

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

“Reasoning behind Reasoning:
Non-monotonicity and Scalar Implicatures”

verfasst von / submitted by

Nelli Matevosyan

angestrebter akademischer Grad / in partial fulfilment of the requirements for
the degree of

Master of Arts (MA)

Wien, 2017 / Vienna, 2017

Studienkennzahl lt. Studienblatt /

degree programme code as it appears A 066 867

on the student record sheet:

Studienrichtung lt. Studienblatt /

degree programme as it appears on the Allgemeine Linguistik: Grammatikthe-
orie und kognitive Sprachwissenschaft

Betreut von / Supervisor:

Univ.-Prof. Dr. Daniel Buring

Acknowledgements

This thesis came to life with the support and help of several people. I want to thank them all for being with me during this long, fruitful but also dreadful process.

First of all, it's my supervisor, Martin Prinzhorn, who has helped me all along, with his wisdom and expertise as well as with the willingness to give me moral support 24/7.

It's my beloved fiancé, Alexei Korolyov, who unwittingly and to his misfortune had to go with me through the whole process and was always there when I needed him.

It's my two colleague friends, Magdalena Roszkowski and Susanne Veil, from the passionate discussions with whom I learned so much about the things I was working on.

And of course, it's my family and my non-linguist friends who encouraged me and couldn't wait for me to reach the triumphant end.

Contents

1	Introduction	4
2	Historical View	6
3	Logic under the microscope	11
3.1	Classical propositional logic and debates around it	11
3.2	Monotonic vs. Non-Monotonic Logic	15
3.2.1	Default Logic	18
3.2.1.1	Closed-World Reasoning Assumption	20
3.2.2	Autoepistemic Logic	21
3.2.3	Erotetic Principle	22
3.3	Conclusion	24
4	Experimental Insights in General Reasoning	25
4.1	From Wason Selection Task	25
4.2	From Probabilistic Reasoning Approach	29
4.3	Conclusion	35
5	Reasoning: where and when?	36
5.1	Structure of Reasoning	36
5.2	Do Reasoning Skills Develop?	41
5.3	Conclusion	45
6	Non-Monotonicity in Linguistics	46
6.1	Introduction	46
6.2	Scalar Implicatures	47
6.2.1	General Observations	47
6.2.2	Empirical Evidence	51
6.2.3	Number words	57
6.2.4	Two types of defeasibility	58
6.3	Conclusion	60

Contents

7	Consequences for the Architecture of Language and Cognition	62
7.1	Conclusion	68

1

Introduction

This study has emerged out of my interest in the intersections of various sciences and branches of human knowledge. I believe that the truth is always there, or at least that's the place where our quest for the truth must begin.

Human cognition is the subject of many sciences, and much debate. And so are its constituent parts, logic and language. In this study I look at the relationship between them. Particularly, I am interested in the big question of whether it's the logic of language that shaped the logic of reasoning, or was it the other way round? The reason this question is worth asking is that our language system seems to share some core characteristics with the logic of our reasoning, such as the ability of both to adapt themselves to new information. I do not aim to answer definitively which influenced which, but rather wish to open a discussion about what this question might mean for the architecture of language and cognition.

In my opinion, the relationship between general human logic and logic in language, as well as the question of which influences the other, may bring new insights to the general study of mind. The study of that relationship is inextricably linked with the bigger dispute between the nativist/modularity approach to language, which postulates that there is innate language-specific endowment, and the empiricism/unitary approach, which claims that our linguistic knowledge is based in general experience and is not domain-specific.

Over the last years there has been a great deal of research on the nature of human reasoning and the aspects of linguistics related to the topic at hand. All this has created a wealth of data and information that I drew upon and analysed for the purposes of this study, which employs therefore a theoretical approach.

The study is structured as follows. It begins with an overview of the history of logic and how logic relates to psychology (Chapter 2). I then (Chapter 3) introduce the concept of classical logic as espoused by Aristotle and those in

1 Introduction

his tradition and provide data that shows that it cannot be applied to everyday human reasoning because of its monotonicity. I then describe different types of non-monotonic logic which can be suited to everyday reasoning.

In Chapter 4, I present empirical data, some of which has served as evidence that humans are not rational; I discuss how this data can lead to different conclusion when viewed in the context of non-normative logic. In Chapter 5, I focus on the development of reasoning, both on the findings of the origins and evolution of reasoning as well as on the developmental processes by children.

Chapter 6 looks at the phenomenon of non-monotonicity as expressed in language. In particular, I focus on the so-called scalar implicatures. I begin by describing the phenomenon and then provide empirical data. The data aims to determine the phenomenon's nature and to look at it from the developmental perspective.

Chapter 7 maps the findings in the previous chapters onto the principal topic of this study, namely the broader question of whether non-monotonicity first emerged in language or in general cognition.

“Ever noticed how supervillains are always trying to change things and superheroes always try to resist the change?”

@TheWeirdWorld. Tweet #ShowerThoughts

2

Historical View

We all think of ourselves as rational creatures. Rational as opposed to each other, and rational as opposed to non-human beings. Rationality is what raises humans over other creatures, what makes us human beings, *Homo truly sapiens*; as such, we are able to reason logically and in a premediated manner while solving problems, avoiding mistakes, planning, making predictions and so on.

The tradition of bringing logic and reasoning together goes back to Aristotle, who famously formulated rules of logical reasoning. Since then being rational meant being logical. The 19th century saw the emergence of *psychologism*, which connected logic and reasoning as processes related to thinking, and therefore being a matter of psychology. In the classical version of psychologism (John Stuart Mill being the most prominent figure here) logic is merged with psychology, in other words logical laws comply with psychological laws.

Mill saw logic as being composed of two parts: art of reasoning, which provides us with the rules according to which we reason, and science of reasoning, which analyses mental processes. His main idea was that these two are connected and that logic can operate only after psychology provides descriptive analysis of reasoning processes:

Its [the Science of Logic's] theoretic grounds are wholly borrowed from Psychology, and include as much of that science as is required to

2 Historical View

justify the rules of the art. (Mill 1865: 359)

However, this matrimony didn't last long - at the end of the 19th century, these ideas came in for criticism from Gottlob Frege and Edmond Husserl, who defended the autonomy of logic. They viewed logic as objective, formal, exact, ideal, pure and a priori - as opposed to psychology, which they regarded as subjective, dealing with human beliefs and a posteriori. On their view, logic is truth, and truth cannot depend on each individual's thinking: "What is true is true absolutely, in itself; the truth is one, identical with itself, whatever may be the beings who perceive it - men, monsters, angels or gods" (Husserl 1901 : 117). It's exact and unambiguous - qualities that empirical sciences like psychology do not exhibit. Psychologism makes logic lose these core characteristics, which is a contradiction in terms.

The antipsychologist ideas of Frege and Husserl dominated Western analytical philosophy across the 20th century. In the middle of 20th century, however, a new twist happened after the celebrated Swiss psychologist Jean Piaget formulated his famous theory of cognitive development. According to the theory the natural construction of logical abilities in growing children begins from the so called *concrete operational stage* (ages (6-7)-11). At this stage, children incorporate inductive reasoning (making generalizations), while deductive reasoning (making predictions) comes later - namely, at the next *formal operational stage* (ages 11-(15-20)). According to Piaget's theory, the acquisition of deductive rules is the peak of human cognitive development. Piaget's theory gave a new direction to the study of thinking: from that point on there followed empirical research of human reasoning. Over time, the results brought in by this research led some researchers to discard the assumption that people reason by logical laws, thereby discrediting the followers of Piaget's theory and enlarging the gap between normative logic and descriptive psychology. Because of this, some psychologists came to the conclusion that humans are inherently irrational.

Accordingly, Peter Wason (1966) using his famous *selection task* discovered a surprisingly striking deviation from the classical formal logic. For Herbert Simon (1957) in his turn human reasoning deviations are an effect of the *bounded rationality* as opposed to the *ideal* rationality. The idea is that our rationality is bound by information processing constraints and various cognitive limitations.

2 Historical View

In the 70s Kahneman and Tversky started developing the *Heuristics and Biases* approach, whereby people's irrationality is accounted for by their poor knowledge of statistics. And since human reasoning appeared to be not normatively correct in everyday life, it was no longer the concern of logic.

This disavowal trend was fed by abortive attempts in the nearby field of artificial intelligence (AI) to successfully imitate human reasoning processes by developing algorithms relying on formal rules of inference (e.g., Bledsoe 1977). These algorithms were so slow that many AI researches had to abandon orderly formal logic in favour of other methods of analysis (e.g., Toulmin 1958; Scriven 1976).

In the end we see that logic, which was initially conceptualised as the product of our mental device that subsequently took centre stage in the cognitive science field, was divorced from its origins by various developments in psychology and philosophy and headed off in a different direction.

More recently, however, a new line of thinking emerged as some researchers started arguing that this newfound incompatibility between logic and reasoning comes from the wrong notion of logic we adopt (see, for instance, Cohen 1981; Oaksford and Chater 1994, 1995; Gigerenzer and Goldstein 1996; Todd and Gigerenzer 1999, Stenning and von Lambalgen 2008, 2012). Their main argument is that classical logic does not constitute a proper framework in which to describe human reasoning behaviour, so we had better seek alternative systems of logic instead. And then within these systems of logic we can be rational or irrational. In other words, humans are irrational only with respect to the systems of logic we reason within.

Among the advocates of this idea are Stenning and van Lambalgen - one a cognitive scientist, the other a logician, who in their book 'Human Reasoning and Cognitive Science' (2012) likewise argued that monotonic logic is not an appropriate framework for explaining human reasoning mechanisms. *Monotonicity* is a characteristic inherent in classical logic, according to which inference depends on the presence of information only. Instead, they argued, human reasoning should be viewed through the prism of non-monotonic logic. *Non-monotonicity* is simply the lack of monotonicity, i.e. in non-monotonic systems inference is sensitive to the absence of information.

In their book, Stenning and van Lambalgen speak of the problems posed by language when interpreting logical tasks; that's because all tasks are formulated

2 Historical View

in natural language so people first need to choose an appropriate type of logic to address the task before proceeding to solve it. The tradition of linking logic with language goes as far back as Wilhelm von Humboldt's work in the 19th century - von Humboldt was one of the first in the Western world to propose that language is not just a random collection of words and sounds but rather a coherent system. Unlike his contemporaries, however, von Humboldt rejected the idea of one unifying grammar that underlied all languages because all attempts to construct such a grammar were based exclusively on European languages. That said, he did believe in the universality of mental structures and held that the rules and laws of thinking translated themselves into the rules that govern our linguistic utterances (SEP: Wilhelm von Humboldt). This connection between logic and language was developed further in the 1960s, when the idea arose that languages have a systematic semantic structure that adheres to certain logical laws. In particular, logician Richard Montague developed a theory of semantics which states that languages abide by the principle of compositionality first introduced by Gottlob Frege - that is, the meaning of the whole is a function of the meanings of its parts and their mode of syntactic combination (Partee 1984). Despite big debates about whether natural languages can be compared to formal languages, Montague made a bold statement in his works of the time saying that he considered it possible "to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory" (Montague 1970 : 222). These ideas and ideas of similar kind gave rise to the study of formal semantics. It was fully established by the beginning of the 1990s.

A further impetus to studying the link between logic and language was given by research in AI that made attempts to formalize knowledge representation and reasoning, and did that by using non-monotonic logics. Semantics followed their example and started making use of non-monotonic logic as a tool for natural language interpretation. In the course of time it has been shown that non-monotonicity can also provide appropriate foundations for any other area of linguistics because many if not all of its areas are characterised by generalisations and exceptions to these generalisations, which is exactly where non-monotonicity comes in handy. It's not the purpose of my research to indicate all its applications; I will focus instead on one case, namely the so-called scalar implicatures. But before we get to them let's first talk about how the study of logic and reasoning

2 Historical View

developed over the years.

*“A key property of intelligence - whether exhibited
by man or by machine - is flexibility.”*

Matthew L. Ginsberg

3

Logic under the microscope

3.1 Classical propositional logic and debates around it

The traditional view of human reasoning has always been that canons of rationality are provided by the theory of classical logic. That is, in order to derive conclusions from different bits of information we translate them into the formal language of classical logic where we then deal with them. Here is an example of a classical syllogism:

- (1) All Armenians are big-nosed.
All big-nosed people are intelligent.
Therefore, all Armenians are intelligent.

What is supposed to be happening when we want to derive the conclusion given in the last sentence is that we substitute the premises by variables and check whether a substitution verifies the premises; if it does and it also satisfies the conclusion, then we have a *valid* argument:

- (2) All x are y .
All y are z .
Therefore, all x are z .

3 Logic under the microscope

So what matters here is the form or schema of the argument, the subject is not important, that is whatever we put instead of nonlogical constants x , y and z , we will reach the conclusion. Logical constants such as *all*, *some*, *not*, *and*, *or*, *if* are characterized as being *topic neutral* or *domain independent*.¹ Reasoning is thought to be correct if the form of the argument constitutes a valid argument. Accordingly, we do not necessarily have to know what are we talking about in order to draw the right conclusion: if you substitute x and y with any unknown or even unexisting matters the conclusion will still hold.

The topic neutral character of reasoning patterns is a fundamental principle of formal logic in general. Naturally, since every argument is translated into one preestablished formal frame, there can be no topic- or subject-dependencies.

The domain in which classical logic fits must satisfy its core settings, namely:

- (1) recursivity: if θ and ψ are formulas, then so are $\neg\theta$, $(\theta \vee \psi)$, $(\theta \rightarrow \psi)$,...;
- (2) truth-functionality: truth-value of a compound is a function of the truth-values of its components;
- (3) bivalency: sentences have only one truth-value, either true or false, with nothing in between (law of excluded middle);
- (4) consequence: an argument is valid if whenever all models of the premises are true, so is the conclusion: conclusions are consequences of their premises.

But does natural reasoning exhibit these formal features? It is easy to see that the reasoning and logical tasks that we face in everyday life do not easily fit into this rigid framework. For one thing, as Stenning and van Lambalgen (2012) state in their book "Human reasoning and Cognitive science", human reasoning is domain-dependent, i.e. the validity of the argument and reasoning in general very much depends on which domain one is reasoning in. That is to say, context plays a crucial role. Domains change depending on discourse, and based on context same words and phrases can correspond to different things in the world. I will come to these ideas again in the next pages when discussing the meanings of vague expressions and in the next chapter when explaining the results of the reasoning

¹Domain here is taken to mean the scope of interpretation or discourse, and all the representations within it such as objects and individuals in this domain and their relationships, together with the formal language we use to reason about them.

3 Logic under the microscope

experiments.

Secondly, it is not always the case that the statement is either true or false. There are also ‘undecided’ or ‘not known to be true or false’ statements, calling the bivalency setting into question. The authors provide the following example:

One example of where this occurs in practice is a computerized primality test which checks whether the input $2^{1257787} - 1$ is a prime number. One could say that, while the program is running, the statement ‘ $2^{1257787} - 1$ is a prime number’ is undecided; but a decision may follow in the end, if the program halts. (Stenning and van Lambalgen 2012: 27)

An even more confusing task, they say, is to formalize the so-called *vague expressions*, like ‘small’, ‘painful’, ‘high’, ‘rich’, ‘usually’, ‘often’. One solution they offer is changing the logic and adding a third truth-value: for the first example it would be ‘undecided but possibly decided at some point later’ one (assigned as u), whereas for the vagueness predicates it would be ‘intermediate between true and false’ one (assigned as $\frac{1}{2}$). This many-valued model of logic is called *Łukasiewicz logic* named after Polish logician and philosopher Jan Łukasiewicz. This many-valuedness is extended in the so-called *fuzzy logic*, which allows for a continuum between an ‘absolutely false’ and an ‘absolutely true’ truth-values. Here is what Łukasiewicz wrote in his ‘Philosophical remarks on many-valued systems of propositional logic’ (1930):

it was clear to me from the outset that among all the many-valued systems only two can claim any philosophical significance: the three-valued one and the infinite-valued ones. For if values other than “0” and “1” are interpreted as “the possible”, only two cases can reasonably be distinguished: either one assumes that there are no variations in degrees of the possible and consequently arrives at the three-valued system; or one assumes the opposite, in which case it would be most natural to suppose, as in the theory of probabilities, that there are infinitely many degrees of possibility, which leads to the infinite-valued propositional calculus. I believe that the latter system is preferable to all others. Unfortunately this system has not yet been investigated

3 Logic under the microscope

sufficiently; in particular the relations of the infinite-valued system to the calculus of probabilities awaits further inquiry. (Łukasiewicz 1930: 173)

Another attempt to characterize the vagueness of frequency adverbs, like ‘usually’ and ‘often’ was the use of a not truth-functional propositional *probability logic*. In a natural language these expressions are used constantly (now, for instance). But what they describe and entail is not obvious. What percentage the probability of the proposition p is high enough to justify the use of the adverb ‘constantly’, with probability here meant as a relative frequency: greater than 60% but less than 90% or exactly 100%?..

Paradoxically, the exact classical logic and the uncertain probability logic are very closely bound to each other. Sometimes probability logic is even seen as an extension of classical logic, if we substitute $L \rightarrow [0, 1]$ of classical propositional logic with probability functions $P : L \rightarrow \mathbb{R}$, which take values from the interval $[0, 1]$, where 0 denotes contradiction (‘never’) and 1 denotes confirmation (‘always’). As Stenning and van Lambalgen say, “probabilistic reasoning is too much tied to classical logic to be able to encompass the wide variety of reasoning that actually occurs” (Stenning and van Lambalgen 2012 : 32).

These were arguments against the idea that every statement is either true or false. The bivalence property in classical logic is connected to the truth-functionality one, according to which sentence components take on truth values in order to determine the meaning of the whole sentence - now is this property sufficient? No, because for that we would have needed to have only statements that obey the law of the excluded middle, i.e. are not known not to be true or false and nothing in between, and as it was just shown that’s not always the case.

Intuitionistic logic offers a infinite-valued frame that omits the principle of the excluded middle. One can say intuitionistic logic is classical logic without the principle of the excluded middle. L.E.J. Brouwer, whom is attributed the intuitionistic logic, expressed his view as follows:

The belief in the universal validity of the principle of the excluded third in mathematics is considered by the intuitionists as a phenomenon of the history of civilization of the same kind as the former belief in the rationality of π , or in the rotation of the firmament about

3 Logic under the microscope

the earth. The intuitionist tries to explain the long duration of the reign of this dogma by two facts: firstly that within an arbitrarily given domain of mathematical entities the non-contradictoriness of the principle for a single assertion is easily recognized; secondly that in studying an extensive group of simple every-day phenomena of the exterior world, careful application of the whole of classical logic was never found to lead to error. (ed. D. van Dalen 1981: 7)

Under intuitionistic tradition the proposition is not true or false, instead it's proven or not proven (or justified/not justified) until given direct evidence. All operations can be defined by means of proofs. For instance:

- (1) a proof for $\theta \wedge \psi$ means having a proof of θ and a proof of ψ ;
- (2) a proof for $\theta \vee \psi$ means having either a proof of θ or a proof of ψ ;
- (3) a proof for $\theta \rightarrow \psi$ means having a function transforming a proof of θ into a proof of ψ ;
- (4) a proof of $\neg\theta$ means that any proof of θ leads to a contradiction.

These were several examples of logics that make it possible to deal with some of the aspects where classical logic seems to fail. Following up on Stenning and van Lambalgen, in the next section I will present an alternative account, whose core difference from classical logic is in its focus on *non-monotonicity*.

3.2 Monotonic vs. Non-Monotonic Logic

As we have seen in the previous section, some formal characteristics of classical logic do not seem to be able to describe natural reasoning. One of the big problems with classical logic is the fact that it does not tolerate any kind of incomplete or uncertain information, but of course the real world and the language used to describe it are brimming with it: sometimes information is unavailable, and sometimes we just need to respond as quickly as possible and don't have time to gather all the relevant information.

By the same token, classical logic is too rigid to cope with added information. Formally, it is the *monotonicity* of the consequence relation of classical logic that

3 Logic under the microscope

underlies it: in all models $\mathcal{M}$ if $\psi_1, \dots, \psi_n \models \phi$ then $\psi_1, \dots, \psi_n, \theta \models \phi$ for any sentence θ . In practice it means that new information cannot make the previous conclusion false, it can only produce additional data - which makes the set of propositions grow ‘monotonically’.

It is not difficult to see that that’s not what happens in everyday reasoning. Consider the following example:

(3) If birds fly, and x is a bird, then x flies.

If we add new information, claiming that x is actually a penguin, we’ll invalidate our previous conclusion and update it to x doesn’t fly. The kind of logic that represents *defeasible* reasoning, i.e. allows to draw conclusions from the given knowledge with a possibility to update it if necessary by adding new information to the premises, is called *non-monotonic* logic.

Classical logic by contrast requires all the information needed in advance to be able to give a valid conclusion. The problem with this approach is not only that all the relevant premises are not always available in advance, but also that there is in fact an infinite amount of them:

(4) $\forall x. \text{Bird}(x) \wedge \neg \text{Penguin}(x) \wedge \neg \text{Chicken}(x) \wedge \neg \text{Ostrich}(x) \wedge \neg \text{Dead}(x) \wedge \neg \text{Baby}(x) \wedge \dots \rightarrow \text{Fly}(x)$

Note here again that the previously mentioned topic-neutral character of formal logical rules does not apply to everyday logic: it does indeed matter whether we are reasoning about the loosely-defined animal class of birds or about specific concepts such as, say, triangles. Human reasoning takes into account the meaning of presented elements and can differ under various circumstances. Speaking about the peculiarities of everyday reasoning, Nassim Nicholas Taleb writes in his ‘The Black Swan: The Impact of the Highly Improbable’:

(...) some problems we can understand in their applications but not in textbooks; others we are better at capturing in the textbook than in the practical application. People can manage to effortlessly solve a problem in a social situation but struggle when it is presented as an abstract logical problem. We tend to use different mental machinery

3 Logic under the microscope

- so-called modules - in different situations: our brain lacks a central all-purpose computer that starts with logical rules and applies them equally to all possible situations. (Taleb 2007: 54)

This is an important point to which I will come back later on when discussing various reasoning experiments.

Any logic that characterizes inductive reasoning is non-monotonic. The bird example above demonstrates dynamic inductive generalization, where one generalizes the rule, assuming that it holds in most/typical cases in which the premises hold. Making generalizations help us to make a model of the world around us. This can be demonstrated by a language learning example: adding the suffix ‘-ed’ in order to form a Past Tense form is easier than memorizing the corresponding form of every English verb. Practically, the most basic problem of language acquisition applies here as well: a child gets only a finite language input (often incorrect) but generalizes from it an infinite one.

In their book Stenning and Lambalgen do not just disclaim the belief that in order to see if an argument is valid, one has to translate it into the formal language of classical logic; they also present a new conception of human reasoning. What they do is distinguish between two stages of reasoning: reasoning *to* and reasoning *from* an interpretation. Their argument goes as follows: because most reasoning tasks are formulated in natural language, reasoners must at first establish a logical form relative to the task’s context and to their goals (reasoning *to* processes). In doing so they choose an appropriate formal language, its semantics, notion of validity, the truth and false values, semantic consequence etc. Once all this is established, reasoners then set about solving the task at hand in accordance with their choice (reasoning *from* processes). The authors argue then that in most cases in standard reasoning tasks, subjects use non-monotonic logical forms. This is rather like going to a hospital: upon checking in, you don’t go straight to the doctor; first you come up to the clerk at registration, who, depending on what bothers you, then sends you to an appropriate specialist.

There are various formal approaches to non-monotonic logic, the three most influential ones being *Default Logic*, *Closed-World Reasoning Assumption* and *Autoepistemic Logic*. I give an overview of them in the next sections. In the last section I briefly present one new approach, *erotetic Principle*, that is based on

some ideas relevant for my topic.

3.2.1 Default Logic

Default logic is the most prominent approach to the non-monotonic logic, probably because of the simplicity of the notion of default. “Monday is a work day unless it’s a holiday.” This universal truth can be represented by the default rule:

$$\frac{\textit{monday} : \neg \textit{holiday}}{\textit{workday}} \quad (3.1)$$

The interpretation goes as following: if there is no information that monday is a holiday, it is reasonable to believe that it is not a holiday and conclude that it’s a work day; but if we get additional information that it is indeed a holiday, we cannot longer assume that it’s a workday, and therefore can’t apply the default here.

As you can see from this example, a default theory consists of a pair of sets (D, W) , where D is a set of defaults, i.e. our assumptions, and W is a set of first-order formulas, i.e. facts. A default rule has thus the following form:

$$\frac{\alpha(x) : \beta_1(x), \dots, \beta_n(x)}{\gamma(x)} \quad (3.2)$$

where $\alpha(x) : \beta_1(x), \dots, \beta_n(x)$, and $\gamma(x)$ are the formulas of the first-order logic, and $x > 0$; $\alpha(x)$ is the *prerequisite*; $\beta_1(x), \dots, \beta_n(x)$ are the *justifications*; and $\gamma(x)$ is the *consequent*. If we believe that the prerequisite is true and the justification is consistent with our knowledge and beliefs then we are permitted to accept the consequent. In other words, based on our default assumptions we believe the consequent to be true.

Default logic reasons about typical situations and facts that are known to be true in the majority of cases: birds typically fly, children typically go to school, swans are typically white etc. The pattern of inference has thus the following form: in the absence of the contrary, assume x to be true. The following quote by R. Reiter, the father of Default Logic, reveals the idea behind it:

Most of what we know about the world, when formalized, will yield an incomplete theory precisely because we cannot know everything -

3 Logic under the microscope

there are gaps in our knowledge. The effect of a default rule is to implicitly fill in some of those gaps by a form of plausible reasoning... Default reasoning may well be the rule, rather than the exception, in reasoning about the world since normally we must act in the presence of incomplete knowledge ... Moreover, ... most of what we know about the world has associated exceptions and caveats. (Reiter, 1978)

The key object of the default theory is *extension*. Virtually an extension represents possible world or set of beliefs based on the default assumptions and given facts. If default rules are conflicting, it can give rise to multiple extensions, as in the well-known ‘Nixon diamond’ example:

$$D = \left\{ \frac{quaker(x) : pacifist(x)}{pacifist(x)}, \frac{republican(x) : \neg pacifist(x)}{\neg pacifist(x)} \right\} \quad (3.3)$$

$$W = \{quaker(Nixon), republican(Nixon)\} \quad (3.4)$$

The first default rule in (3.3) says that quakers are pacifists by default, the second one says that the republicans are not pacifists by default. The facts in (3.4) tell us that Nixon is both a quaker and a republican. There are two extensions of this default theory (D, W) : in one Nixon is a pacifist, in another he is not. The question naturally arises: who on earth is he?

In cases like such, where there is more than one extension, people are called to choose between two strategies of reasoning to get to a conclusion: *credulous* (also called *brave*) reasoning or *sceptical* (also called *cautious*) reasoning. The credulous reasoner picks either one of the extensions on whatever grounds or all of them and accepts their conclusions. So, s/he will believe that Nixon is a pacifist or he is not a pacifist based on their knowledge or intuitions about him, or will believe that Nixon is a pacifist and is not a pacifist at the same time. In contrast, the sceptical reasoner believes in none of them and thus doesn’t draw any conclusions at all. That is, the reasoner will have no judgment about Nixon’s views on violence.

As was shrewdly pointed out by Antoniou (1999), default logic is also the main principle of justice in the Western cultures: “In the absence of evidence to the

3 Logic under the microscope

contrary assume that the accused is innocent” (also called *no-risk reasoning*):

$$\frac{accused(x) : innocent(x)}{innocent(x)} \quad (3.5)$$

Defaults model the most common *Closed-World Reasoning Assumption*, according to which what is not known to be true is believed to be false.

3.2.1.1 Closed-World Reasoning Assumption

Suppose you are going to the dentist’s at 3:30 pm and see that the door is closed and the sign on it says that the consultation hours are from 9 am till 3 pm. Does the doctor also see patients at 3:30 pm so you can get your teeth done today? If you reason within Closed-World Assumption framework (CWA), then you certainly have no chance. Whereas if you rely on the Open-World Assumption for Reasoning (OWA), which is supported by classical propositional logic, there will not be simply enough data to furnish an answer, as no information is given about that time of day. There might be models where the doctor is seeing patients as well as models where he is not. Since the conclusion must hold on all models no conclusion can be made here or, more accurately, any conclusion can be made.

By contrast, when reasoning within the CWA framework we assume that all the positive information has been already specified and what is not specified is negative (lack of knowledge = negative knowledge). This negation is often called *Negation as Failure* in order to distinguish it from classical negation. In logic programming it takes the following form:

$$\tau \leftarrow \psi_1, \dots, \psi_n, \neg\phi_1, \dots, \neg\phi_m, \quad (3.6)$$

where τ is a logical atom and $\psi_1, \dots, \psi_n, \phi_1, \dots, \phi_m$ are logical literals.

The example with the birds from the previous section will go like this:

- (5) fly $\leftarrow$ bird, $\neg$ penguin
- $\neg$ fly $\leftarrow$ penguin
- bird $\leftarrow$ penguin

If the system has been told *bird*, it will return *fly* from the first rule without

3 Logic under the microscope

having to trigger the others, but if it has been consequently told *penguin* it will return *bird* and $\neg fly$ from the last two rules. Birds fly unless it has been explicitly said that they are penguins.

It is useful to ‘close’ the world we are reasoning in because the amount of negative information is much larger than the amount of positive information. If we didn’t ‘close’ the world, we wouldn’t be able to make any decisions, to have any expectations or to be able to forecast. However, this ability has its downsides too. These predictions are only based on the past and known information, which can often result in dramatic outcomes like not being able to predict conflicts, accidents and wars, because if something has worked in the past does not mean it will work in the future (the problem of induction). This is the cost we pay for simplifying the world.

CWA is used in many fields of application where it is sufficient to represent positive information only. For instance, CWA formalization is used in a lot of logic programming and deductive databases. For example, if a hotel booking database does not have an entry for the room 501, this will indicate that no guest has booked it yet. With databases we always assume that they are complete. In general, it is with database systems that CWA goes particularly well.

3.2.2 Autoepistemic Logic

Another important formalism of NML is *Autoepistemic logic*. It is very much related to the default logic and as the name reflects, can be used to represent one’s knowledge about knowledge. Originally it goes back to Moore (1985), who had the idea to represent knowledge as belief. According to him, for instance the absence of knowledge ϕ might cause one to believe $\neg\phi$: *If I had an older brother, I would know about it. Since I don’t, I believe not to have an older brother.* To trigger an autoepistemic power of the logic a belief operator B is used, which signifies “I believe in”: $\neg Bbrother \rightarrow \neg brother$. This pattern of reasoning corresponds to the Negation as Failure principle mentioned in the previous section and takes the following form with the meaning “it is not believed that ψ ” ($\neg B\psi$):

$$(\phi_1 \wedge \dots \wedge \phi_n \wedge \neg B\psi_1 \wedge \dots \wedge \neg B\psi_m) \supset \psi \quad (3.7)$$

3 Logic under the microscope

Researchers have also noted connections between autoepistemic and default approaches. This has deserved attention of a number of studies which made attempts to translate one into the other (Konolige, 1988; Truszczyński, 1991). But despite the obvious similarities between default and autoepistemic approach, the crucial difference between them lies in the notion that default logic is defeasible itself and autoepistemic logic gets its non-monotonicity only from the belief operator.

The exceptional characteristics of the autoepistemic logic that distinguishes it from anything else is its capacity to reason not only of the external actual world but also about the knowledge one has of it.

3.2.3 Erotetic Principle

Although *Erotetic Principle* does not fall into the list of examples of non-monotonic logic, I decided to list it here as it shares similarities with its principles and with ideas of Stenning and Lambalgen (2012).

Erotetic Principle holds that human reasoning proceeds “by updating a model of *discourse* with *mental models* representing the premises” (Mascarenhas 2013 : 12). *Mental models* are mental representations of premises based on their meanings and reasoner’s background knowledge. The idea behind the erotetic principle is that our reasoning takes shape of searching answers to questions posed by premises and this questions are represented in the form of mental models. The answers in their turn are very much dependent on interpretation of linguistic content.

Erotetic Principle

Part I - Our natural capacity for reasoning proceeds by treating successive premises as questions and maximally strong answers to them.

Part II - Systematically asking a certain type of question as we interpret each new premise allows us to reason in a classically valid way.

As you can see from the Part II, while Mascarenhas follows Stenning and Lambalgen’s idea of approaching reasoning from the point of view of formal semantics

3 Logic under the microscope

of natural language, he argues that in the end subjects fall back on classical logic forms. They do so by strengthening or enriching the literal content while not meddling with the other, non-literal meanings:

Given a sequence of premises $P_1, \dots, P_n$, and a conclusion C , begin by calculating the strengthened meaning of each premise, getting the sequence $P_1^+, \dots, P_n^+$. Then, check to see if the conclusion C follows classically from $P_1^+, \dots, P_n^+$.

Mascarenhas shows this line of reasoning based on various reasoning fallacies, such as, for instance, the illusory inference from disjunction, illustrated in the following example:

- (6) P_1 : Either Jane is kneeling by the fire and she is looking at the window or otherwise Mark is standing at the window and he is peering into the garden.
 P_2 : Jane is kneeling by the fire.
Conclusion: Jane is looking at the window.

Schematically this fallacy has the following form:

- (7) P_1 : $(a \wedge b) \vee (c \wedge d)$
 P_2 : a
Conclusion: b

This inference is fallacious because the conclusion is invalid in the scenario where Jane is kneeling by the fire and *not* looking at the window, while Mark is both standing at the window and he is peering into the garden. In this scenario both premises are true while the conclusion is false.

By the erotetic principle, P_1 would be interpreted as a question, something like “are we in a Jane-kneeling-by-the-fire-and-looking-at-the-window situation, or in a Mark-standing-at-the-window-and-peering-into-the-garden situation?”. When given P_2 , the reasoner realises that it fits to the first possible answer and not the second one, which brings them to the fallacious conclusion. In other words, we pragmatically strengthen the meaning of the premise and interpret it in the following way:

3 Logic under the microscope

(8) $P_1^+ : (a \wedge b \wedge \neg c \wedge \neg d) \vee (c \wedge d \wedge \neg a \wedge \neg b)$
 $P_2^+ : a$
Conclusion: b

As a result, we have an inference that will be judged classically valid, not a fallacy.

Thus, Mascarenhas and the erotetic principle tell us that (1) we'd better pay more attention to how language is interpreted, and that (2) we can learn to ask the right questions when we reason in order to produce (classically) correct reasoning.

3.3 Conclusion

In this chapter I've shown why the popular conception of human reasoning being tied solely to classical logic is not adequate. As opposed to classical logic, human reasoning is to a great extent context- or topic-dependent. A vast majority of experimental research uses natural language, and language is vague. While interpreting the language, subjects have to first translate it to a formal logic and then reason in accordance with the rules of that formal logic. Depending on the discourse (and reasoner) the choice of the appropriate logic can be different, and researches argue that in most cases it's not classical logic. In order to cope with vague and incomplete information that we are exposed to every day we often choose non-monotonic systems of logic that allow us to be more flexible in our reasoning.

“Hyponym and Hyperonym walk into a bar.

Hyper asks, want some wine?

Hypo says, sure - red or white?”

Alexei Korolyov (fiancé), in conversation

4

Experimental Insights in General Reasoning

4.1 From Wason Selection Task

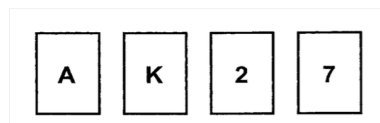
As has been shown, everyday human reasoning does not operate according to classical logic. But because classical logic has always been a canon of rationality, people often appear illogical when faced with logical tasks. The deviations from norms of rationality demonstrated in the tasks aimed at revealing humans reasoning mechanisms have convinced many researchers of human irrationality. These experiments include Wason’s Selection task (Wason 1966; Wason 1968; Johnson-Laird Wason, 1977) and The Conjunction Fallacy (Tversky and Kahneman 1983), on which I will focus. The former falls into the category of experiments which look at human reasoning from the point of view of classical logic standards, whereas the later belongs to more innovative experiments that employ the probabilistic approach. I will show how both approaches - despite giving us a better understanding of human cognitive abilities - are nonetheless not effective enough to describe the real-life picture of human’s rationality.

The most famous experiment aimed at disproving human rationality is Wason’s Selection Task. In this experiment participants are presented with four cards as

4 Experimental Insights in General Reasoning

well as a rule, and have to determine whether the rule is true or false. Here is one variant of it:

Below is depicted a set of four cards, of which you can see only the exposed face but not the hidden back. On each card, there is a number on one of its sides and a letter on the other. Also below there is a rule which applies only to the four cards. Your task is to decide which if any of these four cards you must turn in order to decide if the rule is true or false.



RULE: *If there is vowel on one side of the card, then there is an even number on the other.* (Stenning and van Lambalgen 2012: 44)

Wason has found that subjects performed very poorly on this task. Applying the classical notion of material implication $p \rightarrow q$, participants should only turn cards that are potentially logically incompatible with the conditional rule. Here p would stand for ‘a vowel on one side’, q for ‘an even number on the other’ and ‘ $\rightarrow$ ’ would represent the material implication of classical logic. In this way, the string $p, \neg p, q, \neg q$ would represent the upper arrangement of the cards. Cards that are incompatible with the conditional rule $p \rightarrow q$ would be those with p on one side and $\neg q$ on the other (as a conditional is false only if its antecedent is true and its consequent is false), i.e. the ‘A’ and ‘7’ cards in this case. As you can see below from the figure with the subjects’ choices, most of them responded correctly that the ‘A’ card must be turned over, but many also judged that the ‘2’ card must be turned over, despite the fact that the ‘2’ card is irrelevant as it could not falsify the claim no matter what is on the other side.

However, as many researches pointed out, Wason was too quick to jump to conclusions and judge people irrational. For example, Evans et al. (1996) and Evans & Lynch (1973) attributed the preference in p and q answers to the so-called *matching bias*, a tendency to select cards matching those named in the rule. In their own experiment, two researchers showed that it is challenging for some participants to realise that a $\neg q$ may have any logical relation to the rule

4 Experimental Insights in General Reasoning

p	p, q	$p, \neg q$	$p, q, \neg q$	misc.
35%	45%	5%	7%	8%

Figure 4.1: Typical proportion of people’s choices

“if there is vowel on one side of the card, then there is an even number on the other.” Introduction of explicit rather than implicit negations, however, in which the negative cards bear generic descriptions along the lines of the contrast set (like “a letter which is not a vowel”) does away with matching bias. This finding has also been confirmed by Sperber et al. (1995), who argue that explicit negations reduce the mental effort that goes into the task.

Johnson-Laird, Legrenzi, and Legrenzi (1972) showed that using familiar (i.e. closer to reality) terms instead of the original abstract (descriptive) terms simplifies the task significantly. Griggs and Cox (1982) came up with the same results when presenting subjects with a reasoning task about a drinking age rule: “If a person is drinking beer, then he or she must be over 19 years of age”. The discrepancy in the results can also be due to different interpretations of a conditional rule. The distinction concerns the detection of a rule violation. *Descriptive* conditionals describe states of affairs and can therefore be either true or false, so the rule can be proved or disproved. In this case participants are required to reason about its true value and choose instances that prove it true or false. *Deontic* rules correspond to norms and laws that are true from the start and can thereby be violated but not falsified. Here participants are required to reason about a rule which is always true and choose instances that could violate it. So, if we find someone who is drinking beer and is not over 19 years of age, we can either interpret it as a rule falsification (descriptive) or a rule violation (deontic). A descriptive interpretation may be prompted by a descriptive formulation of the rule, as in “If a person is drinking beer here, then he or she is over 19 years of age”. In this case an underage boy with a pint may be interpreted not as violating the rule, but as falsifying it.

Generally, a deontic interpretation is easier to evoke and process in conditional reasoning processes. This was confirmed by test results where participants showed greater improvement when conditional rules expressed deontic norms (Beller 2008;

4 Experimental Insights in General Reasoning

Cosmides & Tooby 2013; Cummins 1996; Stenning & Lambalgen 2012). That might be because of the so-called *cheater detecting mechanism* possessed by people, according to Cosmides (1989) and Cosmides and Tooby (1989). They argue that this mechanism helps us identify cases of cheating, which - from an evolutionary standpoint - comes in handy in social situations. The two researches developed the so-called *social contract theory*, a computational theory of social exchange between two or more individuals for mutual benefit. The cheater detection mechanism activates when people accept the social contract:

Specifically, cheating can be defined as the violation of a rule established, explicitly or implicitly, by acceptance of a social contract. A social contract relates perceived benefits to perceived costs, expressing an exchange in which an individual is required to pay a cost (or meet a requirement) to an individual (or group) in order to be eligible to receive a benefit from that individual (or group). Cheating is the failure to pay a cost to which one has obligated oneself by accepting a benefit, and without which the other person would not have agreed to provide the benefit (Cosmides, 1985). (Cosmides 1989 : 197)

According to Stenning and Lambalgen, the choice between deontic and descriptive interpretations happens during the reasoning *to* process mentioned in the last chapter - during the formation process of the appropriate logic. At the same time a great number of other parameters are also getting fixed. For instance, the logic of ‘not true’ in classical logic is the same as ‘false’, but interestingly some subjects refuse to see it like that even after an explicit instruction to do so. If they turn the cards ‘A’ and ‘7’, it will only tell them that the rule is not false, but will not make it true. Note that this is not possible in case of a deontic interpretation, as the truth of the rule is already there by default.

Some subjects have shown difficulties in selecting cards without having a chance to turn them over and see what’s on the other side. The confusion is caused by a thought that the cards are dependent on one another: should a subject that have just turned an ‘A’ card and seen an odd number stop or go on to turn the other cards? This is, again, not possible in the case of a deontic interpretation because the violation of one case has no connection with any other possible violations.

4 Experimental Insights in General Reasoning

Different outcomes of these experiments are strong evidence of the fact that people’s interpretation of the tasks varies from subject to experimenter and from subject to subject. Tasks are presented in natural language and natural language is subject to different interpretations, with its complexity and vagueness. Stenning and Lambalgen explain what is meant by interpretation:

Interpretation maps representation systems (linguistic, diagrammatic, etc.) onto the things in the world which are represented. Interpretation decides such matters as: which things in the world generally correspond to which words; which of these things are specifically in the domain of interpretation of the current discourse; which structural description should be assigned to an utterance; which propositions are assumed and which derived; which notions of validity of argument are intended; and so on. Natural languages such as English sometimes engender the illusion that such matters are settled by general knowledge of the language, but it is easy to see that this is not so. Each time a sentence such as “The presidents of France were bald” is uttered, its users must decide, for example, who is included, and how bald is bald, *for present purposes*. (Stenning & Lambalgen 2012 : 43)

The task of interpretation gets even more challenging when put out of context in a laboratory environment. This environment puts the subject in a classroom setting where they are forced to reason in a certain trained way. This is exacerbated if the task is formulated descriptively, while deontic conditional rules alleviate some of the problems caused by the experiment’s environment by making the task appear more relatable, especially if the subject matter is familiar and close to real-life situations. This all is a good illustration of the content-dependency of our reasoning processes, which results in an interpretational variety with different success level.

4.2 From Probabilistic Reasoning Approach

In the previous chapter I said that because of its non-defeasibility classical logic cannot cope with the incompleteness of the world, therefore we had better seek

4 Experimental Insights in General Reasoning

out defeasible systems to help us get by. The first people to contest the notion that defeasible logic is the basis of human reasoning were Oaksford and Chater who proposed that probabilistic theory might be more appropriate approach to describing human reasoning because it deals with uncertainty - they even called probability theory the calculus of uncertainty (Oaksford and Chater 1998).

Probability here is taken to mean degrees of belief. As has been shown above, belief updating, that is defeasibility, is crucial to human reasoning. When viewed from this perspective, the probabilistic approach does seem to account for some cases of human reasoning behaviour which classical logic has problems with. Thus the bird example from above, “If birds fly, and x is a bird, then x flies”, can now be interpreted in terms of conditional probability. So let’s say the conditional probability of that statement is high, say 0.9. So if all that we know is that something is a bird, then $P(\text{Fly}|\text{Bird}) = 0.9$. However, if it emerges that the bird in question is an ostrich, then the conditional probability of the statement would be extremely low, $P(\text{Fly}|\text{Bird} \wedge \text{Ostrich}) \approx 0$ (Oaksford and Chater 1998, 2001).

In general, Oaksford and Chater use the Bayesian approach to account for human reasoning, as it “relates probability and statistics most directly to the problems of belief updating, and hence has the most natural relation to cognitive processing” (Oaksford and Chater 1994, 1998 : 16). With this approach in mind they also provide an explanation of the results of Wason’s selection task. According to them, the selection task is a task of probabilistic *optimal data selection* in Bayesian statistics (Fedorov 1972; Lindley 1956). To solve the task participants have to gain the information needed to discriminate between two hypotheses related to the conditional rule ‘if p then q ’: in one hypothesis there is a dependency between the antecedent p and the consequent q (the dependence hypothesis, M_D); in the other, p and q are independent (the independence hypothesis, M_I). So, the participant’s job is to select cards in such a way as to obtain the most informative data that will reduce uncertainty about which one of these two hypotheses is true. This requires calculating the posterior probability of M_D and M_I - that is $P(M_i|D)$ - given some evidence. This can be done by using Bayes’ Theorem.

4 Experimental Insights in General Reasoning

So, for each hypothesis we have:

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{\sum_j P(D|M_j)P(M_j)} \quad (4.1)$$

As we proceed from the assumption that participants have not the slightest idea which of the hypotheses is true, we assign the prior probabilities to both of them at 0.5 ($P(M_D) = P(M_I) = 0.5$). For the calculation of the above formula we also need the likelihoods of evidence given a hypothesis (for example, the probability of finding ‘A’ (p) if turning ‘2’ (q), assuming the hypothesis M_D , $P(A|2, M_D)$). These can be calculated using the contingency tables in the following figure, modified by Oaksford and Chater (1998):

	q	$not-q$	
Dependence Model			
p	$a(1 - \varepsilon)$	$a\varepsilon$	a
$not-p$	$b - a(1 - \varepsilon)$	$(1 - b) - a\varepsilon$	$1 - a$
	b	$1 - b$	
Independence Model			
p	ab	$a(1 - b)$	a
$not-p$	$b(1 - a)$	$(1 - b)(1 - a)$	$1 - a$
	b	$1 - b$	

Note: $a = P(p)$, $b = P(q)$, and $\varepsilon = P(not-q|p)$.

Figure 4.2: Contingency Tables for the Dependence Hypothesis (M_D) and the Independence Hypothesis (M_I)

Information gain (I_G) is then measured by the difference between the prior and posterior probabilities, i.e. by the change in the “uncertainty-level” after turning over a card:

$$I_G = I(M_i) - I(M_i|D) \quad (4.2)$$

In reality, however, no turning over is taking place, which, as was mentioned in the previous section, is confusing for some subjects. Non-confused subjects have to make their decision according to their expectations about what’s on the other side of the card. This will require calculating the probabilities for any outcome

4 Experimental Insights in General Reasoning

possible for any of the four cards (for example, the probability of finding ‘A’ (p) and also the probability of finding ‘B’ ($\neg p$) if turning the ‘2’ card(q)).

Assuming that a and b are quite small, the p and q cards, the most frequently chosen cards in the selection task, would have a higher *expected information gain* than the others and therefore provide the most reduction in uncertainty about which hypothesis is true. Moreover, Oaksford and Chater (1994) generalize that rare evidence is more supportive than common one, and call it the *rarity assumption*. For example, “red wine” as a statement is more supportive than “wine” because it provides more information gain. Quoting McKenzie (2000 : 7), the rarity assumption “is resumed to exist because of lifelong learning that presence is rarer than absence”.

In that way Oaksford and Chater come to the conclusion that the Information Gain Model justifies and “rationalizes” people’s choices. Consequently, the selection task which was aimed at demonstrating people’s “irrational” behaviour served as a good demonstration of people’s rational data selection strategy:

So, rather than representing a blatant example of human irrationality, performance on this task can be viewed as an example of human *rationality*. Crucially, our account reconciles the paradox between the apparent irrationality of human performance on the selection task and the manifest success of human reasoning in the everyday, uncertain world. (Oaksford and Chater 1998 : 20)

Two other prominent proponents of the probabilistic reasoning approach, Kahneman & Tversky, have noticed, however, significant and extensive violations of the probabilistic paradigm in human reasoning and decision making processes. Firstly, they observed that when dealing with incomplete and uncertain information, specifically under complicated circumstances, people tend to use the so-called *heuristics*, rule-of-thumb, which allows them to evaluate information in a sufficient way. Secondly, they observed that people have certain biases that often cause them to pass over relevant information in favour of irrelevant one. These observations gave rise to a ‘heuristics and biases’ approach and were initially described by Kahneman & Tversky (1975). In 2002, one member of the pair, psychologist Daniel Kahneman, was awarded the Nobel Memorial Prize for

4 Experimental Insights in General Reasoning

describing errors, anomalies and traps in human mind when making decisions (the prize was interestingly in Economic Sciences)

One of the common rule violations described by Kahneman & Tversky is the *conjunction fallacy* in probabilistic reasoning, demonstrated in the famous Linda Problem (Tversky and Kahneman 1983):

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

- (a) Linda is a teacher in an elementary school
- (b) Linda works in a bookstore and takes Yoga classes
- (c) Linda is active in the feminist movement
- (d) Linda is a psychiatric social worker
- (e) Linda is a member of the League of Women Voters
- (f) Linda is a bank teller
- (g) Linda is an insurance salesperson
- (h) Linda is a bank teller and is active in the feminist movement

In Kahneman and Tversky's study, 85% of subjects rated (h) as more probable than (f), which is a violation of a fundamental law of probability which states that the probability of a conjunction must be less than or equal to the probability of its conjuncts, i.e. for all A and B , $P(A \wedge B) \leq P(A)$. Compare this to the previously mentioned rarity assumption, which can be adopted to the Linda Problem in the following way: the appearance of two (or more) events is rarer than the appearance of one (or less).

The fact that most subjects chose (h) over (f) led the authors to the conclusion that people use one of their heuristics, namely the representativeness heuristic, i.e. they attribute the properties of Linda as being representative to feminists and therefore choose the description that describes her as a feminist. And as a result the use of this heuristics causes them to reason incorrectly. Subsequently, however, this view has suffered a similar fate to Wason's Selection Task as it increasingly came under scrutiny. It was argued, for example, that certain linguistic

4 *Experimental Insights in General Reasoning*

factors block the use of the aforementioned conjunction rule. In particular, Morier and Borgida (1984) pointed out that the existence of (h) “Linda is a bank teller and is active in the feminist movement” might cause (f) “Linda is a bank teller” to be interpreted as “Linda is a bank teller and is *not* active in the feminist movement”. In an effort to avoid such an interpretation, the authors included “Linda is a bank teller and is not active in the feminist movement” among the options in the task, but this did not trigger any significant reduction in the conjunction fallacy. This interpretation of (f) might be caused by two factors. First, there is a conversational principle, namely the maxim of quantity, proposed by philosopher Paul Grice (1975), according to which it’s expected that the speaker should be as informative as possible: if Linda were a feminist they would have said so, if they didn’t then she is not a feminist.² Secondly, the exclusive (not-nested) choices like (h) are more transparent and help us better understand each other: which would you prefer after all, wine or a Cabernet Sauvignon?

It can also be the case that people don’t judge probabilities here at all. In fact the conjunction fallacy decreases when the presented information is based on frequencies rather than on making predictions about one person. Thus, Gigerenzer (1991) came up with an updated new version of the Linda Problem, in which, after the original description part, instead of asking his subjects to rank the likelihood of the two statements about Linda, he asked his subjects how many people out of 100 fitting that description are (1) bank tellers and how many are (2) bank tellers and are active in the feminist movement. With the help of that hint, all the participants gave the correct answer, rating (1) higher than (2).

There are problems with the frequency approach, too. One of them is the so-called reference class problem, nicely illustrated by Hájek (1996):

Consider a probability concerning myself that I care about - say, my probability of living to age 80. I belong to the class of males, the class of non-smokers, the class of philosophy professors who have two vowels in their surname, ... Presumably the relative frequency of those who live to age 80 varies across (most of) these reference classes. What, then, is my probability of living to age 80? It seems that there is no single frequentist answer. Instead, there is my probability-qua-

²More about conversational maxims to come in Chapter 6.

4 Experimental Insights in General Reasoning

male, my probability-qua-non-smoker, my probability-qua-male-non-smoker, and so on.

Another problem with the frequency approach - perhaps better illustrated with the classic example of flipping a coin rather than the Linda Problem - is the problem of the single case, i.e. it is hard to see how frequencies can be applied to events which are in principle non-repeatable such as John F Kennedy's assassination, the Euro Cup finals or the Russian Civil War.

4.3 Conclusion

In order to test human's rationality, researchers have put on a series of experiments which ended up convincing them in human's irrationality as results showed an extreme deviation from what was expected by the normative standards of rationality. Using an example of the Wason's Selection task and the Conjunction Fallacy I have shown that irrationalities are only that in respect to the approach that is chosen or is provided when reasoning (that's not to say, however, that people cannot have contradictory beliefs and ways of thinking). The inconsistencies found in Oaksford and Chater's approach, for instance, do not discount their theory completely, they just mean that humans will always appear irrational when judged from some other vantage points. And even within the formal system that one chooses at any given moment (in the *reasoning to* stage, see above) there is always an is/ought problem or competence/performance discrepancy between the stated goal of this or that reasoning activity and its actual results - meaning simply that the person can fail to apply their chosen logic correctly (in the *reasoning from* stage).

There is, consequently, no reason whatsoever to reduce human reasoning to just one approach or one system of logic. And while it is true that humans have been shown to prefer non-monotonic systems of logic in their everyday reasoning as well as in special "laboratory" tasks such as Wason's Selection task, it doesn't rule out the possibility that in certain situations people may well use monotonic logic. The fact that meat eaters by definition eat meat does not mean that they don't eat vegetables; but the opposite cannot be said of vegetarians.

*“I used to lie and cheat,
Lie and cheat and step on people’s feet,
Now I’m standing on this corner,
Salvation is my beat.”*

The Band, “Saved”

5

Reasoning: where and when?

In this chapter I will discuss the mental architecture of logical reasoning and afterwards look at the reasoning from the temporal perspective. These perspectives can contribute greatly to the discussion of the principal topic of this study, namely the origins of non-monotonicity, the question of whether it first emerged in language or in general cognition. I will focus both on the findings in relation to the structure and evolution of reasoning as well as on the developmental processes in children.

5.1 Structure of Reasoning

Reasoning and human psychological mechanisms have been the focus of evolutionary psychology, a relatively new field created in the 1980s and 90s by Leda Cosmides and John Tooby which studies the relation between human minds and human evolution and combines evolutionary theory and cognitive psychology. Much of the debate in evolutionary psychology revolves around the question of whether human reasoning is domain-general or domain-specific.

The concept of domain specificity was first advanced and elaborated on by Jerry Fodor in his ground-breaking book *The Modularity of Mind* (Fodor 1983). According to Fodor, “domain specificity has to do with the range of questions for

5 Reasoning: where and when?

which a device provides answers (the range of inputs for which it computes analyses)” (Fodor 1983 : 103). In other words, domain specificity describes devices, or modules, devoted to tackling a specific range of questions; the narrower the range of questions, the more domain-specific the module. While Fodor declines to give a clear-cut definition of modules, saying he is not in the business of defining his terms, he understands them as “domain-specific computational mechanisms” which are “innately specified, hardwired, autonomous, and not assembled” (Fodor 1983 : 37). The most generally accepted examples of modules are sense modalities like vision and audition, as well as language-processing, face-recognition, and various motor-control systems.

In Fodor’s account, domain specificity is just one of nine features that characterise modules, the others being mandatory operation, limited central accessibility, fast processing, informational encapsulation, ‘shallow’ outputs, fixed neural architecture, characteristic and specific breakdown patterns and characteristic ontogenetic pace and sequencing. I’m not going to go into detail about these because they are of no concern to us at this point, and also because the proponents of evolutionary psychology such as the aforementioned Cosmides and Tooby differ significantly from Fodor in their understanding of modularity. For one thing, Cosmides, Tooby and those who followed them (Sperber 1994; Carruthers 2006) hold that the mind, including reasoning, is modular through and through, while Fodor believed that some reasoning is general-purpose and content-independent. Carruthers, in particular, dismisses most of the properties associated with Fodor’s modules and says: “That leaves us with the idea that modules might be isolable function-specific processing systems, all or almost all of which are domain specific <...>, whose operations aren’t subject to the will, which are associated with specific neural structures (albeit sometimes spatially dispersed ones), and whose internal operations may be inaccessible to the remainder of cognition” (Carruthers 2006 : 12).

One example of such an “isolable function-specific processing system” is the aforementioned *cheat-detecting mechanism*, which was hypothesised by Cosmides and Tooby. The researchers, whose views have come under repeated criticism since they first put forward their theory in the early 1990s, proceed from the assumption that the human mind is a product of natural selection, meaning that in order to survive our ancestors had to solve an increasing number of adaptive

5 Reasoning: where and when?

problems (finding shelter, food, mates etc.). One such adaptation is exactly the cheat-detecting mechanism, which is described as an algorithm or a set of inference procedures specialised for detecting cheaters in situations of social exchange. Doing this has proven evolutionarily crucial because human well-being usually involves a ‘social contract’ between individuals whereby they cooperate for mutual benefit; it is thus imperative to single out those cheating on the contract in order for the cooperation to work properly. The validity of the cheat-detecting mechanism becomes evident when comparing how people perform on abstract (the original Wason selection task) vs familiar tasks (the drinking age rule). That performance on the latter is better is the direct result of putting the person in a situation of observing a “social contract”. As Sugiyama, Tooby and Cosmides (2002) wrote in their research on cognitive adaptations among the Shiwiar of Ecuadorian Amazonia:

Selection pressures favoring social exchange exist whenever one organism (the provisioner) can change the behavior of a target organism to the provisioner’s advantage by making the target’s receipt of a provisioned benefit conditional on the target acting in a required manner. This mutual provisioning of benefits, each conditional on the other’s compliance, is what is meant by social exchange or reciprocation.

Evolutionary biologists have shown through game-theoretic techniques that adaptations for social exchange can be favored and stably maintained by natural selection, provided they include design features that (i) enable them to detect cheaters (i.e., those who do not comply or reciprocate), and (ii) cause them to channel future benefits to reciprocators, not cheaters. (Sugiyama, Tooby, Cosmides 2002 : 11537)

So, for Cosmides, Tooby and their creed the fact that reasoning in deontic norms - i.e. seeking out violations of the rule - is easier to evoke than reasoning in descriptive norms points unequivocally to the domain-specificity and massive modularity of mind (Cosmides 1989; Cosmides & Tooby 2013; Barrett & Kurzban 2006 etc.). On this view, the “adaptive” nature of human reasoning is proof that there is no general capacity for reasoning; rather, reasoning exists in specific modules designed to solve this or that problem.

5 Reasoning: where and when?

There are other researchers who likewise argue for the domain specificity of mind, but who do not link it to modularity. For instance, Cummins (1996) argues that the domain-specific nature of mind is determined by canalisation, an evolutionarily predetermined process that is independent from environmental changes. In his opinion, there exists a specific mental structure devoted to processing deontic norms, which children possess already from an early age, and which is ubiquitous. Opponents of the massive modularity approach argue for the existence of specialised deontic domains, one for social exchange (or reciprocation), another for hazard management and yet another for altruism detection. These domains are believed to be both acquired and common to all cultures. Empirical evidence shows that children express social contract starting from the age of 3-4 years. Thanks to the hazard management mechanism people avoid danger by following certain precautionary rules. The altruism detection mechanism is required in situations where one is behaving altruistically, i.e. taking the price without any benefits. The existence of the altruism detection domain is often called into question. According to Cosmides, “altruists” would not survive in this world, and if they don’t exist, there is no need for an “altruism detector” to evolve. Empirical evidence on this topic has brought forth contradictory results (Oda, Hiraishi, & Matsumoto-Oda 2006; Thompson et al. 2013).

The concepts of the domain specificity and modularity of reasoning are not universally accepted. A large number of scholars, including Fodor himself (Fodor 2000; Chater & Oaksford 1996; Jackson & Griggs 1990; Noveck & O’Brien 1996 etc.), claim that reasoning is by contrast domain-general (in Fodor’s case, partly domain-general); they have consequently put forward alternative explanations for the different performance in deontic and descriptive tasks. To them, reasoning relies on formal logical mechanisms (form rather than content), and the discrepancy in deontic and descriptive performance is grounded in the different inferential processes.

However, the domain-general account runs up against a lot of difficulties. For the domain-general theorists some phenomena are hard to explain - why, for instance, does alternating content make such a profound influence on our reasoning? One way to explain the seemingly domain-specific processes cited by the proponents of evolutionary psychology is by referring to the distinction between the two types of information processing: active and automatic (also called System 1

5 Reasoning: where and when?

and System 2 - more on that in the next section). On this view, when faced with a new task, people are thought to be using domain-general processing; when doing the same task again, however, there is no need anymore to engage in active processing as we can just automatically retrieve previous solutions. Such a shift can be seen in children from learning how to count on fingers to subsequently retrieving answers from memory (Siegler & Shipley 1995).

Stenning and van Lambalgen (2005, 2008, 2012) chose perhaps the safest ground to stand on in this complicated debate, cheering as they do for both the domain-general and domain-specific accounts. While they don't call into question the existence of difference modules, neither do they believe that one type of reasoning should necessarily be bound to one module. In their opinion, "closed world reasoning occurs in such diverse domains as action planning, natural language interpretation and attribution of belief" (Stenning and van Lambalgen 2005 : 20).

As I have described above (see Chapter 3), Stenning and van Lambalgen conceptualise reasoning as a two-stage process - 'reasoning to' and 'reasoning from' - whereby the reasoner first decides what logic corresponds to the proposed domain and then sets about solving the task using that logic. 'Reasoning to' can therefore be said to be a domain-general procedure; it initiates a search for parameters ("in which domain are we exactly?") and generates logical laws suitable for the domain, while 'reasoning from' is an instance of domain specificity ("aha! so in this domain I'm reasoning *this* way"). That said, the authors note that many domains share general reasoning procedures (the authors argue, for instance, that some of the procedures underlying human planning can be used to construct models of linguistic discourse, such as structuring the plot of a story). In the same way, while they agree with the proponents of domain-specificity about the context-dependency of human reasoning, Stenning and van Lambalgen emphasise that this is only true to some extent because as reasoners we tend to generalise rules.

Take the situation with the dentist from Chapter 3, for instance. You show up at the dentist's at 3:30 pm, outside his listed working hours, the door is shut. Even though it's not stated explicitly that he doesn't see patients at this time, using closed-world reasoning you conclude that he doesn't. It's then a natural thing to generalise and apply this to all doctors and similar eventualities, effectively rendering the situation topic-neutral. According to the authors, this again is

5 Reasoning: where and when?

proof that domains share some generalities across the board.

To test the validity of both the domain-specific or the domain-general approaches, Ferrara, Fabrizio et al. (2015) ran a series of experiments. These experiments showed deontic tasks eliciting more correct replies than descriptive ones by a large margin. According to the authors, this was due to two factors: a) the complexity of the descriptive tasks which require additional metareasoning abilities, and most importantly, b) pragmatics and context that helped the participants, as in “the presence of explicit negations, the wealth of information conveyed by the scenarios, the presence of instructions that explicitly demand to look for the potential rule violators” (Ferrara, Fabrizio et al. 2015 : 220). As the authors show, controlling these pragmatic effects worsens the participants’ performance in deontic tasks. Therefore, the results do not fully bear out either approach; what they do show convincingly however is that human reasoning is massively context-dependent - demonstrated by the much better performance in the deontic tasks that employed the aforementioned pragmatic cues.

5.2 Do Reasoning Skills Develop?

In the previous section I gave an overview of some of the research done by evolutionary psychologists, whereas in this section I will zoom in on the individual and review the discussions about reasoning in developmental psychology. Research literature on reasoning competence in children compared to adults is varied, albeit not particularly large.

In order to see if any development happens, we first need to know what to look at, in other words what the process of reasoning involves.

According to Moshman (2004), the key aspect of reasoning - something that sets reasoning apart from making inferences and thinking - is its metacognitive nature. What he means by that is relatively conscious awareness of norms of reasoning and control of logical inferences. Here is how he defines metalogical understanding:

Metalogical understanding - conceptual knowledge about logic - includes (1) awareness of inference as a process that generates conclusions from premises; (2) understanding that some inferences are better than others; (3) knowledge about the logical properties of proposi-

5 Reasoning: where and when?

tions, inferences, and arguments; and (4) conceptualizations of logic as an epistemic domain. (Moshman 2004: 225)

Lack of awareness of inference was noted in children of an early age. Children do draw inferences before they come to understand the norms and properties of logical inferences but their reasoning is weak. For example, the believability of content and validity of form in conditional statements like “If it’s raining, then the street is wet” are strongly connected for them. Starting at about age 6, the dependency diminishes but even some adults have trouble drawing correct inferences if the facts do not add up.

With the awareness of inference comes the understanding that different inferences can lead to different outcomes. In a series of experiments Pillow (2002) showed that both children and adults understand the certainty and logical necessity of deductive inferences, i.e. when the truth of premises guarantees the truth of the conclusion, but when it comes to inductive and other logically not necessary inferences such as guessing etc. adults show a better understanding of differences between those than children. Research indicates that along with logical awareness, the understanding of certainty, necessity and possibility also begins at about age 6 (Ruffman 1999; Somerville, Hadkinson, Greenberg 1979; Tunmer, Nesdale, Pratt 1983).

The main conclusion made by developmental psychologists is that there is a transition from inference to reasoning, and that this transition takes the form of the development of explicit knowledge about inferences. This line of thinking was first advanced by Piaget (1985) who, as mentioned in the Introduction, held that human logical abilities are naturally constructed over time. Unlike Piaget, however, modern-day developmental psychologists do not assume some universal sequence of developmental stages that lead to rationality and maturity.

If norms of reasoning sanction logically necessary inferences, then these inferences are deductive; if they sanction probable but not necessary inferences, then the inferences are inductive. Deductive and inductive inferences are two main types of inference in reasoning. The logical necessity of deductive inferences means that they don’t allow for exceptions, and are thus monotonic. Inductive inferences on the other hand are non-monotonic because they cannot guarantee the truth of the conclusion as new information might come to light that would

5 Reasoning: where and when?

Deductive Arguments Are Monotonic			
Base Argument		Expanded Argument	
(a)	Mammals have property X — Cows have property X	(b)	Whales have property X Mammals have property X — Cows have property X
		(c)	Mammals don't have property X Mammals have property X — Cows have property X
Inductive Arguments Are Nonmonotonic			
Base Argument		Expanded Argument	
(d)	Cows have property X — Mammals have property X	(e)	Whales have property X Cows have property X — Mammals have property X
(f)	Whales have property X — Cows have property X	(g)	Horses have property X Whales have property X — Cows have property X

Note: The additional premise in expanded arguments (b) and (c) neither adds to nor diminishes the strength of the conclusion in base argument (a). All three arguments are valid. By contrast, the expansion of base arguments (d) and (f) into arguments (e) and (g), respectively, increases the strength of the conclusion.

Figure 5.1: Deductive and inductive arguments and the property of monotonicity

render the conclusion false or make it stronger.

Above is a table from Chapter 13 of Handbook of Child Psychology and Developmental Science (Lerner, 2015) demonstrating the workings of deductive and inductive inferences in terms of monotonicity.

Many kinds of our thinking are inductive in nature because, as was amply shown in this paper, our reasoning proceeds under conditions of uncertainty. That's why the development of inductive reasoning is of particular interest to the research of the origins of non-monotonic logic.

In describing the course of development of inductive and deductive reasoning many researchers resort to dual process accounts. Originally proposed by Wason

5 Reasoning: where and when?

and Evans (1975), the dual process theory separates between two distinct cognitive systems with different evolutionary histories, usually marked as *System 1* (or Type 1) and *System 2* (or Type 2):

System 1 processes are fast, automatic and largely unconscious, evolutionarily old and shared with animals;

System 2 processes are slow, controlled, conscious and analytic with a use of working memory systems, evolutionarily recent and unique for humans.

In these theories logical reasoning is ascribed exclusively to System 2 processes. However, Stenning and van Lambalgen (2012) argue that although System 1 processes are fast and automatic, they still proceed according to logical laws (albeit nonclassical) and are defeasible, defeasibility being a property of nonclassical logic. As a prime example of this they offer discourse interpretation: it proceeds according to nonclassical logic and “involves the construction of representations in central working memory” (2012: 124). Discourse interpretation may interchange between both System 1 and System 2 processes depending on the interpretation performance, i.e. when additional computation power is required, System 1 changes to System 2 before System 1 resumes operation.

Those who ascribe logical reasoning to System 2 processes do so by arguing that metacognition and metalogical competence - the key aspect of reasoning, as was previously said - is a complex process associated with System 2 processing. Thus, according to them, System 2 processes are crucial for the development of reasoning. Particularly, for the development of deductive reasoning, as was mentioned earlier in this chapter, adult children and adults are much more likely to conform to such deductive norms as truth preservation, necessity, validity, consistency and monotonicity and they make use of deductive strategies (e.g. search for counterexamples) more often than young children. Regarding inductive reasoning, the fact that older children and adults show better results in inductive inferences than young children is accounted for by greater control over attentional and inhibitory processes and greater involvement of working memory resources, which are carried out through the System 2 processing.

As discussed in Lerner (2015), System 1 processes are not only automatic; because they are often focused on problem-solving, they are also heuristic and

therefore largely domain-specific in nature. System 2 processes, by contrast, are domain-general as they operated on a more abstract, decontextualised level.

5.3 Conclusion

In this chapter I shifted the attention to the question of the underlying structure of human reasoning and to the related topic dealing with the development of the reasoning processes. Empirical evidence shows conclusively that human reasoning is mostly content-dependent, although a certain number of reasoning tasks can be done using general-purpose mechanisms. The question of content-dependency is closely bound up with the question of whether human reasoning is domain-general or domain-specific. For some, content-dependency is enough proof of the domain-specificity of human reasoning, in the sense that specific tasks call for specific reasoning procedures. This stands in contrast to the domain-general view which holds that the same procedures are applied to all tasks across the board. This view explains the content-dependency of reasoning in terms of System 1 and System 2 processes, whereby once you have learned a specific task (System 2) you can apply the solution to all similar tasks, irrespective of the content (System 1).

The key aspect of reasoning is its metacognitive nature. That's why reasoning abilities at the very young age, when children not yet fully developed metacognition, are limited and weak. The development of reasoning comes from the acquisition of drawing inferences with different outcomes. Logically necessary deductive inferences are acquired before logically unnecessary and non-monotonic inductive inferences. Our reasoning develops as we increasingly master System 2 processes.

This discussion will benefit greatly from an analysis of the evolutionary aspects of reasoning, something that I have touched upon at some length when talking about the field of evolutionary psychology created by Cosmides and Tooby. Understanding the origins of human reasoning from an evolutionary perspective can help shed light on the correlations between different cognitive faculties and their developments.

“I’ve changed my mind, it’s what it’s there for.”

John Balance, “Ether”

6

Non-Monotonicity in Linguistics

6.1 Introduction

So far I have been talking about reasoning and (non)monotonicity in the context of general cognition. In this chapter, I will show the link between non-monotonicity and the language system. In relation to the linguistic phenomena the term “non-monotonicity” indicates two factors. On one hand, it is the possibility of canceling statements in the light of new information. Usually this cancellation happens by means of special correction phrases (like ‘in fact’, ‘if not’, ‘or even’). On the other hand, it is the listener’s pattern of reasoning or train of thought, if you like, when s/he makes inferences based on what was said and, more importantly, what was not said. Although such non-monotonic behaviour can be found all around in linguistics, here I will focus on its applications in conversational implicatures, particularly scalar implicatures.

6.2 Scalar Implicatures

6.2.1 General Observations

We do not have to go far away to encounter the linguistic phenomenon of non-monotonicity, let's look at our every-day communication. Consider the following typical example:

- (9) A: I like Mary. She's intelligent and good-hearted.
B: She's intelligent.

From B's utterance one would infer that B thinks that Mary is not good-hearted. But why is that? From the fact that B thinks that Mary is intelligent does not follow that s/he thinks that Mary is not good-hearted. Hence such reasoning follows from what B did not say. It is plain to see that this course of reasoning resembles the Negation of Failure principle, which was introduced earlier, whereby we make assumptions from what is not said and make inferences that do not directly follow from the literal meaning of what is said. Inferences of this type in linguistics are very common and are called *conversational implicatures*.

Conversational implicatures were first described by Paul Grice (1975), who claimed that they play an essential role in discourse interpretation. The main idea behind them is that participants of the conversation have to be responsible to each other and follow some rules to achieve a successful communication. Grice proposed a number of *maxims of conversation* for the speaker, seen as *defaults*, the use or violation of which gives rise to a conversational implicature:

- (1) the maxim of quantity - give as much information as needed, no more
- (2) the maxim of quality - be truthful, do not give false or unreliable information
- (3) the maxim of relevance - be relevant
- (4) the maxim of manner - be clear, brief and orderly, avoid obscurity and ambiguity

If B in the example (9) says that Mary is intelligent that means that according to the maxim of quantity it's all s/he can say about Mary, thus s/he does not

6 Non-Monotonicity in Linguistics

think that Mary is good-hearted, otherwise s/he would have said so trying to be as informative as possible.

And while they have been widely disputed, in the years that followed these maxims have served as an explanation for many linguistic phenomena.

The most commonly studied conversational implicatures based on one of the above maxims, namely the maxim of quantity, are *scalar implicatures* (henceforth SIs). SIs suggest that, according to the maxim of quantity, the speaker uses a particular term on a scale so as to be maximally informative. There are plenty of examples of such scales, here is one from Levinson (1983: 134):

<all, most, many, some, few>
<and, or>
<n, . . . , 5, 4, 3, 2, 1>
<excellent, good>
<hot, warm>
<always, often, sometimes>
<succeed, Ving, try to V, want to V>
<necessarily p, p, possibly p>
<certain that p, probable that p, possible that p>
<must, should, may>
<cold, cool>
<love, like>
<none, not all>

Monotonicity properties of contexts are believed to determine which scalar term leads to a stronger statement. These properties are: *upward monotonicity* and *downward monotonicity*. A predicate p is *upward monotone* (or *monotone increasing*) if for any two sets x and y the following holds:

$$\text{Iff } x \subseteq y, \text{ then } p(x) \subseteq p(y) \quad (6.1)$$

For example, the environment ‘Mary heard_’ is upward-monotone. It is easy to test: suppose we insert ‘Mozart’s Symphony No. 40’ as x and ‘music’ as y , in this case (10a) entails (10b):

6 Non-Monotonicity in Linguistics

- (10) a. Mary heard Mozart's Symphony No. 40.
 b. Mary heard music.

A predicate p is said to be *downward monotone* (or *monotone decreasing*) if the opposite holds:

$$\text{Iff } x \subseteq y, \text{ then } p(y) \subseteq p(x) \quad (6.2)$$

This relation can be found for instance in negative contexts: 'Mary didn't hear _' shows that property which results in (11b) entailing (11a):

- (11) a. Mary didn't hear Mozart's Symphony No. 40.
 b. Mary didn't hear music.

Environments that are neither upward- nor downward-monotone are *non-monotone*.

The use of a weaker term in an upward-monotonic environment suggests that a stronger sentence with a corresponding stronger term would be false. If one utters (12), the listener will infer (13); if that had not been the case, the speaker would have used a stronger term. Here again we are faced with a situation where lack of information corresponds to negative information.

- (12) Some of the boys went to the party. (Levinson 1983: 133)
 (13) Not all of the boys went to the party.

The scalar nature of SIs makes it possible to easily defease them, i.e. an additional fact can cancel them out without any sense of contradiction - a characteristic of inferences in non-monotonic logic. The cancellation happens by means of a certain phrase such as 'in fact', 'if not', 'or even' and alike and followed by an item higher up the scale. The following examples illustrate the defeasibility of 'some' and 'most':

- (14) Some, and possibly all/ in fact all/ if not all, students are coming to the party.
 (15) Most, and possibly all/ in fact all/ if not all, students are tired.

Likewise, it seems that number words behave in a similar way. Consider for example:

(16) Mary ate three ice creams, if not more/ in fact more/ in fact four.

There are several theoretical positions regarding the relation between grammar and conversational implicatures including SIs. *Griceanism* views conversational implicatures as being generated globally, i.e. after an utterance has been assigned a (truth-logical) meaning by grammar. *Neo-Griceanism* is similar to Griceanism in this sense but insists on a reduced or joined number of maxims (Q principle - say as much as you can [given R] and R principle - say no more than you must [given Q]) (e.g. Horn 1984). In corresponding literature these two views are united under the name *Globalism*.

An alternative view hold the proponents of the *Relevance Theory* (Sperber and Wilson 1986), according to which the implicature arises only if it's required to maximise relevance. They emphasize that it happens not by means of pragmatics and numerous maxims, but of a single principle, specifically relevance. The hearer makes a cost-benefit analysis examining all the alternative propositions and chooses the proposition with the highest relevance ratio. According to the ration of the meaning of 'some' in a given context, the implicature may or may not arise.

More recently, one more theory has appeared, namely *Localism* in contrast with Globalism (e.g. Chierchia 2004; Landman 1998). Under this position, SIs are not computed after grammar has done its assigning job, but are derived as part of and in tandem with compositional meaning, whereby conversational implicatures can be generated in subparts of a sentence (words or phrases), not necessarily the entire sentence. Essentially, they argue that SIs are part of core grammar. Chierchia et al. (2004) propose to represent SIs grammatically by adding an exhaustivity operator *O*, whose meaning is similar to the meaning of the operator *only*. On another so-called *Defaultism* view - which nevertheless fits in with the localist view - SIs are default inferences which the listener infers automatically unless there are reasons not to (Levinson 1995, 2000).

According to another fairly new account, namely an *interpretation-based account*, the Neo-Gricean perspective is not complete enough: it needs an additional step of calculating the pragmatically strengthened meaning of the premises on the basis of formal alternatives (Mascarenhas 2013). Although the creator of this theory abstains from choosing between localists and globalists, he admits that his results "translate into equivalent results in a localist theory of scalar implicature"

(Mascarenhas 2013 : 10).

To sum up, there are two major schools of thought here: Pragmatical View vs. Grammatical View. In the next section I will present empirical evidence for their theoretical assumptions.

6.2.2 Empirical Evidence

Numerous experiments have been conducted addressing the question of whether the interpretation of SIs is an automatic and fast process or a slow and controlled one. The claim that SIs are interpreted automatically and fast is central to Defaultism, which states, more precisely, that SIs are generated by default, i.e. blindly - unless something blocks that interpretation.

Here is how Chierchia (2001) explains the Defaultist theory:

On the whole, the theory predicts that sentences have certain default interpretations. Defaults (when incompatible with the context) can be overridden. This happens in two ways. First, if a scalar implicature is absent in a positive context, backtracking takes place. We first try incrementing the context with the strong meaning and when that fails, we revert to the plain one. Second, if a (direct) scalar implicature seems to be maintained in a negative context, accommodation takes place. I.e. the implicature is not added in at the first pass, but accommodated at some point. Besides the purely linguistic evidence we have considered, these expectations might be (dis)confirmed through psycholinguistic experimentation. (Chierchia 2001 : 24)

Evidence from psychological research on human reasoning (Evans and Newstead 1980; Paris 1973) suggests, however, that SIs should not be considered automatic and fast. For example, one study has found that when asked to determine in a pair of pictures whether simple unrelated statements connected with the scalar term ‘or’ (as in ‘The bird is in the nest, or the shoe is on the foot’) are true, two-thirds of participants preferred the inclusive interpretation of ‘or’, which goes against the defaultist view (Paris 1973). Similar conclusions were drawn in experiments with the scalar term ‘some’, in which participants were asked to determine if sentences of the form ‘Some As are Bs’ are true in situations where all As are Bs

(as in ‘Some elephants have trunks’) (Bott, Noveck 2004; Geurts, Pouscoulous 2009). If the process were automatic, ‘some’ would imply ‘not all’ in the absence of evidence to the contrary, whereas in these studies the negative responses were recorded in 59% and 34% of the time respectively.

Breheny et al. (2006) came to the similar results when analyzing reading times of contexts both triggering and cancelling SIs. In their study the participants were presented with sentences of two types (the materials were presented in Greek): in a context where the content of the implicature were relevant (upper-bound context) and in a context where it was not relevant (lower-bound context):

- (17) Upper-bound context: John was taking a university course and working at the same time. For the exams he had to study from short and comprehensive sources. Depending on the course he decided to read the class notes or the summary.

Lower-bound context: John heard the textbook for Geophysics was very advanced. Nobody understood it properly. He heard that if he wanted to pass the course he should read the class notes or the summary.

(Breheny et al. 2006 : 17)

If SIs are derived automatically and later on cancelled that would result in longer reading times of not relevant contexts; if they are derived only in relevant contexts that would result in longer reading times of relevant contexts than not relevant contexts where they are not derived. The authors found that the reading time in the upper-bound contexts was significantly longer than in the lower-bound context, that way confirming the non-defaultist account.

The results of experiments carried out with children also confirm the non-default nature of SIs. The importance of these experiments also lies in the fact that they approach the question from a developmental perspective. Children seem to be even ‘worse’ in interpreting SIs than adults. Like adults, they treat ‘some’ as compatible with ‘all’ and allow for an inclusive-or interpretation of the disjunction, but to a larger degree. These observations come from a number of studies. Firstly, Smith (1980) tested 4-7 year old children: her studies showed that children accepted statements like ‘Some elephants have trunks’ where ‘all’ would have been a more appropriate and informative term. Later Noveck (2001)

broadened Smith’s experiments by testing people at three different ages: 7, 10 and adults. His results showed that children at the ages of 7 and 10 were much more tolerant to true but pragmatically infelicitous sentences than adults (89% of negative responses for 7 years olds, 85% for 10 years olds, and 40% for adults), which can be interpreted as a developmental process, in which people over time improve their pragmatic abilities. Noveck came to similar results for statements with the term ‘might’ and the more informative ‘must’. As stated by Noveck himself (Noveck 2001: 183), “children are rather strong at recognizing when utterances are overinformative relative to the given situation or fact (as in All birds live in cages) but are not so determined when they are required to detect an infelicitous underinformative utterance (as in ‘Some elephants have trunks’).” Later experiments with modified tests also revealed similar results (Papafragou and Musolino 2003; Papafragou and Tantalou 2004).³

As we have seen, none of these results conform to the Defaultism, and in general there is no experimental support for it in literature (see Geurts 2010 for an overview).

Another hot topic between the two opponent schools of thought is context-dependency of scalar inferences.⁴ For localists and defaultists, as their names suggest, SIs are triggered locally or by default dependent on the constituent, not on the context. It does not mean, however, that context is not important; in fact context is always important, but under the latter view, context comes into the picture only after the implicature has been triggered and can cancel it out afterwards depending on the speaker’s goals. This is demonstrated in the Levinson’s (2000) example (18), where (18-c) is triggered but later on cancelled because the respondent understands that the one who puts the question is only interested in whether ‘at least some’ of the documents are forgeries:

- (18) a. Is there any evidence against them?
 b. Some of their identity documents are forgeries.

³One possible explanation proffered by researchers for such behaviour in children is that they might not have yet developed the metacognitive abilities required for the experimental tasks and the interpretation of SIs, e.g. the ability to understand other people’s beliefs, goals, intentions (Theory of Mind) etc. For example, it is questionable whether children are able to compare what the speaker actually said and what s/he could have said instead.

⁴This topic can be seen as related to the previous one as it is tempting to view the derivation of context-independent implicatures as an automatic process.

- c. Not all of their identity documents are forgeries.

However, further empirical research (Noveck and Posada 2003; Bott and Noveck 2004; Breheny and Williams 2006; Zonderman 2009) leans in favour of the globalists (who are in this connection sometimes also called *contextualists*). It shows that in different types of context SIs can be not triggered at all depending on what is relevant, as opposed to Levinson’s belief that they are always triggered but cancelled if (not)necessary.

Among various arguments for the grammatical approach to SIs Chierchia puts forward that SIs cannot be derivable post-grammatically in non-monotonic environments. Consider the following example:

- (19) Exactly two students wrote a term paper or made a class presentation.

Chierchia (2001) argues that sentence (19) can be interpreted both inclusively (i.e. the number of students who did both equals two) and exclusively (the number of students who did one or the other (but not both) equals two), but Globalism did not allow the latter interpretation to derive. It is so because the non-monotonicity of the quantifier ‘exact’ makes it impossible to form a suitable scale of informational strength, hence the implicature cannot be computed. For the localists it is not a problem as the implicature would be generated locally and not post-compositionally.

Another strong piece of evidence in favour of the grammatical approach concerns embedded SIs. Since Cohen (1971) and all the way up to Chierchia et al. (2009) the localists defended the grammatical view on SIs asserting the existence of embedded SIs. According to them, globalist view cannot predict the existence of embedded SIs, because “under this view scalar implicatures are derived on the basis of the maxim of quantity which is only relevant to complete speech acts, and not to constituents” (Chierchia, Fox, Spector 2009: 1). Grammatical view, however, easily allows embedded SIs in virtue of the covert exhaustification operator *O*. For instance, in the scope of the belief ascription, as in (16):

- (20) John believes that some students are waiting for him.

Chierchia claims that in (20) we can interpret the sentence as saying that John believes that *not all* students are waiting for him. Whereas under the globalist

6 Non-Monotonicity in Linguistics

post-compositional assumption the interpretation of this sentence is weaker - it only predicts that John *does not* believe that some students are waiting for him.

Geurts and Pouscoulous (2009), however, present some data which again speaks against the localist view. They compared interpretations of ‘some’ in different embedded and non-embedded contexts in the scope of ‘think’, ‘want’, deontic ‘must’, and the universal quantifier ‘all’. They found that the acceptance rate for the SI decreases under embedding, which is not accountable for under the localist view.

The results of the analysis of SIs within the interpretation-based account posited by Mascarenhas (2013) go along the lines of the localist ideas about the language-specific nature of SIs. Following their ideas, he argues for the evidence of a strengthening or exhaustification operation in natural languages. Mascarenhas differentiates between *primary implicatures* (conclusions about what the speaker is not in a position to assert) and *secondary implicatures* (conclusions about what the speaker believes to be false). Secondary implicatures are derived from primary implicatures via the strengthening procedure. Conjoining the literal meaning of the statement with all of its secondary implicatures gets us the final interpretation of the statement. In 3.2.3 I provided Mascarenhas’s assumption about how that strengthening procedure operates. Here I repeat it again:

Given a sequence of premises $P_1, \dots, P_n$, and a conclusion C , begin by calculating the strengthened meaning of each premise, getting the sequence $P_1^+, \dots, P_n^+$. Then, check to see if the conclusion C follows classically from $P_1^+, \dots, P_n^+$.

The strengthening operation is a purely language-specific procedure and speaks therefore for the localist view.

Figure 6.1 shows two possible scenarios of interpreting SIs based on the various characteristics described above. In the first scenario, SIs are said to be acquired, generated globally and processed slow, which points to their being embedded in pragmatics. The second scenario shows SIs as being inherent, generated locally, and processed fast and in a controlled fashion, which indicates their grammatical attribution. This division is rough and by no means absolute and is rather more of a helping tool.

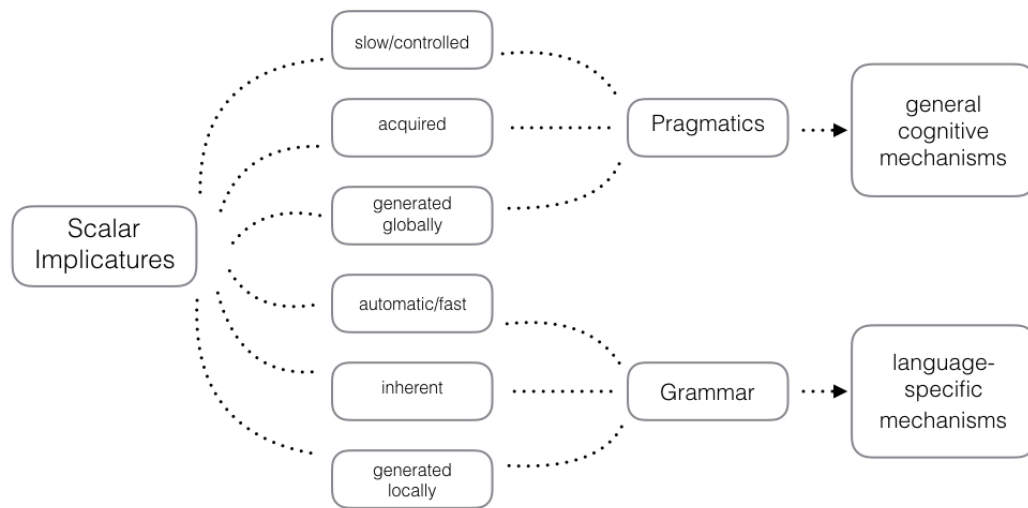


Figure 6.1: Attribution of SIs

As we have seen, however, empirical evidence seems to refute the localist view and most likely comes down in favour of the globalist view. If this is so and classical analysis turns out to be true, though, it will need to overcome a wide range of ever growing challenges thrown up by the localists.

6.2.3 Number words

Number words or numerals are also important to the debate about SIs. As was quickly shown in example (16), repeated here as (21), number words can be canceled too just like other SIs:

(21) Mary ate three ice creams, if not more/ in fact more/ in fact four.

There is an ongoing discussion about the semantics of numerals, as it was noticed that numerally quantified NPs in addition to their ‘exact’ (or upper-bounded) meaning can have an ‘at least’ (or lower-bounded) meaning and even an ‘at most’ meaning, such as in (22), (23) and (24) respectively:

(22) If Mary eats three ice creams, there will be nothing left for me.

(23) If Mary eats three ice creams, she might get sick.

(24) If Mary eats three ice creams, she will still stay thin.

Example (22) is likely to be interpreted in an upper-bounded way as meaning exactly three ice creams given the context, in which case there will be none left. Example (23) by contrast means that at least three ice-creams will be enough to make Mary sick. And the example (24) can be read to mean that the maximum Mary can allow herself to eat in order to stay in shape is three ice creams at a time. This last interpretation is written off by many as another reading of the ‘exact’ interpretation, which is why the main debate is pitched between the lower-bounded and the upper-bounded interpretations.

Some researchers have argued that the difference in interpretations is due to pure lexical ambiguity (Horn 1992), while for others this is more about the nature of number words themselves. Breheny (2008), for example, claims that numbers are exact and the ‘at least’ meaning is just a matter of pragmatical

inference. According to this point of view, meanings are empirical phenomena and all nuancing is merely down to context. Their opponents believe that the ‘at least’ meaning is by contrast the literal, core meaning of the numeral, while the ‘exact’ meaning is derived as an implicature. Depending on how exactly this implicature is derived there are two subcategories, with the Neogriceans on one side, who argue that it’s a purely pragmatically based process, and Chierchia, Fox and Spector (2008) on the other, saying that it’s rooted in grammar.

Panizza and Chierchia (2011) have argued that polarity factors of the context (i.e. downward vs. upward monotonicity) can influence the choice of one or the other meaning of a numeral. Namely, that the ‘exact’ interpretation occurs preferentially in upward monotone contexts, whereas the ‘at least’ interpretation occurs preferentially in downward monotone ones. Panizza and Chierchia transferred this approach from studying other SIs (quantifiers like *some*, connectives like *or*), which display a similar behaviour.

Empirical evidence is as expected controversial. Huang et al. (2013) designed an experiment in which they excluded pragmatic inferences in order to induce purely semantic interpretations of number words. Their results suggest that when SIs are cancelled in the experiment setting, number words are interpreted as referring to an exact number of a set. From this, the authors conclude that number words have exact semantics. What is not explicitly stated, but is to be inferred from their conclusion is that non-exact interpretations (such as ‘at least’) can only arise through pragmatic inferences.

There is also evidence that supports the theory that the ‘at least’ meaning is the initial one when interpreting numerals. For instance, Barner and Bachrach (2010) studied how very young children perceive the numeral ‘one’ and found that they derive its exact meaning by an implicature contrasting it with the numeral ‘two’, which then gets an exact meaning only when the child acquires larger numerals, and so on.

6.2.4 Two types of defeasibility

Several times in this section I pointed out the defeasible nature of scalar (and other) implicatures. Moreover, scalar implicatures allowed us to talk not only about one, but two types of defeasibility. One type is *explicit* defeasibility, the

examples of it are given in (14), (15) and (16). In this case SIs are canceled explicitly by entering additional information. A very different type of defeasibility appears in instances where the implicature is canceled due to the context, as was demonstrated in (17). This type refers to *implicit* defeasibility.

Note that only localists advocate for the latter type of defeasibility, as only according to them implicatures arise automatically as part of grammar. Nevertheless Grice also distinguished between two ways of canceling implicature calling them by similar names - explicit and contextual cancelability. The confusion here lies in terminology. What Grice meant by contextual cancelability is in fact not cancelability but rather not arising of implicature in virtue of background information, like in the following example where the implicature ‘some but not all’ is neither intended nor arisen:

(25) Some people believe in God.

Here is how Grice explains his ideas:

“It may be explicitly canceled, by the addition of a clause that states or implies that the speaker has opted out, or it may be contextually canceled, if the form of utterance that usually carries it [the implicature] is used in a context that makes it clear that the speaker is opting out.” (Grice 1975: 57).

However, as was discussed in the previous subsection, there is no evidence that implicit defeasibility occurs. By contrast, research on explicit defeasibility provided positive results. In his studies Katsos (2007) presented participants short texts with an utterance containing SI followed by an utterance that contradicted either the SI or a semantic entailment of that utterance. It was found that in the cases where a SI was followed by a contradicting SI the coherent rate was significantly higher than when a SI was followed by a contradicting entailment. Here is one example from Katsos’ experiment with SI (22, a-b) and entailments (23, a-b):

(22) a. The director asked his consultant: Who is representing our company at the court hearing? His consultant replied: Turner or Morris. *In fact, both of them are.*

- b. How is our candidate doing in the polls? The campaign manager replied: He has overtaken some of his opponents that have little funding. *In fact, he has overtaken all of them.*
- (23) a. The director asked his consultant: Who is representing our company at the court hearing? His consultant replied: Turner or Morris. *In fact, none of them are.*
- b. How is our candidate doing in the polls? The campaign manager replied: He has overtaken some of his opponents that have little funding. *In fact, he hasn't overtaken any of them.*

Entailment is a relation between two propositions (p and q) where p entails q ($p \models q$) if and only if the truth of p requires the truth of q (Huang 2012, Levinson 1983). An implicature, however, is non-truth conditional, i.e. the truth of q is independent of p. This is why in case of implicature we are able to make inferences beyond the literal meaning, and in case of entailment not.

Note that it doesn't mean that an entailment can't be denied. Of course, one can for instance utter "I had three beers yesterday, no, actually one beer." In fact, it might be even useful to describe entailment as a subject of cancellation and an implicature rather than a subject of correction or specification. Because in case of an implicature "I had three beers yesterday, in fact four", what was implied before the correction phrase still holds after it (because, as extensive fieldwork shows, it is not possible to have four beers without having three before).

6.3 Conclusion

In this chapter I wanted to show that, like any other information interpretation, language interpretation proceeds according to logical laws, and some of these laws are non-monotonic. When we interpret discourse, we engage in reasoning processes similar to those of "reading" bus timetables. These processes include assumptions about social cooperativeness, knowledge and ignorance. What is intriguing about this is whether the roots of these processes are located in the language system itself or whether they are generic for all interpretative operations. For these purposes I discussed the phenomenon of SIs that received a

6 *Non-Monotonicity in Linguistics*

roaring dispute as to whether they are derived within grammar or pragmatics. In the Gricean tradition, SIs are considered to be a pragmatic phenomenon, as something that is computed on a postcompositional level of interpretation. This approach was challenged by the “grammaticalists” who argued that SIs are derived by means of compositional rules that apply to the sentence constituents. Although the compelling arguments of the latter ones called into question many aspects of the traditional Gricean reasoning, the empirical research I provide in this chapter did not support the grammatical view. The last chapter will examine what this means in terms of the origins of non-monotonicity and reasoning.

7

Consequences for the Architecture of Language and Cognition

If you have come so far (or just opened the thesis on this chapter) you probably want to finally know where all the discussion was heading and hear *the* answer to the question raised in the beginning, the one concerning the nature and origins of human reasoning. I myself was wondering the same thing, and in order to make things clearer for myself and the reader I decided to turn this chapter into an essay-like discussion. That way we can free ourselves from the burdens of bold conclusions and use this space for unrestrained discussion.

So, my main question throughout this thesis is, where does non-monotonicity take its root - in language or in general cognition? That question is actually part of that other, bigger, one: namely how is human cognition structured? This has been the bugbear of cognitive science since its beginnings in the mid-1950s, when the field of artificial intelligence was founded in the U.S. and Noam Chomsky rejected the idea that language is a learned habit and proposed instead his theory of mental grammars. But because inner qualities are the least definable ones (especially when you are trying to use those very qualities as tools to define themselves), we are far away from any satisfactory answer as to how the mind really operates.

The question about the whereabouts of non-monotonicity depends on where we stand on the dispute between the nativist and modularity approaches to language and reasoning. If we agree that the mind is domain-general and/or non-modular to the full, then the need to localise non-monotonicity becomes irrelevant: it's everywhere. If we suppose, however, that language and reasoning are independent domain-specific modules, then the question remains very much alive.

Let's look at the language first. As I said above, Noam Chomsky was one of the

7 *Consequences for the Architecture of Language and Cognition*

first theorists to hypothesise that language and human linguistic behaviour are something separate and distinct from general cognition. He saw language as an innate and independent ‘organ’ which records and stores linguistic information as we grow up. From the day we are born, we are exposed to adult language speakers, and all the information we are getting from this communication is being fed into what Chomsky and his followers called the ‘Universal Grammar’. This language organ is therefore an assemblage of basic linguistic principles and rules common to all languages which guide the language acquisition process. These principles are so abstract that we cannot possibly learn and know them by mere association and induction, as was claimed by the then dominant behaviourist ideas. The rationale behind Chomsky’s theory includes among other things the fact that children tend to acquire language far too quickly and effortlessly and without reinforcement, and that a great deal of what we say is not the result of direct input (we can say things we have not been “trained” to say in the same way that we can understand things we have never heard before, provided they comply with grammatical rules etc; the input is finite compared to infinite output):

It seems plain that language acquisition is based on the child’s discovery of what from a formal point of view is a deep and abstract theory - a generative grammar of his language - many of the concepts and principles of that are only remotely related to experience by long and intricate chains of quasi-inferential steps. (Chomsky 1965 : 58)

While Chomsky’s ideas enjoy a wide currency among cognitive scientists, some, including most notably Michael Tomasello (Tomasello 2003), have offered alternative explanations for the mechanics of language acquisition. In Tomasello’s view, we learn language by using a variety of inductive, analogical and statistical learning methods, some of them highly complex. He also argues that children acquire language without the help of any inborn linguistic information or set of communicative behaviours, saying that whatever skills they exhibit in mastering language are not specific to language learning per se and are also employed in other domains (in particular, he talks about skills such as intention-reading and pattern-finding):

The cognitive and social learning skills that children bring to the acquisition process are much more powerful than previously believed,

7 *Consequences for the Architecture of Language and Cognition*

and <...> the adult endpoint of language acquisition comprises nothing other than a structured inventory of linguistic constructions, a much closer and more child-friendly target than previously believed. (Tomasello 2003 : 6-7)

He consequently called his model of grammar ‘cognitive-functional grammar’ or ‘usage-based grammar’ whose rules and structure emerge from its patterns of use rather than from some linguistic core.

Other researchers have taken issue with Chomsky’s understanding of the primary linguistic data - that is, the language-relevant information that we absorb growing up and use to master language - but the fact remains that ever since Chomsky challenged the behaviourist ideas of language acquisition language has largely been thought of a separate and independent unit.

Another question that is important to us here is where the pragmatics stands in the dispute over language modularity. If pragmatics belongs to the grammatical computational system of language, then we could assume that non-monotonicity originates in language (and all its components) or - which is also possible (as everything can be possible where nothing is given) - that the non-monotonicity of language and the non-monotonicity of general cognition are in no way related to each other.

The question about where pragmatics belongs is a very tricky one and bumps into another question, namely what pragmatics actually is and how we define it. Let’s go back to Chomsky again. In spite of giving a special place to grammar, he does not undermine the importance of pragmatics; quite the contrary, he sees it as a central component of linguistic theory. He distinguishes between ‘grammatical competence’, which is an innate system of knowledge and rules, and ‘pragmatical competence’, which is about “knowledge of conditions and manner of appropriate use, in conformity with various purposes” (Chomsky 1980 : 224). This distinction resembles another binary opposition in his theory, that between ‘linguistic competence’ and ‘linguistic performance’, but the relation between these two pairs seems pretty unclear. According to Chomsky, competence is an ideal knowledge of language that the speaker has, whereas performance is the actual use of that knowledge. Both linguistic performance and pragmatical competence are about language in use - does this mean now that they are associated with each other

7 *Consequences for the Architecture of Language and Cognition*

or is pragmatic competence rather a component of linguistic competence which then determines linguistic performance? In an online interview about the nature of pragmatics Chomsky says that it all depends on how we define our terms:

As for the term “competence”, it has the informal meaning of ordinary usage, and also a technical meaning, which is whatever its users assign to it (as in the case of “tensor”, “undecidability” in the technical sense, etc.). I am familiar with only one technical sense, the one I’ve used for many years to mark the conceptual distinction between what Jones knows and what Jones does (competence vs. performance), a distinction that is not controversial. I introduced the technical term in a (probably vain) effort to avoid pointless debate engendered by uncritical use of the informal English notion “knowledge”, or use of technical concepts of knowledge introduced in the philosophical literature that diverge considerably from informal usage (which is fine, as long as it is recognized clearly, in which case the debates dissolve).

If we are using the term “competence” in my technical sense, then pragmatics is not part of a theory of linguistic competence, for uninteresting terminological reasons. If we are using the term “competence” in its ordinary English sense, then I suppose one might say that pragmatics is part of linguistic competence, but the conclusion is again uninteresting, merely a matter of terminology. (Brigitte Stemmer 1999 : 400)

To me the conclusion does seem interesting and even significant because if pragmatics is a component of linguistic competence, then this would suggest that pragmatic competence is an innate and not an acquired capacity (and when saying that I immediately hear a crowd of oncoming protesting linguists behind my door). Although that doesn’t have to be surprising for us because in the previous chapter I discuss debates around the nature of SIs that were first described as being part of pragmatics and later assigned by some researchers to matters of core grammar; although later in that chapter I come to the conclusion that this theory does not check out empirically. Still, most scholars agree that pragmatics is something that is not encoded in grammar, it’s about language in use, which means it’s governed by general cognitive abilities and is acquired over time.

7 Consequences for the Architecture of Language and Cognition

So far we have established - for the purposes of this discussion - that language is a separate module which can be described as a grammatical computational system, and that everything else that is associated with language use lies outside it. Now that we have made this distinction clear, we can ask whether linguistic phenomena that exhibit non-monotonic behaviour such as SIs are part of grammar or pragmatics (i.e. general cognition). In the previous chapter, I provide evidence in support of both claims but it looks like that SIs' pragmatic provenance is the more likely option. As empirical research has shown, SIs are acquired rather than innate, processed slowly and in a controlled way, and are generated globally. If that's true and pragmatics is where SIs occur with their non-monotonicity, then this would speak in favour for the idea that the origins of non-monotonicity and, by extension, reasoning are in our general cognition (given, that is, that pragmatics is outside of grammar).

Here I could finally finish my thesis but before I do there still remain important questions to our debate - what do we mean by general cognition and how does our reasoning function there? This was discussed in Chapter 5, where I talked about the structure of reasoning.

It appears that some of our reasoning is content-dependent and some is content-independent. It is content-dependent because, as was shown in the experiments in Chapter 4, alternation of the context may change the outcome. But it is also content-independent because we can generalise from previous experiences and apply the resulting logic to similar instances in the future, in a sort of learning mechanism.⁵ From this perspective, following Stenning and Lambalgen, it seems legitimate to me to propose that our reasoning is both domain-specific (different domains call for different logics depending on the context) and domain-general (several domains can share the same reasoning mechanisms independent of the context). In other words, it seems that the different cognitive abilities and areas of our mind have a lot in common. Moreover, we can go further and suggest that in order for that commonality to exist there must be some kind of information exchange going on between all those places. Precisely because various domains share the same reasoning mechanisms it seems fair to suppose that information

⁵Interesting insights on this are provided in a paper by Carraher et al. about working children in Brazil who use different, more complicated, computational strategies than those learned in school when selling wares on the street (Carraher et al. 1985).

7 Consequences for the Architecture of Language and Cognition

feeding into one domain can be read off or used in another: one example could be the interconnectedness between language as grammar and language as pragmatics, which in its turn is connected to Theory of Mind⁶, which is connected to something else still, and so on. If that's true, however, this would contradict one of the key properties of Fodorian modules (for Fodor this is the most important one, in fact) - namely, that modules are informationally encapsulated, meaning that information doesn't go in and out of the modules.

For future research it would certainly be interesting and valuable to test that. One possible way to do that would be to trace at what point these reasoning mechanisms begin to develop in different domains and whether these developments coincide in time. This paper has already presented a step in this direction. In Chapter 5, I provided evidence pointing to the late acquisition in children of induction. This can be compared with the similarly late acquisition of SIs; the temporal dependency between these two non-monotonic processes could be an indication of their interconnectedness. On the other hand, the Negation as Failure principle of the non-monotonic logics is present already at a very young age; in fact it seems to be even more pronounced in a person's first months. According to Piaget (e.g. 1954), children don't develop *object permanence* - the awareness that objects continue to exist when they are out of sight - until the age of 8 or 9 months, i.e. for them the object is not only not there if they can't see it, it has actually disappeared (that's why the peek-a-boo game is such a delight for them). But the crucial point here is that this type of non-monotonic inference (object permanence) is exhibited by children much earlier in development than some of the other ones mentioned in this paper, such as, for instance, acquisition of induction. What would also need to be tested here is whether the mastering of one cognitive ability that involves non-monotonic reasoning mechanisms affects another cognitive ability - like when people say that learning chess is beneficial for politicians as it teaches strategic thought.

Another system of our cognition that appears to involve non-monotonic concepts is Theory of Mind (ToM), an ability to understand other people's beliefs and intentions, a kind of cognitive device of empathy. By placing ourselves in the mind of another person we can reason independently of our own beliefs. This

⁶It has been observed that people with theory of mind deficits also show impaired pragmatic abilities (see, for instance, Leslie 1987).

enables us to consider alternative courses of events that are based on the default principle ‘assume a unless there is explicit information to the contrary’, a principle of non-monotonic logics, and make inferences based on what we think other people think and what the information that they communicate to us tells us about what they think. In other words, the process of understanding the actions and words of others is an instance of complex (causal) reasoning guided by our perception and beliefs. This brings us back to Grice and his principles of cooperation; and that is why, as mentioned before, pragmatics relies on ToM. ToM was studied vastly in the literature, and since it can be seen as a kind of non-monotonic reasoning I think that studying its nature and development in children can also contribute greatly to the discussion of the architecture of mind, and non-monotonicity in particular.

7.1 Conclusion

One of the points I’m hoping to make with this paper is that it’s crucial that theories of reasoning should not ignore linguistic theories. Firstly, because the experimental data is in most cases presented in natural language. And secondly, because many linguistic phenomena including SIs provide instances of non-monotonic reasoning.

The fact that empirical evidence shows that SIs are a phenomenon of pragmatics rather than of core grammar strongly suggests that the origins of non-monotonicity, and by extension reasoning in general, come from general cognition and not from the language system (given of course that we agree that pragmatics lies outside of core grammar). Reasoning therefore can be said to have given expression to language.

The follow-up question arises, however: if reasoning is not part of our language system and language still embodies aspects of reasoning, how is all this presented in our mind? The fact that manipulating the context in reasoning experiments can affect their outcomes points to the domain-specificity of reasoning; at the same time, however, people tend to generalise rules and apply them to similar tasks, which also makes our reasoning domain-general. Moreover, the domains must be interconnected with each other to form networks, otherwise reasoning

7 *Consequences for the Architecture of Language and Cognition*

mechanisms in different domains wouldn't exhibit the commonalties that they exhibit. This interconnectedness can also be observed by tracing how the improvement of knowledge in one domain improves the "neighbour" ones: how, for example, the acquisition of SIs affects the acquisition of other phenomena that show non-monotonicity.

Although it was discourse interpretation where non-monotonicity was first observed, I believe it shows up almost everywhere else. In order to find the source of non-monotonicity we just need to seek out cases where default rules fail to apply. To paint the complete picture, however, we must examine all these cases to see whether non-monotonic behaviour exhibited by them has a language-specific or language-independent origin. It will be very interesting to see if the results of such an investigation compare with the conclusions drawn in this paper. Something that I hope will be a concern for future generations of linguists.

Bibliography

- Antoniou, G. (1999). A tutorial on default logics. *ACM Computing Surveys (CSUR)*, 31(4), 337–359.
- Barner, D. & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive psychology*, 60(1), 40–62.
- Barrett, H. C. & Kurzban, R. (2006). Modularity in cognition: framing the debate. *Psychological review*, 113(3), 628.
- Beller, S. (2008). Deontic norms, deontic reasoning, and deontic conditionals. *Thinking & Reasoning*, 14(4), 305–341.
- Bledsoe, W. W. (1977). Non-resolution theorem proving. *Artificial Intelligence*, 9(1), 1–35.
- Bott, L. & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, 51(3), 437–457.
- Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics*, 25(2), 93–139.
- Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? an on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, 100(3), 434–463.
- Brouwer, L. E. J. & van Dalen, D. (2011). *Brouwer’s Cambridge lectures on intuitionism*. Cambridge University Press.
- Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British journal of developmental psychology*, 3(1), 21–29.
- Carruthers, P. (2006). *The architecture of the mind*. Oxford University Press.

Bibliography

- Chater, N. & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. *Mind & Language*, 11(2), 191–202.
- Chierchia, G. (2001). Scalar implicatures. *Polarity Phenomena, and the Syntax/Pragmatics Interface* (typescript).
- Chierchia, G. et al. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. *Structures and beyond*, 3, 39–103.
- Chomsky, N. (1976). 1965. aspects of the theory of syntax. *Cambridge, Massachusetts: MIT Press. Aspects of the theory of syntax*.
- Chomsky, N. (1980). Rules and representations. *Behavioral and brain sciences*, 3(01), 1–15.
- Cohen, L. J. (1971). Some remarks on grice’s views about the logical particles of natural language. In *Pragmatics of natural languages* (pp. 50–68). Springer.
- Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4(03), 317–331.
- Cosmides, L. (1985). *Deduction or Darwinian Algorithms?* PhD thesis, Citeseer.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? studies with the wason selection task. *Cognition*, 31(3), 187–276.
- Cosmides, L. & Tooby, J. (1989). Evolutionary psychology and the generation of culture, part ii: Case study: A computational theory of social exchange. *Ethology and sociobiology*, 10(1-3), 51–97.
- Cosmides, L. & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual review of psychology*, 64, 201–229.
- Cummins, D. D. (1996). Evidence for the innateness of deontic reasoning. *Mind & Language*, 11(2), 160–190.
- Evans, J. & Lynch, J. (1973). Matching bias in the selection task. *British Journal of Psychology*, 64(3), 391–397.

Bibliography

- Evans, J. S. B., Clibbens, J., & Rood, B. (1996). The role of implicit and explicit negation in conditional reasoning bias. *Journal of Memory and Language*, 35(3), 392–409.
- Evans, J. S. B. & Newstead, S. (1980). A study of disjunctive reasoning. *Psychological research*, 41(4), 373–388.
- Fedorov, V. V. (1972). *Theory of optimal experiments*. Elsevier.
- Ferrara, F. M., Esposito, A., Pizzini, B., & Matarazzo, O. (2015). Evolutionary psychology and human reasoning: Testing the domain-specificity hypothesis through wason selection task. In *EAPCogSci*.
- Fodor, J. (2000). Why we are so good at catching cheaters. *Cognition*, 75(1), 29–32.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.
- Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.
- Geurts, B. & Pouscoulous, N. (2009). Embedded implicatures?!? *Semantics and pragmatics*, 2, 4–1.
- Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Gigerenzer, G., Todd, P. M., ABC Research Group, t., et al. (1999). *Simple heuristics that make us smart*. Oxford University Press.
- Grice, H. P. (1975). Logic and conversation. 1975, (pp. 41–58).
- Griggs, R. A. & Cox, J. R. (1982). The elusive thematic-materials effect in wason’s selection task. *British Journal of Psychology*, 73(3), 407–420.
- Hájek, A. (1996). “mises redux”—redux: Fifteen arguments against finite frequentism. *Erkenntnis*, 45(2-3), 209–227.

Bibliography

- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. *Meaning, form, and use in context: Linguistic applications*, (pp. 11–42).
- Horn, L. (1992). The said and the unsaid. *Ohio State University working papers in linguistics*, 40, 163–192.
- Huang, Y. (2012). *The Oxford dictionary of pragmatics*. Oxford University Press.
- Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What exactly do numbers mean? *Language Learning and Development*, 9(2), 105–129.
- Husserl, E. (1901). *Logische Untersuchungen: 1. T. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*, volume 2. M. Niemeyer.
- Jackson, S. L. & Griggs, R. A. (1990). The elusive pragmatic reasoning schemas effect. *The Quarterly Journal of Experimental Psychology*, 42(2), 353–373.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British journal of Psychology*, 63(3), 395–400.
- Johnson-Laird, P. N. & Wason, P. C. (1977). *Thinking: Readings in cognitive science*. CUP Archive.
- Katsos, N. & Bishop, D. V. (2009). The development of informativeness from a speaker’s and a comprehender’s perspective. *Unpublished manuscript, University of Cambridge*.
- Konolige, K. (1988). On the relation between default and autoepistemic logic. *Artificial intelligence*, 35(3), 343–382.
- Koralus, P. & Mascarenhas, S. (2013). The erotetic theory of reasoning: bridges between formal semantics and the psychology of deductive inference. *Philosophical Perspectives*, 27(1), 312–365.
- Landman, F. (1998). Plurals and maximalization. In *Events and grammar* (pp. 237–271). Springer.
- Lerner, R. M., Liben, L. S., & Mueller, U. (2015). *Handbook of Child Psychology and Developmental Science, Cognitive Processes*, volume 2. John Wiley & Sons.

Bibliography

- Leslie, A. M. (1987). Pretense and representation: The origins of theory of mind. *Psychological review*, 94(4), 412.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge university press.
- Levinson, S. C. (1995). Three levels of meaning. In *Grammar and meaning: Essays in honour of Sir John Lyons* (pp. 90–115). Cambridge University Press.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.
- Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, (pp. 986–1005).
- Lukasiewicz, J. (1930). Philosophische bemerkungen zu mehrwertigen systemen des aussagenkalkiils. *Comptes Rendus des Seances de la Societe des Sciences et des Lettres de Varsovie*.
- Mckenzie, C. R. & Mikkelsen, L. A. (2000). The psychological side of hempel's paradox of confirmation. *Psychonomic Bulletin & Review*, 7(2), 360–366.
- Michael, T. (2003). Constructing a language: a usage based theory of language acquisition.
- Mill, J. S. (1865). *An Examination of Sir William Hamilton's Philosophy: And of the Principal Philosophical Questions Discussed in His Writings*, volume 1. Longman, Green, Longman, Roberts & Green.
- Montague, R. (1970). Universal grammar. *Theoria*, 36(3), 373–398.
- Moore, R. C. (1985). Semantical considerations on nonmonotonic logic. *Artificial intelligence*, 25(1), 75–94.
- Morier, D. M. & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin*, 10(2), 243–252.
- Moshman, D. (2004). : From inference to reasoning: The construction of rationality. *Thinking & Reasoning*, 10(2), 221–239.

Bibliography

- Noveck, I. A. (1996). To what extent do pragmatic reasoning schemas affect performance on wason's selection task? *The Quarterly Journal of Experimental Psychology: Section A*, 49(2), 463–489.
- Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2), 165–188.
- Noveck, I. A. & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and language*, 85(2), 203–210.
- Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608.
- Oaksford, M. & Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking & Reasoning*, 1(2), 121–152.
- Oaksford, M. & Chater, N. (1998). *Rational models of cognition*. Oxford University Press Oxford.
- Oaksford, M. & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in cognitive sciences*, 5(8), 349–357.
- Oda, R., Hiraishi, K., & Matsumoto-Oda, A. (2006). Does an altruist-detection cognitive mechanism function independently of a cheater-detection cognitive mechanism? studies using wason selection tasks. *Evolution and Human Behavior*, 27(5), 366–380.
- Panizza, D. & Chierchia, G. (2011). Numerals and scalar implicatures. *Experimental pragmatics/semantics*, (pp. 129–150).
- Papafragou, A. & Musolino, J. (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition*, 86(3), 253–282.
- Papafragou, A. & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition*, 12(1), 71–82.
- Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of experimental child psychology*, 16(2), 278–291.

Bibliography

- Partee, B. & Bach, E. (1984). Quantification, pronouns, and vp anaphora. *Truth, interpretation and information*, (pp. 99–130).
- Piaget, J. (1954). Language and thought from a genetic perspective. *Acta Psychologica*, 10, 51–60.
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chicago Press.
- Pillow, B. H. (2002). Children’s and adults’ evaluation of the certainty of deductive inferences, inductive inferences, and guesses. *Child development*, 73(3), 779–792.
- Reiter, R. (1978). On closed world data bases. In *Logic and data bases* (pp. 55–76). Springer.
- Ruffman, T. (1999). Children’s understanding of logical inconsistency. *Child Development*, 70(4), 872–886.
- Scriven, M. (1977). Reasoning.
- Siegler, R. S. & Shipley, C. (1995). Variation, selection, and cognitive change. *Developing cognitive competence: New approaches to process modeling*, (pp. 31–76).
- Simon, H. A. (1957). Models of man; social and rational.
- Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30(2), 191–205.
- Somerville, S. C., Hadkinson, B., & Greenberg, C. (1979). Two levels of inferential behavior in young children. *Child Development*, (pp. 119–131).
- Spector, B., Fox, D., & Chierchia, G. (2008). Hurford’s constraint and the theory of scalar implicatures. *Manuscript, MIT and Harvard*.
- Sperber, D. (1994). The modularity of thought and the epidemiology of representations.

Bibliography

- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57(1), 31–95.
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and cognition*, volume 142. Citeseer.
- Stemmer, B. (1999). An on-line interview with noam chomsky: On the nature of pragmatics and related issues. *Brain and Language*, 68(3), 393–401.
- Stenning, K. & van Lambalgen, M. (2005). Explaining the domain generality of human cognition. *Integrating the mind*, (pp. 179–209).
- Stenning, K. & Van Lambalgen, M. (2012). *Human reasoning and cognitive science*. MIT Press.
- Stenning, K., Van Lambalgen, M., et al. (2008). Interpretation, representation, and deductive reasoning. *Reasoning: Studies of human inference and its foundations*, (pp. 223–248).
- Sugiyama, L. S., Tooby, J., & Cosmides, L. (2002). Cross-cultural evidence of cognitive adaptations for social exchange among the shiwi of ecuadorian amazonia. *Proceedings of the National Academy of Sciences*, 99(17), 11537–11542.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*, volume 2. Random house.
- Thompson, G. J., Hurd, P. L., & Crespi, B. J. (2013). Genes underlying altruism. *Biology letters*, 9(6), 20130395.
- Toulmin, S. (1958). The uses of argument. cambridge university press. *Cambridge, UK*.
- Truszczyński, M. (1991). Embedding default logic into modal nonmonotonic logics. In *LPNMR* (pp. 151–165).
- Tunmer, W. E., Nesdale, A. R., & Pratt, C. (1983). The development of young children’s awareness of logical inconsistencies. *Journal of Experimental Child Psychology*, 36(1), 97–108.

Bibliography

- Tversky, A. & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141–162). Springer.
- Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, 90(4), 293.
- Wason, P. C. (1966). Reasoning.
- Wason, P. C. (1968). Reasoning about a rule. *The Quarterly journal of experimental psychology*, 20(3), 273–281.
- Wason, P. C. & Evans, J. S. B. (1975). Dual processes in reasoning? *Cognition*, 3(2), 141–154.
- Zondervan, A. (2009). Experiments on qud and focus as a contextual constraint on scalar implicature calculation. *From experiment to theory, presuppositions, negation and scalar implicatures*. Basingstoke: Palgrave.

Abstract

Since Aristotle formulated the theory of classical logic, to be rational has meant to be logical. In the last 60 years, however, this view came under increasing criticism as empirical research showed that humans do not actually act according to the laws of classical logic, and are therefore not rational. More recently, some researchers have suggested that classical logic is not a proper framework in which to describe human reasoning behaviour. It has been argued that while solving a task in natural language people apply different logics, and in many cases they pick some non-classical logic. The main difference between the logics that people pick and classical logic is that the former show so called non-monotonic behaviour, whereby new information may invalidate previous conclusions; classical logic by contrast is monotonic.

Apart from logic and other computational disciplines, non-monotonicity has also been identified in linguistics, and not only in its computational areas. In principle, non-monotonic logics can provide appropriate foundations for any area of linguistics that deals with generalizations and their exceptions - which is probably every area of linguistics. If we agree that this is not a coincidence, there may be two ways of looking at the relationship between non-monotonic logic and linguistic theory: (1) non-monotonicity in human logic and, by extension, in general cognition is the result of the emergence of non-monotonicity in language, or conversely (2) non-monotonicity in language originated in general cognition.

In order to understand this relationship I set out to investigate whether non-monotonicity in language is language-specific, i.e. rooted in grammar, or a matter of general cognition, i.e. related to non-grammatical pragmatics. For this study I base my research on scalar implicatures, an area of linguistics where non-monotonicity plays an important role.

The results show that non-monotonicity is language-independent. Further, I hold that the architecture of our mind is structured in such a way as to allow both domain-specific and domain-general reasoning mechanisms.

Abstract

Seit Aristoteles die Theorie der klassischen Logik entwarf, war rational sein, gleichbedeutend damit, logisch zu sein. In den letzten 60 Jahren wurde diese Ansicht zunehmend kritisiert. Empirische Forschungen haben belegt, dass Menschen nicht den Gesetzen der klassischen Logik gehorchen und somit nicht als rational gelten können. Kürzlich wurde vorgeschlagen, dass die klassische Logik keinen angemessenen Theoriehintergrund darstellt, um natürliches Schließen zu beschreiben. Es wurde argumentiert, dass Menschen zum Lösen einer Aufgabe in einer natürlichen Sprache verschiedene Logiken anwenden und in vielen Fällen nicht-klassische Logik verwenden. Der Hauptunterschied zwischen dieser von Menschen verwendeten Logik und der klassischen Logik liegt darin, das erstere sogenanntes nicht-monotones Verhalten aufweist, worin neue Informationen bisheriges natürliches Schließen unwirksam machen können. Im Gegensatz dazu ist klassische Logik monoton.

Abseits von Logik und anderen computationalen Disziplinen, spielt Nicht-Monotonizität auch in der Linguistik eine Rolle und dort nicht nur in computationalen Disziplinen. Im Prinzip kann nicht-monotone Logik Grundlagen für alle Bereiche der Linguistik bereitstellen, in denen es um Generalisierungen und deren Ausnahmen geht - wovon praktisch jeder Bereich der Linguistik betroffen ist. Unter der Annahme, dass dies kein Zufall ist, bleiben zwei Möglichkeiten, die Beziehung zwischen nicht-monotoner Logik und linguistischer Theorie zu beschreiben:

1. Nicht-Monotonizität in menschlicher Logik und, im weiteren Sinne der allgemeinen Kognition, ist das Resultat von Nicht-Monotonizität in der Sprache;
2. Nicht-Monotonizität in Sprache hat ihre Ursachen in der allgemeinen Kognition.

Um diese Beziehung zu verstehen, gehe ich der Frage nach, ob die Nicht-Monotonizität der Sprache sprachspezifisch ist, das heißt in der Grammatik wurzelt, oder ob sie eine Sache allgemeiner Kognition ist, das heißt mit Pragmatik zu tun hat, die nicht grammatikalisch bedingt ist. Für diese Studie konzentriere ich

Bibliography

meine Forschung auf skalare Implikaturen, einen Bereich der Linguistik, in dem Nicht-Monotonizität eine große Rolle spielt.

Das Resultat zeigt, dass Nicht-Monotonizität sprachunabhängig ist. Des Weiteren bin ich der Meinung, dass die Architektur unseres Verstandes der Art strukturiert ist, dass sie beides zulässt: domänenspezifisches und domänenübergreifendes natürliches Schließen.