The relevance of Relevance Theory to first and second language acquisition

Susan Foster-Cohen

Abstract

This paper begins by outlining the key tenets of Relevance Theory (Sperber and Wilson 1986, 1995), focussing on the Principle of Relevance and notions of cognitive context. It then suggests that this theory can provide a useful perspective on language learner behaviour, both child and adult, a proposal that is supported in the rest of the paper through an examination of a range of published studies. Relevance Theory is shown to go beyond standard analyses of both children and adult language learners' behaviour, providing explanations for otherwise baffling behaviour, and offering insights into recognised issues in acquisition, such as failure in referential communication tasks, focus-on-form, the status of declarative knowledge, and notions of learner effort.

Key words: relevance, pragmatics, acquisition, cognitive, comprehension, assumption, interpretation, inferencing, logical form, effort, noticing, effect, manifest, focus-on-form, declarative knowledge

1. Introduction

It is probably fair to say that the study of developmental pragmatics, which deals with the emergence of pragmatic skills in first and second language learners, has been generally hampered by the lack of a strong theoretical foundation. The study of pragmatics (and developmental pragmatics along with it) has frequently been limited to descriptive cataloguing of behaviours without sufficient attention to explanatory adequacy. This seems particularly true of sociolinguistic/anthropological approaches, which have produced detailed descriptions of how
speakers/hearers achieve different pragmatic goals in different contexts. Also generally lacking have been psycholinguistic approaches that can explain what speakers and hearers know and do.

A notable exception has been the research tradition stimulated by the work of Austin, Searle and Grice, although much of that work, too, has resulted in lists (of speech acts, for example), rather than a principled account of how speakers and hearers arrive at interpretations of utterances in context. Relevance Theory (Sperber and Wilson 1986, 1995), therefore, is a significant contribution to the field because it seems to be a theory of the right kind. It is a cognitive account, in the Gricean mould, of how hearers take what speakers say and derive the message expressed. Although there is still much active work going on developing the theory itself, it is not inappropriate to begin to consider how Relevance Theory (henceforth RT) applies to language acquisition, both first and second. In what follows, I will present the outlines of the theory and then examine some aspects of what we know about language development and so see what kind of light RT can throw on them.

2. Relevance Theory

Relevance Theory, developed by Dan Sperber and Deirdre Wilson, differs significantly from the Gricean account in that, unlike Grice, Sperber and Wilson argue that non-demonstrative inferencing is involved, not just in the interpretation of implicatures from the proposition expressed, but in the derivation of the proposition expressed itself. That is, pragmatic interpretation is involved as soon as the linguistic modules release any structured string of concepts, and hearers construct on the basis of the most accessible (most relevant) context, with the least amount of effort possible, the most likely, i.e. most relevant, interpretation.

For example, take an utterance like *He wants some*. This requires inferential work to be understood. At the very least, we need to know who *he* is, and what it is *he* wants. We also need to know what time frame is involved. In a context in which *He wants some* is to be interpreted as *He wants some tea*, the time frame for the verb of wanting is probably going to be a matter of a few minutes. On the other hand, in a context in which it is to be interpreted as *He wants some success in his life*, the time-frame of the verb is potentially a matter of years. Beyond this, hearers also have to be able to infer other aspects of the speaker's inten-

---

1 In fact, this work was begun some time ago by Smith (1985), but has not been taken up since.
tion: for example, whether the utterance is to be interpreted simply as an observation of a state of mind or perhaps as a request for action (in a context in which the hearer is seen to be circulating the room with an open whisky bottle, topping up the glasses of those who desire more). Clearly, different contexts will produce different interpretations. But we might well ask why we should make the hearer do all this work when a more explicit utterance would need much less work. The answer is that human beings prefer to exert themselves as little as possible. To achieve their desired results, speakers use just as little language as they need for a relevant interpretation to be achieved by the hearer. More language equals more effort for the speaker, and, in the sense of the phonetic string that must be processed, also for the hearer. The desire to avoid effort wherever possible is why exact repetitions, which produce no extra effect, as well as irrelevant utterances, can both cause communication to fail. They call for processing effort without any payoff, no “contextual effect”, in Sperber and Wilson’s terms.

The driving force of interpretation in RT is the Principle of Relevance:


Optimal Relevance: (a) The ostensive stimulus [i.e. the utterance-SFC] is relevant enough for it to be worth the addressee’s effort to process it.

(b) The ostensive stimulus is the most relevant one compatible with the communicator’s abilities and preferences (: 270).

What this amounts to, is that hearers assume that anything said to them (with the intention to communicate) must be relevant to some degree; they assume they can get to a relevant interpretation without unreasonable effort; and they assume that what they think the speaker assumes will have guided him/her in the construction of the message. So if John says to Mary, Would you like some coffee? and Mary says Coffee keeps me awake, how John interprets Mary’s utterance will depend on both the context of the utterance and what he knows/thinks about Mary’s state of mind. For example, if he knows Mary has a large project to finish before the following day, he is likely to interpret her as having accepted the coffee. If he knows/thinks she is tired and wants to get an early night, he is likely to interpret her as having refused it. In other words, different interpretations depend on the contextual assumptions brought to bear on the calculation. (Note that in addition to the sort of assumptions mentioned
above, a whole range of other factors are involved in the formation of any set of contextual assumptions, including information from ongoing interactions, setting, nonverbal communication, etc. What is important is that features of context external to the hearer can only play a role to the extent that they are mentally represented by the hearer.)

The processing for relevance leads to the reinforcement, creation and contradiction of assumptions, thus changing the cognitive context of the hearer. In the coffee example, if John concludes that Mary does not want coffee, his view of Mary as being tired will have been reinforced, and he will have added the assumption that she does not want coffee. If John only suspected that Mary wanted to go to bed, but he wasn’t sure, and she had said Yes, coffee keeps me awake, then she would have reinforced his assumption. If, on the other hand, she had said No, coffee keeps me awake, his assumption would have been overturned in favour of one that she has work to do. Note that, speakers usually help hearers reach the intended conclusion by particles (such as the “yes” or “no” above) and by hesitations, facial expressions, etc. which accompany the utterance. Notice also that the interpretations I have considered here by no means exhaust the possible interpretations that could be made if the relevant assumptions are brought to bear.

Relevance Theory, thus, treats context as cognitive context: as the contents of the speaker and hearer’s minds, i.e. the set of assumptions that they are each independently capable of bringing to bear on the creation and interpretation of messages. These assumptions are derived from physical setting, inferences from previous utterances, and from the utterance currently being processed. They vary in their strength depending on the reliability of the source.

The RT notion of comprehension, it is important to note, does not require mutual knowledge. It is not necessary that speaker and hearer share the same knowledge in order for communication to be successful. It is only necessary that the cognitive context called up by the hearer, as a result of the utterance expressed by the speaker, overlaps sufficiently with that assumed by the speaker as being necessary for the interpretation, for communication to be successful. This is one reason why communication is risky, why it can fail, but also why conversations between inept speakers (e.g. learners) can actually succeed against the odds. On many occasions a fairly low degree of overlap of contexts may be sufficient for communication to occur, or at least to appear to occur.

Finally, and relatedly, RT does not require any kind of prior or overt agreement between speaker and hearer: no cooperative principle, no agreement to follow maxims. The Principle of Relevance does all that work. The Principle of
Relevance simply says that when a speaker says something, with the clear intention to communicate, the hearer will assume that there is a relevant interpretation to be had, which will deliver something worth hearing with a minimal amount of effort.

3. First language learner behaviour

We can now ask in what sense a theory of pragmatic interpretation can be said to be a theory either of learner behaviour or of the process of language acquisition. I will consider the learner behaviour question first and then, more tentatively, the acquisition question. Whether a Relevance Theoretical account can help us understand learner behaviour at any given stage of language acquisition (first or second) is, at one level, really asking quite simply whether learners (child or adult) are normal communicators. Not surprisingly, the answer will be affirmative. Linguistic pragmatics looks the way it does, RT argues, because it reflects the general functioning of the human brain as a signal processor. At an important level, understanding communication is only a special case of understanding anything. So, it would be surprising if children or second language learning adults did something different from each other or from adult native speakers.²

There is, then, a distinct advantage in using an account of language interpretation intended to account for adult behaviour because it means we can make the initial, strong, empirically-testable assumption that learners are operating with the same system as competent adults, and that when something goes wrong or is differently interpreted by learners, we can look for its causes in the precise calculation made, rather than in the nature of the mechanisms. With this assumption in hand, we can now look at existing knowledge about learner behaviour and see what evidence for a relevance account we can find³.

My first encounter with RT was in about 1993 when I was looking for an account of pragmatics that I could use to explain why tests of children’s knowledge of the binding facts seemed to produce such inconsistent results (Foster-Cohen

² There is a caveat I would add, however, which is that RT assumes a “theory of mind”, and in this respect children are to some extent different from adults, although it is now clear that many of the key aspects of “theory of mind” are in place by three years of age. However, I argue (Foster-Cohen 1999) that the theory of mind is a separate component that provides input to the calculation of relevance, rather than part of relevance itself.

³ My discussion here concerns normal language development, but readers are referred to Leinonen, Letts and Smith (2000) for a discussion of RT and language disorder in children.
1994). It was clear to me that the tests were radically different from each other pragmatically, and it turned out that the predictions RT made about the infelicitousness of these paradigms seemed to correlate well with the children's level of success on the binding tests.

Specifically, RT predicts that certain features of testing paradigms, such as test sentences that are irrelevant given the context, gratuitous repetitions of information, and violations of relevance-driven informational structure, all contribute to poor performance. A comparison of three paradigms and their results showed that the more they violated these constraints, the worse the results they got.

More recently, a survey of the literature of children's pragmatic behaviour has revealed a range of studies that provide interesting support for the theory. I will briefly consider three studies based on completely different theoretical foundations (Axia 1996, Lloyd, Mann and Peers 1988, Babelot and Marcos 1999), and then one from within the traditional Gricean paradigm (Ackerman 1981). In each case I will try to show that an RT account has more explanatory power than the existing analyses.

Axia (1996) studied Italian children's pragmatic behaviour when they were trying to persuade their mothers to buy them things. The children change their methods from simple demands of the Mummy please buy me that type to negotiating for cooperation by making reference to the hearer's needs and interests, as for example in Granny, look at the books. Weren't you looking for one? Within RT we can reinterpret the same observation by saying that with age the child is able to make better assessment of the hearer's needs and preferences (cf. part (b) of Optimal Relevance, above), something that is necessary to negotiate for cooperation.

This is a simple recast of Axia's results, but an RT account can also help us see the developmental progression in terms of how much inferencing children expect the hearer to make, and the sophistication of the "contextual effects" (the "pay-off") to be had from such inferencing. Utterances of the type Granny, look at the books; weren't you looking for one? reflect an ability to invite a hearer to invest in inferencing which takes them beyond the proposition expressed to the most relevant interpretation (i.e. one in which the child is expressing her own desire, rather than commenting on the desire of her grandmother). Incidentally, parents' input language, particularly the use of conditionals to children, also provide ample modelling of this kind of inference invoking, precisely because it allows strong statements to be inferred without the conflict risked by using imperatives (Foster-Cohen and Konrad 1997).

Once Axia's results are seen in this light, we need to know more about the
length of utterances produced at different ages since increased length of utter-
ances increases the effort required and has to be counterbalanced with the con-
textual effects to be derived. It also raises questions about the number and com-
plexity of inferences that have to be drawn in each case by the adult. Clearly this is an avenue for further research.

Lloyd, Mann and Peers (1998) is a study of children’s referential communication skills, using the standard screen set-up with two participants each with an identical array of pictures. The task was one in which children had to identify references to clowns: references that could be to any one of several similar but not identical clowns, where more information must be sought to be sure which clown is being referred to, and so, for instance, clown with buttons is indeterminate between clown with buttons and hat and clown with buttons and no hat. Lloyd, Mann and Peers suggest that while younger children are not able to process the indeterminacy which they actually term “ambiguity”, but see Smith (1985) for clarification of this issue, older children and adults, on the other hand, are able to respond appropriately.

However, the statistical evidence is actually not that they cannot do the task, but that they usually do not respond appropriately: about one third of the time, children of all ages, and adults, pick a clown quickly without asking further ques-
tions. Everyone jumps to the first relevant interpretation in the context, a sig-
nificant proportion of the time. Thus, young children can and do respond ap-
propriately to referential tasks, a fact that has been confirmed by Beduizenhout and Sroda (1998), and has been shown to be disrupted in language disordered children by Leinonen and Letts (1997), from an explicitly RT perspective.

Babelot and Marcos (1999) studied the responses of French children (aged 1;7-2;5) to directives in a play situation. A variety of linguistically different kinds of directives (imperatives, need statements, hints, etc.) were addressed to the children in contexts which were either relevant to the directive (e.g. the adult said, I really need the horse when they were playing with farm animals), or irre-
levant to the directive (e.g. the adult said, I really need the horse) when they were playing a different game and the farm animals were lying elsewhere in the room.

What Babelot and Marcos found was that when the directive occurred in a relevant context, it was almost always interpreted as a request for action, i.e. the child gave the named object to the adult, whether or not that was what the di-
rective had demanded. When the directive was not relevant to the context, how-
ever, the children showed they were quite capable of paying attention to the lin-
guistic form of the utterance, and if the directive was of the kind Where’s the horse?, of indicating where it was, even though in the so-called “relevant” con-
text, *Where's the horse?* tended to result in the child giving the horse to the adult. This suggests to me that children naturally jump to the conclusion that an action is required if the context gives them that option. The particular action is the most relevant interpretation given the context.

Importantly, this experiment also shows that even very young children can, and do, process the precise linguistic structure of the utterances addressed to them, at least if they are forced to by the fact that the first relevant interpretation is not possible and they have to put in more effort to derive relevant (contextual) effects. This is important because it supports RT’s contention that inferencing operates on the output of the linguistic system, and that speakers of all ages are able to construct, at least partial, logical forms from their linguistic processing before pragmatic processing takes over.

Brian Ackerman, in his (1981) study of children’s reactions to maxim conforming and maxim violating utterances, within the standard Gricean perspective, expressed surprise that kindergartners do not discriminate between “conforming” utterances and so-called “rule-violating” utterances, and, particularly, that some children resort to a perseverative strategy in which they give the same answers all the time (assigning all utterances to the same speaker). However, within RT, which explicitly argues against the “rule-violating” idea and in favour of a framework in which all utterances are processed as best they can be for relevance, the behaviour of Ackerman’s children becomes unsurprising.

Where Grice would have predicted a difference between “conforming” and “rule violating” utterances, because children would be argued to know the maxims and obey the cooperative contract, RT would predict no such distinction. Rather, utterances vary in the effort-effect equation in ways that cross-cut Grice’s and, therefore, Ackerman’s categories, and, thus, for a range of reasons, utterances may overload children’s systems, leading to a timing out of the effort-driven processing mechanism, which causes them to give up early, before they have actually derived an interpretation. Under those conditions, it is not surprising that some children resort to a “perseverative idiosyncratic response strategy” in which they give the same answer to everything, even to ones that Grice would have predicted to be easy to interpret.

Through the RT lens, we can suggest that older children, who are more successful at Ackerman’s task of responding to different categories of utterance, recognise the extra work that “non-contingent”, “redundant” and apparently “irrelevant” utterances take to get a reading, and are able to apply the extra effort to process them. But we need not say that these older children are able to recognise the rule violations implied within standard Gricean theory for these
kinds of utterances. Rather, they are operating under the assumption that what
the interlocutor says is worthy of processing for its relevance, and it’s just a
question of how quickly they can come to a possible interpretation consistent
with the context, and of whether they run out of processing energy before they
do so. Ackerman’s main results, as well as the things he finds puzzling fall out
from a Relevance view of his paradigm.

4. Second language learner behaviour

What now about second language acquisition? As with children, learners do
not necessarily start their pragmatic processing with the same logical forms as
adult native speakers, and one particularly interesting area of research concerns
what happens to utterance interpretation when a second language learner mis-
parse or only partially parses an utterance. The learner whose interlanguage
does not yet contain some of the lexical items to be decoded in an utterance, or
whose syntactic parsing produces a non-native analysis, will be feeding the prag-
matic part of the comprehension process with a logical form that may be seri-
ously faulty.

In his discussion of learners’ approach to input and what they actually
process as intake, Van Patten (1996) suggests that learners have specific prefer-
ences for content words and will only process non-meaningful grammatical
items if the processing load is low enough that they can assign attention to them,
and thus detect them. If this is correct, then we can generate specific hypothe-
ses about the nature of the logical forms generated/decoded by language learn-
er, and try to track the consequent impoverishment through later levels of in-
terpretation. This is something that has been attempted, within another frame-
work, by Stephen Ross.

Ross (1997) is an analysis of learners’ behaviour on a listening test, in which
learners had to pick out icons from an array on the basis of sentences they
heard. For example, the test tape plays Before the development of air travel, this
was the only method of transport available for people wishing to make journeys
overseas. The right icon was a boat, but learners often picked the aeroplane icon
because they made an incorrect inference from the words air and overseas, in the
absence of being able to process the rest of the language. Here the learners
failed to derive the proposition expressed because the rest of the language could
not be processed.

4 A range of connections between RT and SLA are discussed in Foster-Cohen (2000), which
overlaps in some respects with the current discussion.
Notably, they sometimes got the right icon anyway, even if they were not, in fact, processing correctly. For example, a listener correctly chose an artist's palette having heard *An artist might use this for painting*, even though s/he admitted in the protocol that s/he only clearly heard the word *artist* and guessed. Both these reactions are predictable from RT, as they show the effect of inaccurate or incomplete logical forms fed to the pragmatic interpretation system. This is clearly an area where more tight experimental work is called for.

It is important to emphasise that because relevance-based processing cannot be turned off, faulty logical forms are still processed pragmatically, but the inferential processes may produce results that differ significantly from those which a native speaker might derive. On the other hand, inferential processes might also cover up the deficiencies in the original logical form. For example, the ability to access an appropriate context may allow a learner to fill in gaps appropriately, and override faulty syntax in the derivation of the proposition expressed, as in the example of the correct answer to the *artist* question.

All this raises the question of how much effort second language learners expend in relation to native speakers. In a situation such as that just described, we could say learners are putting in less effort because they have less linguistic material to work with. On the other hand, they could be seen as putting in more effort because they have to fill in more on the basis of inferencing work. In fact, there seems to be evidence for both more effort and minimal effort by non-native speakers.

Research showing that comprehension is improved for learners when there is repetition and redundancy (Pica, Young and Doughty 1987) suggests learners are willing to put in more effort, despite the fact that repetition and redundancy are usually indicative of additional processing work. If learners perceive that more contextual effects can be derived (as a result of fuller comprehension), then redundant information is actually not redundant for those learners, and therefore not irrelevant.

Other evidence of the "extra effort" put in by non-native speakers is the often noted overcoding behaviour of lower level non-native speakers, who are perceived to be overly verbose or to be "waffling" (Edmonson and House 1991). Such behaviour ought to be a clear violation of the expectations of RT. One can only assume, therefore, that there are indeed extra effects to be had for the extra effort that such learners put in. Perhaps learners are driven by perceptions or fears of miscommunication, which can only be overcome by such efforts, even if they result in extra effort for the native-speaker hearer. Perhaps recognising that they do not have the lexical expressions and formulae, which al-
low for lower levels of explicitness, they have little choice but to construct overly full-blown alternatives. Cheng and Warren (1999) have suggested, in fact, that one of the noticeable features of even quite advanced learners is that they don't know how to leave inferencing up to the hearer, producing more elaborate utterances with fewer fillers such as some, it and one than is normal for native speakers.

On the other side of the coin, Ying (1996), who has looked at minimal attachment of prepositional phrases, has argued that the principle of least effort is working well in learners, leading them to make certain faulty syntactic attachments in preference over others. Evidence of the minimisation of effort in second language studies can also be seen in non-native speaker interactions when comprehension trouble spots are allowed to persist until they become relevant.

Thus, Wagner and Firth (1997), for example, analyse telephone conversations between people working in European companies who must communicate on technical matters in their foreign languages where they often do not have the appropriate expressions needed. Instead of putting in the effort at the first mention of such expressions to make sure it is clear to both parties what is meant, the expression is often left for some time, repeated but still unclear in the conversation until such time as it becomes evident that the interaction is going to stall unless, and until, the particular word is sorted out. This suggests that if sufficient comprehension can be achieved for the interaction to move forward, no extra effort needs to be put in. When, and if, it becomes clear that this will not do, it becomes relevant to put in extra effort for the extra contextual effects that are likely to be gained.

The key question is obviously to try to work out when minimal effort wins out over extra effort, and what differentiates learners' responses in the two cases: what is it about some situations, some learners, some tasks, and some types of language that lead to extra effort or to minimisation of effort?

5. Relevance Theory and acquisition

Finally, let me briefly address the question of whether RT can, in any sense, be seen as a theory of language acquisition. Because RT contains within it an inference generator, designed to update the hearer's view of the world by adding, supporting or cancelling assumptions, one could argue that to the extent that assumptions about language are capable of being added to the knowledge store in this way, the inference generator contributes to language acquisition. However, if it does, it would seem to do so only in a limited way, and only in terms of de-
declarative, metalinguistic, knowledge of language. To that extent it might be that
RT could contribute to accounts by those for whom declarative knowledge is an
important part of the story: Towell and Hawkins (1994), for second language ac-
quisation, for example.

We might then ask questions about how such assumptions are derived and
the strength of the evidence for them. Assuming that learners come to believe
certain facts about a language on the basis of a variety of different kinds of
information (being told the rule by a teacher, reading it in a book, having a lan-
guage fact modelled by a teacher or another adult, having it modelled by an-
other student, observing it in another context from analysis of native speaker
language, etc.), the prediction would be that the fixation of belief about these
particular facts is affected by the input to the device, the reliability of the as-
sumption, and that, in turn, affects any new assumptions derived by using those
assumptions as premises. I would, however, not want to suggest that adopting a
role for inference generation in second language learning means that all second
language needs to be acquired this way. It is entirely compatible with RT to say
that there is knowledge of language that operates entirely within the linguistic
modules and might be Universal Grammar driven, and other knowledge which
is acquired by the actions of the central processor.

Returning to the idea that the pragmatic processor may be involved in the
acquisition of at least declarative knowledge, I would like to suggest that a pe-
culiarly RT idea might be of use here. I have already discussed how processing
messages leads to the addition of assumptions to the cognitive context, but
Sperber and Wilson introduce a notion of “manifestness” which is weaker than
(and thus includes) the stronger notions of what is “known” or “assumed”.
What is “known” is mentally represented with a guarantee of truth; what is “as-
sumed” is mentally represented but with no guarantee of truth; what is “man-
ifest”, however, is what one is capable of inferring or of perceiving. So, it is man-
ifest to you that Julius Caesar and Noam Chomsky never played billiards to-
gether (Sperber and Wilson’s example), even though you have never thought
about it before. A car passing in the street may also be manifest to you, even
though you are not actually paying attention to it. This notion of manifestness
is important to Sperber and Wilson because with it they can avoid the require-
ment of mutual knowledge, and rely instead on a requirement of shared cogni-
tive environments, which are the set of all facts that are “manifest” to both
speaker and hearer.

I think this notion of manifestness has potential in second language research
for contributing to the debate over “noticing” as a key to second language
acquisition (Schmidt 1990, 1994, Van Patten 1996, Doughty and Williams 1998). There are various proposals for how attention to and detection of forms are crucial for learners to make acquisitional use of input. The idea that information may be less than known (i.e. manifest) but yet capable of being re-examined, reinforced, and strengthened into a representation may help to explain how learners may be able to raise to the status of knowledge information about language, which, until it was relevant to do so, remained in some sense inchoate. In other words, for the second language learner, linguistic information may be there as part of the manifest context, and may be brought into the comprehension process as and when necessary/possible.

This is an idea that bears some further development, I think. It would imply that declarative knowledge of language could go through stages of acquisition from manifest assumption to genuine knowledge. It would allow “knowledge” to be restored to its rightful place as a representation with a guarantee of truth. And it would allow us to raise the question of where, along the weakly to strongly manifest continuum, “noticing” takes place, with the possibility that, as Schmidt (1990) originally defined it, something is noticed when it is far enough along the continuum to coincide with conscious awareness. One would also predict that the trigger for whether something is conscious or not is relevance.

The problem with learning language from the input is that it calls for dual processing of incoming utterances of the meaning, which cannot be turned off, and of the form, which is an effort. The work of Lyster (1998) on recasts and repetition in L2 classroom discourse suggests the problems this creates, because it shows that the comprehension mechanism, whose goal is to reach a relevant interpretation as quickly and easily as possible, gets in the way of the kind of focus on form that we are talking about. Lyster argues that students do not necessarily see negative feedback for what it is because they are processing for meaning and not for metalinguistic feedback. A point that Van Patten (1996) also makes in his discussions of input processing. Susanne Carroll (1995), using an explicitly RT approach, argues that the perception of negative evidence, for what it is, runs completely counter to relevance-driven processing. In some contexts, of course, one can increase the likelihood of negative feedback being so perceived; this is precisely what focus on form teaching is intended to do. I would say that focus on form can be seen from a Relevance Theory perspective as a question of, as it were, persuading the learner that extra effects/payoffs are to be gained from the additional effort of paying attention to the form as well as the meaning, but it will not happen unless it is relevant to do so.
6. Conclusion

When one takes a new theoretical perspective, and applies it to new territory, it can do three kinds of work. Firstly, and at its weakest, it can simply restate what has been said before in new terms. This may be useful, but only if the new paradigm has already been accepted for stronger reasons. Secondly, it can recast existing knowledge so that observations realign themselves along different coordinates. Under these conditions, new questions can be posed of existing data or new data sought in order to clarify the new view. Thirdly, and finally, a new perspective can open up genuinely new questions, which uncover new knowledge and advance our understanding. I hope to have gone some way, at least, in showing the RT has the potential for all three stages of validity.

References


Foster-Cohen, Susan and Erika Konrad (1997). “If you’d like to burn your mouth, feel free! A relevance-theoretic account of conditionals used to children”. In Marjolein


