Logic, Fallacy, Reasoning and Superficial Resemblance

H.M. Hubey, Department of Computer Science
Montclair State University
Upper Montclair, NJ 07043

Abstract
Perusal of the literature of historical linguistics (and electronic mailing lists) shows a bewildering array of claims, counterclaims, and opinions based on what appear to be some heuristics (rules of thumb). It is not clear if the motion in intellectual space of linguistics is in the direction of tightening up the heuristics to create a science, or loosening up the heuristics possibly under the belief that one heuristic is as good as another. Perhaps the fact that there is motion in all directions is a sign of a period of great agitation (and hopefully ferment). In the big scheme of things there appear to be slow motion advances toward the use of statistics, probability theory and other mathematics while many are cleaning up the raw data. However, there does not seem to be much discussion of the most basic and fundamental heuristic of comparative (historical or diachronic) linguistics; what makes languages related, how and why . It is the purpose of this short paper to look at the facts purely from a logical point of view.

1. Basic Ideas

There are two categories of reasoning;

1. Under Certainty
2. Under Uncertainty

The only method of reasoning under certainty is logic, both propositional logic and predicate logic. Under uncertainty we have

2.1. Probabilistic (stochastic methods)
2.2. Fuzzy Logic

Fuzzy logic is the new kid on the block, and although has shown itself to be useful in quite a few areas is still mistrusted by those trained in the classical ways. Under probabilistic methods we have such areas as statistical inference, Bayesian reasoning, correlation-regression analysis, and stochastic processes. What kind of logic is used by historical linguists? It is an embarrasing truth that probably nobody knows what kind of reasoning is used by historical linguists. Only in one place can there be found an honest, and clear statement of the algorithm/rules used by historical linguists [Crowley,1997], and in only one other place is there a criticism of what has transpired to date [Lass,1998]. All the rest treat the subject as a cross between a literary narrative and black magic. There is yet another kind of reasoning apparently indulged in by physicists, engineers and other physico-mathematical scientists and that seems to be “model building”. As the social scientists increase their “theory” making, the physical and especially the mathematical scientists have given up the use of the word “theory” and started to substitute simply “model”. The main reason is the indiscriminate multiplication of “theories” created by anyone who can string a few hundred words together. In the same way as Tolstoy felt that all happy families were happy in exactly the same way while there were many different ways for families to be unhappy, there can be only one correct model of historical linguistics while there may be many incorrect ways.

2. Logic and Fallacy

The most important thing to know about logic is that it is a special addition to what is known as Boolean Algebra (Laws of Thought). Boolean algebra like high school algebra has two mathematical operations; addition, +, and multiplication, *. But the variables are only allowed to have the values, True or False (often written as 1 and 0). The rules of addition and multiplication are simple:

P	Q	+
0	0	0
0	1	1
1	0	1
1	1	1

P	Q	*
0	0	0
0	1	0
1	0	0
1	1	1

Figure 1: *Addition* (OR) and *Multiplication* (AND) Tables of Boolean Algebra. These are also called *Truth Tables*. In logic books, one often sees the symbol $\vee$ for OR and $\wedge$ for AND.

As can be seen, multiplication is identical to multiplication as we know it. The only difference in the addition is that we have the seemingly strange addition 1+1=1. However, if we recall that these are to be interpreted as truth values, then the equation is nothing more than T+T=T or in natural language “True OR True=True”. This definition of OR is the inclusive-OR which distinguishes it from the exclusive-OR which is written as XOR, $\oplus$. The reason for the terminology is that $1 \oplus 1=0$ meaning that the case when both variables are True is excluded from being true in the truth table. If we want to obtain Propositional Logic from Boolean Algebra we have to introduce another connective (aside from AND, OR and XOR), and this is the $\Rightarrow$, the *implication* (often called *material implication*). A statement of form $P \Rightarrow Q$, is read as one of

P implies Q
If P, then Q
Whenever P, then Q
P can be true only if Q is true
P is a sufficient condition for Q
Q is a necessary condition for P

and means that whenever P is true, then Q is true. It does not say that P is false when Q is false. That would be an “equivalence” (which is the complement of XOR) not an implication. It should be made clear here that there is no *causality* in the logical implication. That is, if we claim that $X \Rightarrow Y$, we do not claim that X causes Y although it could. They might both be caused by the same third variable and thus might co-occur. However, even if X does cause Y or if we understand the implication this way, it still does not mean that X is the only cause of Y, but at most one of the causes.

We can make up sentences by assigning different meanings to the variables. For example let P=*I fall into the river,* and Q=*I get wet,* then $P \Rightarrow Q$ is read as *"If I fall into the river, then I [will] get wet"* or '*Whenever I fall into the river, I get wet"*. Variables also have *complements* (*negations* or *denials*) so that $\neg P$ is "I do not fall into the river". Sometimes it is convenient to write sentences in more compressed notation so that a sentence such as "Ravens are black" can be written as $R \Rightarrow B$. Here we could have defined the variables as *R=an object is a raven,* and *B=the object is black* so that the statement is more like "If an object is a raven, then it is black". Now that the *contrapositive* of the statement $P \Rightarrow Q$ is the statement $\neg Q \Rightarrow \neg P$ and it is equivalent to the former, simply stated differently. A common mistake made by students of logic is to assume that $P \Rightarrow Q$ and $Q \Rightarrow P$ are equivalent. This is called the *Affirmation of the consequent fallacy* [see for example, Rosen[1996]]. We can show this quite clearly by using the definition of 'material implication' which is given in Figure 2.

P	Q	$P \Rightarrow Q$
0	0	1
0	1	1
1	0	0
1	1	1

$\neg Q$	$\neg P$	$\neg Q \Rightarrow \neg P$
1	1	1
0	1	1
1	0	0
0	0	1

P	Q	$Q \Rightarrow P$
0	0	1
0	1	1
1	0	0
1	1	1

Figure 2: Definition of Implication and the equivalence of the contrapositive

We notice here that probably the most important development in the philosophy of science of this century is embedded in this definition. In other words, *the only time the statement can be false is if there is a counter example*. If, for example, someone makes the claim $R \Rightarrow B$, the only time we can say that this statement is false if we can find a raven that is not black. This is the 'falsificationism" of Popper. It was stated clearly by Einstein "No experiment can prove me right, but a single experiment can prove me wrong." This idea has been twisted out of shape in linguistics, as something like "the negative cannot be proven" [Crowley,1996]. Unfortunately there are worse examples of this fallacy. We find this kind of reasoning constantly in linguistics. One of the examples of this (because it involves mathematics and is thus accepted without critical analysis by many) is the work done by Ringe[1992,1995]. The essential form of reasoning involved is this: Let *X= some random process had occurred in the development of language*, and *Y=a particular form of data would have been obtained*. Now Ringe does some mathematical work (i.e. calculates some numbers from the Binomial Distribution) and thus argues $X \Rightarrow Y$. There is nothing wrong with this part naturally. But then he argues, that this justifies $Y \Rightarrow X$ which is to say that since the data that he found fits that particular form, then this particular random process must have occurred. But of course, this is nothing but the *affirmation of the consequent fallacy.* However if we did use probabilistic reasoning then we would have to calculate the probability that this particular set of data was obtained as a result of the occurence of that particular process.

3. Modus Ponens, Modus Tollens and Related Fallacies

Modus Ponens is an *inferencing rule* which is formalized as $P(P \Rightarrow Q) \Rightarrow Q$ which says in words what was argued above. For example, "*if whenever I fall into the river I get wet*" is true, (i.e. $P \Rightarrow Q$ is true), AND (recall that AND is multiplication) "*I have fallen into the river" is true* (i.e. P is true) meaning that $P(P \Rightarrow Q)$ is true, "*then I must be wet*" must be true (i.e. Q must be implied by $P(P \Rightarrow Q)$ thus $P(P \Rightarrow Q) \Rightarrow Q$. Sometimes this is written on several lines to show the steps

Premise	$P \Rightarrow Q$
Fact	P
Therefore	Q

(a) Modus Ponens

Premise	$P \Rightarrow Q$
Fact	$\neg Q$
Therefore	$\neg P$

(b) Modus Tollens

Figure 3: Inferencing Rules of Logic

Right next to *Modus Ponens* in Figure 3 is *Modus Tollens*, another valid way of arguing. As an example of Modus Tollens we have "*whenever I fall into the river and I get wet*" (is true, i.e. $P \Rightarrow Q$ is true), *BUT* (this is actually *AND* in logic, and hence multiplication) "*I am not wet"* (i.e. Q is false meaning that $\neg Q(P \Rightarrow Q)$ is true), then P must be false i.e "*I could not have fallen into the river"*. In Ringe's example above, his argument could have also been of the correct form (Modus Ponens) $(\neg Y(X \Rightarrow Y) \Rightarrow \neg X)$ meaning "*if the process was X then it would have given rise to a particular form of data Y, but we do not have this kind of data, therefore it was not process X*". But Ringe argues $(Y(X \Rightarrow Y) \Rightarrow X)$, that is to say , "*if the process was X we would have obtained Y; we have Y so then the process must have been X"*. Clearly, there could be other processes that could have given rise to the data. In order to properly use an argument of this type one must use a statistical inferencing scheme, which is of the form "*there are N processes including process X that could have caused the data Y to be obtained, and the probability that this data was caused by X is very high.*"

The *Denial of the antecedent* fallacy is an argument of the form $\neg A(A \Rightarrow B) \Rightarrow \neg B$. This is the converse of the fallacy of *Affirmation of the Consequent.* This argument then is of the form "*If I had fallen into the river I would have gotten wet, however I did not fall into the river, therefore I cannot be wet.*" Of course, there are many other ways to get wet than falling into the river.

Another fallacy related to modus ponens or modus tollens inferencing scheme is *converting a conditional* which is of the form; $(A \Rightarrow B) \Rightarrow (B \Rightarrow A)$. This fallacy is similar to the *Affirmation of the Consequent,* but phrased as a conditional statement. Here we would argue "*If I fall into the* *river I get wet, therefore, whenever this is true then it must be true that whenever I am wet I must have fallen into the river.*" or "*If the process A causes the observed data B is true, therefore whenever this is true the observed data B implies that the process A must have happened.*" It is yet another way in which Ringe's method can be viewed if we view it only from the perspective of logic. Note that this fallacy is different from *Non Causa Pro Causa* (see below in section 6).

4. Proof, Induction, Abduction and Circular Logic

So if Ringe is not arguing logically, then is he arguing from probabilistic point of view? For example, there are arguments of this type [Lass,1998: 334] in which a case is 'inferred' from a rule and a result;

Rule: *The beans from this bag are white.*
Result: *This bean is white.*
Case: *This bean is from this bag.*

Obviously, this is illogical and fails to convince. This form of reasoning was called abduction by Lewis. Abduction, according to Sebeok (1983: 9) is supposed to enable us to formulate a general prediction, but with no warranty of a successful outcome. One reason why this is done may be because we see so much regularity (a very dangerous word, like 'natural') around us that we expect what we see to be a part of a general pattern and immediately try to generalize. In logic something like this would be called a *hasty generalization* fallacy. It is conceivable that those with some knowledge of statistical inference might see it superficially similar to this case:

Fact: *The beans from this bag weigh 40 gm on average.*
Fact: *The beans from that bag weigh 60 gm on average.*
Conclusion: *This bean (which weighs x grams) is most likely from bag.*
or
The probability that this bean is from this bag is p.

First, the detailed answer in this case is based on some assumptions: (i) there are at least 2 bags from which the bean could have been obtained, (ii) the bean is definitely from one of the bags in question, and (iii) that the bean was sampled (randomly) from one of the bags. Both conditions are necessary for statistical inferencing to work. If they were not, we could go into one of the bags and purposely pick the most unlikely bean and thus throw off the probability calculation. Or there could be a zillion other bags of beans from which the bean in question could have come.

The biggest problem in science, first of all, is that nothing can be 'proven' outside of mathematics. A proof in mathematics proceeds using deduction (or induction) from basic truths called axioms. Axioms for arithmetic may be statements such as "x+y=y+x" or "xy=yx", called 'commutativity'. These cannot be "proven". No matter for how many integers we try this (e.g. 2+3=3+2, etc) we can never be sure that it will work for every pair of integers and we cannot try it for every pair of integers since there are an infinite number of them. But there does not seem to be any reason why it should fail. Therefore we accept them as *truths without proof*. Then we prove other things from these axioms (which are assumed or accepted to be true) called theorems. Of course, the axioms need not be true since they need not refer to anything in the real world. We can create any set of axioms which are not obviously inconsistent and then create theorems from them, and these theorems will be true in that ‘system’ (the system defined by the axioms). These are called *formal* or *axiomatic* systems. The rest of the sciences are far from this state of affairs. In a field like physics (or its applied fields, the various branches of engineering) the axioms are called *postulates* or *laws*. Now can we accept statements such as those below as ‘laws’?

1. *The sun will always rise in the east.*
2. *Water always boils at 100C at STP.*
3. *All birds have feathers.*

This is the famous *problem of induction* first clearly written about by Hume. There are, unfortunately, gradations of truth in the sciences. Unless the laws of the universe change #2 will probably be true everywhere on earth or on other planets. Barring some catastrophe such as an astoroid colliding with the earth #1 probably will also remain true. As far as birds go, the question is partly empirical and partly definitional. We might simply decide to call birds without wings something other than bird. In less concrete cases, we could face yet another logical fallacy called *equivocation* (see below). There are many problems of this type in historical linguistics. A form of this occurs constantly in linguistics matters when someone starts off or insist that “It has not been proven”. In logic related fallacies are: *Converse accident / Hasty generalization* where a general rule is formed by examining only a few specific cases which aren't representative of all possible cases. A hasty generalization might be said to occur when someone argues; *“Math was tried in linguistics in the year 18xx by Z and it did not work, therefore it will not work.”* To see how ridiculout this argument is we need to simply compare this to this statement *“My friend tried to fix his car yesterday with a hammer, and could not do it, therefore cars cannot be fixed with tools”*.

4.1) Audiatur et altera pars

Often, people will argue from assumptions which they don't bother to state. The principle of *Audiatur et Altera Pars* is that all of the premises of an argument should be stated explicitly as shown above in the case of Modus Ponens and Modus Tollens. It's not strictly a fallacy to fail to state all of your assumptions; however, it's often viewed with suspicion.

4.2) Circulus in demonstrando

This fallacy also called *circular argument* means that the conclusion is assumed as a premise. One can see such useless argumentation in allegedly profound sentences such as;

1. *Language is what sets humans apart from other animals.*
2. *Languages are spoken only by humans.*
3. *[Human] [L]anguage is innate [to humans].*

Statements such as the above occur constantly and are essentially useless unless we accept them as a vehicle by which students learn to think more deeply about the problem. Almost all of the statements derivable from #1 can be seen to be circular (see below) since when most people say ‘language’ they really mean ‘human language’. Substituting the latter phrase into sentence #1 shows how ridiculously circular the whole argument is. Statement #3 is just as meaningless. Not only is it circular it is also a particularly nasty example of *equivocation*. The whole thing hinges on the meaning of ‘innate’. If it merely means that only humans can speak (i.e. speak a human language) it is a universally known fact, much like the fact that trees don't walk. If it means that there are specific language neurons, it is not true. But if the meaning is to be understood to mean that the articulatory apparatus, the brain capacity to learn complex syntax, to memorize hundreds of thousands of words, and convey messages about the environment, and even about abstract ideas is possessed, so far as we know, only by humans on earth, then it is again another obviously true (and nonsignificant) fact. A particularly odious fallacy of this type is committed often by historical linguists in this form:

1. IE languages are genetically related because method-X proves it is so.
2. Since method-X says that IE languages are genetically related, and we know that the IE languages have been proven to be genetically related, then method-X works.
3. Therefore method-X is the only method that can prove genetic links amongst languages.

There are so many things wrong with this line of no-argument that one is appalled that it is continuously (*argumentum ad nauseam*) repeated on practically every electronic mailing list with nary a complaint from anyone.

A particularly atrocious form of this kind of argument is of the form;

1. The words in the list from languages M and N have are not cognates but merely phonetic resemblances.
2. Therefore languages M and N are not genetically related.

The first argument by asserting that the words are not cognates already assumes that the languages M and N are not genetically related so that the conclusion #2 is really the same as the premise #1 which is that the comparanda are not cognates. Yet this form of argument can be seen in discussions. The clearest form of the comparative method can be seen in Crowley[1997]

> ...you have to look for forms in the various related languages which appear to be derived from a common original form. Two such forms are cognate with each other, and both are reflexes of the same form in the protolanguage [Crowley,1997:88].
> ....
> In deciding whether two forms are cognate or not, you need to consider how similar they are both in form and meaning. If they are similar enough that it could be assumed that they are derived from a single original form with a single original meaning, then we say that they are cognate [Crowley,1997:89].

Furthermore, the word 'resemblance' is really the inverse of perceptual distance which is a function of phonetic distance so that it is distance that we need to do science, not the invention of more words [see appendix]. The method's fuzzy rules are explained clearly in Crowley[1997]

> Having set out all of the sound correspondences [SC or RegSC] that you can find in the data, you can now move on to the third step, which is to work out what original sound in the protolanguage might have produced that particular range of sounds in the various daughter languages. Your basic assumption should be that each separate set of sound correspondences goes back to a distinct original phoneme. In reconstructing the shapes of these original phonemes, you should always be guided by a number of general principles:
>
> (i) Any reconstruction should involve sound changes that are plausible.

> (ii) Any reconstruction should involve as few changes as possible between the protolanguage and the daughter languages [Crowley,1997:93].... It is perhaps easiest to reconstruct back from those sound correspondences in which the reflexes of the original phoneme (or protophoneme) are identical in all daughter languages. By principle (ii) you should normally assume that such correspondences go back to the same protophoneme as you find in the daughter languages, and that there have been no sound changes of any kind [Crowley,1997:93].
>
> (iii) Reconstructions should fill gaps in phonological systems rather than create unbalanced systems.... Although there will be exceptions among the world's languages, there is a strong tendency for languages to have 'balanced' phonological systems. By this I mean that there is a set of sounds distinguished by a particular feature, this feature is also likely to be used to distinguish a different series of sounds in the language. For example, if a language has two back rounded vowels (i.e. /u/ and /o/), we would expect it also to have two front unrounded vowels (i.e. /i/ and /e/) [Crowley,1997:95].
>
> (iv) A phoneme should not be reconstructed in a protolanguage unless it is shown to be absolutely necessary from the evidence of the daugher languages [Crowley,1997:98].

Here it is seen clearly that cognacy and geneticity are obviously almost identical for the purpose of argument about the relationship of the languages in question. If two languages are genetically related the comparanda are cognates ($G \Rightarrow C$), and if the comparanda are cognates then the two languages are genetically related ($C \Rightarrow G$). This is a logical equivalence since $X \Rightarrow Y$, and $Y \Rightarrow X$ is equivalent to $X \Leftrightarrow Y$ which says that X and Y are equivalent.

We often read in mailing lists or USENET discussions that we cannot compare daughter languages, but instead are required to reconstruct the protolanguage before we can make geneticity judgements, or ideas to this effect. This is simply bad reasoning or confusion which is related to circular reasoning since it emanates from a miscomprehension of what the comparative method (heuristic, i.e. a rule of thumb) does. For example, let R represent the relation which we can use to obtain the sounds/words of a daughter language L_m from the proto-language L_0, so that $L_m=R(L_0)$. Let S represent a similar relation for another daughter language (i.e. $L_n=S(L_0)$). If the sound changes are completely regular (or even mostly regular) then we can invert this so that we can obtain $L_0=R^{-1}(L_m)$ so that $L_n=S(R^{-1}(L_m))$ meaning the we can relate the languages to each other via regular sound change again. Indeed the protolanguage is reconstructed from the daughter languages so the creation of genetic relationships, family trees and protolanguages are a parts of a same single iterative algorithm. There is nothing to prevent anyone from comparing present day languages to each other because most of the time that is all the data we have. The protolanguage is merely a shorthanded representation of the relationships that the presumed daughter languages have with one another.

5. Incapacity Arguments

Many of these have to do with something that is closely related to how science is allegedly done but it isn’t. They are based on premises that might not be true, but are taken to be true. Among these are argument from ignorance, or argument from lack of imagination, or even argument from authority, and related fallacies. Many of them are based on a false understanding of how we obtain scientific knowledge, and how science is supposed to be done.

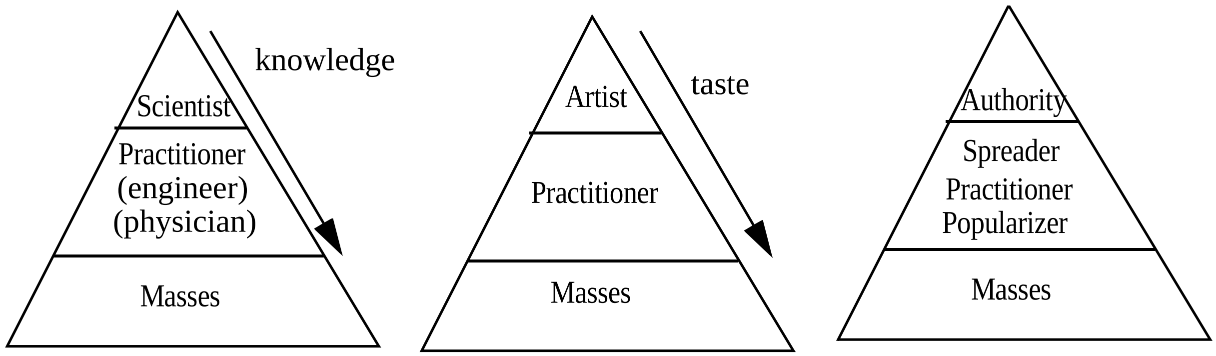

Figure 4: Knowledge/Information is supposed to “flow down” from the scientists just like the flowing down of styles, and taste from those artists and experts in taste. In both cases, the apex of the triangle represents ‘authority’ and the masses allegedly only follow.

In the real world none of us can be experts at everything. Therefore we are constantly asking others who presumably know more than us for advice. But this normal scheme of things is merely the way society operates because we have not yet been able to do better but it is not how science is done or should be done. In informal social discussion (not scientific discussion) it is generally thought (at least by some) that appeal to authority isn't always completely bogus; for example, it may be relevant to refer to a widely-regarded authority in a particular field, if you're discussing that subject. However, this is not how science is done. If we are to appeal to Newton’s laws, we should know some physics. In any case, for example, the way physics is done, a freshman physics student sits in a lab and confirms that some of the ‘laws of physics’ which he/she is learning do indeed turn out to be the way it is written in the textbook. From there the mathematics carry it forward into more complexity. The only serious problem, discussed above is the problem of induction. That is why the scientific laws of any aspect of the universe are best treated as tendencies, propensities, or events with various levels of uncertainty. It just turns out that the laws of physics, those dealing with nonliving things without volition turn out to be easier to fathom than those dealing with living volitional sentient beings. Practically speaking it makes the job of the scientist dealing with living entities much more difficult which is all the more reason to have more powerful analytical tools in the toolkit. Therefore social scientists should know much more mathematics than physical scientists such as engineers, computer scientists, or physicists.

Unfortunately, there is a concept called *triage* which is usually at work in human affairs. The most naked form of it occurs in wartime treatment of the wounded. The scarce resources (medical treatment in this case) must be apportioned in a way as to do the most good. The wounded are split into three groups. Those with the light wounds will not die even if they are not treated, so they do not get immediate treatment. Those who suffered extremely grave wounds will probably die even if they are treated. Meanwhile the middle group which could have been saved if treated immediately will die as the most seriously wounded are being treated. Therefore it is in the middle group that immediate treatment does the most good since it is they who will most likely survive if treated and die if not treated. Society’s rationing of goods and services basically follows this triage model. The energy of the state is concentrated in the fields in which the return is maximum which means that the physical sciences and technology get the greatest funds, and it is these people who also know from centuries of past experience that mathematics does the most good in these fields. Therefore, the social sciences are left to those who unfortunately somehow are not told about the power of the tools that created modern civilization. The result for the present in some fields is almost a total inability to comprehend or use even the most basic and fundamental mathematical tools in the arsenal of the scientists. Of course, one can easily make the argument that nobody ever died due to a lack of poetry, therefore not pouring money into the social sciences is again a mark of putting the energy of a society where it will do “most good”. It is probably due to reasons like these that many of the ‘incapacity arguments’ are used, too often in linguistics and allied fields of study. There really is no excuse in the 21st century not to know at least about logic and how philosophy of science has affected the [other] sciences, and how science is done.

5.1) Argumentum ad nauseam etc.

This is the incorrect belief that an assertion is more likely to be true, or is more likely to be accepted as true, the more often it is heard. It might have been named as *Argument by Repetetion* because, it only gets its force from the fact that it gets repeated often enough by enough people. Related this fallacy is the *Argumentum ad numerum***,** and also *Argumentum ad Populum*. It consists of asserting that greater the number of people who support or believe a proposition, the more likely it is that that proposition is correct. The more people repeat this on discussion lists devoted to problems of historical linguistics, the more it is supposed to be believable and closer to truth, whereas nothing changes simply via repetetion except for the fact that it engenders disgust, hence the name ‘ad nauseam’. Of course, it is possible in some probabilistic sense that it might work out on a majority of cases, but logic does not deal with propensities or tendencies or probability theory. An infamous example of this form of argument is *“About x% of the languages are agglutinating so that says nothing about geneticity*” or “*So many languages are isolating that it can’t be genetic*”. Since we argue often in historical linguistics via analogy, let us compare this argument to “About 80% of the people in the world are white/Europoid so that cannot be genetic.” This is obviously quite false. Being white is genetic and is inherited from parents who are white.

Very closely related to this *Argumentum ad verecundiam* which is simply an *Appeal to Authority*, and uses admiration of a famous person to try and win support for an assertion. It is particularly odious if the authority figure is simply using one of the other fallacies, such as *argumentum ad numerum*, or *argumentum ad populum*, or *argumentum ad antiquitatem*, or even using an earlier authority figure, such as Saussure. For example, if a nobody uses this form of argument by appealing to say Trask [1996], it is a simple argumentum ad verecundiam. The reason these arguments are used regularly in discussion is likely because the rules of the heuristic (the comparative method), and why it works are not understood so that the discussants find it easier to invoke authority (usually an author of a book). Unfortunately, if two authors do not agree, then the argument usually reduces to some kind of a vote on how many authorities agree with each other. This is due to a misunderstanding of how science is supposed to be done (please see the first part of the section).

5.2) Argumentum ad ignorantiam

Argumentum ad ignorantiam means "argument from ignorance". The fallacy occurs when it's argued that something must be true, simply because it hasn't been proved false or, equivalently, when it is argued that something must be false because it hasn't been proved true. Note that this fallacy doesn't apply in a court of law in the USA, where you're generally assumed innocent until proven guilty. Also, in scientific investigation if it is known that an event would produce certain evidence of its having occurred, the absence of such evidence can validly be used to infer that the event didn't occur, but again this would most likely use probabilistic arguments which are beyond the scope of this paper. The same argument is being used when someone says "*I don't see how that can be true"* because *"I can see how it can be true so it must be true*" is not an argument and arguing its opposite can't be either. In other words the often heard refrain "absence of evidence is not evidence of absence" is true in logic but not necessarily in probability theory, so words like logic, reason, rationality etc. should be used with caution and with precision.

5.3) Argumentum ad antiquitatem

This is the fallacy of asserting that something is right or good simply because it's old, or because "that's the way it's always been." the natural implication here being that we cannot do any better because if there was such a way, it would have already been done and is the opposite of Argumentum ad Novitatem. This is one of the most famous ones, if not the most famous in historical linguistics. We are simply told *"this is the way we do it in historical linguistics"* or *"this is the way we have been doing it for the past N years*", or *"this is the comparative method of linguistics and that is how it works, and that's all there is to it"*. It seems to be the standard argument of those who do not understand the reasoning behind the heuristic called the 'comparative method'. It is inexcusable for a linguist not to understand why the method allegedly works, if it does. Variants of this can be found almost any day of the week in some mailing list or USENET newsgroup. Many books devoted to the principles of historical linguistics claim this, more or less, by omission. The two books that are quite clear and honest about the methods of historical linguistics and which are exceptions are those of Crowley [1997] and Lass [1998].

5.4) Shifting the burden of proof

The burden of proof is always on the person asserting something. Shifting the burden of proof, a special case of *Argumentum ad Ignorantiam*, is the fallacy of putting the burden of proof on the person who denies or questions the assertion. The source of the fallacy is the assumption that something is true unless proven otherwise.

5.5) The "No True Scotsman..." fallacy

Suppose someone asserts that no Scotsman puts sugar on his porridge. A counterexample could be to point out that Angus likes sugar with his porridge. But then it is only too easy to say "*Ah, yes, but no true Scotsman puts sugar on his porridge.*" This is an example of an ad hoc change being used to shore up an assertion, combined with an attempt to shift the meaning of the words used original assertion. One can argue again by pointing out that "experts" (real linguists) agree on something, but then when pointed out that X is a linguist, one might retort "*you call him a real linguist*", or "*you call that real linguistics*". This is often used with phrases like "*You call what Greenberg did historical linguistics?*" It would seem that it is precisely those things in language that can be changed slowly piecemeal and via copying (borrowing) that should not be used as markers in geneticity, but are used because it is said that some of them are resistant to borrowing, but again this is mere circularity and not backed up by any objective evidence, and made to look even better by the creation of ad hoc artifices. For example, it is now well known (see for example, Hauser,1997 with evidence stretching back to the 1970s, and even popping up more recenly in a newsweekly, US News and World Report, 1998) that infants are attempting to repeat sounds made by their parents when they babble words which are considered in standard linguistics jargon as 'infant talk', and yet this ad hoc idea invented to shore up the state of IE studies is still being repeated in books and journals.

6.1) Fallacies of Complexity, Distance and Measurement

This class of fallacies has to do with putting things out of order, or allegedly taking things apart (analysis) or putting the pieces together (synthesis) in ways in which the assumptions are buried inside seemingly complex and complicated machinations. For example, someone might say that linguistics is so complex that mathematics is useless. Doing science is a complex undertaking and until someone finds a way to do it some other way, we will always be using analysis & synthesis. We take apart a complex object to examine the pieces to understand how they work, and then using this knowledge we put back the pieces to try to understand how the big object works. Often times, this is called reductionism in an effort to use strawman arguments (see below) against it.

6.1.1) Plurium interrogationum / Many questions

This fallacy occurs when someone demands a simple (or simplistic) answer to a complex question. For example, on the one hand linguists attempt to shove every language into some family on the basis of the historical method, and on the other hand deny that something like ‘dialect’ can be defined at all. Often linguists deny that anything in linguistics can be measured but simply by creating language families they are, in fact, using ‘measurement’ or ‘distance’ at least on an ordinal scale (please see appendix). On the basis of these ‘distances’ among languages, a family tree can be constructed, but allegedly the same ‘distances among languages’ cannot be used to create a metric for whether the languages are mutually intelligible or whether they are so close to each other that they should be called dialects.

6.1.2) Model Building and Non causa pro causa fallacy (false cause fallacy)

The fallacy of Non Causa Pro Causa occurs when something is identified as the cause of an event, but it has not actually been shown to be the cause. In this we argue $\neg A(A \Rightarrow B) \Rightarrow \neg B$ however it is not true (or has not been shown to be true) that $A \Rightarrow B$. This is a very serious problem for model building in the sciences. Every must of course, attempt to find causes for events. Probably the biggest problem now in historical linguistics is to attribute causes of change; *exogeneous vs endogeneous*. When is a language descended from an ancestor and when is it a ‘borrowed language’ (i.e. new language learners)? The whole method hangs by a very thin thread, and this thread happens to be that a small proportion of words in languages is resistant to change. The formalization of this concept is due to Swadesh, but other lists like it such as the Yakhontov 35-65 list can be found. We are stuck with Swadesh lists and others like it because allegedly the other possibilities are too complex to be tractable by mathematical methods but not too many of them have even been tried for no other reason than the application of triage in society. Those who can work in other fields, and those who work in this field can’t. Sad but true. One wishes sometimes that it was only sad and not shameful when one is faced with blatant attempts by those in the field to resort to methods more suitable for politicians instead of scientists, but truth cannot be found any other way except the scientific way. If it can be then it is purely serendipitious.

6.1.3) Equivocation

This occurs when a keyword is used with two or more different meanings in the same argument. This is used quite often in historical linguistics for example, when someone argues that some word correspondences are merely phonetic similarity when phonetic similarity has not been defined or variant definitions are being used. A popular version of this occurs when a word or a phrase is used whose meaning is apparently not clear, so that even the ramifications or implications of the definition are not clear. For example, someone makes a blanket statement with all earth-shattering seriousness “phonetic similarity has nothing to do with geneticity” without realizing that if there is a function d(x,y) that measures the distance between phonemes or words/morphemes, and that this function is the inverse of the similarity function s(x,y) then the claim is equivalent to “phonetic distance has nothing to do with geneticity”. But if we consider any language at all, say English to English, and then compare the Swadesh list, or the Yakhontov list or any other meaningful comparison upon which geneticity might be based, we will find, not to our surpise, that the words are identical which means that the phonetic similarity is 100% and the phonetic distance is zero between each pair of comparanda. Obviously, this is how we know that the two languages being compared are identical languages. What kind of a method would say that English is not genetically related to itself?

6.2) Discretization and Contiguity Fallacies

6.2.1) Fallacies of Composition and Division

One *Fallacy of Composition* is to conclude that a *property shared by the parts of something must apply to the whole*. Sometimes it does and sometimes it does not. These concepts have been formalized in thermodynamics by the division into *extensive* and *intensive* variables or parameters. If a system (say a container of gas) is split into two equal halves, then the values of the extensive parameters will be halved i.e. $X_1=X_2=X/2$ (e.g. volume of gas, or entropy of gas) but the values of the intensive parameters will stay unchanged i.e. $x_1=x_2=x$. (e.g. temperature of the gas, or its pressure). If a language was split in half then its lexicon would be halved, but would its morphology also be halved? It is the intensive parameters of languages that survive longer because they are the parameters that truly give it its character. Our human species’ characteristic shape is an intensive system property. Even if the human population was split into a million pieces those characteristics which define us as human would still persist because it is those characteristics that make use human. Even if an arm or a leg is missing a human can still be recognized as a human because many characteristics go into the definition of these intensive parameters that make us human. Similarly lopping off words from sentences still leaves much semantics in the sentence, and still leaves the flavor of the language in the mangled sentences for these are the intensive parameters of the human languages.

Further along these lines are the fallacies that are called *sorites paradoxes* in logic (see below). We should also note that people often mean to imply the distinction between extensive and intensive when they use the words *quantity* vs *quality*, for quality is quite easily quantifiable, and often is in physics, engineering and economics. Extensive and intensive variables; system characteristics, vs piecemeal change (i.e. changing the vocabulary, lexicon and keeping the syntax) are all a part of the various fallacies of composition and division. The *Fallacy of Division* is the opposite of the *Fallacy of Composition*. Like its opposite, it exists in two varieties. *The first is to assume that a property of some thing must apply to its parts*. The other is to assume that a *property of a collection of items is shared by each item.*

6.2.2) The slippery slope argument: Sorites paradoxes

This fallacy is a very complex one and still unresolved because it hits right at the heart of the weakness and limitations of logic and points in the direction of what other mathematical methods we must use in order to resolve the issues that are involved in it. The classical case is an argument of the type;

1. A man with a penny is not rich.
2. If a man with n pennies is not rich, then a man with n+1 pennies is not rich.
3. Therefore no matter how many pennies a man may have, he is not rich.

A more modern version and closer to home can be seen in Faletta [1990]. There is an aquarium with a tadpole in it. A VCR tapes the tadpole until it turns into a frog. Suppose there are 1 million frames from the beginning to the end. We are sure that the 0th frame shows a tadpole, and the millionth frame shows a frog, but there is no frame N such that this frame is a tadpole and the next frame is a frog. Similarly if we considered only the lexicon in geneticity, and if we had a dictionary of language X at 10-year intervals for 10,000 years we might see that it is no longer the same language after 10,000 years but we would not be able to point to a year in which it was language - X and language-Y ten years later. This obviously has serious repercussions for geneticity and family trees. There is another version of this called Neuraths' Boat, in which planks of a boat keep getting replaced. At what time is it no longer the 'same' boat? If we had been cannibalizing other boats to add planks to this board, then when is the boat no longer the same boat? If we had used parts from another boat, does *this boat* at some time become the 'other' boat? If we complicated this problem slightly by slowly changing the shape of the boat, and also making changes to other parts of it such as adding an engine, replacing the engine, etc then we are closer to the problems faced by historical linguists. One can also see the fallacies of composition and division and equivocation also in these examples, and these are in fact probably the ones in which explicit mathematical models are needed. Obviously such models require measurement and the concept of distance. Resistance to the concept of a language changing to another one, or gaining two ancestors, or even n ancestors purely on the strength of repetetion of a mantra like "this is how historical linguistics is done" then becomes one of the fallacies such as *argument from authority*, or *Argumentum ad antiquitatem*.

6.2.3) Multifurcation

Also referred to as the "black and white" fallacy, bifurcation (which is a special case of multifurcation) occurs if you present a situation as having only two alternatives, where in fact other alternatives exist or can exist. This also has affinities with the sorites paradoxes, and *plurium interrogationum* (many questions) fallacies. This form of reasoning occurs repeatedly and constantly in discussion in language change. For example, there are only several ways in which language changes can take place: divergence, pidgins, creoles. Finally, the concept of a 'mixed language' has began to take shape, but again, it is yet another discrete class thrown into the several classes. We can, in fact, view language changes as a continuum and then view the various classes as extreme cases of the various possibilities. If we do this, then we also do not have to put up with the genetic, areal and borrowing relationship amongst languages since these are also examples of multifurcation. Much of this is to be expected since naming abstract or concrete objects and classification is a very big part of the social sciences (and also the life sciences). It is true that there is much naming and classification in chemistry but the scheme is quite orderly and regular, and one does not need to create latin names for compounds hodge-podge. Measurement is a very important part of science, and classification is at the bottom of the heap in the measurement scales (please see appendix). Since almost all of linguistics is about discrete objects, sets and classifications, there are many choices. However, since this is about historical linguistics only some applicable ones should be given. The simplest one is the claim that sound change is “discrete”. Obviously, phonemes are discrete so how else could sound change be? Just as good are classifications of relationships into discrete categories genetic, areal, borrowing, etc. Yet another set of classifications is into the types isolating, agglutinating, fusional etc. We could just as easily consider these to be idealized versions of a continuum in the interval [0,1].

7. Ad Hominem and Miscellaneous Fallacies

These fallacies revolve around irrelevant items being introduced into an argument. The irrelevancy can be made in many ways, including attacking the person, his/her character, his/her background, or simply be diverting the discussion, appealing to the audience, etc. A form of this argument is related to *Argument from Authority* when the person says that he is a linguist and that some other person is not, and hence the argument is invalid.

7.1) Red herring

This fallacy is committed when someone introduces irrelevant material to the issue being discussed, so that everyone else's attention is diverted away from the points made, towards a different conclusion.

1. Isn’t it true that some mathematician had problems with the Marilyn and the Goats problem?
2. Mr. X asked me about Y, but I don’t have time here to discuss it; I sent him email.
3. But Ms. Y is not a linguist.

8.2) Straw man

The straw man fallacy is when you misrepresent someone else's position so that it can be attacked more easily, then knock down that misrepresented position, then conclude that the original position has been demolished. It's a fallacy because it fails to deal with the actual arguments that have been made.

8.3) Tu quoque

This is the famous "you too" fallacy. It occurs if you argue that an action is acceptable because your opponent has performed it.

1. You are only repeating some idea from X.
2. So you are only repeating some idea from Y.

First of all, probably many ideas that could arise in the study of historical linguistics have already been explored by at least one linguist, even if linguists are not ready to defend every view like philosophers, so that it is unlikely for anyone to be completely original. Furthermore, science is not art and is not on the look out for originality (i.e. being different for its own sake) therefore it is easy to see how ideas can be repeated. If, on the other hand, the person repeats it without seeming to have an understanding of the issues involved then it is the appeal to authority or *argumentum ad numerum*, or *argumentum ad populum*, or *argumentum ad antiquitatem*. But an argument of the type

1. You are playing to the audience.
2. So you are too.

does not have the same complexity as the earlier one and is simply a personal attack, and is therefore a special case of *Argumentum ad Hominem.*

Conclusion

The best and only way to escape committing logical fallacies is to practice doing science the way it is supposed to be done. This can be done by laying the foundations and the basis of the working of the comparative method, to understand its limitations, why and how it works.Otherwise historical linguistics will be merely a memorization of interesting facts, and even more interesting games, at best a heuristic rule for finding some patterns in languages, but never a science.

Appendix: Measurement Scales

Nominal Scale:

On this scale, we can only say things like "x is y" i.e. this is the classification scale. We can say that "x is a member of the set X" which is the mathematical/logical way of saying it. But we cannot even rank things on this scale. A dog which can distinguish between food and non-food is doing categorization and can work at the nominal scale, but is it doing science?

Ordinal Scale:

These are rankings. We can use the > sign for the ranking. We can say x>y, or y>z but we cannot add or subtract these numbers.

Interval Scale:

We can add these numbers (and subtract). But we cannot multiply. The Fahrenheit and Celsius scales are interval scales.

Ratio/Absolute Scale:

These numbers can be multiplied (and divided). We can exponentiate and take logs, etc etc. Measurement of length is done on a ratio scale. However, the Fahrenheit and Celsius scales are not ratio scales and those number cannot be multiplied. Their products have no meaning.

Now, the concept of a "distance metric" or "distance" for short, puts any social science at the top scale where the numbers can be added, subtracted, divided, multiplied, logs taken, and exponentiated. In other words it is on its way to becoming a full-fledged science. Is this concept not worth looking into? Nay, is this concept not worth seizing immediately?

References

Anttila, R. (1989) Historical and comparative linguistics, John Benjamins Pub. Co., Amsterdam

Campbell, L. (1998) Historical Linguistics: An Introduction, MIT Press, Cambridge, MA.

Crystal, D. (1992) An Encyclopedic Dictionary of Language and Languages, Blackwell, Cambridge, MA.

Crowley, T. (1997) An Introduction to Historical Linguistics, Oxford University Press, New York.

Dixon, R. (1997) The Rise and Fall of Languages, Cambridge University Press, New York.

Faletta, N. (199) The Paradoxicon, Wiley, New York.

Foley, J.(1977) Foundations of Theoretical Phonology, Cambridge Univ. Press, Cambridge.

Fox, A. (1995) Linguistic Reconstruction: An Introduction to Theory and Method, Oxford University Press, New York.

Hauser, M. (1997) The Evolution of Communication, MIT Press, Cambridge, MA.

Hubey, H.M. (1999) Computational and Mathematical Linguistics, Lincom-Europa, Muenchen.

Hubey, H.M. (1999a) Mathematical Methods in Historical Linguistics; their use, misuse and abuse, submitted to Journal of the IQLA.

Katicic, R. (1970) A contribution to the general theory of comparative linguistics, Mouton, Hague.

Lass, R. (1997) Historical linguistics and language change, Cambridge University Press, New York.

Lehmann, W. (1962) Historical Linguistics: an Introduction, Holt, Rinehart, and Winston, NY.

Lyons, J. (1968) Introduction to Theoretical Linguistics, Cambridge University Press, New York.

Malmjaer, K. ed (1991) The Linguistics Encyclopedia, Routledge, London.

McMahon, A. (1994) Understanding Language Change, Cambridge University Press, New York.

Ringe, D. (1992) On Calculating the Factor of Chance in Language Comparison, The American Philosophical Society. Transactions of the American Philosophical Society, vol. 82, Philadelphia.

Ringe, D. (1995) ‘Nostratic’ and the Factor of Chance, Diachronica XII:1.55-74.

Thomason, S. and T. Kaufman (1988) Language Contact, Creolization, and Genetic Linguistics, University of California Press, Los Angeles, CA.

Trask, R.L. (1996) Historical Linguistics, Arnold, New York.

Trask, R.L. (1996) A Dictionary of Phonetics and Phonology, Routledge, London.

USN&WR (1998) Baby Talk, June 15.