

# Galen's Critique of Rationalism and Empiricism, and its Relevance for Modern Linguistics

Andrew R. Plummer  
The Ohio State University  
[plummer@ling.osu.edu]

[Annual meeting of the North American Association for the History of the Language Sciences (NAAHoLS); Portland, OR; January 6-7, 2012]

## Abstract

Methodology in Linguistics, since the late 1950s, has largely followed the model put forward in Chomsky's early adaptations of the Structuralist approach (Chomsky, 1957, 1965, 1966, 1975; Chomsky and Halle, 1968), though misapplication of the model has resulted in a methodological dichotomy not unlike that between the early Rationalist and Empiricist approaches to anatomy that emerged from Hippocrates's medical philosophy. In Galen's critique of the Rationalist and Empiricist sects (see Cosans, 1997), he sought to unify the approaches by eliminating shortcomings in each, and synthesizing what remained. I discuss here the applicability of his critique to the current dichotomy.

## 1 Preamble

- 1.1 Methodology in the "core areas" of Linguistics, since the late 1950s, has largely followed the model put forward in Chomsky's early adaptations of the Structuralist approach (Chomsky, 1957, 1965, 1966, 1975; Chomsky and Halle, 1968), though substantial oversimplification of the methodological model has led to its persistent misapplication.
- 1.2 Chomsky's "Galilean style" of theorizing about linguistic phenomena, wherein data that potentially refute a theory are often dismissed if the theory is believed to provide "deep insight" concerning some array of phenomena, has been sharply contrasted with more recent "data-driven" approaches to linguistic modeling, which focus primarily on statistical analysis and inference based on massive amounts of observed tokens (see Sampson, 2001, for further discussion). This challenge to the "Galilean style" of theory-driven modeling is meant to correct for data-based shortcomings, yet it brings with it other, rather severe shortcomings well-known in the philosophy of science (see Chomsky, 2011).
- 1.3 A coarse parallel can immediately be drawn between the theory-driven/data-driven dichotomy in modern Linguistics and that between the early Rationalist and Empiricist approaches to anatomy that emerged from Hippocrates's medical philosophy; a dichotomy whose falsity was in part recognized by the philosopher/physician Galen nearly 2,000 years ago (Cosans, 1997).
- 1.4 In Galen's critique of the Rationalist and Empiricist sects (see Galen, 1985; Ottoson, 1984; Cosans, 1997, for discussion), he sought to unify the approaches by eliminating shortcomings in each, and synthesizing what remained, stressing, in particular, the importance of experimentation.

## 2 Present Aims

- 2.1 I will present a (very) high-level elaboration on the coarse parallel in hopes that Galen’s insight into the false dichotomy of his time provides a bit of guidance with respect to that of our own. The focus will be primarily on syntax, though the main points carry over for other subdisciplines without much modification. Specifically, I briefly review:
1. a few key methodological concepts from the theory-driven approach promoted by Chomsky, along with a few of the main problems pointed out by its critics.
  2. a few key methodological concepts from the data-driven approach put forward as an “alternative” to the theory-driven approach, along with a few of the main problems this approach has always faced.
  3. the main points of the Rationalist/Empiricist dichotomy, toward refinement of the parallel with the theory-driven/data-driven dichotomy.
  4. Galen’s critique of the Rationalist/Empiricist dichotomy, focusing on the relevance of Galen’s appeal to experimentation as a key aspect in his own approach. I then consider this appeal with respect to the modern dichotomy, and discuss a few key points on how to use it to move forward.

## 3 “Theory-driven” approaches

- 3.1 Chomsky’s early adaptations of the Structuralist approach (Chomsky, 1957, 1965, 1966, 1975; Chomsky and Halle, 1968) brought about significant changes to the activities of linguistic inquiry as they were in the 1940s and early 1950s. We focus on the following two fundamental aspects of the adaptations:
- i) The shift in focus on what constitutes linguistic phenomena away from corpora and utterance tokens and toward native speaker “intuitions” concerning the grammaticality of utterance types, initiating a return to a more mentalistic/psychological conceptualization of linguistics and its objects of inquiry.
  - ii) The corresponding shift in linguistic modeling, focusing on the underlying psychological aspects that yield native speaker grammaticality judgments, and modeling these aspects using “generative mechanisms” borrowed from logic and abstract algebra (as well as some previous work in linguistics and philosophy of language).
- 3.2 Concerning the shift in phenomena:
- (a) The shift necessitated the development of methodology for investigating aspects of native speaker grammaticality judgments. The primary method put forward by Chomsky was simply: i) present a native speaker with a token of some relevant utterance type, ii) have the subject report on their introspection regarding the grammaticality of the utterance type.
  - (b) Chomsky’s defense of his methodology involved appealing to advances in scientific methodology engendered during the early Enlightenment (see Chomsky, 2002, for a review). One aspect of the appeal with quite significant consequences was the adoption of the “Galilean style” of theorizing, wherein data that potentially refute a theory are often dismissed if the theory is believed to provide “deep insight” concerning some array of phenomena.

- (c) Problems in Practice:
  - i. Researchers employing Chomsky’s method tended to disagree on what their own introspections determined to be grammatical, and, using Chomsky’s methodological defense, justified the dismissal of judgments that were unfavorable to their own theories (see Labov, 1975). The method also came under fire due to recognized differences in introspective reports and behavior (see Labov, 1975).
  - ii. Another common criticism is that theory-driven researchers create “fantasy data” that has nothing to do with real human language, e.g., Lesk’s (1988) review of Garside et al. (1987) (quoted in Sampson (2001)):
 

“Why is it so remarkable to have a book whose analysis of language is based entirely on actual writing?...It is a great relief to read a book like this, which is based on real texts rather than upon the imaginary language, sharing a few word forms with English, that is studied at MIT...a testimony to the superiority of experience of over fantasy.”
- (d) Problems in Principle:
  - i. A higher-level view of the data problems mentioned above reveals a more general problem in principle concerning psychological entities. Things like native speaker intuitions are thought to be “unobservable,” and many believe that science ought to restrict itself to “observable” things that can be objectively, and interpersonally evaluated.
  - ii. Even more generally, problems with introspection in psychological and philosophical pursuits, has been recognized for a long time. The notion of confabulation, and the relevance of attention have been discussed quite a lot recently, and are potential problems that lead to quite radical positions in psychology and philosophy, though we will keep our discussion to linguistics, and to one particular position, one that suggests a return to corpora and utterance tokens, since they are “observable.”

### 3.3 Concerning the shift in modeling:

- (a) By the mid 1970s, technical matters concerning the mathematical objects called “grammars” had already come to widespread attention given their applicability in the theory of programming languages. Criticism of Chomsky’s formulation of transformational grammars (inter alia) in, e.g., *Syntactic Structures* and *The Logical Structure of Linguistic Theory*, prompted more intense focus on the mathematics of grammars and related technicalia within Linguistics departments and the computational linguistics sectors of computer science departments.
- (b) During the 1980s and ’90s, a substantial amount of syntactic research activity consisted of a largely theoretical effort to refine the mathematical grammars and investigate their formal properties. This is still common practice in some corners of current syntactic research, which proceeds in accordance with desiderata drawn from adjacent formal disciplines such as logic and computer science, foregoing attention to psychology.
- (c) When the psychological “referents” of grammars are relegated or ignored, however, syntactic research becomes somewhat fossilized around a limited collection of “canonical data,” and what follows is a mathematical arms race to account for the

very specific aspects of the canonical data set. This problem is not confined within the theory-driven approaches, as we discuss in the next section.

## 4 “Data-driven” approaches

- 4.1 In reaction to the theory-driven modeling, “data-driven” approaches, which focus primarily on statistical analysis and inference based on (massive amounts of) observed tokens, have dramatically increased in popularity. Self-styled as competing with theory-driven modeling, data-driven modeling is meant to correct for the data-based and model-based shortcomings of “theory-driven” linguistics.
- 4.2 The syntax-based work collected in Sampson (2001) represents the more extreme end of the data-driven philosophy, strongly promulgating the message of keeping linguistic investigation staunchly centered on analysis of “everyday language” that is interpersonally “observable.” I focus herein on the extreme position taken by Sampson, rather than a more moderate one, due to how serious this position is currently being taken by the linguistics community.
- 4.3 According to Sampson (2001), the theory-driven approach, and particularly Chomsky’s instantiation of it, is not in accordance with good science, and his reasons for this are essentially the main arguments against theory-driven linguistics presented above. Sampson’s (2001) alternative is a “common sense” approach to linguistics, which he calls “empirical linguistics,” adopting the following simple positions:
  - i) Linguistic data consists of observable tokens of natural language recorded in some fashion, essentially corpora such as the Wall Street Journal corpus and the Brown corpus. Moreover, linguistics decidedly “is not about the psychological mechanisms underlying [grammatical] structure” (p. 12).
  - ii) Linguistic modeling consists of some kind of statistical modeling based on the tokens in corpora. It is important to note that grammars over corpora are statistics in the same way that token frequency counts are.
- 4.4 Concerning the data:
  - (a) Now, the data-driven approach is supposed to remedy all the bad things about the theory-driven approach listed above. Given the very strong claims about the virtues of this approach, it is worth seriously considering how well it does in doing so, especially since it seems to bring with it other, rather severe shortcomings long-known in the philosophy of science (see Chomsky, 2011).
  - (b) Problems in Practice:
    - i. Just as theory-driven researchers “manipulate” grammaticality judgments in order to support their own theories, data-driven researchers manipulate the statistical models in a variety of ways to improve their F-scores and inflate error reduction rates. This is done in a variety of ways depending on the statistical model being used (e.g., in a Bayesian model, adjusting of priors), or the aspects of the structure the statistical model is operating over (e.g., postulating basic categories as needed, such as a “dollar sign category” when parsing the WSJ corpus).

- ii. Corpora certainly do not have all the linguistic data one could ever want. An oft-told anecdote used to illustrate this point is that the Brown corpus, a very large one, does not contain the word very common word “weasel.” More generally, there are lots of linguistic phenomena that we might want to study which do not appear in any corpus, and cannot in principle, e.g., the Lucky’s Hat corpus.
- (c) Problems in Principle:
  - i. Even at (what is assumed to be) the most basic level, labeling a corpus presupposes a helluvalotta theory-based segmentation concepts and theory-based determination of similarity across segments. In order to segment and categorize Korean stops you need the abstract, theoretical concept “VOT”. In order to separate and categorize Korean fricatives you need even more advanced theory-driven concepts. In short, we don’t even know that we’re considering the right things when we limit our inquiry to “observables,” in Sampson’s sense.
  - ii. “Common sense” notions of vowels and consonants don’t help very much when segmenting a Korean corpus into sound segments, and there is little reason to believe that current “common sense” syn/sem notions are all that useful for present (and future) corpus-based syn/sem. Grammatical structure itself, extensively used by Sampson, is an example of a something that, 70 years ago, was “unobservable” in corpora, and really did not accord with common sense. What else are we going to find useful that isn’t observable right now? How are the non-common-sense things like grammatical structure revealed to us?

#### 4.5 Concerning the modeling:

- (a) A lot of computational syntax/semantics modeling is based on the “bake-off” style of progress: a computational/statistical model that “gets better numbers” than another computational/statistical model, on a comparable task, is considered the better model.
- (b) As a result, researchers tend to all use the same corpus for their modeling efforts so that such a numbers-based comparison can be made. This leads to a kind of overfitting of models to data that is similar in nature to the overfitting of mathematical models to the “canonical data” in the theory driven approaches.

## 5 Mediate Summary

- 5.1 Both the theory-driven and data-driven approaches seem to have significant problems in practice, and in particular aspects of modeling. Some of the problems in practice seem to be problems with the human aspect of science – researchers are going to “take liberal interpretations of data” and “fudge the numbers” – and similarly the modeling problems are always going to exist. As we will see in Section 6, they have existed since antiquity, and can be mitigated, more or less.
- 5.2 The problems in principle for the theory-driven approach, however, seem more manageable than those of the data-driven approach, since, as we argue in Section 7, they fall within the scope normal scientific advancement, due in large part to understanding revealed through experimentation, as Galen recognized.

## 6 Rationalism and Empiricism, Briefly

6.1 Hippocrates was a Greek physician in the Golden Age of Athens, the Fifth Century BCE, when philosophy was therein flourishing. Two of his major thrusts were the organization of extant medical knowledge, and the development of medical practice. Hippocrates is credited with both incorporating philosophy in medicine, and distinguishing medicine from philosophy.

- i) The former was done through the use of abstract categorization, particularly for classification of diseases and other organisms into genera and species.
- ii) The latter was done by “empirica,” achieved through study of the body in detail, as it is “the beginning of all medical knowledge.”

6.2 Two major approaches to medicine formed based on a sharpening of these two aspects of Hippocrates’s teachings:

- Rationalism – Medical knowledge is achieved through abstraction and deduction. Abstraction over symptoms and circumstances, followed by some deductive process, yields a method of treatment.
- Empiricism – Medical knowledge is achieved through experience, and held in memory for future application. Prior circumstantial experience yields a method of treatment in future, similar circumstances.

6.3 High-level parallels between Rationalism and Empiricism, and theory-driven and data-driven approaches to linguistics:

- The Rationalist approach essentially involved taking some observable phenomena and constructing an abstract story involving unobservables to yield a deeper understanding of the observables. This corresponds fairly closely to Chomsky’s use of “observable” utterances in deriving information and constructing stories about “unobservable” native speaker intuitions.
- The Empiricists insisted that “everyday” experience in treatment of wounds and ailments was the main concern of the physician, and no greater reality or explanation need be investigated. This corresponds fairly closely to the data-driven focus on experienced, “everyday language” tokens as the exclusive content of linguistic inquiry.

6.4 Criticisms of Rationalism and Empiricism and their primary parallels with criticisms of theory-driven and data-driven approaches to linguistics:

- The Empiricists criticized the Rationalist practice of incorporating unobservables in their development of treatments, and the tendency to dwell on the reasoning aspect of the treatment development. This corresponds to the data-driven criticism of the theory-driven linguist’s focus on the unobservable psychological mechanisms, and the tendency of modelers to dwell on the more mathematical aspects.
- The Rationalists criticized the Empiricist approach as unsound, claiming that determination of similarity of observables is baseless without abstraction. Two wounds

may look similar but indicate very different conditions, while two wounds may look very different, but be symptoms of the same condition. Moreover, there may be “wounds” that are not be apparent at all (including the case of preventative treatment). This corresponds to the theory-driven criticism of the data-driven focus on experienced tokens and their lack of “common sense” similarity, and sparsity.

## 7 Galen’s Synthesis and its Modern Application

- 7.1 Galen (AD 129 - 199/217) was a prominent physician and philosopher, serving as physician to Marcus Aurelius, and making substantial contributions to physiology and neurology. Galen wrote several critiques of both the Rationalist and Empiricist sects. Believing that the division was detrimental to medicine, he sought to unify the approaches by eliminating shortcomings in each, and synthesizing what remained.
- 7.2 Galen suggested an approach to medicine that began with a significant amount of empiricist training, including clinical experience and study of histories of clinical experience, and was to be followed by a significant amount of rationalist training in reasoning from the collected experience to abstraction, and putting the abstraction to use in explaining the experience.
- 7.3 More than attempting to synthesize, Galen recognized the falsity of dichotomy between the two approaches. Rather than viewing them as mutual exclusive, competing approaches, he argued that the two approaches could coexist harmoniously, even if his own synthesis was not adopted. He recognized that a physician could achieve practicable competence through experience, without necessarily having the theoretical expertise he himself recommended. Hence Galen’s treatise “That the best physician is also a philosopher” (see Brain, 1977, for a translation). Interestingly, the computational linguist Mark Johnson has already pointed out something similar for the theory-driven/data-driven dichotomy.
- 7.4 Importantly, Galen recognized the importance of experimentation to the rationalist approach (Cosans, 1997). It was believed by the empiricists that anatomical experiments that isolated some part of the body in order to investigate its properties or functions was useless to the practicing physician, as the study of a body in such a state was too far removed from experience with actual patients. More generally, the empiricists believed that the “special experience” created in experimentation was too special to tell the experimenter anything useful about the body.
  - (a) This is one aspect of Galen’s approach that was far ahead of its time. The value of the “special experience” created by experimentation and its use in understanding the world would not be fully appreciated until the early Enlightenment.
  - (b) Chomsky recognized the value and use of creating this kind of special experience in linguistics, which is the reason for the “fantasy data” coming out of MIT, as it was created for (a very simple kind of) experimentation, meant to reveal something about the world, and decidedly not meant to correspond to any “real data” that one would find in a corpus.

- (c) The fantasy data may be substantively criticized, but not due to its nature as special experience. The substance of the criticism lies in the crudity of its representation and reliability in investigating the unobservable phenomena. Recognition of this, and efforts toward improvement are simply normal science, and exactly how progress is made (e.g., Boyle’s criticism of the methods for investigating the Aristotelian elements).

- 7.5 Sampson (2001) states that the data-driven empirical linguist “cannot wait for the outcome of abstruse future psychological investigations” (p. 84) to provide the representations needed for their modeling tasks, and must proceed with representation imposed by the modeler. This is a reasonable position to take, but it doesn’t mean that the psychological investigations are useless or ought to cease. Introspection is a difficult thing to get a hold of, and it’ll be some time before we have decent psychological representations, but scientists are making progress in “easier” areas such as the cognitive psychology of perception (see discussion in Jack and Roepstorff, 2004).
- 7.6 The “mental rotation” results derived from the classic experiments by Shepard and Metzler (1971) are a good example of “unobservable” equivalence judgments that became “observable” through the development and use of fMRI.

## References

- Brain, P. (1977). Galen on the ideal of the physician. *South African Medical Journal*, 52:936–938.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Chomsky, N. (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper and Row.
- Chomsky, N. (1975). *The Logical Structure of Linguistic Theory*. Springer.
- Chomsky, N. (2002). *On Nature and Language*. Cambridge University Press.
- Chomsky, N. (2011). Language and Other Cognitive Systems: What is Special about Language? Albertus-Magnus Professorial Lecture. Video available at: <http://international.uni-koeln.de/3963+M55e9db57069.html>.
- Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Cosans, C. E. (1997). Galen’s critique of Rationalist and Empiricist anatomy. *Journal of the History of Biology*, 30:35–54.
- Galen (1985). *Three Treatises on the Nature of Science*. Hackett Publishing Company. Translated by Richard Walzer and Michael Frede, with an introduction by Michael Frede.
- Garside, R., Leech, G., and Sampson, G., editors (1987). *The Computational Analysis of English: A Corpus-based Approach*. Longman.
- Jack, A. and Roepstorff, A., editors (2004). *Trusting the Subject? Volume 2*. Imprint Academic.
- Labov, W. (1975). Empirical foundations of linguistic theory. In *The Scope of American Linguistics*, pages 77–133, Belgium. The Peter De Ridder Press.
- Lesk, M. (1988). Review of garside, leech, and sampson 1987. *Computational Linguistics*, 14:90–91.
- Ottoson, P.-G., editor (1984). *Scholastic Medicine and Philosophy*. Bibliopolis.
- Sampson, G. (2001). *Empirical Linguistics*. Contium.
- Shepard, R. and Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171:701–703.