relevance to a listener.

Interpretational ambiguity due to failure to control for a possible alternative explanation of the results may be overcome in future studies as sketched in the discussion.

The optimistic hope for a replication of the results in a more sophisticated design leaves TTT as promising naturalistic setting of further research on RT.

### Acknowledgements

I would like to thank Melanie Jaeger for her stoicism in helping to collect the data reported. Thanks are also due to Anita Ianucci, Barbara Schauenburg, Björn Meder, Kai J. Jonas, and Margarete Boos (in strictly alphabetical order) for help with this work.

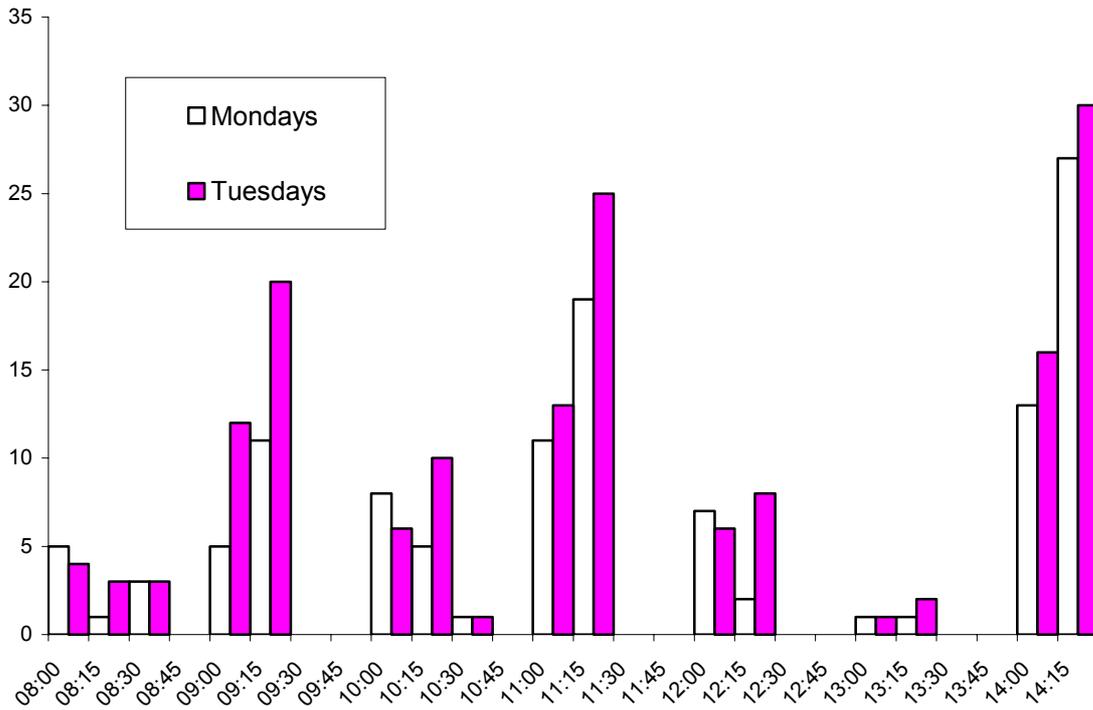
## References

- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37-46.
- Bezuidenhout, A. & Cutting, J. (2002). Literal meaning, minimal propositions, and pragmatic processing. *Journal of Pragmatics*, *34*, 433-456.
- Byrne, R. & Johnson-Laird, P. (1989). Spatial Reasoning. *Journal of Memory and Language*, *28*, 564-575.
- Chaiken, S. & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York, NY: Guilford.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with Wason's Selection Task. *Cognition*, *31*, 187-276.
- Gibbs, R. & Moise, J. (1997). Pragmatics in understanding what is said. *Cognition*, *62*, 51-74.
- Giroto, V., Kimmelman, M., Sperber, D., & van der Henst, J.-B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, *81*, B69-B76.
- Grice, H. (1975). Logic and Conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts*, (pp. 41-58). New York: Academic Press.
- Grice, H. (1978). Further Notes on Logic and Conversation. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 113-127). New York: Academic Press.
- Happé, F. (1993). Communicative Competence and Theory of Mind in Autism: A Test of Relevance Theory. *Cognition*, *48*, 101-119.
- Higgins, E. (1996). Knowledge Activation: Accessibility, Applicability, and Salience. In E. Higgins & A. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 133-168). New York, NY: Guilford.
- McCann, C., & Higgins, E. (1992). Personal and Contextual Factors in Communication: A Review of the 'Communication Game'. In G. Semin & K. Fiedler (Eds.), *Language, Interaction and Social Cognition* (pp. 144-172). London: Sage.
- Johnson-Laird, P., Legrenzi, P. & Legrenzi, M. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, *63*, 395-400.
- Nicolle, S. & Clark, B. (1999). Experimental pragmatics and what is said: a response to Gibbs and Moise. *Cognition*, *69*, 337-354.
- Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608-631.
- Oaksford, M. & Chater, N. (1995). Information gains explains relevance which explains the selection task. *Cognition*, *57*, 97-108.
- Shannon, C. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Sperber, D., Cara, F., & Giroto, V. (1995). Relevance theory explains the selection task. *Cognition*, *52*, 3-39.
- Sperber, D. & Wilson, D. (1986). *Relevance: Cognition and Communication*. Oxford: Blackwell.
- Sperber, D. & Wilson, D. (1987). Précis of Relevance: Communication and Cognition. *Behavioral and Brain Sciences*, *10*, 697-754.
- Sperber, D. & Wilson, D. (1995). *Relevance: Cognition and Communication*. Second edition. Oxford: Blackwell.
- Sperber, D. & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language*, *17*, 3-23.
- van der Henst, J.-B., Carles, L. & Sperber, D. (2002). Truthfulness and Relevance in Telling The Time. *Mind & Language*, *17*, 457-466.

- van der Henst, J.-B., Sperber, D. & Politzer, G. (2002). When is a conclusion worth deriving? A relevance based analysis of indeterminate relational problems. Thinking and Reasoning, 8, 1-20.
- Wason, P. (1966). Reasoning. In B. Foss (Ed.), New Horizons in Psychology (pp. 135-151). Harmondsworth: Penguin.
- Wilson, D. & Sperber, D. (2002). Truthfulness and Relevance. Mind, 111, 583-632.

Figures

Figure 1: Distribution of starting times of courses around Platz der Göttinger Sieben at the University of Göttingen on Mondays and Tuesdays (absolute frequencies of a total of N=280).



Note: it is a wide spread custom at German universities to indicate the starting time as the full hour but actually start it at a quarter after (so called "academic quarter"), thus the predominance of courses starting at X:15 might still be underestimated.

## Tables

Table 1: Conversational maxims according to Grice (1975).

<u>Supermaxim</u>	<u>Submaxims</u>
Quality	1. Do not say what you believe to be false. 2. Do not say that for which you lack adequate evidence.
Quantity	1. Make your contribution as informative as is required (for the current purposes of the exchange). 2. Do not make your contribution more informative than is required.
Relation	1. Be relevant.
Manner	1. Avoid obscurity of expression. 2. Avoid ambiguity. 3. Be brief (avoid unnecessary prolixity). 4. Be orderly.

Table 2: Number of answers of the three types by location and time interval, percent per location  $\times$  time interval cell in parentheses.

R: rounded answers with minute portion; X: unrounded answers with minute portion; K: short, unambiguously unprecise answers; C: cases observed on campus; D: cases observed downtown;  $\Sigma$ : marginal sums; see text for additional details.

	time interval						$\Sigma$		
	earlier			later			R	X	K
	R	X	K	R	X	K	R	X	K
C	108 (86.4)	10 (8.0)	7 (5.6)	96 (72.2)	21 (15.8)	16 (12.0)	204 (79.1)	31 (12.0)	23 (8.9)
D	90 (82.6)	8 (7.3)	11 (10.1)	80 (76.9)	7 (6.7)	17 (16.3)	170 (79.8)	15 (7.0)	28 (13.1)
$\Sigma$	198 (84.6)	18 (7.7)	18 (7.7)	176 (74.3)	28 (11.8)	33 (13.9)	374 (79.4)	46 (9.8)	51 (10.8)

Table 3: Proportions of Inaccurate Answers (I), values of S and  $N_{\text{corr}}$ , artificial S' and N' (see text for details on these values).

		time interval									
		earlier					later				
		I	S	$N_{\text{corr}}$	S'	N'	I	S	$N_{\text{corr}}$	S'	N'
location	campus	.9014	91.4	101.4	91	101	.8084	88.6	109.6	89	110
	downtown	.9105	81.4	89.4	81	89	.9192	79.6	86.6	80	87

## Footnotes

<sup>1</sup> The notion of what is "explicitly" or "literally" said is not unproblematic (see Gibbs & Moise, 1997; Nicolle & Clark, 1999 and Bezuidenhout & Cutting, 2002 for discussion of this matter). In the present context however, an intuitive and unsophisticated sense of these terms should not lead astray, the problem will hence not be further treated.

<sup>2</sup> The difference was preserved (97% vs. 75%) after correcting for cases in which a "rounded" answer was actually not rounded but an accurate reflection of a time showing on the watch of the askee, e. g. maybe it was exactly 2:05pm on her watch. In these cases it cannot be decided whether participants would have given an answer accurate to the minute, had their watch showed a time with a minute portion that is not a multiple of five. Since the actual time on participants watches were not recorded however, a correction had to be made based on an estimate. These cases should amount to one fifth of a random sample of times taken accurate to the minute and be excluded from analysis. The exact procedure of correction will be discussed below in the method section, since it will provide the basis for the measure used in the present study.

<sup>3</sup> This argument of course rests on the assumption that the distances between arbitrary times with minute portions dividable by five without rest are already or better encoded and hence can be retrieved easily while the difference between times one or both of which have a minute portion not dividable by five without rest involves some degree of calculation. Of course, this matter awaits empirical investigation but is taken on good faith as an auxiliary hypothesis in the present context.

<sup>4</sup> Even if they were trying to achieve such an effect, why should they do so only shortly before the appointment, but not earlier? Also, the case called "flouting" by Grice (1975) must provide for some hint that a conversational implicature is intended: "... He may flout a maxim; that is, he may blatantly fail to fulfill it." (p. 49, emphasis in the original). Clearly, no such hint is given when an askee gives a rounded answer.

<sup>5</sup> This result however is burdened by the same problem as the analysis using S' and N': the adjusted numbers are non-integers and thus have to be rounded in the same way.

<sup>6</sup> The author explicitly acknowledges the general weaknesses of observational studies and, by contrast, the splendid unambiguosness experimental studies afford. However, the

impressively naturalistic setting of the TTT is presently viewed as deserving some more exploitation with more sophisticated observation methods and designs.

<sup>7</sup> Additionally in the city of Göttingen there are single University buildings scattered across the entire town such that even if one were sure that someone is or is not going to the campus proper, they could still be going to another university building to attend a lecture or seminar there. That this was a problem because participants in the downtown group also had an appointment in either earlier or later time intervals, is not plausible, since in that case, there should have been no difference between the campus and downtown conditions. Also in that case, the assumption that the person asking for the time intended to make it to a lecture in time would have been less manifest, which essentially reinstates the original difference in manifestness of the assumption that the person asking needs a more precise answer because it is relevant.

<sup>8</sup> Fisher's exact text was used in the analyses involving the A/L distinction because cell frequencies were well below 5 in several cases.

<sup>9</sup> If this happened, the case was of course excluded from analyses, see above.