The 'Galilean style in science' and the inconsistency of linguistic theorising

András Kertész

Abstract

Chomsky's principle of epistemological tolerance says that in theoretical linguistics contradictions between the data and the hypotheses may be temporarily tolerated in order to protect the explanatory power of the theory. The paper raises the following problem: What kinds of contradictions may be tolerated between the data and the hypotheses in theoretical linguistics? First a model of paraconsistent logic is introduced which differentiates between week and strong contradiction. As a second step, a case study is carried out which exemplifies that the principle of epistemological tolerance may be interpreted as the tolerance of week contradiction. The third step of the argumentation focuses on another case study which exemplifies that the principle of epistemological tolerance must not be interpreted as the tolerance of strong contradiction. The reason for the latter insight is the unreliability and the uncertainty of introspective data. From this finding the author draws the conclusion that it is the integration of different data types that may lead to the improvement of current theoretical linguistics and that the integration of different data types requires a novel methodology which, for the time being, is not available.

Keywords: Chomsky, Galilean style, epistemological tolerance, theoretical linguistics,

1. Introduction

The problem of whether linguistic theories count as empirical is one of the central foundational issues that have been discussed since the advent of theoretical linguistics in the twentieth century. Although the discussion of this topic has been continuous, its intensity and its focus have varied. Currently, the empiricalness problem is manifested in a controversy on the reliability of linguistic data and their capability of supporting or refuting the hypotheses of theories.[1] The problem of linguistic data is especially crucial with respect to grammaticality judgments. As we know, during the past five decades the *primary data type* used to support the hypotheses put forward by linguistic theories has been grammaticality judgments rooted in the native speaker's introspection as the *primary source of data*. However, in current debates the reliability of introspection as a data source and, consequently, the tenability of linguistic hypotheses built on grammaticality judgments stemming from introspection, have been seriously questioned. In this debate, the symptoms of a possible crisis may be detected, because some of the views put forward maintain that the development of theoretical linguistics has been slowed down by the problematic nature of linguistic data in general and that of grammaticality judgments rooted in introspection in particular (see e.g. Featherston 2007, 2009, Schlesewsky 2009, Schütze 1996 etc.).

The present paper tackles the reliability of introspection and grammaticality judgments from a point of view which is of utmost importance but which has not yet been considered in the literature: namely, the methodological consequences yielded by the logical structure of contradictions between the data and the hypotheses.

During the history of theoretical linguistics, the contradictions between the data and the hypotheses have been evaluated in two extreme ways. It was Chomsky who introduced one of the extremes into theoretical linguistics as a component of what he means by the 'Galilean style in science'. In Chomsky's interpretation, the 'Galilean style in science' consists of three components (for discussion see Botha 1983, Brame 1984, Boeckx 2006, 2010, Freidin 2007, Freidin – Vergnaud 2001, Mukherji 2010, Riemer 2009).[2] The first is the conviction that the way nature works can be captured by mathematical structures. According to the second, one may reach a deep understanding of the universe by abstracting from what can be observed on the surface of phenomena. The third is what the literature calls 'epistemological tolerance' and considers the key component of Chomsky's idea of the 'Galilean style in science' (Botha, 1983, Klausenburger 1983, Riemer 2009):

> "What was striking about Galileo, and what was considered very offensive at that time, was that he *dismissed a lot of data*; he was willing to say *'Look, if the data refute the theory, the data are probably wrong.'* And the data that he threw out were not minor. […] Euler, Gauss, and so on […] just said: 'We'll live with the problems and do the mathematics and some day it will be figured out', which is essentially Galileo's attitude towards things flying off the earth. That's pretty much what happened. During the first half of the nineteenth century Gauss, for example, was creating a good part of modern mathematics, but kind of intuitively, without a formalized theory, in fact with approaches that had *internal contradictions*. […] now we understand it, but for a long

---

[2] According to a certain part of the literature what Chomsky seems to mean by this term has little to do with the views Galileo had. See for this e.g. Botha 1983 and Riemer 2009 as well as the historiographic literature cited there. Nevertheless, in this paper it is not our task to check whether Chomsky interpreted Galileo correctly or not. What matters is that, independently of its origin, the 'Galilean style in science' refers to a set of methodological principles which Chomsky expects to be followed in theoretical linguistics.

period, in fact right through the classical period, the systems were informal and even *contradictory*. […] And what's true of mathematics is going to be *true of everything*. […] The recognition that that's the way science ought to go *if we want understanding*, or the way that any kind of *rational inquiry* ought to go – that was quite a big step and it had many parts, like the Galilean move towards discarding recalcitrant phenomena […] *That's all part of the methodology of science*." (Chomsky 2002: 98-102; emphasis added)

Referring to Weinberg (1976), from the late 1970s on Chomsky has put forward his view of the 'Galielan style in science' in different contexts (see also Chomsky 1980a: 2, 1980b: 9–10, 2004: 25, 170–172). If we relate the above quotation to linguistic theories, we obtain the following formulation of *the principle of epistemological tolerance*:

(PET) In a linguistic theory
   (a)   contradictions between the data and the hypotheses may be *temporarily tolerated*,
   (b)   in the hope that the *later development* of the theory will resolve them, and
   (c)   in order to protect the theory's *explanatory* principles.

However, a couple of decades earlier, in his seminal book first published in 1943, which played a decisive role in the development of theoretical linguistics, Louis Hjelmslev maintained the opposite view:

"The description *shall be free of contradiction (self-consistent)*, exhaustive, and as simple as possible. The requirement of freedom from contradiction *takes precedence* over the requirement of exhaustive description. The requirement of exhaustive description *takes precedence* over the requirement of simplicity." (Hjelmslev 1969: 11; emphasis added)

Hjelmslev calls this methodological stance the 'empirical principle' but adds that in the light of epistemological considerations it may be renamed. For the purposes of the present paper, in analogy to the principle of epistemological tolerance, we will call it the *principle of epistemological intolerance* and introduce it as follows:

(PEI) Non-contradiction is the *most important* methodological requirement of linguistic theorising; therefore, contradictions between the data and the hypotheses *must not* be tolerated.

In the case of both extremes, the author considers his own principle to be generally valid for the whole of theoretical linguistics. For example, Hjelmslev (1969: 11) maintains that (PEI) applies to "every theory". Chomsky (2002: 101-102), as the above quotation witnesses, illustrates (PET) by examples taken from mathematics and emphasises that "what's true of mathematics is going to be *true of everything*". (PET) has been applied implicitly in the everyday practice of linguistic research, but is also reflected on explicitly in different subfields beyond syntax: for example, in word formation and semantics (Botha 1988), in phonology (Hale – Reiss 2008), or in historical linguistics (Hale 2007). Therefore – although, in order to prepare the problem to be raised in a pointed way, we started from quotations by Chomsky and Hjelmslev – the subject matter of the present paper is not restricted to the differences between the methodology of Hjelmslev's glossematics and Chomsky's generative linguistics, but affects *the foundations of theoretical linguistics in general*.

Namely, against the background sketched, the question arises whether theoretical linguists should follow (PET) or (PEI). Since this is a metascientific question, it is the philosophy of science that is expected to provide points of departure for finding the answer.

According to the Analytical Philosophy of Science – which played a decisive role in shaping our understanding of scientific inquiry in the 20th century – one pillar of rational inquiry is the principle of non-contradiction in the sense of classical two-valued logic which says that no statement can be both true and false. Contradictory systems of statements cannot be rational, because in such systems any statement can be inferred and therefore they result in *logical chaos*. A scientific theory which includes a contradiction is not able to achieve the primary aim of inquiry insofar as it cannot make claims about the world.[3]

However, there is also another point of departure, because, as we know, the Analytical Philosophy of Science is by now out of date. New developments have systematically rejected its tenets. Among others, current trends in the philosophy of science have re-evaluated the structure and function of contradictions in scientific theories as well (see e.g. the papers in Meheus ed. 2002). First, contradictions may differ with respect to their structure and there are *different kinds* of contradiction. Second, *not* all contradictions are *harmful*. Different kinds of contradiction may have different functions in scientific theorising. Third, new systems of logic have been elaborated which allow for certain kinds of contradiction without being exposed to logical chaos. Such logics are called *paraconsistent* (for overviews see e.g. Priest 2002, Priest – Beall – Armour-Garb (eds.) 2004).

The comparison of these two possible points of departure offered by the philosophy of science shows that the question we have raised has no simple answer. Rather, it should be reformulated as follows:

(P)   *What kinds* of contradiction between the data and the hypotheses *may be* tolerated in linguistic theories and what kinds *must not*?

(P) is the central problem of the present paper. However, due to its complexity, we will not be able to solve it directly, but we might arrive at its solution by subdividing it into a series of more easily accessible sub-problems, the solutions to which will eventually yield the solution to (P).

Since 'contradiction' is a logical category, as a first step we have to find a logic capable of modelling different kinds of contradiction between the data and the hypotheses. Thus, *the first sub-problem* is as follows:

(P1)   What logic is capable of providing a solution to (P)?

As the solution to (P1), in Section 2 we will outline a logic which distinguishes between strong and weak contradiction. After having briefly discussed the logic to be applied, we will have to decide whether the principle of epistemological tolerance can be re-interpreted as one or both of these kinds of contradiction. Accordingly, the *second sub-problem* is:

---

[3] See for example:
"The requirement of consistency plays a special role among the various requirements which a theoretical system, or an axiomatic system, must satisfy. It can be regarded as the first of the requirements to be satisfied by *every* theoretical system, be it empirical or non-empirical. […] the importance of the requirement of consistency will be appreciated if one realizes that a self-contradictory system is uninformative. It is so because any conclusion we please can be derived from it. Thus no statement is singled out, either as incompatible or as derivable, since all are derivable." (Popper 1980 [1959]: 91-92; emphasis as in the original)

(P2)   Can the principle of epistemological tolerance be re-interpreted as the tolerance of weak contradiction between the data and the hypotheses?

In Section 3 we will discuss a case study which serves to illustrate that the answer to this question is: yes. As the next step, the *third sub-problem* will be raised:

(P3)   Can the principle of epistemological tolerance be re-interpreted as the tolerance of strong contradiction between the data and the hypotheses?

Section 4 will be devoted to a case study, too, which, however, will show that the solution to (P3) is: no.

The solutions to the sub-problems (P1)-(P3) will lead to the solution of our basic problem (P) in Section 5. Nevertheless, the solution thus obtained raises a new problem, the far-reaching methodological consequences of which will be dealt with in Section 6.

Before tackling the first sub-problem, a terminological remark must be made. To avoid misunderstandings, it is worth clarifying in what sense we will use the notion of grammaticality judgment as a data type. Consider (1):

(1)   (a)   The sentence
                It is unclear what shocked whom
                is grammatically correct.
        (b)   The sentence
                It is unclear whom what shocked
                is ungrammatical.
        (c)   *It is unclear what shocked whom.*

The *statements* in (1)(a) and (b) are *data* whose *source* is the introspection of one or more informants. Such statements claim about a sentence that it has or has not the property of being grammatically correct. The *sentence* in (1)(c) does not count as a datum. In order to emphasise the structure of grammaticality judgments as data, instead of the conventional notation which distinguishes between starred sentences marking ungrammaticality and unstarred ones standing for grammatical correctness, in the examples we will follow the pattern of (1)(a) and (b), and explicitly decompose judgments into a sentence and the property 'grammatically correct'/'ungrammatical'. In this way – since both the data to be discussed and the hypotheses at issue are treated as statements, and since, from a logical point of view, 'contradiction' is a relation between statements – it will be legitimate to examine possible contradictions between the data and the hypotheses of the theory in question.

2.   The solution to (P1): Two kinds of contradiction

The main issue which paraconsistent logics tackle is how, despite the presence of a contradiction, logical chaos can be avoided. At present, there are several paraconsistent logics available which suggest different approaches to this issue. In order to obtain a solution to (P1), we have chosen Rescher – Brandom (1980).[4]

Rescher and Brandom's logic presupposes a Kripke-semantics which the authors modify in two respects, among others. The first modification, which is important from the

---

[4] See Kertész (2004) for another application of Rescher – Brandom (1980) to linguistics.

point of view of our problem (P), is the introduction of the operation of *superposition* on the set of possible worlds.

(2)   The superposition of the possible world w_1 and w_2 is a possible world w in which a statement p is true if and only if it is true either in w_1 or in w_2.

In other words, p is true in $w_1 \cup w_2$ if and only if it is true *at least* in one of the component worlds. The notation is $w_1 \cup w_2 = w$, where '$\cup$' stands for the operation of superposition. See the truth-table below.

p is true in w_1	p is true in w_2	p is true in $w = w_1 \cup w_2$
+	+	+
+	−	+
−	+	+
−	−	−

Table 1

As a result of (2), in w it may be the case that both p and $\sim p$ are true – however, separately in the two different component-worlds of w. p & $\sim p$ is not true in w because it is neither true in w_1, nor in w_2.[5]

The second modification applies to the semantic interpretation of the valid inference principle of classical logic. The latter says:

(3) $p_1, p_2, ..., p_n \vdash q$

(3) is ambiguous, because it can be interpreted semantically in two different ways. According to the first, a valid inference can be obtained if the premises are *distributively* true:

(4)   If
        (a) $p_1, p_2, ..., p_n \vdash q$ is a valid inference principle of classical logic, and
        (b) p_1 is true in w, p_2 is true in w, ..., p_n is true in w,
        then
        (c) q is true in w.

The second way to interpret (3) semantically requires that the premises should be *conjunctively* true:

(5)   If
        (a) $p_1, p_2, ..., p_n \vdash q$ is a valid inference principle of classical logic, and
        (b) p_1 & p_2 & ... & p_n is true in w,
        then
        (c) q is true in w.

(4) allows that p and $\sim p$ serve as premises in an inference and, accordingly, logical chaos may emerge. However, since statements contradicting each other cannot be conjunctively true, (5) does not permit inferences from contradictory premises. Therefore, Rescher and Brandom

---

[5] '&' stands for the logical constant 'conjunction' and '~' for 'negation'.

reject the distributive semantic interpretation of the valid inference principle of classical logic and accept the conjunctive interpretation in (5).

(2) and (5) jointly facilitate the distinction between two kinds of contradiction.

A contradiction is called *weak* if two requirements are met. First, in a possible world w such that $w_1 \cup w_2 = w$, both p and $\sim p$ may be true; thereby, p is true in w_1 and $\sim p$ is true in w_2 or vice versa. That is, the statements contradicting each other obtain in different possible worlds – among different conditions, in different contexts, against different background assumptions. Second, since p and $\sim p$ must not be used simultaneously as the premises of inferences, logical chaos can be avoided at the outset.

In contrast, we will speak of *strong* contradiction if these two requirements are not met. The notion of strong contradiction corresponds to the notion of contradiction in classical logic (Rescher – Brandom 1980: 15-26).

Accordingly, we obtain the following solution to (P1):

(SP1) (a) The logic as characterised by (2) and (5) is capable of solving (P), because it distinguishes between two kinds of contradiction.

(b) Weak contradiction is harmless, while strong contradiction is harmful.

Our next task is to apply (SP1) to the principle of epistemological tolerance. Since (PET) neither excludes its interpretation as the tolerance of weak, nor as that of strong contradiction, we have to examine cases of both kinds of contradiction in theoretical linguistics.[6]

3. The solution to (P2): A case study

We will illustrate the solution to (P2) by an instructive case study focusing on Zubizarreta's (1982) seminal contribution to Government-Binding Theory. Among others, Zubizarreta analyses the syntactic properties of Spanish verbs like *poder* (modal verb meaning 'can', 'able to', 'may') in sentences such as (6):

(6) *Juan puede visitar a   María*.
John can   visit   to   Mary.
'John can visit Mary'

Zubizarreta asks the following question:

(7) Is *poder* a main verb or a non-main verb in Spanish?

In order to answer (7), Zubizarreta (1982: 135-141) applies a series of tests. For example, one of the tests examines how *poder* behaves with respect to null complement anaphora.[7] Below *A1* stands for a set of data made use of in the test. *A1* includes, among others, the following grammaticality judgments:

---

[6] From this point on we will use the term 'contradiction' without an attribute in the sense 'either strong or weak contradiction' in order to refer to cases in which it has not yet been decided whether the relation between two statements rests on strong or weak contradiction.

[7] "Null complement anaphora refers to an elliptical construction in which a VP or IP complement of a verb is dropped. What is omitted can normally be understood from the preceding clause or context.

(1.9) (a) They asked him to pay at least lip-service to the principle, but he refused." (Huang 2000: 5).

A1 =   (a)   The sentence
          Juan   podría   visitar a   María   y   Pedro también   podría [...].
          John   could   visit to   Mary   and   Peter also   could […].
          'John could visit Mary and Peter could also'.
          is grammatically correct.

   (b)   The sentence
          Juan   ha   visitad a María   y   Pedro también   ha [...].
          John   has   visited to Mary   and   Peter also   has […].
          'John has visited Mary and Peter did also.'
          is ungrammatical.

The datum in (a) judges as grammatically correct the sentence in which both clauses include *poder* and the second clause contains null complement anaphora. If, however, instead of *poder* the auxiliary *haber* appears in the same positions, then, as witnessed by (b), the sentence is ungrammatical. From such data Zubizarreta infers the generalisation that *poder* behaves like main verbs and unlike auxiliaries in that it permits null complement anaphora. Thus, *A1* supports the hypothesis *H1*:

H1 = *Poder* is a main verb.

Another test concerns clitics. In the test Zubizarreta used the set of data *A2*:

A2 =   (a)   The sentence
          Pedro le   pudo   hablar   personalmente.
          Peter DAT.CL could   to.talk   personally.
          'Peter could talk personally.'
          is grammatically correct.

   (b)   The sentence
          Pedro le   prometió   hablar personalmente.
          Peter DAT.CL   promised   to.talk   personally.
          'Peter promised to talk personally.'
          is ungrammatical.

The generalisation the data in *A2* suggest is: In contrast to main verbs, *poder* allows cliticisation. Therefore, *A1* supports the hypothesis *H2*:
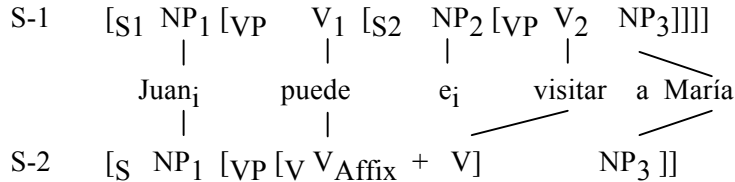
H2 = Poder is a non-main verb.

If we compare *A1*, *A2*, *H1* and *H2*, then it will be easy to realise that the result of Zubizarreta's argumentation is contradictory. Namely, *A1* and *A2* are distinct and there is no criterion on the basis of which *H1* or *H2* could be rejected. Accordingly, in the theory there are three contradictory pairs of statements. First, the data in *A1* contradict the hypothesis *H2*. Second, the data in *A2* contradict the hypothesis *H1*. Third, *H1* and *H2* also contradict each other.[8]
        Against the background of (PET) and (PEI), one might think that there are two alternative ways to restore the consistency of the theory. In accordance with (PET), one way is to decide either for *H1* or for *H2* and to discard the other. At the same time, the data supporting the hypothesis chosen are retained while those contradicting it are, though

---

[8] Please note that both a datum in the sense introduced in Section 1 and hypotheses are statements. Therefore, from a logical point of view, it is fully legitimate to speak of a contradiction between them.

registered, not integrated into the theory. The other way is to follow (PEI): to deem the theory unworkable. However, Zubizarreta (1982: 161) suggested a third option. She assigned two parallel structures to the sentence in (6) and represented the double structure of the sentence by a double-tree. In (8) the structure S-1 corresponds to *H1* and the the structure S-2 to *H2*. [9]

(8)

$$S\text{-}1 \quad [_{S1} \text{ NP}_1 [_{VP} \quad V_1 [_{S2} \text{ NP}_2 [_{VP} V_2 \text{ NP}_3]]]]$$

$$\text{Juan}_i \qquad \text{puede} \qquad e_i \qquad \text{visitar} \quad a \text{ María}$$

$$S\text{-}2 \quad [_S \text{ NP}_1 [_{VP} [_V V_{\text{Affix}} + V] \qquad \text{NP}_3]]$$

(8) yields the following answer to the question asked in (7):

(9)   The two structures corresponding to *H1* and *H2* are simultaneously present, but under different conditions.

After we have briefly summarised Zubizarreta's argumentation, in accordance with the sub-problem raised in (P2) now we have to examine how to reconstruct (9) with the help of the logic introduced in Section 2.

It is easy to observe that (9) meets the requirements defining weak contradiction. First, as we have seen, *H1* and *H2* rest on different conditions. In certain cases – e.g. if there is no clitic – *poder* behaves in accordance with *H1*; in other cases – e.g. if there is a clitic – according to *H2*; but under the same conditions it never behaves in both ways. Second, the conjunction of *H1* and *H2* is true neither with respect to *A1* nor to *A2*. Consequently, logical chaos can be avoided.

Thus, the relationship between *H1* and *H2* as put forward in (9) can be reconstructed as follows:

(10)   (a)   Let *H1* and *H2* be given.
        (b)   *H1* and *H2* are incompatible.
        (c)   Let us assume that $w_1 \cup w_2 = w$, where *w* is the whole of Zubizarreta's theory.
        (d) w_1 includes *A1* and w_2 includes *A2*.
        (e)   Since *H1* is consistent with *A1*, but contradicts *A2*, *H1* will be true in w_1 and false in w_2. Since *H2* is consistent with *A2*, but contradicts *A1*, *H2* will be true in w_2 and false in w_1.
        (f)   Both *H1* and *H2* are true in *w*, because each of them is true in one of the component worlds of *w*. However, *H1* & *H2* is false in w_1, in w_2, and, therefore, in *w* too, because neither w_1 nor w_2 includes both *H1* and *H2*.

Accordingly, Zubizarreta's theory constitutes a paraconsisent system. What solution to the sub-problem (P2) does this finding suggest?

Since (SP1) says that weak contradiction is harmless, and in (10) Zubizarreta's theory was reconstructed as a paraconsistent system that tolerates weak contradiction, we obtain the following solution to (P2):

---

[9] Zubizarreta (1982: 160-174) characterises *poder* and similar verbs as syntactic affixes which modify their verbal complement. However, this is not relevant for our line of reasoning.

(SP2)   The principle of epistemological tolerance is acceptable if it is re-interpreted as the tolerance of weak contradiction.

In what respects does the re-interpretation differ from PET?

The first component of (PET) concerns the treatment of data contradicting the hypotheses of the theory at issue. According to (PET)(a), the tolerance of contradictions means that the counterexamples "should be carefully noted, but it is often rational to put them aside […]" (Chomsky 1980a: 2) – that is, such data should be temporarily *excluded* from the theory. In contrast, if the relationship between the counterexamples and the hypotheses is interpreted as weak contradiction, then it will be possible to *integrate* these data into the theory.

The second component of (PET) is the relationship between the development of the theory and the resolution of the contradiction. On the one hand, as witnessed by (9), (SP1) is compatible with (PET) insofar as the consequences of strong contradiction can be avoided "pending further study" (Chomsky 1980a: 2) i.e. the later development of the theory. On the other hand, unlike (PET), the re-interpretation suggests that the later development of the theory *does not lead to consistency* in the sense of classical logic, but to *weak contradiction*.

The third element of (PET) is the protection of the theory's explanatory power ("principles of a certain degree of explanatory power", Chomsky 1980a: 2). On the one hand, in the sense of (PET)(c) the protection of explanatory principles means that despite the emergence of a contradiction, the theory must not be given up. Rather, one of the two contradicting sets of data, together with the corresponding hypothesis, has to be retained, while the other must be "put aside" (ibid.). However, then the data "put aside" will remain *unexplained*, because the hypothesis explaining them is dropped, too. On the other hand, as (9) shows, the re-interpretation proposed in (SP2) protects the explanatory principles of the theory in another way: as a result of the fact that the relationship between contradicting hypotheses is interpreted as weak contradiction, *both data sets can be explained*, because both hypotheses – in the case study *H1* and *H2* – are retained.

The fourth component is the fertility of the principle of epistemological tolerance. On the one hand, Chomsky (2002: 102) considers the tolerance of contradictions between the data and the hypotheses "quite a big step". (PET) is indeed more progressive than the naïve falsificationism of (PEI), which would require the abandonment of even the workable parts of a theory if a contradiction emerges. On the other hand, in the case study the "big step" was that in Zubizarreta's theory the tolerance of weak contradiction functioned as an effective *heuristic tool*, because it led to an acceptable answer to the question in (7), whereas (PET) could not have achieved this. The reason why we chose Zubizarreta (1982) as the object of the case study is that it was the first work to make use of double trees. Further applications of this theoretical tool justified the heuristic fertility of weak contradiction, because double trees were successfully applied in a series of later theory versions (see e.g. Baltin 1987, É. Kiss 1987, Saddock 1991, Zubizarreta 1987; see also Moravcsik 2006: 45-61, 2010).[10]

4.   The solution to (P3): A case study

The solution to sub-problem (P3) will be exemplified by a case study on Haider (2009).

Since Chomsky (1973) – in various frameworks, using a great variety of techniques – the literature has continuously discussed the question of under what conditions the *wh*-subject may remain in situ and under what conditions it has to be fronted in different languages. In

---

[10] Nevertheless, it goes without saying that despite the heuristic fertility of the tolerance of weak contradiction the desideratum of consistency as a regulative principle of linguistic theory formation must not be given up.

one of his contributions to the debate on the nature of linguistic data, Haider (2009: 80) mentions the following examples in which *what* is the *wh*-subject and *whom* the *wh*-object:

(11) (a) The sentence
      It is unclear <u>what shocked whom</u>
      is grammatically correct.
     (b) The sentence
      It is unclear <u>whom what shocked</u>
      is ungrammatical.

Haider (2009: 80) formulates the hypothesis supported by (11) in a very simplified form as follows:[11]

(12) In English a *wh*-subject must not remain in situ.

However, the German data are different from those in (11):

(13) (a) The sentence

Es	*ist*	*unklar,*	<u>*was*</u>	<u>*wen*</u>	*schockierte*
It	is	unclear	what-SBJ	who-OBJ	shocked

      'It is unclear what shocked whom.'
      is grammatically correct.
     (b) The sentence

Es	*ist*	*unklar,*	<u>*wen*</u>	<u>*was*</u>	*schockierte*
It	is	unclear	who-OBJ	what-SBJ	shocked

      'It is unclear what shocked whom.'
      is grammatically correct.

---

[11] It may be useful to explain briefly in what sense generative linguists assume that the *wh*-subject in the sentence in (11)(a) does not remain in situ but has to be fronted by *wh*-movement. For example, using the framework of Government-Binding-Theory, Haegeman 1994: 398) analyses the sentence

Who will arrive first?

according to the „vacuous movement hypothesis" as follows:

[$_{CP}$ Who$_i$ [$_{IP}$ t$_i$ will arrive first]]?

That is, „who" has been moved from IP to CP leaving behind the trace t_i coindexed with it. Haegeman elucidates the vacuous movement hypothesis by maintaining that „[…] the effect of short movement […] cannot be observed on the surface string, t_i having no phonetic content. Movement transformations whose effects cannot be observed are referred to as instances of vacuous movement". Then, in this framework, Haider's example in (11)(a) could be analysed in the following way (for the analysis of an analogous example see Haegeman 1994: 508):

It is unclear [$_{CP}$ what$_i$ [$_{IP}$ t$_i$ shocked whom]]

Here the *wh*-subject has been moved from IP to CP, too. Although the technical details of the analysis have changed continuously, despite serious counter-arguments (see e.g. Chomsky 1986), the transformational analysis of such structures has been widely accepted.

For the technical discussion of the hypothesis in (12) in the framework of Chomsky's Minimalist Program, see Haider (2004, 2005, 2011, and the literature cited there). The technical details of the analyses of the data like in (11) within different versions of generative grammar are not relevant for the point to be made in this paper.

The hypothesis is:

(14)   In German a *wh*-subject may remain in situ.

The generalisation is:

(15)   In English the relationship between the in-situ position of the *wh*-subject and that of the *wh*-object is asymmetric, while in German it is symmetric.

In the literature there have been several attempts to explain the generalisation in (15). Haider's suggestion connects (15) with the syntactic types of languages. Presupposing that the basic word order in English is VO and in German OV, he puts forward the following assumption (Haider 2004, 2009: 79):

(16)   In VO languages the relationship between the in-situ position of the *wh*-subject and that of the *wh*-object is asymmetric, while in OV languages it is symmetric.

Among others, from (16) it follows that in analogy to German, further OV languages, such as for example Dutch, permit that the *wh*-subject remains in situ. However, as Haider (2009: 79) reports, Gisbert Fanselow questioned (16). To support his criticism of a preliminary version of Haider (2004), he sent a questionnaire to 22 native Dutch linguists asking them to judge the grammaticality of 10 sentences containing *wh*-subjects and -objects in different constructions. The result was that in a series of relevant cases the informants did not judge these sentences unanimously. Let us illustrate the situation by the examples in (17)-(18).
        The Dutch sentences in which the *wh*-subject does not remain in situ are grammatically correct according to all informants:

(17)   The sentence
        Ik   weet   niet   wie   wat   gekocht   heeft
        I   know   not   who-SBJ what-OBJ bought   has
        'I don't know who bought what'
        is grammatically correct.

    But according to some of the informants, the *wh*-subject may also remain in situ, while others judge such sentences ungrammatical:

(18)   (a) The sentence
            Ik   weet   niet   wat   wie   gekocht   heeft
            I   know   not   what-OBJ who-SBJ bought   has
            'I don't know who bought what'
            is grammatically correct.
        (b) The sentence
            Ik   weet   niet   wat   wie   gekocht   heeft
            I   know   not   what-OBJ who-SBJ bought   has
            'I don't know who bought what'
            is ungrammatical.

The contradiction between (18)(a) and (b) is seemingly trivial, because it merely illustrates the everyday experience of linguists that there may be differences in the informants'

judgements of the same sentences. But Haider (2009: 80) draws non-trivial conclusions from this commonplace experience: the two judgments in (18)(a) and (b) cannot be maintained simultaneously, but the theory is not capable of deciding between them, either, and this *undermines* the workability of the theory.

Why does the simple problem illustrated in (18) lead to such a devastating conclusion? We will answer this question by applying the logic introduced in Section 2.

Let us assume that the set of data *A1* includes (17) and (18)(a), that is *A1 = {(17), (18)(a)}*. Then the data in *A1* will support the hypothesis *H1*:

H1 = In Dutch the *wh*-subject may remain in situ.

However, nothing speaks against deciding for the data set *A2 = {(17), (18)(b)}*. Accordingly, the grammaticality judgments in *A2* will support *H2*:

H2 = In Dutch the *wh*-subject must not remain in situ.

It is easy to realise that three contradictions emerge. The data in *A1* contradict *H2*; those in *A2* contradict *H1*; and there is a contradiction between *H1* and *H2*, too.

Thus, the situation appears to be similar to the first case study. However, this impression is incorrect, for the following reason.

In the first case study it was not the contradictory evaluation of the grammaticality of the sentences at issue that triggered the two contradicting hypotheses. Rather, there the two hypotheses were supported by tests which yielded two distinct sets of conditions. The grammaticality judgments pertaining to the Spanish sentences were free of contradiction at the outset.

In contrast, in the Dutch examples it was *the same sentence* the grammaticality of which was judged contradictorily and it was the *same source* – namely the native speakers' introspection – which the judgments stemmed from. Therefore, the relation between *A1* and *A2* is contradictory, too. As regards the contradiction between *H1* and *H2*, it is not possible to reconstruct it as weak contradiction, because the two hypotheses cannot be associated with two different possible worlds which represent two different sets of conditions and which are connected by superposition. Thus, the relation between *H1* and *H2* is *strongly contradictory* (self-contradictory).

Consequently, the logical reconstruction of the above example is as follows:

(19) (a) Let *H1, H2, A1* and *A2* be given.
      (b) *H1* and *H2*, as well as *A1* and *A2* are incompatible, respectively.
      (c) Let *w* be the possible world representing the theory.
      (d) In *w* both the grammaticality judgments included in *A1* and those included in *A2* are true.
      (e) *H1* is true in *w*, because it is consistent with *A1*, and *H2* is true, too, because it is consistent with *A2*.
      (f) Therefore, the conjunction of *H1* and *H2* can be inferred in *w*.
      (g) Therefore, *w* is strongly inconsistent.

Now we are able to answer the question of why the problem which (18) raised leads to devastating consequences. Our answer is that Haider's conclusion seems to refer to the *logical chaos* which has been triggered by strong contradiction and which the theory cannot handle. Accordingly, the following solution to the sub-problem (P3) suggests itself:

(SP3) The principle of epistemological tolerance is not acceptable if it is re-interpreted as the tolerance of strong contradiction.

If the relationship between Dutch data and the hypotheses at issue makes the theory strongly contradictory and this relationship leads to the destructive consequences mentioned, then the next question we should ask is: What is responsible for the emergence of strong contradiction? Although Haider does not use the terminology of the logic we introduced in Section 2 and does not raise this question, either, the answer can be found in his paper:

> "The theoreticians must acknowledge that the practice that proved successful in the pioneering phase of the past decades, namely *introspection* and eclectic feed-back from *informants*, has reached its *limits*." (Haider 2009: 97; emphasis added)

The rhetoric of this quotation witnesses a significant change of perspective. First, in accordance with the present methodological presuppositions of theoretical linguistics as well as the principle of epistemological tolerance (PET), Haider ought to have said that it is the theory's technical apparatus that decides whether (18)(a) or (18)(b) should be accepted. Thereby, it should be admitted that the set of data strongly contradicting that set of data which has been decided for cannot be explained until the later development of the theory makes its explanation possible. As opposed to this, in the quotation Haider questions the adequacy of the data source and the data type on which the findings of the theory have been based so far. Second, theoretical linguists continuously encounter problems similar to (18) without, however, being deterred from the application of the theory. Haider's sophisticated reflection on a problem which in the linguists' everyday practice has been unreflectingly considered to be trivial signals a basic change of perspective, too. Namely, this reflection has immediately revealed that in reality the problem at issue is non-trivial in a way and to an extent which substantially undermine the effectiveness of linguistic theorising.

Consequently, the question to be asked is how strong contradiction of the kind reconstructed in (19) can be avoided. The answer is: If strong contradiction must not be tolerated and the reason why strong contradiction obtains is the uncertainty of introspection as a data source and grammaticality judgements as a data type, then this data source and this data type have to be changed in order to avoid strong contradiction. Accordingly, the next question is: How?

The first possibility is to use another data source instead of introspection. Haider suggests the replacement of introspection by instrumental experiments. But Haider's answer is not convincing. One reason is that, although it is indeed the case that theoretical linguistics relying on data stemming from introspection has reached its limits, the complete rejection of the latter would lead to discarding many of the seminal findings of the past decades. Moreover, just like the use of the native speakers' intuition, even the results of instrumental experiments can be interpreted differently, and, what is more, may lead to contradictions, too. Finally, one of the most important insights which current debates on linguistic data and evidence have yielded is that despite the uncertainty and the unreliability of introspection, no other data source is workable without making at least partial use of the latter.[12] Neither in corpus linguistics nor in experimental linguistics can introspection be entirely dispensed with.

Therefore, the second option should be considered seriously: the application of several different data sources – such as corpora, different kinds of experiments, and introspection – and the integration of several different data types – such as search results, the results of experiments and grammaticality judgments. The outcome of recent discussions seems to

---

[12] See the literature referred to in Section 1.

support this conclusion independently of our above reasoning. For example, Schlesewsky (2009: 177) maintains that

> "the key to gaining new insights in the future and the hope of finding more adequate models of language, whether in psycho-/neurolinguistics or in linguistic theory, lies solely in the interaction and the combination of [data processing] methods." (My translation, A.K.)

This rhetoric refers to a change of perspective, too, but in a different way from Haider's. Like Haider, Schlesewsky also draws attention to the fact that current models of language are not adequate; however, he concludes that the tool for making them more adequate is the integration of different data types obtained by the application of different data processing methods.

5.   The solution to (P)

With the above solution to (P3) we have reached an important turning point in our line of argumentation. Namely, the consequences of the principle of epistemological tolerance as revealed by its logical reconstruction yielded two important findings which are considerably different from what (PET) maintains. Among others, (PET) claims that the contradictions between the data and the hypotheses can be avoided by the *future refinement of the theory*. As opposed to this, the first case study illustrated that in certain cases – due to the *immediate refinement of the theory* and the use of *all* the data at one's disposal – the relationship between the data and the hypotheses can be represented as weak contradiction. In addition, the second case study has shown that in particular cases strong contradiction can be avoided by *the extension of data sources and data types*. Thus, we have to realise that it is not only the *theory's* further refinement through which the destructive consequences of strong contradiction can be avoided, but also through transgressing the limits of introspection and grammaticality judgments by a more subtle reflection on the nature of data. Accordingly, we obtain the following characterisation of the current state of the art in theoretical linguistics:

At present, in theoretical linguistics there is a significant shift of focus in progress. While the past decades have been *theory-oriented* insofar as the continuous revision of the technical details of theory formation was focused on, and the peculiarities of the data have not been the subject of special interest, today a *data-oriented* perspective seems to be emerging. This perspective centres on the re-definition and re-structuring of data sources as well as the integration of data types.

Now we are in a position to infer the solution to our starting problem (P) from the solutions to the sub-problems (SP1)-(SP3):

(SP)   Theoretical linguistics
      (a)   may tolerate weak contradiction between the data and the hypotheses, but must not tolerate strong contradiction; and
      (b)   requires the consideration of new data sources and the integration of data types stemming from different data sources.

6.   Conclusions

Let us look back at our reasoning. We raised (P), which is a relevant foundational problem of theoretical linguistics. We subdivided (P) into three sub-problems, solved the sub-problems one by one, and from their solutions inferred the solution to (P). Accirdingly, we carried out the task we set in Section 1.

However, although at first sight (SP) convincingly follows from previous steps in our line of reasoning, in its above form it is not acceptable. Namely, it includes the desideratum of a more sophisticated treatment of data sources and data types than has been practiced in theoretical linguistics so far, but it does not clarify *how* such a desideratum can be met. The present methodology of theoretical linguistics is *not compatible* with (SP)(b). This methodology – which assumes introspection to be the main data source and grammaticality judgments to be the main data type – obviously does not include the *heuristics* that should be applied in order to answer the following open questions closely related to the desideratum mentioned in (SP)(b):

(20)   (a)   How can the contradictions between hypotheses supported by different data types which go back to different data sources be accounted for?

(b)   What argumentation strategies should be applied in order to make decisions between alternative hypotheses supported by different data types and different data sources?

(c)   By what argumentation strategies should former decisions be re-evaluated in the light of data sources considered later?

(d)   How does the consideration of new of data sources and the integration of data types affect the details of the *theoretical* framework?

Thus, by raising these open questions we have arrived at the final conclusion of the present paper: *theoretical linguistics needs a new methodology*. The data-oriented perspective requires the elaboration of new strategies of theory formation. Therefore, (SP) has to be modified as follows:

(SP')   Theoretical linguistics

(a)   may tolerate weak contradiction between the data and the hypotheses, but must not tolerate strong contradiction;

(b)   requires the consideration of new data sources and the integration of data types stemming from different data sources,

and

(c)   also requires the development of a new methodology for theoretical linguistics compatible with (a) and (b).

The most important task for theoretical linguistics to be accomplished in the future is to make serious attempts at *the substantial renewal of its methodological foundations*.[13]

---

[13] According to Schlesewsky (2009: 177), there is no methodology at our disposal capable of integrating different data types and considering different data sources. However, Kertész – Rákosi (in press) explicitly tackles questions like those in (20). Its aim is to reveal the new heuristics of linguistic data processing mentioned in (20) which can be directly applied to linguistic problem solving.

Acknowledgement

References

Baltin, Mark R. 1987. Degree Complements. In: Huck, Geoffrey J. – Ojeda, Almerindo E. eds., *Discontinuous Constituency*. Academic Press, Orlando. 11–26.

Boeckx, Cedric 2006. *Linguistic Minimalism: Origins, Concepts, Methods and Aims*. Oxford University Press, Oxford.

Boeckx, Cedric 2010. Linguistic Minimalism. In: Heine, Bernd – Narrog, Heiko eds., *The Oxford Handbook of Linguistic Analysis*. Oxford University Press, Oxford. 485–505.

Borsley, Robert D. ed. 2005. *Data in Theoretical Linguistics* [= *Lingua* 115: 1475–1665].

Botha, Rudolph P. 1983. On the 'Galilean Style' of Linguistic Inquiry. *Lingua* 58: 1–50.

Botha, Rudolph P. 1988. *Form and Meaning in Word Formation: A Study of Afrikaans Reduplication*. Cambridge University Press, Cambridge.

Brame, Michael 1984. Universal Word Induction vs. Move α. *Linguistic Analysis* 14: 313-352.

Chomsky, Noam 1973. Conditions on Transformations. In: Anderson, Steven – Kiparsky, Paul eds., *A Festschrift for Morris Halle*. Academic Press, New York. 232–286.

Chomsky, Noam 1980a. On Binding. *Linguistic Inquiry* 11: 1–46.

Chomsky, Noam 1980b. *Rules and Representations*. Blackwell, Oxford.

Chomsky, Noam 1986. *Barriers*. MIT Press, Cambridge, MA.

Chomsky, Noam 2002. *On Nature and Language*. Cambridge University Press, Cambridge.

Chomsky, Noam 2004. *The Generative Enterprise Revisited*. Mouton de Gruyter, Berlin – New York.

Featherston, Sam 2007. Data in Generative Grammar: The Stick and the Carrot. In: Sternefeld, Wolfgang ed. 269–318.

Featherston, Sam 2009. A Scale for Measuring Well-Formedness: Why Syntax Needs Boiling and Freezing Points. In: Featherston, Sam – Winkler, Susanne eds., 2009. 47–73.

Featherston, Sam – Winkler, Susanne eds. 2009. *The Fruits of Empirical Linguistics. Vol. 1: Process*. de Gruyter, Berlin – New York.

Freidin, Robert 2007. *Generative Grammar: Theory and its History*. Routledge, London.

Freidin, Robert – Vergnaud, Jean-Roger 2001. Excuisite Connections: Some Remarks on the Evolution of Linguistic Theory. *Lingua* 111: 639–666.

Haegeman, Liliane M.V. (1994). *Introduction to Government and Binding Theory*. Wiley-Blackwell, Oxford.

Haider, Hubert 2004. The Superiority Conspiracy. In: Stepanov, Arthur – Fanselow, Gisbert – Vogel, R. eds., *The Minimal Link Condition*. Mouton de Gruyter, Berlin. 167–175.

Haider, Hubert 2005. How to turn German into Icelandic – and derive the VO-OV contrasts. *The Journal of Comparative Germanic Linguistics* 8:1-53.

Haider, Hubert 2009. The Thin Line Between Facts and Fiction. In: Featherston, Sam – Winkler, Susanne eds., 75–102.

Haider, Hubert 2011. Exceptions and Anomalies. In: Wiese Heike – Horst Simon eds. *Expecting the unexpected – Exceptions in Grammar*. de Gruyter, Berlin – New York. 325-334

Hale, Mark 2007. *Historical Linguistics: Theory and Method*. Wiley-Blackwell, Oxford.

Hale, Mark – Reiss, Charles 2008. *The Phonological Enterprise*. Oxford University Press, Oxford.

Hjelmslev, Louis 1969. *Prolegomena to a Theory of Language*. The University of Wisconsin Press, Madison.

Huang, Yan 2000. *Anaphora. A Cross-Linguistic Study*. Oxford University Press, Oxford.

É. Kiss, Katalin 1987. *Configurationality in Hungarian*. Akadémiai Kiadó, Budapest.

Kepser, Stephan – Reis, Marga eds. 2005. *Linguistic Evidence. Empirical, Theoretical and Computational Perspectives*. de Gruyter, Berlin – New York.

Kertész, András 2004. *Philosophie der Linguistik. Studien zur naturalisierten Wissenschaftstheorie.* Narr, Tübingen.

Kertész, András – Rákosi, Csilla (in press). *Data and Evidence in Linguistics. A Plausible Argumentation Model.* Cambridge University Press, Cambridge.

Klausenburger, Jürgen 1983. Review of Botha, R., 'On the 'Galilean Style' of Linguistic Inquiry'. *Language* 59: 434.

Meheus, Joke ed. 2002. *Inconsistency in Science*. Kluwer, Dordrecht.

Moravcsik, Edith 2006. *An Introduction to Syntactic Theory*. Continuum, London – New York.

Moravcsik, Edith 2010. Conflict Resolution in Syntactic Theory. *Studies in Language* 34: 636-669.

Mukherji, Nirmalangshu 2010. *The Primacy of Grammar*. MIT Press, New York.

Penke, Martina – Rosenbach, Anette eds. 2007. *What Counts as Evidence in Linguistics?* Benjamins, Amsterdam – Philadelphia.

Popper, Karl R. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge & Kegan Paul, London.

Popper, Karl R. 1980 [1959]. *The Logic of Scientific Discovery*. Hutchinson, London etc.

Priest, Graham 2002. Paraconsistent Logic. In: Gabbay, Dov. M. – Guenthner, Franz eds., *Handbook of Philosophical Logic*. Kluwer, Dordrecht. 287–393.

Priest, Graham – Beall, J. C. – Armour-Garb, Bradley eds. 2004. *The Law of Non-Contradiction*. Clarendon Press, Oxford.

Rescher, Nicholas – Brandom, Robert 1980. *The Logic of Inconsistency. A Study in Non-Standard Possible-World Semantics and Ontology*. Blackwell, Oxford.

Riemer, Nick 2009. Grammaticality as Evidence and as Prediction in a Galilean Linguistics. *Language Sciences* 31: 612–633.

Saddock, Jerrold M. 1991. *Autolexical Syntax. A Theory of Parallel Grammatical Representations.* The University of Chicago Press, Chicago – London.

Schlesewsky, Matthias 2009. Linguistische Daten aus experimentellen Umgebungen: Eine multiexperimentelle und multimodale Perspektive. *Zeitschrift für Sprachwissenschaft* 28: 169–178.

Stefanowitsch, Anatol – Gries, Stephan Th. eds. 2007. *Grammar without Grammaticality*. [= *Corpus Linguistics and Linguistic Theory* 3/1: 1–129].

Sternefeld, Wolfgang ed. 2007. *Data in Generative Grammar* [= *Theoretical Linguistics* 33/3: 269–413].

Schütze, Carson T. 1996. *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. The University of Chicago Press, Chicago – London.

Weinberg, Steven 1976. The Forces of Nature. *Bulletin of the American Academy of Arts and Sciences* 29: 13–29.

Winkler, Susanne – Featherston, Sam eds. 2009. *The Fruits of Empirical Linguistics. Vol. 2: Product.* de Gruyter, Berlin – New York.

Zubizarreta, Maria Luisa 1982. *On the Relationship of the Lexicon to Syntax*. Ph.D. Diss. MIT, Cambridge, Mass.

Zubizarreta, Maria Luisa 1987. *Levels of Representation in the Lexicon and in the Syntax*. Foris, Dordrecht.

Author Biography

András Kertész is Professor of German Linguistics and Director of the Centre for Philosophy and Linguistics at the University of Debrecen. In 2001 he was elected as a member of the Hungarian Academy of Sciences. His main fields of research are the philosophy of linguistics and theoretical linguistics. He published 10 monographs, 8 collections of papers and about 130 articles. His recent monographs include *Cognitive Semantics and Scientific Knowledge* (Amsterdam: Benjamins, 2004), *Philosophie der Linguistik* (Tübingen: Narr, 2004) and *Data and Evidence in Linguistics* (Cambridge: Cambridge University Press, in press; co-author: Csilla Rákosi).
http://web.t-online.hu/andraskertesz/angol.htm