Csilla Rákosi: On the rhetoricity of psycholinguistic experiments
Argumentum 10 (2014), 533-547
Debreceni Egyetemi Kiadó

Csilla Rákosi

On the rhetoricity of psycholinguistic experiments*

Abstract

Experiments have to be objective and intersubjectively controllable, and the experimental report must not make use of rhetorical tools that aim merely at persuading the reader but it has to allow the reader a direct access to the experimental evidence. At the same time, however, the reliability of psycholinguistic experiments does not seem to stem from an impersonal and straightforward linkage between "empirical facts" and hypotheses. Rather, it depends crucially on the peculiarities and the plausibility of the argumentation put forward in the experimental report, on its persuasiveness and its convincing force. The present paper aims at resolving this problem that I call *the rhetorical paradox of psycholinguistic experiments*.
Keywords: psycholinguistic experiments, rhetoric, plausible argumentation, philosophy of science

1 Introduction

According to Geeraert's diagnosis, one important step that cognitive linguistics should take in order to reach the status of a scientific enterprise, is the application of empirical methods used successfully within other branches of cognitive science:

> Cognitive Linguistics, if we may believe the name, is a *cognitive science*, i.e. it is one of those scientific disciplines that study the mind […]. It would seem obvious then, that the methods that have proved their value in the cognitive sciences at large have a strong position in Cognitive Linguistics: the experimental techniques of psychology, computer modelling, and neurophysiologic research. (Geeraerts 2006: 28; emphasis as in the original)

Thus, the recent development, namely that reference to psycholinguistic experiments is regarded as one of the most powerful tools in argumentations in favour of, or against, theories in cognitive metaphor research, might be interpreted in such a way that in this field of cognitive linguistics, similarly to cognitive psychology, the idea of treating experimental results as strong evidence for, or against, theories is prevalent. This means that experiments in cognitive metaphor research can, or at least should, be characterised along similar lines as in psycholinguistics, insofar as

* Work on the present paper was supported by the MTA-DE Research Group for Theoretical Linguistics and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. I am grateful to Georg Seel for improving my English.

> […] there is a common, commonly accepted way in psycholinguistics of settling theoretical disputes: experimentation. Given a number of conditions, experimental results decide between competing analyses, and psycholinguists predominantly accept the experimental paradigm as the cornerstone of their discipline. (Geeraerts 2006: 26)

This authority is usually based on the view that experiments allow for confronting hypotheses of theories *directly* with empirical evidence. In this vein, experiments have to be objective and intersubjectively controllable, and apply feasible, well-established procedures providing completely reliable experimental data:

> The conditions that need to be fulfilled to make the paradigm work are in principle simple: the experiment has to be adequately carried out, and it has to be properly designed in order to be distinctive with regard to the competing theories. That is to say, you need good experimental training (knowledge of techniques and analytical tools), and you need the ability to define a relevant experimental design. The bulk of the effort in psycholinguistic research, in other words, involves attending to these two conditions: setting up adequate designs, and carrying out the design while paying due caution to experimental validity. (Geeraerts 2006: 26)

The experimental report has to transmit these characteristics and must not make use of rhetorical tools aimed merely at persuading the reader. From this it follows that the reliability of experiments is supposed to be *inversely proportional* to the rhetoricity of the experimental report.

If, however, we take a closer look at papers dealing with psycholinguistic experiments in cognitive metaphor research, we never actually see the "raw" (numerical) data capturing some observation of linguistic behaviour and a chain of deductively valid inferences leading to the result of the experiment and the latter's confrontation with some hypotheses or theories. Instead, a typical experimental report seems to be a highly complex argumentation which is not strictly deductive. It usually contains, among other things, the following components (not necessarily in this order):

– main tenets, explanatory power, and other strengths of the preferred theory;
– central hypotheses and weak points of the rival theories;
– description of a phenomenon in connection with which the theory and its rivals propose different predictions;
– motivation and description of the experiment to be conducted and conjectures about its outcome;
– details and shortcomings of earlier similar experiments;
– description and results of control experiments aimed at ruling out some known possible systematic errors;
– no "raw data" (individual measurements) at all;
– some excerpts from the stimulus material used;
– type and upshot of statistical analyses;
– presentation of considerations concerning the interpretation and reliability of the results;
– if there seem to be shortcomings in the experiment, then a second experiment is proposed, carried out and its results are analysed, too;
– the impact of the conducted experiment on the theory at hand and its rivals;
– proposals for further inquiry in the given topic etc.

It is plain to see that the relationship between the "raw data" (that is, the complete set of individual measurements) and hypotheses of the linguistic theory or theories at issue cannot be completely reconstructed on the basis of the information provided in the experimental report. Consequently, far from being direct and transparent this relationship is quite fragmentary.

Csilla Rákosi: On the rhetoricity of psycholinguistic experiments
Argumentum 10 (2014), 533-547
Debreceni Egyetemi Kiadó

Despite this, it is the experimental report on the basis of which one decides whether the given experiment is a reliable data source. Compelling, lucid and reasonable experimental reports are regarded as indications of good, reliable experiments and, conversely, poor, shaky, faulty experimental reports lead to the rejection of the experiment itself as well. Therefore, the authority of psycholinguistic experiments does not stem from an impersonal and straightforward linkage between "empirical facts" and hypotheses. Rather, it seems to depend crucially on the peculiarities and plausibility of the argumentation put forward in the experimental report, on its persuasiveness and its convincing force. From this we obtain that the reliability of experiments is *directly proportional* to the rhetoricity of the experimental report.

Thus, our considerations have led to a paradox:

(RPE) *The rhetorical paradox of psycholinguistic experiments*:
The reliability of psycholinguistic experiments as data sources is both directly *and* inversely proportional to the rhetoricity of the experimental report.

The present paper will focus on the following problem:

(P1) How can (RPE) be resolved?

If we examine the two contradictory members of (RPE), two promising starting points present themselves:

1) While the view which considers rhetorical tools unnecessary and worthless is a methodological rule, the opposite view refers to the practice of linguistic research. This should motivate us to raise the question of whether the first view is an *adequate* norm, and, if not, then *under what circumstances* is the practice of presenting the results of experiments in cognitive metaphor research acceptable? That is, the criteria for judging the rhetoricity of experiments should be revealed.
2) The two contradictory views use the concept 'rhetoric' in different senses. The first view reduces it to *irrational tricks and manoeuvres*, erroneously claiming that the experiment provides reliable results. In sharp contrast to this, the second view allows room for interpreting 'rhetoric' as *rational argumentation* that may be fully legitimate and should be an important constituent of scientific experiments.

These findings impel us to transform (P1) to (P2):

(P2) What kind of 'argumentation' is allowed in reports on psycholinguistic experiments and what functions does it fulfil?

(P2) is a metascientific problem that, according to the current literature on experiments in the philosophy of science, cannot be solved on the basis of general and abstract, a priori philosophical considerations, but only with the help of detailed analyses and the evaluation of research practice.

In order to provide a possible solution to (P1) and (P2), we will proceed as follows. Section 2 will be devoted to a brief historical overview of the manifold relationship between experiments and rhetoric/argumentation. It will show that the first view, according to which experimental results should be free from any kind of argumentation, is a contingent historical product – one of the rhetorical/argumentative practices that were applied to secure the authority of experiments in science. In Section 3, we will outline a metatheoretical framework which can be supposed to solve (P2) and (P1). Finally, in Section 4 we will present our solutions to the two problems.

2 A brief history of the relationship between rhetoric/argumentation and scientific experiments

The argumentative/rhetorical tools[1] applied in experiments as well as the role they fulfil have undergone several changes during the history of science.[2] This is mainly due to the variety of ways in which science has been practised and reflected upon.

Experiments were first applied in the 17th century when it became clear that pure reasoning, speculation, passive observation, and reference to ancient authorities or to religious dogmas were no longer capable of providing relevant information about nature. Artificial situations were created, but the usability of experiments and the acceptability of the results were fiercely debated:

> The new "experimental philosophy" was greeted with scepticism on two different grounds. Its critics pointed out two difficulties with regard to experimentation. First, in contrast to the phenomena that could be observed with the unaided senses, the phenomena created by experiment were neither familiar nor accessible to everyone. Second, it was unclear why the manipulation of nature by means of instruments would reveal, rather than distort, its workings. Those difficulties were two aspects of the same issue, namely the authentication of experimental results; an issue which had to be resolved before experimentation could become a proper foundation for natural philosophy. (Arabatzis 2008: 160)

Solely to imagine the processes that might take place under the given circumstances was felt to be unsatisfactory. Thus, the authority of experiments had to be secured by the conduct of the experiment, that is, by the replacement of thought experiments by real ones, as well as by diverse *rhetorical tools*. One method was that scientists listed the names of prominent people who had been present at the experiment. Thus, *the authority of the experimenter and the witnesses* played a decisive role in the appraisal of the experimental results. This was, of course, the application of the earlier rhetorical strategy of reference to authorities to the new method. Another strategy, elaborated Boyle, was also a rhetorical tool insofar as he gave a vivid and detailed account of every phase of the experimental process in order to make the reader, as Cantor (1989: 163) coins it, a *virtual witness*. This strategy has a medieval counterpart as well: narratives in medieval chronicles, as Schiffrin (1981: 59) explains, applied the shift to the historical present as an argumentation technique functioning as an "internal evaluation device: it allows the narrator to present events as if they were occurring at that moment, so that the audience can hear for itself what happened, and can interpret for itself the significance of those events for the experience."[3] Nevertheless, the application of this tool was a very important shift in the role of 'witnessing', since the source of the rhetorical power of the experiment was no longer based on authority but on the (real or only conceived) *reproducibility* of the procedure.

A later, far-reaching move was the use of an *impersonal style* by relying on the dominance of mathematical tools, schematisation and formalisation. This was intended to create the impression that the authority of experimental results stems directly from nature, without (subjective) human

---

1 In this section, the terms 'rhetorical' and 'argumentative' will be used in a pre-explicative sense.
2 Cf. Cantor (1989: 162ff.), Gooding (2000: 117ff.), Arabatzis (2008: 159ff.).
3 For more on this, see Nagy C. (2014).

intervention and interpretation. This style, however, has also led to *fragmentariness*: several details of the experimental process being dismissed from the experimental report.

Despite this, of course, it was often the case that the presented arguments (calculations, formal/mathematical methods, the experimental design, the interpretation of the results) or the devices applied have been criticised. Therefore, the experimental reports have been also *extended* to new elements such as the identification of possible systematic errors and the description of the measures taken for their prevention, statistical analyses enabling the elimination of effects of unavoidable sources of noise, etc. These arguments have grown in importance and have become regarded as decisive factors in judging the acceptability of experimental results. At the same time, the description of the experimental procedure became more theory-oriented.

Thus, the experimental report is considerably *richer* than the experimental procedure itself was, but, at the same time, it remains strongly schematic and informationally *reduced*. Namely, the experimenter selects the relevant moves and events which are accounted for in the experimental report; she/he has the privilege of deciding what counts as an accidental, contingent mistake which may remain unmentioned and, on the other hand, what has to be regarded as a systematic error that has to be reported together with its correction. A highly instructive and often cited example for the *gap* between the "real" happenings in the laboratory and their accounting for in the experimental report is Millikan's celebrated, historic oil drop experiment:

> Yet extant laboratory notebooks also sometimes indicate more interesting mismatches between laboratory practice and published reports. Holton, for example, has drawn attention to Robert Millikan's selection of acceptable results for his oil-drop experiment. During one series of experiments Millikan omitted well over half of his results, retaining data from only 58 drops out of a total of about 140. Against some runs he annotated comments such as "*Beauty*. Publish this surely, *beautiful*", whereas in other cases he dismissed the run with "*Error high* will not use", or some such remark. His reasons for accepting some runs and not others are complex: sometimes parts of this apparatus did not appear to function properly, on other occasions the result was not sufficiently close to the emergent value for *e*, the electronic charge. […] contrary to the manuscript evidence, Millikan announced in his paper: "*It is to be remarked, too, that his is not a selected group of drops but represents all of the drops experimented on during 60 consecutive days*…". (Cantor 1989: 159; emphasis as in the original)

There are, of course, norms – partly formulated explicitly, partly only implicit – governing experiments as well as experimental reports. The fulfilment of the former, however, cannot be checked directly, but only indirectly, with the help of the latter. Nevertheless, as Cantor points out,

> […] a laboratory notebook and a published journal article are two very different literary forms, serving different purposes and subject to different conventions. The published version should not be viewed simply as a tidied up version of the laboratory notes, since the former contains many conventional elements that would find no place in the latter. The publication is a retrospective narrative, an impersonal, passive reconstruction which draws attention to those theories, tests and data which are considered appropriate for consumption by the scientific community. (Cantor 1989: 160)

This has significant consequences for the evaluation of experiments as data sources. It is the *scientific community* that decides whether the experimental results are reliable and epistemologically useful, that is, whether they can be used for theory testing, explanation, elaboration of new theories etc. This decision is based not on the analysis of the experimental procedure itself but only on the *judgement of the experimental report produced by the experimenter*. From this it follows that the rhetorical power of the latter is a decisive factor in this case, too.

Although this historical overview is somewhat fragmentary, it clearly shows that the norms related to the acceptability of experiments have undergone several changes. Moreover, there are different norms in different branches of science which are often contested. The same holds true of the rhetorical/argumentative aspects of experimental reports as well. That is, the structure and the rhetorical/argumentative tools applied in experimental reports are *social products* as well.

At this point, of course, the question arises of whether it is possible to elaborate a metatheoretical model of scientific experiments that is capable of accounting for the relationship between the experimental process and the (argumentative) experimental report. This model has to allow for *an evaluation of the reliability of the experimental process on the basis of the arguments presented in the experimental report*.

3 A possible metatheoretical framework

3.1 Components of the experimental process

According to current literature on the philosophy of science, experiments are remarkably complex entities. They comprise several ontologically diverse components such as:

– *experimental design*: a comprehensive preliminary description of the process of experimentation – practically, a thought experiment, providing conjectures about the outcome of the experimental procedure;
– *experimental procedure*: a material procedure where an experimental apparatus is set up, and its working is monitored and recorded under controlled circumstances;
– *a theoretical model of the phenomena investigated*: a description of low-level theoretical constructs (phenomena) that may be relevant in judging hypotheses about high-level theoretical constructs or require theoretical explanation;
– *a theoretical model of the experimental apparatus*: explanations about how phenomena are created or separated from the background, which of their properties can be detected with the help of the equipment, and why it can be supposed that the perceptual data produced by the apparatus are stable and reliable;
– *authentication of perceptual data*: evaluation of the outcome of the experimental procedure, and the experimenter's decision as to whether the experimental apparatus has been working properly so that perceptual data (records of measurements, photographs etc.) are stable and reliable; checking of whether sources of noise have been ruled out, or their effect can at least be eliminated with the help of statistical methods;
– *interpretation of perceptual data*: establishing a connection between the perceptual data gained and the phenomena investigated. It has to be decided whether the former are relevant, real and reliable in relation to the latter;
– *presentation of experimental results*: since experiments are not private but public affairs aimed at supplying data for scientific theorising, it is not only the results of the experiment which have to be put forward, but also every element of the experimental procedure that is judged relevant to the evaluation and acknowledgement of the results. Therefore, the experimenter has to present an argumentation that conforms to certain norms. It should contain all information that may have any significance for the scientific community in enabling them to decide whether the experimental results are reliable and epistemologically useful, that is, whether they can be used for theory testing, explanation, elaboration of new theories etc. To this end, relevant pieces of information have to be selected and arranged into a well-built chain of arguments leading from the previous

Csilla Rákosi: On the rhetoricity of psycholinguistic experiments
Argumentum 10 (2014), 533-547
Debreceni Egyetemi Kiadó

problems raised through the description of the experimental design and the experimental procedure to the evaluation (authentication and interpretation) of data. Thus, experimental data should be suitable for integration into the process of scientific theorising.

From this short characterisation it is clear that experiments involve a highly complex network of activities, physical objects, argumentation processes, interpretative techniques, background knowledge, methods, norms, etc. The reliability of the outcome of an experiment depends on the reliability of its *components* as well as the *fit* between them and pre-existing knowledge.

A very important insight of the current literature on scientific experiments is that neither single experiments nor repetitions of the experimentation process are capable of yielding ultimate and unquestionable results. It is not only the previous considerations and the planning of the experiment which are fallible – the control of the experimental process and the evaluation of the results are to some extent unavoidably uncertain as well. Therefore, experiments are *open processes* in the sense that, in possession of new pieces of information, they may be continued, modified, or even discarded.

From this it follows that the experimental process should be viewed as a *search for the best fit achievable* between the experimental design, the theory of the experimental apparatus, the process of experimentation, the perceptual data gained, the authentication and interpretation of the latter, the theory of the phenomenon investigated, etc.[4] To find this fit, one has, in most cases, to *turn back* to earlier stages of the experimentation process and modify some component. Every component can be *revised* and the revisions have to be repeated again and again till there is mutual support among the constituents. See Figure 1.

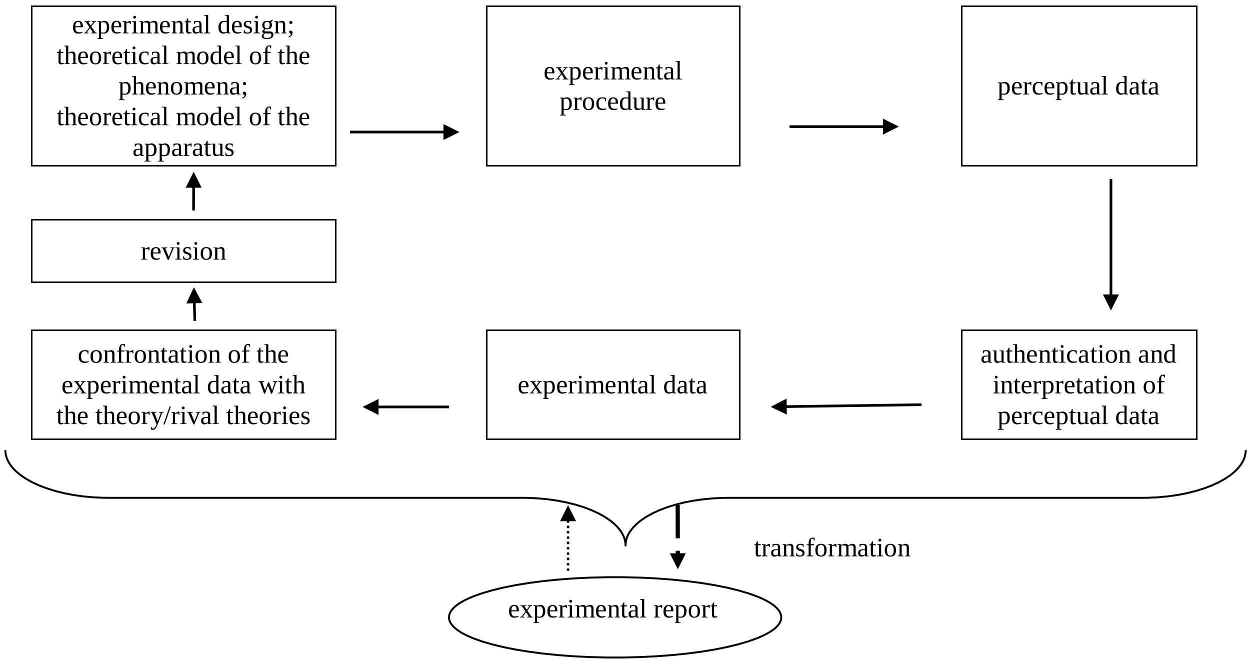

Figure 1: The structure of experiments

4 See, for example, Pickering (1989), Hacking (1992: 56).

3.2 Argumentative aspects of experiments

Let us take a closer look at the process of searching for the best fit among the components of the experimental process *from the point of view of the experimenter*.

Hypotheses used in the experimental design, the theoretical model of the phenomena and of the apparatus make up the starting point of the experimental process. They are not true with certainty but they are supported to some extent by theoretical considerations, by earlier experiments, or are simply (reasonable) conjectures. They allow for a rough estimation of the outcome of the experiment. After the experimental procedure, in possession of the perceptual data, this preliminary guess may be strengthened. Nevertheless, it may happen that the perceptual data cannot be interpreted properly, or they seem to be in conflict with the predictions. In such cases, the reliability of the previously accepted hypotheses also has to be revised.

The interpretation and authentication of the perceptual data may also indicate shortcomings in the experimental procedure, in the experimental design, or in the theoretical model of the phenomena or of the apparatus. Therefore, all facets of the experiment conducted have to be re-examined, and, if it seems to be necessary, control experiments have to be carried out, or the experimental design has to be modified and the experiment repeated. Moreover, even the interpretation or the authentication of the perceptual data itself may be faulty and be in need of modification.

From this it follows that revealing the *connections* between the statements capturing different aspects of the experimental procedure and their *analysis,* as well as the *comprehensiveness* of the checks and cross-checks are of crucial importance.

This characterisation of the experimental process will motivate us to raise the hypothesis that *experiments are cyclic processes organised and conducted by an argumentation process* which tries to clarify the relationship among hypotheses of the experimental design, the theoretical model of phenomena, the theoretical model of the experimental apparatus, the theory under test and its rivals, as well as statements describing the events of the experimental procedure, or which capture the results of the interpretation and authentication of perceptual data etc. This motivates us to modify Figure 1 in the following way.

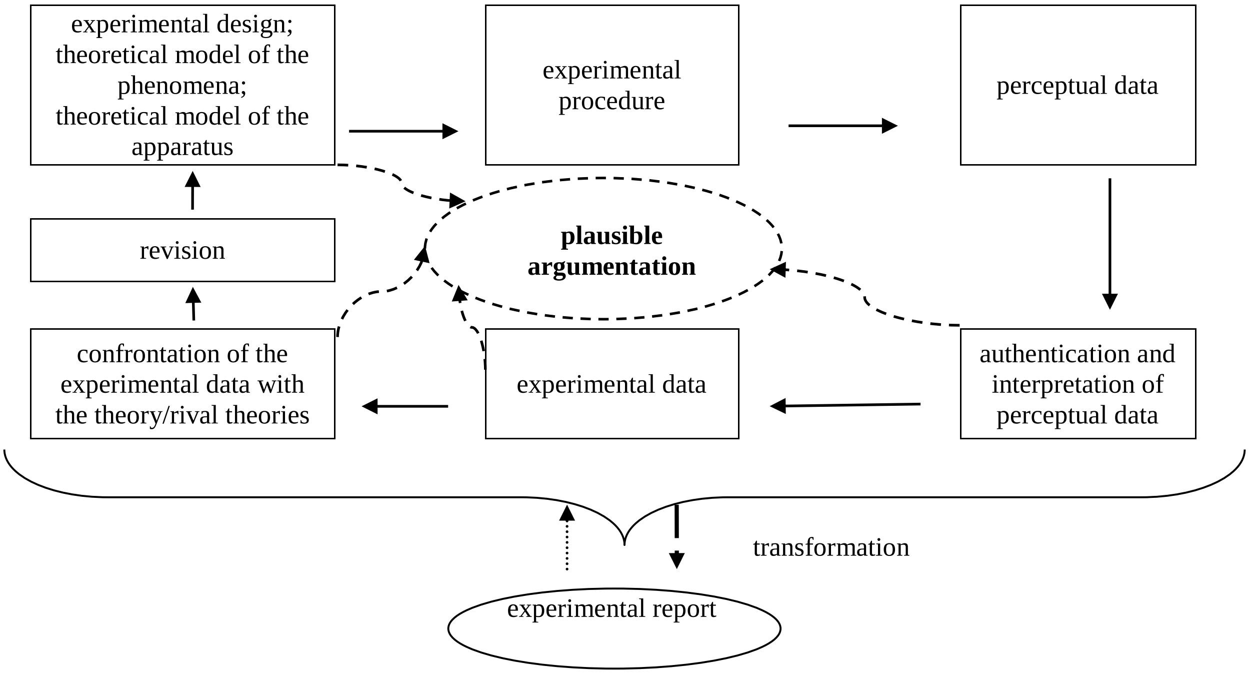

Figure 2: The structure of experiments – revised

The argumentation process organising the conduct and control of the experiment is a central issue in judging the reliability of the experimental results. This argumentation process does not consist of deductive inferences because it takes into consideration the uncertainty of the statements as well. It is not public; rather, it is a private affair of the experimenters. Despite this, it is indirectly influenced by the public norms applicable to experiments. Thus, experimenters have first to convince themselves of the reliability and acceptability of the outcome of the experiment and, after this, they have to persuade the scientific community as well. This means that the researcher has to transform this *private argumentation process* into a *public experimental report*.

If we turn to the *reader's perspective*, we can establish that the evaluation of experimental results can only start from the experimental report, which is an edited, transformed version of the non-public argumentation process. While the latter is part of an ontologically complex process of scientific experiment, the former is purely argumentative. It is a *mixture*: it contains elements or traces of the original argumentation process as well as the argumentative tools needed to make this reduced set of information coherent, comprehensible and persuasive for the reader.

3.3 Modelling of the argumentative aspects of experiments

At this point, of course, the question arises of which metatheoretical tool enables us to reconstruct and evaluate the argumentation process which governs the experimental process and the experimental report, as well as the relationship between them. A purely logical analysis would not suffice because it is formal and is not capable of grasping information related to the uncertainty and reliability of statements. This motivates the application of the *p-model*, elaborated in Kertész & Rákosi (2012).[5]

[5] For further applications of the p-model to experiments, see Rákosi (2011a, b, 2012), Kertész & Rákosi (2012: Part IV), Kertész & Kiefer (2013), Kertész & Rákosi (2014).

3.3.1 The uncertainty of information in experiments: plausible statements

The p-model – following Rescher (1976) – does not interpret scientific hypotheses as propositions but assigns them a structure consisting of an *information content* and a *plausibility value*. The plausibility value of a statement indicates the extent to which its information content is supported, is made reliable by a given *source*, and as a result, to the extent to which we are ready to accept it.

In connection with psycholinguistic experiments, the following plausibility rankings may be applied – of course, this list is only a sample of possible rankings:

$|p|_D =$ 0, that is, the statement *p* has neutral plausibility according to the experimental design abbreviated as *D*, if the experimental design does not allow the risk of even a rough estimate for the plausibility of the statement *p*, and neither *p* nor its negation is supported by *D*;

$|p|_K =$ 0.2, that is, *p* has low plausibility according to the experimenter's knowledge abbreviated as *K*, if *p* is the experimenters' previous, untested and vague conjecture about the outcome of the experiment;

$|p|_{E_1} =$ 0.4, that is, *p* has a rather low plausibility according to an earlier experiment abbreviated as E_1, if *p* results from an experiment, but some possible sources of noise which may cause systematic errors have not yet been ruled out with the help of control experiments;

$|p|_{E_2} =$ 0.6, that is, *p* has a rather high plausibility according to an experiment E_2, if *p* results from a well-designed experiment with a thorough authentication of the perceptual data;

$|p|_T =$ 0.8, that is, *p* has a high plausibility according to a theory abbreviated as *T*, if *p* is a central, generally accepted hypothesis of the given theory which has already been tested with the help of linguistic, corpus linguistic, psycholinguistic etc. investigations;

$|p|_M =$ 1, that is, *p* can be regarded as true with certainty on the basis of a mathematical theory *M*, if *p* is a mathematical theorem proven in *M*.

It has to be stressed that low plausibility values do not mean a statement is improbable but rather that it has a relatively small, limited amount of plausibility (reliability, acceptance). In such cases, the source votes expressly for the given hypothesis. If a source is against a hypothesis then it makes its negation plausible and the given hypothesis *implausible* or even false with certainty; that is, in such cases $0 < |{\sim}h|_S \leq 1$.

The concept of 'plausibility value' allows us to represent and compare the acceptability (reliability) of statements such as previous conjectures, perceptual data, experimental data, hypotheses of linguistic theories, hypotheses about linguistic phenomena, etc. The experimenter's hypothesis about the correctness of the experimental design or about the flawless functioning of the measuring devices can also be only plausible but not certainly true. From the experimenter's point of view this means that the non-public argumentation process which organises and conducts the experimental process deals with uncertain, fallible pieces of information. From the reader's perspective this means that the experimental report consists of plausible but, in most cases, not certainly true statements. Moreover, the concept of 'plausibility' makes it possible to compare the plausibility value which can be assigned to statements on the basis of the identification of their source on the one hand, and the value which they receive in the experimental report on the other. If the latter values are higher than the former, then this indicates an unwarranted overestimation of the plausibility of certain hypotheses or data and leads to a fallacious argumentation.

3.3.2 Sources of plausibility

We distinguish direct and indirect sources. In the case of *direct sources*, the plausibility of the statement at issue is evaluated with respect to the reliability of its source, as above. *Indirect sources* yield the plausibility value of the given statement on the basis of the plausibility of other statements – that is, via *plausible inferences*. Plausible inferences take into consideration not only the logical structure of the premises and the conclusion but their plausibility values and semantic structure as well. They always rest on a semantic relation: for example, causality, analogy, similarity, sign, necessary or sufficient condition, part-whole relation etc., and are not necessarily deductively valid.

The perfect identification of the direct and indirect sources from which the plausibility of the data and other hypotheses in experimental reports originate makes it possible to check and re-evaluate the plausibility of the statements at issue. Specifically, the reconstruction of the plausible inferences (indirect sources) applied in the experimental report may reveal latent background assumptions that are implausible instead of being plausible or of neutral plausibility. It may happen that an inference relies on a hypothesis that is solely a conjecture but which on closer examination turns out to be implausible or false. In such cases the conclusion loses its plausibility as well – and the same holds true of the inferences that made use of the conclusion of this inference as a premise. This kind of reconstruction may be especially useful for the authentication of the perceptual data as well as for establishing a link between the experimental data and the hypotheses of a theory. In both cases the connection between the perceptual data and the experimental data and between the experimental data and theoretical hypotheses relies mostly on deductively invalid plausible inferences that make use of latent background assumptions.

3.3.3 Conflicting information in scientific experiments: p-inconsistency

An important property of the above concept of plausibility is that it allows a statement to be plausible on the basis of some sources and implausible on the basis of others at the same time. Such cases are called *p-inconsistencies*.

Thus, a hypothesis may be made plausible by an experiment as a source but implausible by another one. Similarly, different theories may judge the acceptability of a given scientific claim differently, or an experiment may refute a prediction, etc. – leading to different cases of p-inconsistency.

The decision between conflicting hypotheses cannot be reduced to the mechanical comparison of their plausibility values. Instead, one has to evaluate statements along with the reliability of the sources making them plausible, their relationship to other statements, to the related methodological norms and so on – that is, the system of relations of the rival hypotheses has to be revealed and compared as a whole. Such constellations are called *the p-context*.

According to the p-model, inconsistencies must not be viewed as fatal failures but indications that either the experiment or the theory at issue (or even both) is in need of some kind of modification. Thus, conflicting experimental results, contradictions between predictions and experimental data, inconsistencies between the hypotheses of a theory and the results of an experiment, and other discrepancies among the components of the experimental process are concomitants of experiments. Nevertheless, there are always several possible causes of a conflict, whose identification may require several attempts. In most cases, inconsistencies are not resolved by simply giving up one of the conflicting statements but more comprehensive revisions are needed that may affect further components of the experiment as well.

3.3.4 Solutions and the resolution of p-inconsistencies

In order to resolve a p-inconsistency, one has to *re-evaluate* the p-context. A *solution of a p-inconsistency* is achieved if a p-context has been arrived at in which (a) the statement in question is unanimously supported or opposed by the sources – that is, it has become either plausible or implausible (or even certainly true or false) on the basis of *all* sources in the given p-context –, or (b) the statements causing inconsistency are represented separately and this separation is systematic and well-motivated.

It is possible, however, that a p-inconsistency has several solutions. This necessitates the introduction of the notion of the *resolution of a p-inconsistency*. This means that one finds a solution of the given p-inconsistency which is, when compared with other solutions, the best on the basis of a particular set of accepted criteria, and according to the information available for us in the given p-context. It may be the case, however, that in a given information state one can only show that for the time being there is no resolution achievable.

It is of vital importance that inconsistencies are not put aside without finding a solution which makes it possible to separate the conflicting statements, at least provisionally. Instead, one has to try to elaborate and compare as many solutions as possible in order to find the best solution available under the given information state.

The reliability of an experiment as a data source is largely determined by careful and strict identification of the inconsistencies among its components, by the number, variety, and comprehensiveness of the investigated solutions as well as by the choice of the resolution of the conflicts revealed during the experimental process. Since the p-model describes several strategies of inconsistency resolution, its application may contribute to the elaboration and conduct of better experiments in linguistics.

3.3.5 Cyclic revisions in scientific experiments: plausible argumentation

To achieve the solutions or the resolution of a given p-inconsistency, one needs a *heuristic tool* that makes it possible to re-evaluate the p-context and to find and compare the solutions to its problems. This heuristic tool is what we will call *plausible argumentation.* In simple terms, plausible argumentation is the transformation of a problematic p-context into one that is no longer (or at least, less) problematic. This involves the successive re-evaluation of a problematic p-context by the elaboration of possible solutions to its problems, the evaluation of the alternative solutions and the comparison of the latter. Its aim is the detection of all available solutions and the decision as to which of them is to be accepted as the resolution of the given p-problem.

The above characterisation of plausible argumentation indicates that the argumentation process is basically not linear but *cyclic,* because the re-evaluation of a problematic p-context usually does not lead immediately to an unproblematic one but may raise new problems. This may require the revision of previous decisions, the assessment of other alternatives etc. Therefore, throughout the argumentation process one returns to the problems at issue again and again, and re-evaluates the earlier decisions about the acceptance or rejection of

Csilla Rákosi: On the rhetoricity of psycholinguistic experiments
Argumentum 10 (2014), 533-547
Debreceni Egyetemi Kiadó

statements, the reliability of the sources, the plausibility values of the statements, the workability of methodological norms, the conclusions previously reached by inferences etc.

The p-model's concept of 'plausible argumentation' allows us to interpret both the argumentation organising and conducting the experimental process and the experimental report as pieces of plausible argumentation. The experimental report should not simply summarise and make public the results of the former but make it possible for the reader to continue the non-public argumentation process. That is, a good experimental report is informative enough to allow the reader to add new argumentation cycles to the non-public argumentation process.

Thus, for example, the reliability of an experiment crucially depends on the question of to what extent the experimental data may be supposed to be free of systematic errors. In experiments on metaphor processing, by the application of an offline measure, participants might have made use of conscious strategic considerations distorting the results, or semantic priming effects lead to faulty results, etc. Therefore, when the experimenter suspects or reveals the presence of such a factor, then he/she has to carry out control experiments and/or revise the experimental design and start a new cycle of revision. Nevertheless, one cannot check and rule out the presence of every possible systematic error. The set of the factors that might have influenced the outcome of the experiment is always open. From this it follows that even a good experiment may contain systematic errors that can be revealed only later by some other member of the scientific community. Thus, good experiments are characterised not only by the thoroughness of the elimination of possible errors but are also inspirational and motivate the search for more complex explanations of the investigated phenomena. They pave the way for new experiments that take into consideration further factors and for the elaboration of more refined theoretical models.

4 The solution to (P1) and (P2)

On the basis of our analyses presented in Section 3, we obtain the following solution to (P2):

(S2) Experiments have a *dual argumentative character*. The experimental process is organised and conducted by *a non-public* plausible argumentation process that is then transformed into the experimental report, that is, a *public* piece of plausible argumentation. This transformation can be regarded as acceptable if it does not change the plausibility value of the hypotheses of the original, non-public argumentation. This can be achieved if the experimental report contains all information that might be relevant for the evaluation of the steps of the experimental process. This means that the reader has to be in a position to reconstruct and continue the non-public argumentation process to the greatest extent possible. In this way, the reader can be made *a virtual participant of the creation, analysis and evaluation of the perceptual data*. This means that he/she has to be equipped to reconstruct not only the experimental procedure but the interpretation and authentication of the perceptual data and the significant steps of the argumentation process organising the elaboration of the experimental results.

This solution to (P2) makes it possible to resolve the rhetorical paradox of psycholinguistic experiments as follows:

(S1) The reliability of psycholinguistic experiments as data sources can be judged on the basis of the argumentation theoretical analysis of the experimental report with the help of the p-model. The p-model provides tools for, among others

(a) revealing the *sources* from which the plausibility of the data and other hypotheses used in experimental reports originate. In this way, the points where plausibility values enter the argumentation process can be identified and their reliability evaluated;

(b) representing the *acceptability* of statements as plausibility values. In this way, it can be determined which sources make the statements in experimental reports plausible or implausible, and to what extent;

(c) *determining the plausibility value of conclusions of inferences* with premises that are not true with certainty but only plausible to some extent. Thus, not only the impact of direct sources on the plausibility of the statements can be represented, but the impact of plausible or implausible statements on each other as well;

(d) *comparing and summarising* the plausibility value of hypotheses stemming from different sources. Therefore, the dynamism of the change in the plausibility of data and hypotheses can be accounted for;

(e) *representing the emergence of inconsistencies* and the *strategies* applied to their resolution as plausible argumentation processes. Organising all information at hand (data, sources, hypotheses, inferences, methodological tools, methodological norms etc.), with the introduction of the notion of 'plausible argumentation', the p-model tries to model the comprehensive effect of every change in the informational state;

(f) *comparing and evaluating different solutions* to the problems revealed via differentiating between effective (plausible, cyclic and prismatic) and ineffective (fallacious, circular) argumentation.

References

Arabatzis, T. (2008): Experiment. In: Psillos, S. & Curd, M. (eds.): *The Routledge Companion to Philosophy of Science*. London & New York: Routledge, 159-170.

Cantor, G. (1989): The rhetoric of experiment. In: Gooding, D., Pinch, T. & Schaffer, S. (eds.): *The Uses of Experiment*. Cambridge: Cambridge University Press, 159-180.

Geeraerts, D. (2006): Methodology in cognitive linguistics. In: Kristiansen, G., Achard, M., Dirven, R. & de Mendoza Ibáñez, F.J.R. (eds.)(2006): *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlin & New York: de Gruyter, 21-49.

Gooding, D.C. (2000): Experiment. In: Newton-Smith, W.H. (ed.): *A Companion to the Philosophy of Science*. Malden & Oxford: Blackwell, 117-126.

Kertész, A. & Rákosi, Cs. (2012): *Data and Evidence in Linguistics: A Plausible Argumentation Model*. Cambridge: Cambridge University Press.

Kertész, A. & Kiefer, F. (2013): From thought experiments to real experiments in pragmatics. In: Capone, A., Lo Piparo, F., & Carapezza, M. (eds.): *Perspectives on Philosophy and Pragmatics*. Berlin, Heidelberg & New York: Springer, 53-86.

Csilla Rákosi: On the rhetoricity of psycholinguistic experiments
Argumentum 10 (2014), 533-547
Debreceni Egyetemi Kiadó

Kertész, A. & Rákosi, Cs. (2014): Thought experiments and real experiments as converging data sources in pragmatics. In: Kertész, A. & Rákosi, Cs. (eds.): *The Evidential Basis of Linguistic Argumentation*. Amsterdam & Philadelphia: John Benjamins, 221-270.

Nagy, C.K. (2014): Methods and argumentation in historical linguistics: A case study. In: Kertész, A. & Rákosi, Cs. (eds.): *The Evidential Basis of Linguistic Argumentation*. Amsterdam & Philadelphia: John Benjamins, 71-102.

Nickles, T. (1989): Justification and experiment. In: Gooding, D., Pinch, T. & Schaffer, S. (eds.): *The Uses of Experiment*. Cambridge: Cambridge University Press, 299-333.

Rákosi, Cs. (2011a): Metatheoretical reconstruction of psycholinguistic experiments. Part 1. *Sprachtheorie und germanistische Linguistik* 21(1), 55-93.

Rákosi, Cs. (2011b): Metatheoretical reconstruction of psycholinguistic experiments. Part 2. *Sprachtheorie und germanistische Linguistik* 21(2), 159-187.

Rákosi, Cs. (2012): The fabulous engine: Strengths and flaws of psycholinguistic experiments. *Language Sciences* 34(6), 682-702.

Csilla Rákosi
MTA-DE Research Group for Theoretical Linguistics
H-4010 Debrecen
Pf. 47
rakosics@gmail.com