==========
Language in India www.languageinindia.com **ISSN 1930-2940 Vol. 20:4 April 2020**
==========

Scientific Evidence Bias in Linguistic Theory

Sari Hamoud, M.A.
Ph.D. Research Scholar
Center for Linguistics
Jawaharlal Nehru University
New Delhi, India
sarihammoud@gmail.com

==========

Abstract

Linguistic theory is taught at centers and departments of linguistics throughout the world with little or no attention to the Eurocentric and logocentric aspects of its evidence and examples. This apparent yet unescapable scientific bias is discussed from two points within the philosophy of science. First, general scientific evidence bias is discussed. Then, the difference between Internal versus External linguistic evidence is presented. The paper highlights the role of integrationist approaches within linguistic theory and how their input adds generalizability to linguistic models when certain standards are met.

Keywords: Linguistics, Eurocentric and logocentric aspects, philosophy of science, integrationist approaches, lesser-studied languages.

1. Introduction

It is assumed that modern scientific theories are objective and falsifiable (in comparison to unquestionable subjective beliefs), and they are designed specifically to explain phenomena based on unbiased reproducible evidence. However, this is not entirely accurate. In this paper, I will examine this assumption by taking examples from linguistic theory focusing mainly on the conflict that exists between external and internal evidence vis-à-vis the functional and formal approaches to language, and I will attempt to characterize its impact on the field of language study.

2. Is Science Biased?

Kuhn (1970) explains that science development does not happen in upward linear progression, and that its development is not always fully objective nor rational. He gives illustrative examples that show progress happening through revolutions. Scientists topple old theories in favor of new ones. This change keeps reoccurring in a cyclical fashion whereby science goes from normal status into crisis mode, and eventually a new paradigm replaces the older one.

That fact of the matter is that science is normally full of dead-end theories, false starts and unpublishable failed trials/experiments. That is, progress is not necessarily incremental. Kuhn (1970) mentions that "normal science" has to be practiced in the manner of problem solving with a limited degree of creativity under *unquestioned* assumptions. That is, the dominant paradigm within the concerned discipline sets the rules for allowed vs ignored questions, acceptable statistical p values and experimental designs, and it also creates the standard for the interpretation of results of various research experiments (Field, 2013).

However, anomalies do inevitably start showing in the results of research that contradict a popular paradigm. Thus, works that are unable to explain new data are either developed or more likely dropped for the lack of publishable power. At this stage, the anomalies could be attributed to human errors in conducting experimentation. In this proposed scenario, nevertheless, anomalies will continue to re-occur and accumulate leading to what Kuhn (1970) describes as "crisis science" and scientists start debating publicly the inadequacy of the paradigm. They will not, however, throw out a falsification-challenged paradigm until they have a better one, which is opposed to what Popper (1959) suggests. It might even take hundreds of years for a new paradigm to be adopted by the community within crisis science. For example, the shift between Newtonian physics and Einstein's took more than 200 years. That's when scientific revolution lead to what could be described as a paradigm shift (DiSalle, 2006).

Despite their balanced foundation, rationalism and evidence do not play an integral role in how a new paradigm is adopted by the scientific community. Kuhn (1970) uses the term "conversion" as if revolutionary scientific change is more like a political or religious sudden change in comparison to a rational progressive persuasion process in which evidence dominates. In fact, he mentions that "the competition between paradigms is not the sort of battle that can be resolved by proofs" (160). However, both Popper (1959) and Kuhn (1970) acknowledge that scientists who engage in dogmatic normal science where basic assumptions are never questioned are not going in the right direction.

On the other hand, a certain level of fixed assumptions is needed for science to progress. For instance, during a period of normal science, scientists have a set of beliefs that do not change by scholarly criticism, especially when it comes from the new practitioners in the field. There is a rational reason for this; If one is fully open to criticism, one cannot make any progress within the paradigm nor refine their own research thrust within it. Without this, no one will have time at all to focus on their work, spending all of it on disproving the onslaught of objections or reevaluating the basics of the paradigm.

The view of how science progresses that was illustrated by Kuhn (1970) can be explained by parallel examples from the field of linguistics. For instance, the earliest recorded scientific work on language was done by Pāṇini in India back in the 4th century BCE. Bod (2013) mentions that he is considered to be the father of linguistics. Indeed, Pāṇini has done extensive work on Sanskrit grammar, and his use of formal generative rules practically paved the way into their integration in the discipline on various levels. That is, his work is not limited to syntax. In fact, it encompasses phonology, morphology, and semantics.

Linguistics, however, did not progress in Pāṇini's research paradigm. The reason is purely historical; up until the 20th century, most of the work that was done in the discipline in Europe was under the umbrella of philology which mainly focused on comparing different languages and investigated the historical perspective of language development. The work of Pāṇini reached the European continent at a later stage. For instance, European scholars were introduced to Indian linguistics through the comparative work of Franz Bopp in the eighteenth century. However, it could be argued that Pāṇini's work was influential in the paradigm shift after Ferdinand de Saussure who seems to have been influenced by him in the late nineteenth century. Saussure's structuralist account of how language works sets the standard for modern linguistics and was a natural evolution of the work that was done by Pāṇini. However, this has not always been the case.

During the twentieth century, language was considered by psychologists and anthropologists as a form of behavior as evident in Skinner (1957). However, this paradigm was questioned since Chomsky (1959) criticizes viewing language development as a form of conditioned behavior. Chomsky's review of "Verbal Behavior" led to a revolution in the discipline and inaugurated the cognitive movement in linguistics and other fields. At the time, Chomsky was a young ardent defender of language innateness paradigm who logically seems to have won many arguments whenever his view was questioned, while Skinner never acknowledged Chomsky nor formally responded to his review with a formal defense.

This has led to the strong adaptation of the Chomskyan paradigm in the discipline within a relatively short period of time. Moreover, Chomsky's work on Universal Grammar and Transformational Grammar, and later on, the Minimalist Program utilized a nativist approach to language and specifically examined the internal aspect of language rationally using a formal approach to the degree that some started accusing him of excluding external evidence from the picture and dividing the field into functionalists and formalists.

However, it can be argued that Chomsky's main contribution to the field is that he was able to introduce a paradigm-changing theory that adheres to the standards of Popper (1959) in terms of being concise, specific, and above all, falsifiable. In the 1970s, Pateman (1987) indicated that

the field was haunted by the influence of crisis science which led some to even question whether linguistics was scientific at all.

Therefore, Chomsky's division of labor between Internal and External language should not be seen as a biased attempt to ignore the role of meaning and extralinguistic elements of context. In fact, if one examines his work closely, one will find that he does not deny the role of language use within context nor the importance of meaning. Moreover, Chomsky keeps updating and changing his views on language which is the natural course of science albeit certain assumptions are maintained in terms of the paradigm basics and the type of evidence to be used in examining questions that are allowed under Universal Grammar and the Internalist nativist approach to language.

Therefore, we cannot claim that science nor Chomsky as an example from linguistics is biased per say. There is, however, a certain degree of selectivity within the discipline in terms of data that is regarded as exemplary of what language aspect that is under the scope of investigation. The assumptions and the selective nature of the scientific approach is inbuilt to it and only functions to maintain its progress as was mention above.

The following section attempts to illustrate the difference between external and Internal evidence in language study while trying to explore further the possibility of justifying bias in scientific selectivity.

4. Internal versus External Evidence in Linguistics

Zwicky (1980:598) mentions that within the discipline of linguistics there exists a generally accepted distinction between external and internal evidence. He starts by pointing out that what is "usually classified as internal are data on the cooccurrence and alternation of linguistic elements in some language, as well as such systemic considerations as formal simplicity, economy, and the like." On the other hand, he maintains that external evidence encompasses the data that comes from use of language, its patterns of acquisition, different dialects, and how it is affected by historical change or language pathology.

Furthermore, Zwicky (1989) mentions that the division first began as a result of the need for demarcation of linguistics from other disciplines that study language structure and use such as psychology and sociology. Thus, evidence of language use in society might be interesting to the linguist, but variability of language does not actually help explain how language works in the mind across the species.

On the other hand, the generativists themselves do not fully agree on what evidence should be used when it comes to language study as an entity. For some, language, as Zwicky (1989) explains, is independent of its speakers, contextual culture and even use. He describes this as a "strongly antireductionist, autonomistic **bias** that defines external data as outside the class of data to be explained by linguists" (599). However, this point of view is not very popular. What is common among linguists is that a good language analysis must always take into consideration that language is psychologically real and is internalized in the minds of the speakers (600). It does not mean that we can do without external evidence. Zwicky (1989) cites an example from English: the formation of plural nouns in its morphology. He explains that the base plural morpheme cannot be assigned unless we look at how English has developed historically, and we also need input from how children begin to formulate plurals as they are acquiring the language. He also points out that arguments from external evidence alone are not enough. However, he mentions that phonology in comparison to syntax favors external evidence. He cites two reason. One is the finite domain of phonology that is phonemes and words. The second is the clearly demarked domain that exhausts internal evidence and explanations and invites more data to support an assumption. On the other hand, he argues that syntax with its infinite number of sentences has a generative bias that requires focus on further developing our understanding of the realm of internal evidence and does not leave space for alternative analyses and the use of external evidence.

In brief, Zwicky (1989) is calling for the utilization of both types of evidence towards the improvement of the working paradigm. However, as we shall see in the next section. The two types of evidence are utilized in a biased fashion by the scholars in the field which has cemented the divided between functionalists and formalists.

Before moving to the next section, however, I would like to further mention one more aspects which explain why bias exists when it comes to linguistic evidence. It is simply a matter of choosing examples from the written language most of the time: the bias towards written language versus the spoken one. Derrida (1976) calls this bias a logocentric approach that assigns to writing more attention than it should. Linell (2004) argues that despite acknowledging the primacy of speech over writing, most linguists have approached language analysis using theories and methods that are designed specifically to examine *written* language. In his book, he presents a number of reasons for the biased study of written language and the over-generalization into spoken. For example, he points that the cultural stereotype gives a high status to written, whereas spoken language is not considered real language (10). He also gives an interesting alternative to the common practice paving the way to an integrationist approach to language study that highlights the examination of language as an interface between the speakers' cognitive ability and the people's experience when it comes to its use in their lives.

Moreover, Linell (2004) is not the only voice of dissent in the field who is arguing for an integrationalist approach as can be seen below in (Pateman, 1987; Itkonen, 2005; and Newmeyer, 2010, among others) who seem to support it. However, before we examine their views on this matter, we will continue looking at different types of bias in language study.

5. Different Types of Bias in Linguistics

Another type of bias is between observational versus experimental data. It emanates from the fact that much research in linguistics aims at what Bhaskar (2013) calls as "exceptionless generalization" that have no exceptions which are difficult to arrive at even under highly controlled laboratory experiments (233). Although experimentation in this context helps to isolate the underlying mechanism that generates language structure, it seems obvious that such approach is incomplete by itself, since it ignores any element outside the laboratory such as observational data which are considered either irrelevant or mere confounding variables. Hence, the lack of ecological validity.

Furthermore, Pateman (1987) stresses the need for developing an approach that utilizes both experimental and observational evidence in linguistics. He maintains that in real life scenarios, speakers tend to behave much more freely in comparison to laboratory conditions. For instance, they might ignore grammatical errors and concentrate on meaning which is dependent on context and not on the structure of language. He also indicates that one should take into consideration other factors such as memory and psychological state of the speaker, which are considered external. This approach can be considered to be part of realist vis-à-vis positivist approach that insists on isolation of objects of study and setting certain predictions of how their underlying mechanisms operate producing language as we know it. He, therefore, emphasizes the need to consider "naturally occurring experiments" (9).

Moreover, Itkonen (2005) pushes for a synthesis between what he calls Intuition-based study of norms and observation-based study of corpora. He explains that the "variation-eliminating idealization of a more or less homogeneous langue is indeed a practical necessity" (358). He also mentions that if variability in language was taken into consideration, then the only comprehensive description of language will be a statistical one. However, we know that traditional grammarians or syntacticians rarely use statistics when they describe or prescribe language. That is, they avoid describing the actual linguistic variable *behavior* and tend to focus almost entirely on the structure of the system of language. On the other hand, Itkonen (2005) indicates that the distinction between these two levels of description can be termed categorical as opposed to the statistical one, and the history of linguistics is without a doubt unfairly divided when it comes to this aspect. Moreover, this distinction also opens the door to examining the relation between intuition based norms and observation of actual utterances in corpora. He also maintains that the whole matter of linguistic

bias is a degree of selective difference between qualitative data and quantitative data. Nevertheless, Itkonen (2005:306) also shows that in order to capture the true nature of language, we should not be dealing with the two types of evidence as "absolute black-or-white dichotomies". He adds that the Internalist non-empirical rational analysis of language based on intuitions must be complemented by empirical considerations that account for behavioral variability and external evidence that encompasses language use. However, he also confirms that intuition is sufficient on its own when the linguist is dealing with a clear case. This point, however, could also be put into question, for it reduces the degree of objectivity of his stance. If the analysis is based on the so-called intuition, then how can bias and subjectivity be controlled?

Itkonen (2005: 306) maintains that some linguists exaggerate the importance of laboratory data to the degree of misusing it. He indicates that this happens when some consider that the methodology of natural science "must at any cost be imported into grammar". He, therefore, warns from the disastrous over-use of inessential corpora which does a disservice to the analysis. He argues that balance should be maintained between two levels of knowledge: norm and action. It is just another way to describe what Chomsky refers to as linguistic competence and performance. Itkonen (2005) agrees with Chomsky that competence is higher in terms of priority than performance. However, he accuses positivists of denying this two-level nature of linguistics. He also maintains that there is an undeniable Darwinian influence on the field of language study that emanated from how some feel "hopelessly inferior to the representatives of 'hard sciences'" (366). He makes a very good point. For long, the discipline of language study has been beset on all sides by the lack of uniform methodology that adheres to scientific naturalism making its status weaker. However, the object under study in linguistics cannot be compared to physics. Their objects are well defined and invariable in comparison to the immense variability of language and its users. Here again, Chomsky's Internalist approach plays an integral role in establishing linguistics as a field of science because by limiting the scope of the discipline to the invariable I-language, it regains the status of respectable branch within, we should remember, the current "normal science" paradigm. In fact, Pateman (1987) defends Chomsky's actions by arguing that his contribution to the field established the generative, and therefore, scientific nature of the discipline. Moreover, the Chomskyan abstractions, Pateman continues, were perfect for describing the psychological reality of language and, therefore, avoids falling in the traps of empiricism. However, again one would argue that Chomsky's exclusion of language use angered many in the field such as Linell (2013), and especially the within the field of semiology as explained by Harris (1995).

Itkonen (2005) makes an interesting point in his integrationalist approach. He posits that the distinction between rules of grammar or theoretical aspect on the one hand, and the sentences-rules or the atheoretical on the other, should be understood under the notion that the theoretical and atheoretical are merely in a continuum. He stresses, however, that rule-sentences do not lend

themselves directly to the notion of falsifiability due to their prescriptivist normative nature. The following section is concerned with how functionalist and formalist views differ and why we need a true integrationist approach in the study of language.

6. The Formalist, Functionalist, Integrationist Future

Newmeyer (2010) explains the difference between formal and functionalist explanations of language and attempts to bridge the gap between them by insisting that they are complementary to one another. He calls the divide between the two camps as the "greatest rhetorical conflict" (1). He first explains that the terms "formal" and "functional" are themselves used inconsistently within the two approaches. For instance, some consider an approach to be formal IFF it adheres to the paradigm of mainstream syntactic theory that excludes data from discourse. However, other approaches, such as Head Driven Phrase Grammar, considers an explanation formal if it gives a detailed unambiguous account that is formalized. For functionalists, meaning is central to the explanation of linguistic phenomena which is an alternative approach to formal linguistics. Newmeyer (2010) writes:

> 1) An explanation is *formal* if it derives properties of language structure from a set of principles formula in a vocabulary of nonsemantic structural primitives.
>
> 2) An explanation is *functional* if it derives properties of language structure from human attributes that are not specific to language (2).

Newmeyer (2010) continues to illustrate the nature of formal explanation by giving examples of how ungrammaticality of a sentence is accounted for by UG principles and parameters. However, he also indicates that the constant use of deductive nomological mode leads formalists into trouble if none-standard dialects are subject to the same analysis. That is, what is ungrammatical in the standard might be perfectly allowed in the dialect. For example, in Ebonics, the principles of agreement between the subject and verb do not behave like in Standard English, so the question, "You workin?" is a perfect paraphrase of the Standard "Are you working?". Newmeyer (2010) mentions that "it might be the case that the grammar of this particular dialect contains a covert movement operation that leads to the *Wh*- Criterion being satisfied, despite appearances to the contrary (4). He does not deny that other external factors might also justify why such sentence is produced even though it is not generated by the rules of the grammar of the language under question. I should point here that there seems to be a kind of hidden bias towards the standard languages, especially European ones like English. Indeed, much of the work that has been done on generative linguistics has centralized every model around the examples that illustrate the structure of English. This does not mean that formal approach has not given any insights to other languages. For instance, Newmeyer (2010) gives an example from negation across

languages. He explains that although one cannot establish the claim of the existence of a standard processing procedure for semantics and discourse that works across the board among different languages, when negation takes place, we can generally fall back on the formal explanation to understand what is happening to the structure.

7. Conclusion

Is linguistics Biased? Well, its development has not always been fully objective nor rational, yet within each paradigm inside the cyclical movement, there is consistency. However, when a new paradigm takes over, it removes much of the foundations of its predecessor. For instance, Chomsky led the campaign for a paradigm shift at a time when linguistics was going through a dark period of behavioral positivist empiricism that was full of anomalies. His exclusion of the role of meaning and language use was seen biased by many, and this led to the division of the field between formalists and functionalists. However, the Chomskyan bias could be seen as a justified move in the favor of the discipline as a whole. The division of labor was a must to reach a stage of scientific standard that adheres to the norms of the scientific community. The division within linguists themselves, however, is not justifiable. The divide resulted in two camps that had a negative bias toward the type of data and evidence to be used to answer language most pertinent questions. We saw in this paper that the word bias in certain contexts could either bear a positive meaning and a fair use as in "selective to fulfill certain paradigm requirements" such as the example of Chomsky above, or it could have a negative meaning in which it conveys "unjustified prejudice". There are many examples of unfair bias. One was the focus on writing and assigning more importance to it than to speech as indicated by Linell (2004). Another example is the competition between paradigms during the phase of crisis science where bias and none-scientific factors play a role in adopting a new paradigm. Several voices (Pateman, 1987; Zwicky, 1989; Itkonen, 2005; and Newmeyer, 2010, among others) in the field called for an integrationist approach where both external an internal evidence are taken into considerations when investigating Internalized Language. Furthermore, bias in linguistics can be seen towards experimental vis-à-vis observational data. Although Pateman (1987) does not mention it, the section that is concerned with the hierarchy of the linguistic method alludes to the state of bias in terms of linguistic hierarchy in the discipline itself. That is, the hierarchy that exists in the linguistic society that puts syntacticians at the top of the pyramid and places applied areas such as sociolinguists and language teaching at the bottom, which is another example of unfair bias. The solution to the problem could begin by recognizing the special nature of language itself as being an object which does not lend itself easily to natural science method of analysis and to acknowledge the role of every field within linguistics as being equally important and contributive to the explanation. Hence, a good approach should assume that every stratum of formalist and functionalist research formulates a continuum that has to be utilized as a whole if true progress is to be made in this field.

==========

References

Bhaskar, R. (2013). *A realist theory of science*. Routledge.

Bod, R. (2013). *A new history of the humanities: The search for principles and patterns from antiquity to the present*. Oxford University Press.

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language*, *35*(1), 26-58.

Derrida, J. (1976). *Of grammatology* (GC Spivak, trans.).

DiSalle, R. (2006). *Understanding space-time: The philosophical development of physics from Newton to Einstein*. Cambridge University Press.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.

Harris, R. A. (1995). *The linguistic wars*. Oxford University Press.

Itkonen, E. (2005). Concerning the synthesis between intuition-based study of norms and observation-based study of corpora. *SKY journal of linguistics*, *18*, 357-377.

Kuhn, T. S. (1970). BOOK AND FILM REVIEWS: Revolutionary View of the History of Science: The Structure of Scientific Revolutions. *The Physics Teacher*, *8*(2), 96-98.

Linell, P. (2004). *The written language bias in linguistics: Its nature, origins and transformations*. Routledge.

_____. (2013). *Chomskyanism–from innovation to irrelevance.*

Newmeyer, F. J. (2010). Formal and Functional Explanation. *University of Washington, University of British Colombia, University of Fraser University (manuscript)*, 1-25.

Pateman, T. (1987*). Language in mind and language in society: Studies in linguistic reproduction.* Clarendon Press.

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson.

Skinner, B. F. (1957). "Chapter 1: A Functional Analysis of Verbal Behavior". *Verbal Behavior*. Acton, MA: Copley Publishing Group.

Zwicky, A. M. (1980, January). " Internal" and" External" Evidence in Linguistics. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1980, No. 2, pp. 598-604). Philosophy of Science Association.

==========