

To appear in
R. MARQUEZ-REITER and M.E. PLACENCIA (eds.) (2004)
Current Trends in the Pragmatics of Spanish
Amsterdam, John Benjamins

This is a pre-print version

15

Norms and Principles. Putting Social and Cognitive Pragmatics Together*

Victoria Escandell-Vidal

1. Introduction

We all probably agree that the use of language cannot be characterised in terms of grammar only, and therefore a complementary theory is needed to account for linguistic performance: this theory we call *pragmatics*. However, a quick look at current bibliography is enough to realise that under the same label a wide variety of approaches coexist, each with its own framework, basic tools and interests.

When we say that pragmatics deals with everything that determines the use of language, we are tacitly admitting that actual uses should meet other ‘external’ requirements — requirements different from mere grammaticality. If so, a number of related questions immediately arise: What is the exact nature of these requirements? What is their theoretical status? How are they to be described? What place should they occupy in a pragmatic theory? Answers to these questions have been offered in quite different terms, and notions such as *norms*, *conditions*, *rules*, *postulates*, *maxims*, *strategies*, *laws*, *heuristics*, *principles* have been invoked to account for the underlying regularities that seem to govern linguistic use.

The purpose of this paper is to show that some of the various perspectives mentioned above are not necessarily incompatible, and that a design for pragmatic theory can be suggested that encompasses different approaches in a single, unified model. After briefly commenting on the sources of the current diversity, I will present the bases for a modular approach to pragmatics. Different components will be identified, their specificity deriving from the sort of contribution they make to the interpretive process. The distinction between *representation* and *computation* — a distinction borrowed from Artificial Intelligence — will prove crucial.

* This is a revised version of a plenary paper given at the First UK Symposium on Hispanic Linguistics (School of Language, Law and International Studies, University of Surrey, April 2002). I am very grateful to the organisers (Rosina Márquez-Reiter and María Elena Placencia) and the audience for stimulating debate, and also to Aoife Ahern, Manuel Leonetti, Begoña Vicente and an anonymous reviewer for helpful comments. This research has been partially supported by a grant from the Spanish Department of Education and Culture, DGICYT PB98-0707, “Gramática e interpretación en la teoría de la relevancia”.

2. Two ways of doing pragmatics

Though it is true that a large number of different perspectives exist side by side in the domain of pragmatics, two main general trends can be identified: socio-cultural approaches and cognitive approaches. These represent the two poles that attract most of the current research in the field. Of course, they are not totally homogeneous models; rather, each group draws together several views and proposals, sometimes with bitter discrepancies between them. But despite these differences, they share defining features, such as research objectives and basic tools, which explains why they all bear some family resemblance.

On the one hand, there is a branch of pragmatics concerned with 'external' factors, i.e. with those aspects of the selection and interpretation of linguistic form that happen to be determined by social and cultural factors. This includes politeness studies, research on conversational styles, rhetoric, discourse genres and register. On the other hand, we find pragmatics dealing with 'internal' factors, such as the cognitive bases for linguistic performance, the inferential processes leading to the final interpretation, or the interface relationship between grammar and pragmatics.

Now, I want to claim that the sort of explanations elaborated by socio-cultural pragmatics can be conceived of as *norms*, whilst those envisaged by cognitive pragmatics are better understood as *principles*.

2.1 Socio-cultural pragmatics

Let's consider socio-cultural pragmatics first. Its main task is to identify and characterise the *norms* that underlie the use of language of a given social group.

But what are *norms*? The term can denote two different, though related, notions: on the one hand, it can refer to common practice — normal, usual or habitual behaviour; on the other hand, it can refer to a set of rules and regulations, of socially approved patterns of behaviour. One is the idea of *norm* as a *custom of use*; the other is the notion of *norm* as a *convention of use*. This apparently neat distinction, however, is not always easy to actually apply: often, common practice derives from socially accepted conventions, so no clear borderline can be easily drawn between them; and, even if there is no identifiable, previously established convention, behaviour that strays from the normal path tends to be rejected, so the need for social affiliation also reinforces normal behaviour. Yet, the distinction usually has some psychological reality: when norms are conventions, the individual usually perceives them as something external, socially imposed, that requires formal and explicit instruction; when norms are customary practices, they represent internalised, unconscious patterns that the individual follows without even noticing that he is complying with an unwritten model. The sense in which I am interested for my current purposes is that of widespread tendencies of behaviour, no matter whether they stem from explicit or implicit social conventions, or whether they represent conscious or unconscious patterns of behaviour.¹

The task of identifying norms is achieved by analysing the behaviour of a population sample in a series of situations with identifiable explanatory social variables, and drawing generalisations from the distribution of frequencies obtained. Think, for instance, of politeness studies: they try to establish the norms underlying linguistic behaviour, as shown by the speakers' preferences in a number of different situations

¹ See section 5.2 for further discussion.

(requests, apologies, refusals) In this way, an overall picture of a community's customary practice can be sketched that makes it possible to compare and contrast it with that of different cultural groups.² Other culturally determined aspects of linguistic use can also benefit from this kind of approach, such as the preferences in discourse structure and organisation³ or routines in turn-taking and conversational styles.⁴

Generalisations based on the qualitative and quantitative analysis of the data, thus, contribute to the understanding of language use by bringing norms to light, but they cannot fully predict future action.

2.2 Cognitive pragmatics

Cognitive approaches represent a different way of doing pragmatics. They aim at identifying *principles*. A *principle* is the formalised expression of the behaviour of a system. It is not a statistical generalisation, but a causal, mechanical explanation, a general law with the universal status of the laws of Physics.

Models on the cognitive side have put a lot of effort into discovering the principles that govern different aspects of use and understanding of language. Though probably not everyone would agree, I would include Grice at least as a precursor of cognitive approaches. His Cooperative principle, "Make your 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 you are engaged" (Grice, 1989: 26), was intended as a general law on human rationality, and exhibits the rigidity that distinguishes principles from other sorts of regularities, such as maxims: it is precisely because its validity should be maintained that maxims are flexible and defeasible.⁵ Grice himself emphasises: "I would like to be able to think of the standard type of conversational practice not merely as something that all or most do *in fact* but as something that is *reasonable* for us to follow, that we *should not* abandon" (Grice, 1989: 29; his italics).

The same can be said about Sperber and Wilson's two principles of relevance (Sperber and Wilson, 1995: 260): the first, or cognitive, principle ("Human cognition tends to be geared to the maximisation of relevance") is a generalisation about the way in which our minds work; the second, or communicative, principle ("Every act of ostensive communication communicates a presumption of its own optimal relevance") has to do with the special way in which intentional communicative behaviour is interpreted (i.e. by assuming that its processing will not impose on the hearer any gratuitous effort, one not properly balanced by an equivalent amount of cognitive effects). Unlike Grice's, Sperber and Wilson's principles are always in force, so implicatures can never arise as a consequence of flouting them.

Levinson's (2000) Q-, M- and I-principles are a different attempt in a similar vein: they are intended to establish general conditions both on the speaker's choice and on the recipient's interpretation.⁶ They are always in effect, but since they occupy

² Blum-Kulka, House and Kasper (1989); Oleksy (ed.) (1989); Kerbrat-Orecchioni (1990-(1994); Wierzbicka (1991); Watts, Ide and Ehlich (eds.) (1992); Placencia (1992); Sifianou, (1992) Gass and Neu (eds.) (1993); Haverkate (1994); Scollon and Scollon (1995); le Pair, (1996) Gass (1999); Fukushima (2000); Márquez-Reiter (2000); Placencia y Bravo (2002).

³ Berman and Slobin (eds) (1994); Connor (1996); Vázquez (2001).

⁴ Mey (1993); Gallardo Paúls (1996); Tuson Valls (1997).

⁵ A sort of hierarchical relationship is also observable among maxims, in that one maxim can be flouted to comply with a higher one. This happens, for instance, when the maxim of quantity is violated in order to abide by the maxim of quality.

⁶ "The Q-principle:

different places in a hierarchy, the operation of one of them can make others appear as no longer valid. By focusing on the general conditions that govern interpretation, Levinson's purpose is to restrict the role of individual, once-off inference, and establish the basis for default interpretations.

Cognition-oriented theories⁷, thus, try to find the biological or cognitive bases that underlie communicative behaviour. By virtue of their very status, principles go beyond observed phenomena, and can make causal predictions about unattested facts. Their predictive power is a must for any theory. However, they are difficult to discover and formulate, partly because of their degree of abstraction.

3. A single theory?

We have briefly considered the two main ways of doing pragmatics. Now, one could wonder why it should be so. If the various aspects of linguistic use could be reduced to a single set either of norms or of principles, a more economical theory could replace the existing division. This move would seem clearly desirable, provided that it did not result in any substantial loss in descriptive or explanatory power. But, is this possible? If you are expecting an affirmative answer, I have both good and bad news.

The bad news is that the answer must be negative: reducing one of the competing approaches to the terms of the other does not seem an adequate solution. The attempts to explain typical socio-cultural phenomena by resorting to universal principles and inferential aspects in terms of conventions both run into serious difficulties.

Politeness studies represent the best-known case.⁸ Some proposals have been put forward claiming that politeness can be accounted for in terms of universal principles, namely those responsible for implicatures and indirectness: politeness arises as the effect of an implicature, so the more indirect a form is, the more polite.⁹ Cross-cultural studies, however, soon revealed that proposals invoking rational principles could not explain the existing cultural diversity.¹⁰ For instance, the schema *Could you please ___ ?* is a polite form of request for many cultures, the invoked explanation being that asking

Speaker's maxim: Do not provide a statement that is informationally weaker than your knowledge of the world allows, unless providing an informationally stronger statement would contravene the I-principle. Specifically, select the informationally strongest paradigmatic alternate that is consistent with the facts.

Recipient's corollary: Take it that the speaker made the strongest consistent with what he knows (...)" (Levinson, 2000: 76)

"The M-principle:

Speaker's maxim: Indicate an abnormal, non-stereotypical situation by using marked expressions that contrast with those you would use to describe the corresponding normal, stereotypical situation.

Recipient's corollary: What is said in an abnormal way indicates an abnormal situation, or marked messages indicate marked situations (...)" (Levinson, 2000: 136)

"The I-principle:

Speaker's maxim: the maxim of Minimization. "Say as little as necessary"; that is, produce the minimal linguistic information sufficient to achieve your communicational ends (bearing Q in mind)

Recipient's corollary: the Enrichment rule. Amplify the informational content of the speaker's utterance, by finding the most *specific* interpretation, up to what you judge to be the speaker's m-intended point, unless the speaker has broken the maxim of Minimization by using a marked or prolix expression. (...) (Levinson, 2000: 114).

⁷ See also Horn (1984), Blakemore (1992), Carston (1999), (2002a), (2002b), among many others.

⁸ I cannot present the whole line of argumentation here. A more detailed discussion can be found in Escandell-Vidal (1995 and 1996).

⁹ Searle (1969 and 1975); Lakoff (1973); Brown and Levinson (1987).

¹⁰ Matsumoto (1988); Watts, Ide and Ehlich (eds.) (1992), *Multilingua* 12 (1993); Scollon and Scollon (1995); and most of the references in footnote 1.

whether the interlocutor can perform a certain action is a way of requesting without an overt imposition. However, it does not work that way in many other cultures. According to Wierzbicka (1991), it would be interpreted as a standard question in some Slavic cultures; and even worse, it would be perceived as very impolite by Thais, since it would imply that you doubt their ability to do things, as reported in Kerbrat-Orecchioni (1992).

Attempts to characterise linguistic behaviour in terms of conventions of usage, leaving no room for any sort of universal processes, would count as a move in the opposite direction. I am not sure that anyone would totally endorse such a radical position.

To sum up, a reduction of all aspects of language use to a set of norms or a set of principles does not produce the expected result; on the contrary, it gives rise to important descriptive inadequacies.

4. A way to integration

The fact that a unified theory with a single kind of basic notion cannot escape from major drawbacks does not mean, however, that any attempt in this direction should necessarily be doomed to failure. The good news is that there is a way of building up a single theory and benefiting from its advantages while avoiding the disadvantages mentioned above. The secret is to abandon reductionism.

The idea I want to explore is the following. The two main ways of doing pragmatics show very strong preferences in many respects: each has a particular perspective on its subject (language use as a social product *versus* language use as a cognitive ability); each analyses a particular set of phenomena ('external' or social *versus* 'internal' or psychological); each puts up its explanations in different terms (either by identifying norms, or by postulating principles). Now, to the extent that these two approaches do not come into conflict with each other — or, if they should do, to the extent that any clash of interests could be solved — what such a disparity suggests is that they actually do not represent two ways of doing pragmatics, but rather two different research sub-fields.

If this is correct, social and cognitive approaches are not two conflicting views, but two different components of a larger system.¹¹ Incorporating them into a single theory does not imply reducing one to the other, but integrating both of them in a wider architecture as autonomous, though interacting, devices.

I want to make it very clear that my proposal is not aimed at suggesting an ecumenical solution, or a form of tolerant eclecticism where various models combine freely, regardless of the consequences for the overall theory; quite on the contrary, what I would like to suggest is that there is a way of integrating social and cognitive perspectives into a single, coherent framework. The remainder of this paper is devoted to sketching how this can be done.

4.1 A modular approach

Anyone can see that the factors determining the way in which language is used are many and diverse. A fruitful strategy for analysing a very complex product is conceiving of it as the result not of a single, large and complex system, but rather of a

¹¹ This view has been defended in Escandell-Vidal (1996 and 1998).

set of smaller, interacting subsystems.

The approach I will suggest is a modular one. The idea of a modular pragmatic theory is not new. The best-known proposal is that of Kasher (1991), in which a distinction is suggested between different components, each devoted to a specific class of phenomena: basic speech acts (*core-pragmatics*), ritualised speech acts (*amplified core-pragmatics*), rational action and implicatures (*central pragmatics*), turn-taking (*talk-in-interaction pragmatics*), deixis (*interface pragmatics*), and non-literal uses (*extended pragmatics*).¹²

There are many differences between Kasher's proposal and mine, but the most important one has to do with the way in which subsystems are distinguished from each other. In his model, a component deals with a particular sort of phenomena: implicatures, conventionalised speech acts, non-literal uses. At first sight, this might seem reasonable. However, it is difficult to imagine how such a model would work, since none of those phenomena can be identified *before* processing. Consider, for instance, the difference between a literal and a non-literal interpretation: it does not lie in the features of the input, but rather appears as the output of a process of interpretation taking into account information from many other sources as well. Since no *a priori* distinction can be established between them, the system would be unable to decide to which component it should assign the processing of the stimulus.

What I want to suggest is that the distinction between components should be established in terms of domains.¹³ The rationale for a domain is that it should be dedicated to the processing of a particular class of data. It can deal with information of a specific kind and format only; non-complying data are simply filtered out. Each domain processes only the sort of information to which it is sensitive: for instance, the visual module is sensitive to shape, colour and light patterns, and other simultaneous, but non-complying stimuli are simply ignored. Thus, access to a particular component is restricted by conditions on the input. This is precisely the approach that I would like to argue for: the criterion for identifying components must be established in the form of relevant prerequisites (not *ex post facto* distinctions). Though I will not pursue the matter in detail, the way in which I envisage the differences between the components will become clear later.

4.2 Different abilities, different systems

If different components are postulated, there must be a criterion for deciding what counts as a component. Even if we put aside all the systems and processes that make up the articulatory and auditory bases of speech, verbal communication still involves different abilities. Knowledge of grammar, of course, must play a major part, but is not enough to account for the way in which we use language — this was our starting point.

Cognitive-oriented approaches have overtly stressed the role of inference in natural language understanding: inference bridges the (necessary) existing gap between encoded meaning and utterance interpretation in context. Inference is a general ability that makes it possible to obtain new pieces of information by combining existing ones, following patterns of reasoning.

The role of inference in current pragmatics is such that some cognitive theories even equate pragmatics with inference. Relevance Theory sets the boundary line between semantics and pragmatics according to the sort of process involved: decoding processes

¹² See also Pléh (2000) for a different proposal.

¹³ See, among others, the papers in Hirschfeld and Gelman (eds.) (1994).

are a matter of semantics; inference processes belong to pragmatics.¹⁴ Recent developments have even taken one step further, and restrict the domain of pragmatics to accounting for the inferential system specialised in human communicative behaviour — a system different from other similar common-purpose faculties, such as ‘theory of mind’.¹⁵ Be it a general ability or a communication-specific one, the inferential faculty is specialised enough as to be a separate, autonomous system.

There is, on the other hand, some evidence that a social ability should also be distinguished. It guides our behaviour in social interaction and makes it possible for us to adapt to the situation and to the audience’s social profile. According to Jackendoff (1992), its main task is that of developing a coherent picture of the self in society. Its specificity comes, among other things, from the fact that the formation of social concepts does not follow the usual steps for ordinary perception and categorisation, since it invokes elements and relationships with no direct perceptual motivation.

Social and grammatical faculties develop as the result of a process by which individuals acquire the pattern(s) of their community. The inferential faculty, on the other hand, is universal in the sense that it is not dependent on cultural habits. Its maturation produces similar results in all humans, regardless of their native language and culture.

So far we have considered three abilities that are a part of the systems involved in human communication. Even in a superficial characterisation, such as the one I have just offered, they show clear defining features. The natural hypothesis is that each should have its own set of properties and be subject to different sorts of regularities.

4.3 Mechanisms and representations

How may we account for the difference between culture-dependent (social and grammatical) and culture-independent (inferential) faculties?

Let us consider inference first. Inference in language understanding takes pieces of information obtained from linguistic decoding as their input and combines them with other contextual assumptions to yield new pieces of information,¹⁶ but the system itself does not contribute any information of its own. What inference does is to process representations: therefore, it is a *processing device*, i.e. a computational mechanism that operates on representations taken from different sources. Inference is an ability shared by all mankind, and therefore a good candidate for a universal system.

As for the social ability, one could be inclined to think that it can never be a universal system, given that it is culture-dependent. But, on the other hand, there is evidence that a specific system underlies the formation of social categories, which must be common to all humans. Thus, this ability exhibits a dual nature, with both a universal and a culture-specific side.

An explanation can be suggested along the following lines: when speaking of

¹⁴ Carston (1999).

¹⁵ Sperber and Wilson (2002); Carston (2002).

¹⁶ Consider the following dialogue:

A:— No tendría que ir hoy: habrá un atasco tremendo...

He shouldn't go today: there will be a tremendous traffic jam...

B:— No te preocupes por eso: Juan vive en Majadahonda.

Don't worry: he lives in Majadahonda

Apparently, B's answer is totally irrelevant... unless a piece of information is added to the context, namely the assumption that people who live in Majadahonda (a dormitory town near Madrid) are used to traffic jams, so Juan would not be particularly upset. The selection and addition of exactly this content depends on inference and, in the present case, world knowledge.

social ability, we are in fact putting two different things under the same label. The social ability is, on the one hand, a universal device for social categorisation designed to produce a set of representations on social relations and interaction; the system is 'tuned' to the culture settings, but the device itself is universal. On the other hand, the social ability can also be understood as the set of representations generated by the system, their content being determined by the culture-specific input data.

If this view is correct, we can easily account for the difference between universality and culture-specificity in terms of a well-known distinction taken from Artificial Intelligence: the one between *representation* and *computation*, that is, between *data* and *operations* on data. Mechanisms are, thus, universal, while internalised representations are culture-specific.

At this point *norms* and *principles* take the leading role again. From the perspective I am adopting here, *norms* and *principles* express two different sorts of regularities: *norms* are generalisations on representations, i.e. on data; *principles* are generalisations on computational mechanisms, i.e. on operations. Thus, they both are needed in a system in which computational devices and representations are the basic components.

5. Overview of an integrated model

The distinction between *computation* and *representation* provides a way to account for the different status and the different contribution of each component. A unified pragmatic theory must accommodate processing devices and representations and give norms and principles their proper place. In the following sections I will try to sketch a possible way to do this. My proposal will be necessarily speculative — and probably rather unorthodox. I do not aim at presenting a fully elaborated model, but rather at identifying the desiderata of each component and suggesting a particular articulation between them.

5.1 The inference system

A theory of inference is a theory of a universal device; hence its main task is that of identifying the general principles that govern inferential processes. Three basic issues must be addressed: first, what are the input criteria, i.e. the requirements imposed on the information that enters the system; second, what are the operations performed by the mechanism; and finally, what are the output conditions, i.e. what requirements must the processed entities meet before exiting the system. No existing theory has addressed these issues in these particular terms, but still some partial answers can be borrowed from those on offer.

Various proposals about the nature of inferential processes have been put forward. Sperber and Wilson (1990) committed themselves to a view of spontaneous inference as a non-demonstrative process including deductive devices, though not the whole range of available deductive patterns are used. Levinson (2000: § 1.6) supports the idea that non-deductive (i.e. non-monotonous, non truth-preserving patterns, such as induction, abduction, practical reasoning, default logics, etc.) should be largely preferred. It is perfectly plausible that different situations trigger different sorts of strategies. Much work is still needed before one could take sides on this matter. In any case, identifying and restricting the main reasoning patterns should be a major task for a theory of pragmatic inference.

However, this is not enough. Inference is a very powerful, recursive process, so any conclusion obtained can be the input for a new cycle of processing. If this is correct, the question is: when does the mechanism stop? Is there any principle restricting its operation? Relevance Theory can offer some tools for a suitable answer. Consider the *relevance-theoretic comprehension procedure* (previously known as the *Criterion of consistence with the principle of relevance*), which reads as follows:

- a) Follow a path of least effort in computing cognitive effects. In particular, test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.
- b) Stop when your expectations of relevance are satisfied.

(Sperber and Wilson, 2002: 18)

The generalisation under b) can be conceived of as a principle imposing restrictions on the recursive operation of inference. Relevance is a balance between positive cognitive effects and processing effort.¹⁷ Thus, clause b) states what I would call the *output condition*: when the expected level of relevance is reached the system stops.¹⁸ If the expected level of relevance is obtained, the system delivers an interpretation; if not, then the process is set in motion again and new assumptions are added to it, until the expectations are fulfilled.

So far we have considered the issue of the operations performed by the system and the output conditions. What about the input conditions? Linguistic and extra-linguistic, contextual pieces of information are combined to produce an interpretation; however, not just any piece of non-linguistic knowledge is used. How do we select them? A theory of inference must set the requirements to be met by the pieces of information that enter the system; in other words, what we need is a restrictive theory of *context*.

The *relevance-theoretic comprehension procedure* can again provide an answer: *accessibility*. From the whole set of internalised assumptions of an individual only a few are accessible at a given time: those stored in working memory. A piece of information can be made accessible, or more accessible, through linguistic decoding and situational activation; put in other words, accessibility is a function of linguistic and situational information over working memory. Only more salient, activated assumptions are allowed into the inferential system. Thus, accessibility involves both a restriction of the search space and a hierarchical organisation of the available information.

Obviously, the principles governing reasoning patterns, and input and output conditions must be further detailed and elaborated, but the articulation I have suggested can serve as a starting point. The working of the mechanism can be schematically represented as in figure 1.

¹⁷ Sperber and Wilson (1986/1995 and 2002).

¹⁸ Incidentally, it is not difficult to restate the results of Levinson's (2000) heuristics in these terms: for instance, a prolix, marked or unusual expression requires a greater processing effort, which has to be rewarded with extra effects, hence the search for an unusual reading.

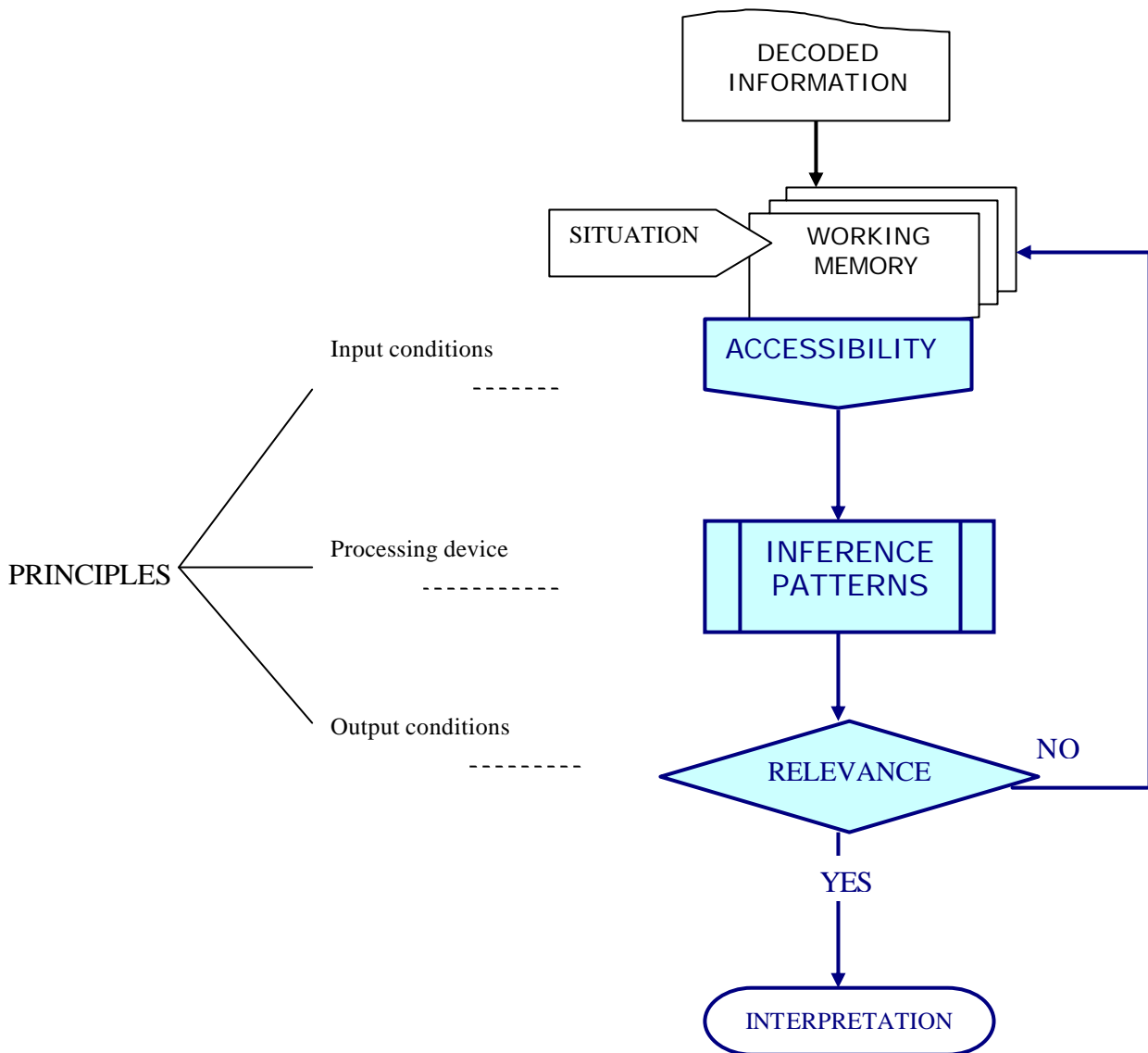


Figure 1: The Inferential System

5.2. The social system

What we usually call *culture* is a collection of ways of thinking and behaving that members of a group learn as a result of the process of socialisation and that determine their beliefs and behaviour. Once acquired, they tend to remain stable and are hardly modified, let alone removed, both at the individual and the social level.¹⁹

How do cultural representations spread? Sperber (1994) provides an account in terms of *epidemiology of representations*.²⁰ Though adequate for a number of cases, this

¹⁹ Strauss and Quinn (1997); Janney and Arndt (1992); Jackendoff (1992); Kasper (1992).

²⁰ "Consider a social group (...). Each member of the group has, in his or her head, millions of mental representations, some short-lived, others stored in long-term memory and constituting the individual's 'knowledge'. Of these mental representations, some — a very small proportion — get communicated repeatedly, and end up being distributed throughout the group, and thus have a mental version in most of

view is biased towards explicit aspects of culture, that is, those depending on the reiterated communication of representations. However, not all aspects of culture are learnt from explicitly expressed representations. As Bourdieu (1980) notes, our internalised knowledge of cultural practices is acquired from practices themselves, as a result of a generalisation on observed behaviours.

The *habitus*, a product of history, produces individual and collective practices — more history — in accordance with the schemes engendered by history. It ensures the active presence of past experiences, which, deposited in each organism in the form of schemes of perception, thought and action, tend to guarantee the ‘correctness’ of practices and their constancy over time, more reliably than all formal rules and explicit norms. (Bourdieu, 1980: 55)

True, common practice can also receive explicit reinforcements — as when we tell children what to do; but this does not rule out the trend towards correlating actions and situations.²¹

The component devoted to forming social representations would have to be articulated in a way that could be compatible with its dual nature as a processing system and as a set of representations. Thus, on the one hand, it should describe the general principles that govern what, following Jackendoff (1992), I will call the *social categorisation* system, and its input and output conditions; on the other, it should identify the set of social representations generated by the system, which can be described in terms of norms.

The social categorisation system would be sensitive to socially dependent features, such as power, distance, age, sex or degree of imposition, among others. These are, in fact, the social categories usually used in politeness theories. They are universal as basic notions, but their particular content is not: any language would exhibit uses dependent on the status of the partners, for instance, regarding their relative age or power, though the exact way in which this would determine the choice of linguistic forms would vary from culture to culture. Once initialised (i.e. once its parameters are set for the first time), the working procedure would remain the same. The social component has both a long-term and a short-term task. The short-term, on-line task is that of analysing and categorising incoming pieces of behaviour. The way in which each extralinguistic situation is characterised (according to variables such as setting, relationship to the partner, purpose, etc.) is crucial to the readjustment of communicative activity: if a situation is understood as one of, say, requesting, the sequence of actions will tend to adjust to the internalised corresponding schema; otherwise, it will receive a negative evaluation. As Bourdieu puts it:

...the *habitus* tends to generate all ‘reasonable’, ‘common-sense’, behaviours (and only these) which are possible within the limits of these regularities, and which are likely to be positively sanctioned because they are objectively adjusted to the logic characteristic of a particular field, whose objective future they anticipate. At the same time, ‘without violence, art or argument’, it tends to exclude all ‘extravagances’ (‘not for the likes of us’), that is, all the behaviours

its members. When we speak of *cultural representations*, we have in mind — or should have in mind — such widely distributed, lasting representations” (Sperber, 1994: 33).

²¹ Hofstede (2001) presents a similar view, according to which social values are acquired in the family and further reinforced during institutionalised education. Different cultures give priority to different values, which explains cultural diversity.

that would be negatively sanctioned because they are incompatible with the objective conditions.

Bourdieu (1980: 55-56)

The long-term task is that of building and updating a set of related representations on socially accepted behaviour. So, we can envisage the social categorisation system as a cognitive sub-system devoted to extracting tacit generalisations from practice. These are not compulsory laws or universal abstractions, but *norms* in the sense defined above, i.e. expressions of statistically usual behaviour. As such, they join the database that makes up the memory store.

Thus, the social subcomponent can be better explained if seen as a processing system that produces both a set of lasting, internalised structured representations and an on-line conceptualisation of current situation. This can be summarised in Figure 2:

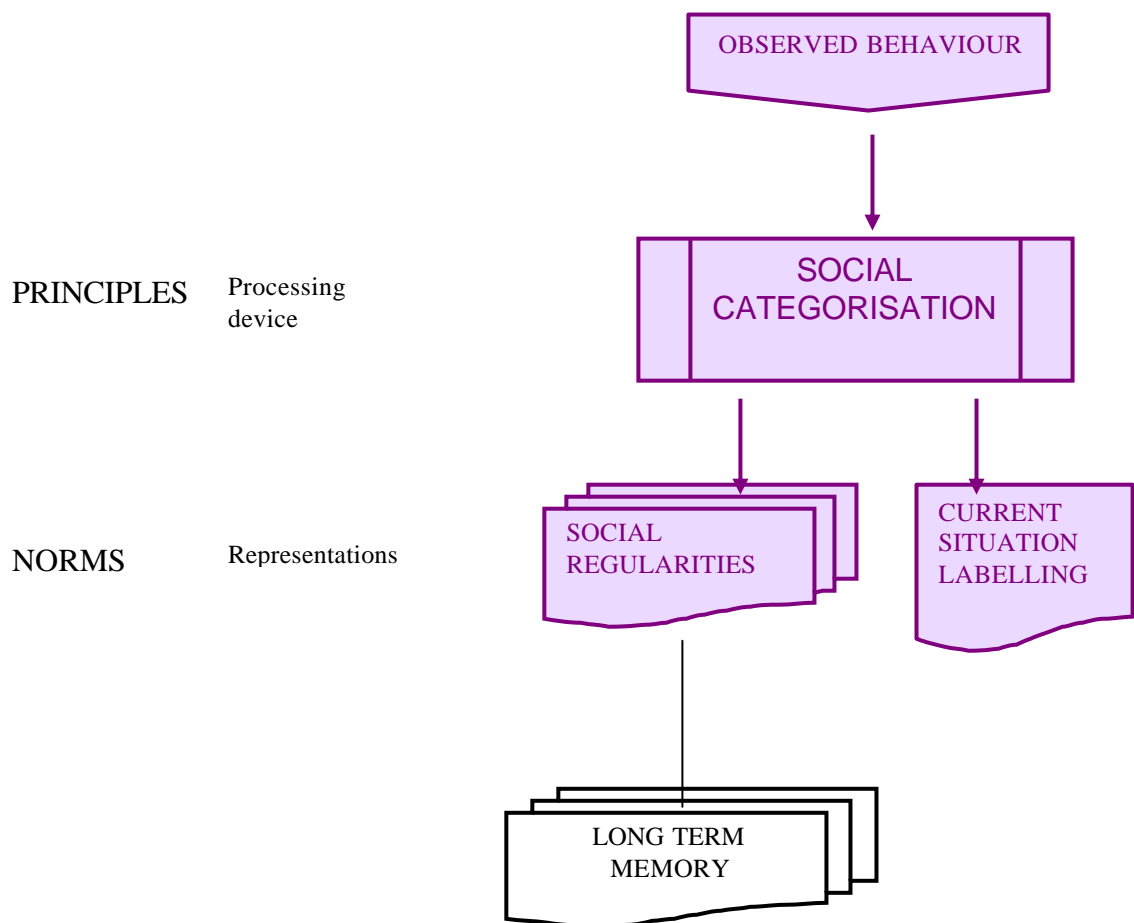


Figure 2: The Social Categorisation System

5.3. Integrating the subcomponents

In the previous sections, I have sketched a brief outline of two components (social and inferential), their tasks and subtasks and their constraints; however, for an integrated theory the main point is their articulation. How are these components related

to each other?

Since inference and social categorisation are mechanisms and social norms are representations, this question amounts to asking how mechanisms and representations integrate. And the answer is easy: representations are the input and the output of mechanisms.

Think of the social categorisation system. Like any other categorisation device, it takes data from perception and elaborates on them, by organising, structuring and labelling them, to turn them into a coherent picture. Its output is a set of representations, some to be used in the on-line processing, some to be stored in the long-term memory. These two kinds of output representations are used in different ways. When activated, old representations stored in the long-term memory access the working memory; when made salient and accessible (i.e. when complying with the input conditions), they can enter the inferential system if needed.

As for the on-line task, the categorisation of the current situation is kept in the short-term memory for a while. This conceptualisation (a set of representations), in turn, imposes adjustments on the input and output conditions of the inferential system. In fact, it acts on accessibility (the input condition) by activating, i.e. by bringing to the working memory the norms associated to the current situation. Thus, these norms are ready to be used in the inferential processing.²² For instance, when entering a shop, the norms and schemata that have to do with the relationship between the customer and the shop assistant are immediately activated and create culturally determined expectations.

The analysis of the situation also acts on the output condition, by adjusting the level at which the expectation of relevance will be reached: for instance, we set it at different points for a chance encounter and for a press conference. Thus, if an utterance meets the basic expectations, relevance is straightforwardly satisfied and inferential processing stops: all behaviours complying with the norms usually go unnoticed (just like a grammatical sentence does). But, on the contrary, if the utterance does not satisfy the expectations, it triggers some extra computation cycles and adds new representations until an interpretation fulfilling the expected level of relevance is obtained.²³ In a sense, norms act as filters: they make salient everything that does not conform to them.

Think of the inferential system now. It takes as its input a set of representations, both from linguistic (decoded) and non-linguistic sources, both of internal and external origin, including those regarding the current situation, that have been made accessible. The system processes them and produces a new set of representations: the interpretation. In addition, it has to be noted that the integration between social and inferential systems does not merely go one way. Inferred conclusions are representations too and, as such, they can be treated as new input data for social categorisation as well, thus providing new samples of social behaviour. This is also the necessary feedback linking again the two components.

The articulation of these two components can be represented as in Figure 3:

²² See Escandell-Vidal (1996) for a more detailed discussion.

²³ In Escandell-Vidal (1998) a fully exemplified argumentation is provided.

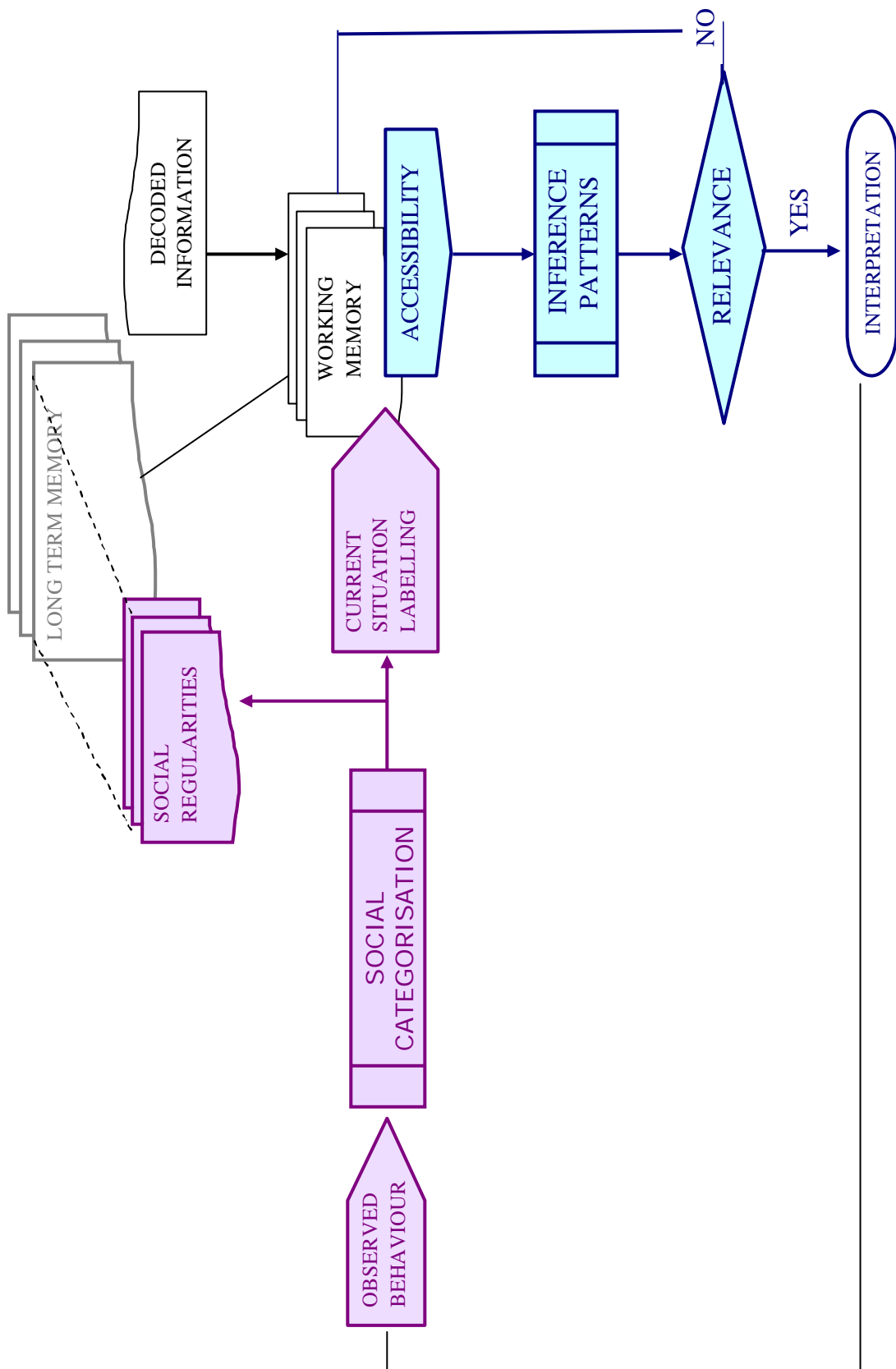


Figure 3: The Architecture of an Integrated Pragmatic Theory

6. Some consequences

I have sketched a proposal in very general terms. In fact my concern has been to suggest an *architecture* for an integrated pragmatic theory, rather than a particular theory. I have tried to put the emphasis on the main desiderata, on the identification of different components and their articulation, not of the actual norms and principles. Thus, the sort of design I have suggested is compatible with a number of different models, each with its own primitives and explanations. I do not intend to argue for a specific theory here; however, some tentative suggestions have been made regarding particular aspects. Of course, the proposal has to be fleshed out in much more detail, but nevertheless, I think that, even if sketchy, it can still serve as the basis for the derivation of various interesting implications.

The approach I have suggested makes it possible to overcome one of the most serious obstacles for pragmatics: the paradox of the level of theorising. The problem is the following. We assume that decoding cannot account for the interpretation of an utterance: no algorithm can fully determine it; contextual information must also be taken into account. The context of interpretation varies from utterance to utterance: no two contexts are totally identical. But if this is so, then we are left with ever different, unique, one-off tokens, which resist scientific generalisations: if pragmatics should operate at this level only, it would irremediably be doomed to failure. The problem is, thus, that science cannot operate at the personal level. To escape this problem, a level of abstraction should be reached so as to find invariants. There are two ways of doing this: first, by moving to the sub-personal level — a level of mechanical necessity, free from individual desires or intentions; second, by moving to a supra-personal level — a level of statistical regularities, not subject to the individual's will either. The first strategy amounts to identifying principles, generalisations governing cognitive systems that are common to all humans. The second one focuses on norms, i.e. on statistical regularities. They both are combined in the suggested architecture.

Some more practical consequences can also be drawn. The first one has to do with the debate about the existence (and, if so, about the status) of *pragmatic competence*. In a model that distinguishes between computation and representation, the answer is straightforward: a competence exists whenever we can find a consistent set of internalised representations (i.e. a body of knowledge). So in the present case, there IS something we can call *competence*, but it is limited to pieces of knowledge, not to mechanisms. Cultural norms and patterns of behaviour qualify as the objects of a pragmatic competence to the extent that they are acquired, internalised and stored. Thus, different individuals can show various degrees of mastery — they can be more or less competent — but the mastery of a particular set of norms does not guarantee a similar skill in other domains: for example, an individual with a good tacit command of the norms governing conversation in non-formal settings can be absolutely incompetent in formal settings. Such differences arise as the result of a poor or insufficient exposure to the relevant situations, not as a consequence of differences in the processing systems themselves. Mechanisms, on the other hand, show a similar performance, regardless of the level of competence or education of individuals; this suggests that the term *competence* is not applicable here.

This has implications for second language learning as well. Cultural misunderstandings usually come up when communicative habits are transferred from one language/culture to a different one. Acquiring communicative competence in a foreign language would involve re-setting the social categorisation mechanisms and

generating a new set of norms, according to the cultural standards of a different community. Though teaching methods make every effort to provide the students with relevant generalisations, this does not seem to be enough. The reason may be the following. Native speakers internalise social patterns directly from practice, by generalisation over a whole range of situations, and therefore these acquire the status of probabilistic norms; at the same time, practice tunes the social categorisation system and determines its on-line output, i.e. the current situation labelling. Students, however, are presented with examples followed by explicit explanations in the form of rules without exception. Now what explicit explanations do is to add new representations to long-term memory but they can hardly modify the working of the social categorisation system once and for all. The result is that even if the relevant generalisations for a new culture have been learned, these are not automatically related to the social categorisation labelling, so the current situation will be categorised as in the native culture and will still trigger the activation of its old corresponding norms, bypassing the newly acquired ones. Thus, explanations can hardly replace the original settings of the social categorisation faculty.

7. Modelling the mind: the lost connection

My proposal has been elaborated within the standards of symbolic cognitivism and particularly by adopting two basic assumptions: first, the distinction between *representation* and *computation*; second, the idea that the mind can be conceived of as a symbolic system that processes and stores internal symbolic representations (i.e. representations made of a combination of symbols according to the system's own specific rules). Knowledge is the set of stored representations (no matter what its actual format is), and learning entails adding new representations to the existing database and updating its contents. Using stored pieces of knowledge involves inferential processes, which combine available representations, both external and internal, to yield new assumptions.

Symbolic cognitivism, however, has been recently challenged by a new model in Artificial Intelligence: connectionism.²⁴ Its starting point is a computational architecture that tries to emulate the working of neurons. Cognitive activity involves a set of parallel, distributed and dynamic computational processes carried out by a high number of small processing units. These units have no intrinsic content, but each can exhibit a different degree of activation; also the connections linking each unit to others can be different. Thus, two processes differ in the level of activation of each unit and in the weight of the connections between the units. There is no database, no explicit representation of knowledge, as representationalism claims; as a consequence, learning cannot entail adding new pieces of information to a memory storage, but rather adjusting the connectivity patterns for each sort of input until fixing them.

Apparently, connectionism obtains good results in all those areas in which classical cognitivism runs into difficulties, such as performing activities and adapting to unexpected circumstances. Now, if pragmatics has to do with linguistic activity and linguistic activity is precisely a matter of adapting to circumstances, maybe — one could wonder — we have chosen the wrong, representationalist, paradigm and we had better reconsider our approach.

The decision, I think, crucially depends on our objectives. There is every

²⁴ Rummelhart et al. (1986); Laks (1996); Clark (1997), among others.

likelihood that the brain will work at the neuron level as connectionists claim, and also that the computational implementation will have to be carried out by using a neural network architecture. Yet, there is an aspect in which representationalism is still far superior: that of its ability to help us describe, depict and understand phenomena. Paradoxically, a disconnection between descriptive capacity and causal explanation results from adopting a connectionist view: when we try to describe external regularities in behaviour, we come up against the problem that there is not a parallel, internal explanation in terms of underlying processes; similarly, when we consider the regularities in the computational processes, these do not correlate with external regularities. Put in other words, whilst a neural network can satisfactorily emulate a certain sort of behaviour, no relevant, interesting generalisation can be drawn on the basis of internal processes, activity flow and weight of connections only. As Laks (1996) admits, no direct link can be established between phenomena and the way they arise.

The moral of the story is: we must relativise the status of our explanations. When we postulate the existence of different cognitive systems and mental processes, or when we explain certain phenomena as a result of certain internal representations — as I have done —, we must keep in mind that what we are producing is a functional and descriptive characterisation that can help us understand and explain how things are, but we should be careful not to mistake the explanation for the facts; this does not mean, however, that our explanations must be necessarily wrong and consequently abandoned — at least, not as long as they provide us with suitable views on difficult issues. Our theories are merely convenient representations, useful tools for better capturing the intricate complexity of our own behaviour, not scale models of our brain. Thus, we can suggest a model of reality without claiming the reality of the model.

8. Conclusion

In the previous sections I have discussed some aspects of the divorce between social and cognitive pragmatics. I have tried to show that they could never reduce to one another. In determining what is adequate to a given situation different factors are involved: some of them are general principles and their operation is not affected by culture; others are norms that stem from socially repeated practices. Yet, they all play a determining role in a characterisation of communicative activity. I have tentatively put forward an integrated model in which both social and inferential components can occupy their rightful place, and I have tried to draw a provisional outline of how it could work.

Of course, I might be wrong. I would be prepared to admit that the solutions I have favoured are not the only ones, not even the best ones; that the architecture of my model calls for substantial changes; that the details of my proposal can be improved to better capture the essence and diversity of the communicative activity.

And still, there is a point in which I would like to believe I am right — otherwise, I would not dare impose on your time: pragmatics has reached an outstanding level of development, both in social and inferential approaches; now, it is time to work on an integrated theory, one in which the two existing perspectives could fruitfully combine. Such integrated model would have to contain partial theories on the various capacities that take part in linguistic activity and make an explicit proposal about the way in which they are linked together. Only in this way will we be able to account at the same time for what is universal and what is culture-specific; only in this

way will we be able to edge closer to a better, fuller and richer understanding of human communication.

References

- Berman, R. A. and D. I. Slobin (eds.)
1994 *Relating Events in Narrative: A Crosslinguistics Developmental Study*. Hillsdale (N.J.): Erlbaum.
- Blakemore, D.
1992 *Understanding Utterances*. Oxford: Blackwell
- Blum-Kulka, S., J. House, and G. Kasper (eds.)
1989 *Cross-cultural Pragmatics: Requests and Apologies*. Norwood: Ablex.
- Bourdieu, P.
1980 *The Logic of Practice*, Cambridge: Polity Press
- Brown, P. and S. Levinson
1987 *Politeness. Some Universals in Language Use*. Cambridge: Cambridge University Press.
- Carston, R.
1999 "The Semantics/Pragmatics Distinction: A View from Relevance Theory". In *The Semantics/Pragmatics Interface from Different Points of View*, K. Turner (ed.), 85-125. Oxford: Elsevier.
2002a "Linguistic meaning, Communicated Meaning and Cognitive Pragmatics". *Mind and Language* 17 (1-2): 127-148.
2002b *Thoughts and Utterances*. Oxford: Blackwell
- Clark, A.
1997 *Being There. Putting Brain, Body and Mind Together Again*. Cambridge (Mass.): MIT Press.
- Connor, U.
1996 *Contrastive Rhetoric: Cross-cultural Aspects of Second Language Writing*. Cambridge: CUP.
- Escandell-Vidal, V.
1995 'Cortesía, fórmulas convencionales y estrategias indirectas'. *Revista Española de Lingüística* 25: 31-66. Available at <http://www.uned.es/sel/pdf/ene-jun-95/25-1-31-66.pdf>.
1996 "Towards a Cognitive approach to Politeness". In *Contrastive Semantics and Pragmatics (Vol. II: Discourse Strategies)*, K. Jaszczolt and K. Turner (eds.), 629-650. Oxford: Pergamon. Also in *Language Sciences* 18 (3-4): 629-650.
1998 "Politeness: A Relevant Issue for Relevance Theory". *Revista Alicantina de Estudios Ingleses* 11 (Special issue devoted to Relevance Theory): 45-57.
- Fukushima, S.
2000 *Requests and Culture: politeness in British English and Japanese*. Bern: Peter Lang.

- Gallardo Paúls, B.
1996 *Análisis conversacional y pragmática del receptor*. Valencia: Episteme.
- Gass, S.
1999 *Interlanguage Refusals: A Cross-cultural Study of Japanese-English*. Berlin: Mouton-DeGruyter.
- Gass, S. and J. Neu (eds.)
1993 *Speech Acts across Cultures*. Berlin: Mouton-De Gruyter.
- Haverkate, H.
1994 *La cortesía verbal. Estudio pragmlingüístico*. Madrid: Gredos.
- Hirschfeld, L. A. and S. A. Gelman.
1994 *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press.
- Hofstede, G.
2001 *Culture's Consequences. Comparing Values, Behaviors, Institutions and Organizations Across Nations*. London: Sage.
- Horn, L.
1984 "Toward a New Taxonomy for Pragmatic Inference: Q-based and Rbased Implicature". In *Meaning, Form and Use in Context.. Linguistic Applications*, D. Schiffrin (ed), 11-42. Washington: Georgetown University Press.
- Jackendoff, R.
1992 *Languages of the Mind*. Cambridge (Mass.): MIT Press.
- Janney, R. W. and H. Amdt
1992 " Intracultural Tact vs Intercultural Tact". In *Politeness in Language. Studies in its History, Theory and Practice*, R.J. Watts, S. Ide, and K. Ehlich (eds.), 21-41. Berlin: Mouton-De Gruyter.
1993 "Universality and Relativity in Cross-cultural Politeness Research: A Historical Perspective". *Multilingua* 12: 13-50.
- Kasher, A.
1991 "Pragmatics and Chomsky's research program". In *The Chomskyan Turn*, A. Kasher (ed.), 122-149. Oxford: Blackwell.
- Kasper, G.
1992 "Pragmatic Transfer". *Second Language Research* 3: 203-231.
- Kasper, G. and S. Blum-Kulka (eds.)
1993 *Interlanguage Pragmatics*. Oxford: Oxford University Press.
- Kerbrat-Orecchioni, C.
1990-1994 *Les interactions verbales* (3 vols). Paris : Armand Colin.

- Lakoff, R.
1973 "The Logic of Politeness, or Minding your P's and Q's". *Proceedings of the Ninth Regional Meeting of the Chicago Linguistic Society*. 345-356.
- Laks, B.
1996 *Langage et cognition. L'approche connexionniste*. Paris: Hermes.
- Leech, G.
1983 *Principles of Pragmatics*. London: Longman.
- Levinson, S.
2000 *Presumptive Meanings. The Theory of Generalized Conversational Implicature*. Cambridge (Mass.): MIT Press.
- Márquez Reiter, R.
2000 *Linguistic Politeness in Britain and Uruguay*. Amsterdam: John Benjamins.
- Matsumoto, Y.
1989 "Politeness and Conversational Universals. Observations from Japanese". *Multilingua* 8: 207-221.
- Mey, J.
1993 *Pragmatics. An Introduction*. Oxford: Blackwell.
- Multilingua* 12 (1993)
- Oleksy, W. (ed.)
1989 *Contrastive Pragmatics*. Amsterdam: John Benjamins.
- le Pair, R.
1996 "Spanish requests strategies: A cross-cultural analysis from an intercultural perspective" In *Contrastive Semantics and Pragmatics (Vol. II: Discourse Strategies)*, K. Jaszczolt and K. Turner (eds.), 651-670. Oxford: Pergamon.
- Placencia, M.E.
1992 "Politeness in Ecuadorian Spanish and British English", *Language Learning Journal* 6: 80-82.
- Placencia, M.E. and D. Bravo
2002 *Actos de habla y cortesía en español*. Munich: LINCOM Europa.
- Pléh, C.
2000 "Modularity and pragmatics: Some simple and some complicated ways", *Pragmatics* 10,4: 415-438.
- Rummelhart, D, et al. (eds.)
1986 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge (Mass.): MIT Press

Scollon, R. and S. Scollon

1995. *Intercultural Communication: A Discourse Approach*. Oxford: Blackwell.

Searle, J.

1969 *Speech Acts*. Cambridge: Cambridge University Press.

1975 "Indirect Speech Acts". In *Speech Acts (Syntax and Semantics, 3)*, P. Cole and J. L. Morgan (eds.), 59-82. New York: Academic Press.

Sifianou, M.

1992 *Politeness Phenomena in England and Greece. A Cross-Cultural Perspective*. Oxford: Clarendon Press.

Sperber, D.

1994 "The modularity of thought and the epidemiology of representations". In *Mapping the Mind: Domain specificity in cognition and culture*, L. A. Hirschfeld and S. A. Gelman (eds), 39-67. New York: Cambridge University Press.

2000 "Metarepresentations in an evolutionary perspective". In *Meta\$ representations: A Multidisciplinary Perspective*, D. Sperber (ed.), 117-137. Oxford: Oxford University Press.

Sperber, D. and D. Wilson

1986/1995 *Relevance. Communication and Cognition*. Oxford: Blackwell.

1990 "Spontaneous Deduction and Mutual Knowledge". *Behavioral and Brain Sciences* 3.1: 179-184.

2002 "Pragmatics, Modularity and Mind-reading". *Mind and Language* 17 (1-2): 3-25.

Strauss, C. and N. Quinn

1997 *A Cognitive Theory of Cultural Meaning*. Cambridge: Cambridge University Press.

Tuson Valls, A.

1997 *Análisis de la conversación*. Barcelona: Ariel.

Vázquez, G. Project ADIEU (Akademischer Diskurs in der Europäischen Union)

<http://www.sparchlabor.fu-berlin.de/adieu>

Watts, R.J., S. Ide and K. Ehlich (eds.)

1992 *Politeness in Language. Studies in its History, Theory and Practice*. Berlin: Mouton-De Gruyter.

Wierzbicka, A.

1991 *Cross-Cultural Pragmatics. The Semantics of Human Interaction*. Berlin: Mouton-De Gruyter.