DOCUMENT RESUME

ED 401 766 FL 024 275

AUTHOR	Embleton, Sheila
TITLE	Comments at Symposium "Against Multilateral Comparisons."
PUB DATE	5 Jan 95
NOTE	11p.; Paper presented at the Annual Meeting of the Linguistic Society of America (69th, New Orleans, LA, January 5-8, 1995).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Comparative Analysis; *Contrastive Linguistics; *Language Classification; *Language Research; *Linguistic Theory; Probability; Research Methodology; Statistical Analysis; *Structural Analysis (Linguistics)

ABSTRACT

The comments presented here were made after the presentation of four papers and commentary by two other symposium participants. They address issues in language comparison and classification. First, comments are made on the papers ("An African Test Case in Comparative Methodology," "The Mathematics of Multilateral Comparison," "Testing a Basic Evaluation Metric," and "Multilateral Comparison and Linguistic Geography"), then more general comments are made on the utility of statistical tests, and of mathematical or probabilistic techniques in general, in linguistics. Issues discussed include sampling, true independence of variables, validity of underlying assumptions, inter-group differences, and error types associated with null hypotheses. Following this, factors supporting use of statistical methods are reviewed. A number of quotations and pieces of information seen as useful in a discussion of these issues is appended to the text. Contains six references. (MSE)

LSA New Orleans, January 5, 1995

Symposium "Against Multilateral Comparisons"

Comments by Sheila Embleton (York University, Toronto, CANADA)

[Following 30-minute papers by Lionel Bender ("An African test case in comparative methodology"), William Poser ("The mathematics of multilateral comparison"), Donald Ringe ("Testing a basic evaluation metric"), and Johanna Nichols ("Multilateral comparison and linguistic geography"), as well as commentary by Alan Kaye (read by Bender) and William Baxter]

I should probably state at the beginning that I only had two of the papers sufficiently in advance to really properly/adequately read them (Bender, Ringe), and received Nichols' most interesting and rich paper only a few days ago, and never received Poser's paper, just his handout this afternoon. In any case, this will contribute to my comments being brief. As Bill did, I should probably also state at the outset that I come here, or to this question of long distance relationships, with an "open mind". I'm not for example an Amerindianist (although I've had a few brief encounters with a couple of the languages), not a Nostraticist (although I have some training or experience in IE and Uralic), but a specialist in mathematical, more specifically statistical and probabilistic, techniques and modelling in linguistics, specifically in historical linguistics, dialectology, and stylistics. I am, as Karl Teeter recently (Dec 16/94) characterized himself on LINGUIST, "neither a 'splitter' nor a 'lumper', just an old-fashioned believer in facts and proof". So I have no axes to grind, no hidden agendas, etc. — and am interested, happy, and otherwise satisfied with whatever result comes out of honest open inquiry in this field. The hour is late, and as I said already, I'm not as prepared as I would have liked to have been — and in any case, I have the impression that a lot of what there is to say on these topics has been said already, tonight or elsewhere, or else is so obvious as to not be worth saying or repeating. I would also like to keep my part short so as to allow time for the audience to ask questions and contribute to the discussion. In fact some of you may be crying out to speak ...

First a few quick comments on the papers, and then some more general comments.

<u>Bender</u>: Not much for me to comment on here, and I certainly can't comment on any of the African part. I agree with him that there seems to be a polarization in the field, and I deplore that. He points out a fallacy in the Multilateral Comparison probabilistic reasoning, namely that the argument is usually stated in a case in which a similarity has already been noticed in two or more languages, but that the situation of interest is actually one in which one BEGINS with a set of languages and THEN looks at a particular item on the word-list, and THEN starts looking for matches. He also correctly points to the problems attendant on "wide latitude", both for phonological matching and semantic matching, the latter largely unsolvable — and how the probabilities of finding a match

quickly escalate. One problem that he didn't address (or else I missed it) was another factor that tends to escalate the probabilities, the way the "method" is usually applied — the precise set of languages is often rather unconstrained too, in that the practitioner often allows him or herself to find a "match" among any of a number of languages belonging to say one sub-family. Many of the points he raises have come up in Ringe's paper, or Ringe's other work, so don't really need addressing here, with my limited time — e.g., the fact that one needs to realize "that every language has its own set of phoneme frequencies and therefore every pair of languages has a different set of paired phoneme frequencies".

Poser: I didn't have this paper in advance. There's a sentence in his abstract that goes right to the heart of something that has bothered me for long, but I haven't had an opportunity to comment on in public, only private — and in fact 20 seconds ago in my brief remarks on Bender's paper. "The mathematical argument used to support multilateral comparison ... ignores the fact that if an equation need not include all the languages in the universe of comparison, the number of possible subsets entering into an equation can be very large, which greatly increases the probability of chance matchings". For me, the other most significant point that he made (and Ringe also referred to briefly), hardly new for statisticians but one which needs emphasizing repeatedly for linguists, is that any part of the whole edifice that you've built up can be the weak link that has caused the statistical test to tell you that you're dealing with a non-random situation, that you have a so-called significant result. Now, you often have your own ideas/preferences as to what this might be, because you have your own agenda, and probably want to assume that it means that the data are not random. But it can be, and often unfortunately is, that the assumptions of the test are invalid or in some way not met. Until this can be ruled out, you can't just confidently assume that you have significant non-randomness in the data.

Ringe: Ringe makes the point, not original with him, also made by Bender tonight, but worth emphasizing over and over, that "comparing approximate synonyms, or using matchings between whole classes of sounds, substantially increases the incidence of chance resemblance". He also makes the point, perhaps not original with him either, but less often repeated, that if we are willing to accept matchings of e.g. initial consonant with "consonants after the first-syllable vowel", we radically increase the incidence of chance resemblance. I find particularly useful his randomized testings — it reminds me of some different and earlier work by Oswalt — but it's invaluable as a useful heuristic, and in getting a feel for the data.

There exist practical reasons for binarity, although Ringe and others have been criticized for binarity, made fun of by "long-rangers" — it's a necessary approximation/short-cut given the degree of complication in the formulas necessary. If you are going to do these more complicated approximations that allow for different phoneme distributions and frequencies in different languages, there is no other way — except to get into less precise approximations, simulations, etc.

— Jacques Guy, Robert Oswalt, Bill Baxter, etc. Both are useful approaches. In particular though these types of simulations are great for getting a feel for the system, the interrelationship of parameters, how a change in latitude allowed in phonological matching alters the chances of matching and so on — I've always been a great advocate of simulations, especially when run in large numbers, and I continue to be.

<u>Nichols</u>:

As mentioned earlier, I didn't have a lot of time to go over her paper, and there's certainly lots in there, but I can offer a few comments. She sums up the essence of the whole debate rather succinctly very early on: "the kind of evidence that is diagnostic for genetic relatedness is evidence that could not be expected to recur elsewhere by accident. This much seems to be universally agreed. Less well understood is how to decide whether some pattern or form could or could not be expected to recur elsewhere in the world's languages by accident". I am puzzled by a number of things in her paper, but perhaps they are not best questioned now — for example, again early on, I don't understand phrases such as "diagnostic of one individual language", or just why it is that she multiplies the rate of occurrence of a feature in one individual language, namely 1 in 5000 or .0002, something which should be deterministic, by the level of significance, e.g. 0.05, as a "margin of safety", to get .00001 and so on (page 3). I think there are also problems in the calculations involving the 4-consonant word PIE *widhew "widow" (section 2.1.1), which is of course not even transparently 4 Cs to those not up on their IE. And remember that at the depth and breadth one is likely to have to work at, it's unlikely that one <u>would</u> be thinking of this as 4 C's. Also, one has already used one's "extra" knowledge of the language to even know that v=w, dh=dd=d, etc., unless one is explicitly allowing that as similarity latitude in the phonological matching. One should also make an attempt to use the phonological frequencies appropriate to the language (as Ringe does), or the whole calculation really does become quite loose. In fact, it would probably be heuristically useful to combine some of Ringe's calculations with some of Nichol's approach, and get some really good heuristics. Although it may still be all right as a very rough ballpark figure, which is what she is more or less attempting to do. [I could also make the same comment about some of the calculations related to age in section 3, page 15.] I am also puzzled by some of the calculations involving "arbitrary" vowels and consonants. Particularly actually arbitrary consonants (e.g. page 7) — if arbitrary in the sense of "any consonant will do", then why multiplying by a probability .5 rather than 1? Unless .5 is to represent presence vs. absence, which it may do, given the existence in the cited set of *maa*, where there is no consonant between the two a's. There are simpler ways of doing the probabilistic computations in 2.3 [handout (11) and (12)] (chances of no heads after a given number of tosses). But I don't mean to nitpick — regardless of this, her results (page 9 [handout (12), (13), (14)]) that if you allow yourself e.g. 5 shots at finding a match, by searching around the semantic field, you are getting pretty good chances of <u>finding</u> a match, are of course correct, not to mention sobering. Some of her demonstrations in the other direction, if one can term it that, are equally important and equally sobering — e.g. (page 11 [handout (16)]) the "small lexical search" for personal pronouns. (page 13) Although I agree in general with her claim that "individual-identifying evidence can prove genetic relatedness", I do not agree that it is a corollary, or follows logically, that "any putative genetic grouping NOT backed up by individual-identifying evidence should be ... rejected", although I do agree that regarding it as "hypothetical or speculative" is reasonable enough. More on this below, in my more general comments.

Now the more general comments.

I would like to make a few remarks about the utility of statistical tests, and perhaps mathematical or probabilistic techniques in general, for linguistics. Not because I think that linguists are naive about statistics — some are and some aren't — or because I think that the use of statistics will cause anyone to be convinced of anything, or to switch sides in any debate, such as the current one, but because often doing statistical tests properly forces you to make explicit, and maybe even re-examine, some of your assumptions, makes you change/re-consider assumptions, see some things in new perspectives. Some examples:

— You are forced to consider seriously the size of your sample, and whether it's adequate. Also what to do about missing data points, or data that is hard to categorize in some way (e.g., languages where it's hard to determine a basic word order type).

— You are forced to consider whether the features you are examining are statistically independent, something required for most statistical tests to be applicable/valid (e.g., chi-square).

— You are forced to lay out <u>all</u> the assumptions underlying your test, and (as mentioned earlier) to realize that a "significant" result can often simply mean an invalid assumption has been made, not that the data are significant in some way.

— Checking the significance of inter-group differences forces you to explicitly look at intra-group differences (you may see new things!), to make sure that inter-group differences really are bigger than the intra-group differences.

— It forces you to consider that there are after all TWO types of error that you can make with regard to a null hypothesis — that statisticians refer to as Type I and Type II errors. Essentially that rejecting a true hypothesis (Type I error) and accepting a false hypothesis (Type II error) are both errors. This forces you to think carefully about which is actually the greater/worse error — or at the very least to realize that both <u>are</u> errors. In short, it clarifies your thinking. I think this particular fact, that there are TWO types of error, is not adequately addressed in linguistics in general, nor specifically in the consideration of long distance relationships. We seem to spend all our time worrying about one type of error, Type II error — phrased differently, we are being very careful not to accept the null hypothesis when it's false, and seem to worry less about rejecting a null hypothesis that might be true. Statisticians would take this as a sign of conservatism, by the way. Usually such types of conservatism are considered reasonable by statisticians only in some contexts, where the "cost" of an error is very high — e.g. studies of the <u>safety</u> of a new medicine (as opposed to studies of the <u>effectiveness</u> of a new medicine). In other words, when a certain type of error is devastating and must be avoided at all costs. I'm not saying that we shouldn't be conservative — just that we should be aware of what we are doing, and of the other type of error. And I'm not going to go into the whole issue of decision-theory, and just how one can try to optimize the expected gains and losses vis-à-vis both types of error.

So, quite apart from what the statistics may (or may not) tell you/others, it's good for a general re-examination of assumptions, goals, conclusions, new perspectives in general.

People often ask me what WOULD constitute statistical proof, or beg me to set up some formula for them which would e.g. calculate the probability of chance resemblance in these etymologies/long-distance look-alikes. These issues have already been touched upon by Ringe, but I'll touch on them briefly again. Just think of all that has to go into such a formula, to even come close to approximating reality — and remember, if you don't put in all these factors, people will criticize you for it, claim your formula is invalid because it doesn't include whatever, and therefore disregard your conclusions, particularly if they don't fit into THEIR views. So, you have to have not just the phoneme inventory, but also information on the phoneme frequencies, their phonotactic restrictions/cooccurrence restrictions on their distribution, should probably take account of factors related to persistence/universality, maybe factors related to acquisition, maybe the TYPE of morpheme it's in, and so on. And that's just the phonology. What about when you get to the meaning side of the whole thing? There simply is not the theoretical apparatus or even plain practical knowledge here that there is in phonology. Lexicostatistics/glottochronology handled this semantic problem by allowing no latitude whatsoever — which is of course easy to criticize, for all sorts of reasons — and it certainly <u>was</u> criticized! — but the most obvious is the sort of thing like missing English *hound* as being cognate to German *Hund*, because you insist on only considering *dog* in English. But as soon as you allow some latitude, NOBODY is going to agree on just how much latitude, and you've got yourself a whole new can of worms, and a whole new set of reasons as to why people won't accept your method, if they don't like what your method concludes. [I could add parenthetically that semantics is the trickiest part even when using "traditional" methods, even e.g. in very traditional approaches to reconstruction or etymology.] In any case, such formulas quickly become hopelessly complex -- and then they criticize Don for not tackling more than TWO languages! And another thing you quickly learn when you work in statistics/mathematical methods — whatever formula is eventually constructed, most especially if it looks complicated, will be in general met with one of two reactions. People may be very impressed and immediately convinced by whatever you have to say, but much more likely — it won't convince people at all. Any reasonable formula, with any pretense of accuracy, will be so complicated that skeptics will take it as hocus pocus, obfuscation from "the other side". — Mary Clayton, 1993, *Language* 69: 604, in another context (review of an NWAV volume): "[S]tatistical methods can be seductive. They always produce an answer, leaving even the naive or dull of mind with a feeling of accomplishment. Whether that answer is valid, important, or relevant lies in the skill of the linguist — not only in one's prowess in manipulating numbers, but even more in the knowledge, insight, and imagination that one brings to the initial formulation of questions and to the interpretation of the resulting data".

Or perhaps, in the words of Henry Clay (US statesman and orator, 1777-1852), "Statistics are no substitute for judgment".

Of course, after hearing all this from me, you will probably wonder just why it is that *I* persist in doing mathematical methods! Am I particularly stubborn, or thick-skinned? I'm sure there are those that would accuse me of that, but I go back to my first reason, the one that I offered you earlier as the reason for doing statistical tests at all — the fact that it sharpens up your assumptions, goals, etc, and your general approach to the problem at hand. And I would say that our papers here tonight, whatever else you might think of them in general or in detail, have at least done that admirably. And I will come back to another use of statistical methods later ...

Another point — one which is relevant to any statistical application in linguistics, but has special relevance to some of our topics tonight. It's hardly a new idea, and you would find it in any basic statistics course, but people sometimes lose sight of it, maybe just in the heat of the argument, so I'll say it again here. Suppose you have:

A is related to A' with .99 probability

B is related to B' with .99 probability

C is related to C' with .99 probability, and so on.

Each of these individually looks pretty secure, at 99% certainty. But note that the probability that A is related to A' AND B is related to B' is .99 x .99 = .9801. And the probability that A is related to A' AND B is related to B' AND C is related to C' is .99 x .99 x .99 = .9703, etc. With 4 such relationships, the probability of all 4 being correct is .9606, with 5 .951 and so on, with 11, .8953. So the important point, in our context tonight, is that as you increase the size of the number of hypotheses, you get TWO things — both the additional weight/security of a "package" of hypotheses, but concomitantly an increase in the chance of any one item (or even more than one item) being wrong. It is absolutely important to recognize the difference between rejecting the package and rejecting one item. It's also important to note that you don't know WHICH one (or more than one) item without a <u>careful</u> <u>painstaking</u> examination of the data.

To speak now just briefly about one point of more specific relevance to diachronic linguistics and to the reconstruction of trees/relationships ... Languages change over time — no matter what one wants to say about rates of changes or types of change, they CHANGE, as an undeniable fact.

There is a progressive loss of the data that we depend on for reconstruction. So at large time depths, it is absolutely inevitable that one is going to have to be dealing with residues (of the pre-existing similarity) that are so small that "chance" and "borrowing" and "universals" and any other "non-genetic" factors that anybody can think of are going to loom large, be very important — possibly even dominant. We have seen tonight some important attempts to overcome this problem (e.g. Nichols), but no matter what, it still comes down to making the best of what is in effect a bad situation, trying to detect the signal amongst the considerable noise, by using the most sophisticated techniques available to us. Or as Ringe put it — "Reality is intractable. Get used to it".

To go back for a moment to another use for statistical methods ... They can be good, especially in huge uncharted fields with a wealth of data, for hypothesis generation and/or hypothesis testing. As a general claim, made by many others, I would endorse that for mathematical methods in general in linguistics, whether it be historical linguistics, dialectology, stylostatistics, etc. They are good for generating <u>provisional</u> hypotheses, PENDING the results of full-scale painstaking investigation by traditional and more detailed methods. They are NOT a quick and easy short-cut to a FINAL result. Thus whatever you may think of Multilateral Comparison, it can only produce provisional results, in my view. Bender already said this, by saying it's "not really a method of doing genetic language classification", "it is a pre-theoretical step preceding Comparative/Historical ... Reconstruction, which is the real method". If our mathematical methods are good, those interim results should be good, and may eventually be shown to agree with the final consensus (if there ever is such a thing). But if they're bad, the interim results will probably also be bad (unless of course your data are so robust, the trends are so strong, that no matter what lousy method you use, the results will come out right anyway). [There were allusions to this in Bender's comments on Greenberg's Nilo-Saharan work, which as I've said before, I'm in no position to judge. And actually also in Nichol's comment at one point "Good evidence, in short, is very robust".] To return to the production of interim results and hypothesis generation... What is the harm in this? Isn't hypothesis generation <u>always</u> good? Can there be harm in this? Well, yes, there <u>can</u> be harm. Bad interim results can end up diverting a lot of research time that might have been better spent in other pursuits, not barking up the wrong tree. And another way in which interim results can be harmful is that it may prematurely generate a "received" view — which then means that "the truth" will have an even harder time getting itself established, because it will first have to combat this false "received" view. And I would like to emphasize that from this point of view statistical methods and the conclusions reached by statistical methods are no different from any others in linguistics — methods are always open to improvement, and conclusions reached by any method, statistical or not, are ALWAYS provisional, subject to revision in the light of further evidence.

Another, perhaps more philosophical point — statistics and probability will probably never outright convince anybody anyway. (Why else would people still buy lottery tickets?) Suppose I tell you that language A is 95% certain to be related to language B. That might be good enough for some of you, perhaps many of you. Anything that's 95% likely can probably be shown without statistics anyway, by the way ... But those of you who for whatever reason don't want language A to be related to language B, or maybe are just very conservative by nature, will point to the 5% probability that I am wrong, and prefer not to accept the relationship. And supposing I move the cut-off to 99% or even 99.9% — it's not likely to have any real effect anyway, on those who for whatever reason, valid or invalid, don't want to be convinced. Statistics is unlikely to ever be able to PROVE anything in the real sense (the sense in the rest of mathematics!) of PROVE — which would require 100%. So what do we mean by "establishing proof", anyway? "Beyond a reasonable doubt"? But then we are back to fighting over just how much doubt can be allowed, or 5% vs. 1% vs. .1%, etc. Maybe instead of chasing elusive proofs we should instead look towards establishing the hypothesis which, at this particular point in our investigations and our state of knowledge in general, most fully satisfies as many criteria as possible. Maybe we should start using phrases such as "count as evidence for" and "count as evidence against" rather than the absolute terms "prove" and "disprove". Or, looked at another way, maybe we need a third possibility, besides decreeing languages either "related" or "not related" — we need a category for "possible, or promising, but not proven". Compare the Scottish legal system, which allows verdicts of "guilty", "not proven", and "not guilty" — so with "not proven" in addition to the more familiar "guilty" and "not guilty". We would need something similar for more distant comparisons — at the moment we have to say either "yes" or "no", and don't seem to allow ourselves to say "maybe", or "this looks promising, and bears further investigating". Compare also Raimo Anttila's analogy to medicine — doctors don't just attempt to treat patients who are (almost certain) to recover. Similarly, our methodology should be applicable to any problem, and should produce some sort of prognosis, not just be able to deal with the cases that are most certainly related or most certainly unrelated.

I would just like to conclude with a plea for open minds — that open discussion from both sides should ideally continue until a consensus is reached.

****

Useful information and quotes, in case any need for this arises in the discussion

— Justeson & Stephens (1980) give distributions for statistical assessment of apparent resemblances (multiple phonetic and semantic resemblances). Still doesn't/can't adequately treat phonological side (inadequate on inventories, frequencies, let alone phonotactic/positional questions), let alone semantic side.

"As the criteria for phonetic resemblance are weakened it becomes more likely that a single form on one list will resemble more than one on another. This obviously increases the probability of getting a chance cognate for that item, and the expected number of chance cognates rises accordingly. The same argument holds as criteria for semantic agreement are relaxed. In both cases, multiple resemblances alter the combinatorial model" (42). "Thus we can expect the number of chance cognates to increase approximately in proportion to the average size of the similarity sets" (43). "[this paper] quantifies the dramatic decrease in the likelihood of chance cognation under mass comparison and its rapid increase when criteria for phonetic or semantic similarity are weakened to the point that many items on one list are similar to more than one on another" (45).

— David Sankoff 1973 (in Sebeok, p. 95): "Swadesh himself repeatedly indicated that he considered these methods additions, not replacements, with regard to other methods of historical linguistics, and that interpretation of a particular case should always use all lines of evidence available".

— Starostin's 35-word list, due to Yakontov (according to Laurent Sagart and Bill Baxter), has the following meanings: "blood", "bone", "die", "dog", "ear", "egg", "eye", "fire", "fish", "full", "give", "hand", "horn", "I", "know", "louse", "moon", "name", "new", "nose", "one", "salt", "stone", "sun", "tail", "this", "thou", "tongue", "tooth", "two", "water", "what", "who", "wind", "year". He claims that there are statistical limits to borrowing within the basic vocabulary, and that therefore genetic relationships can be deduced from basic vocabulary retention. Dolgopolsky was using 15 (see Shevoroshkin & Markey 1986), in order of decreasing stability, based on 140 languages: "1st person marker", "two", "2nd person marker", "who/what", "tongue", "name", "eye", "heart", "tooth", "verbal negative", "finger/toe nail", "louse", "tear [noun]", "water", "dead". Dryer used 20, at least in 1987: "I", "you sg.", "who", "two", "three", "not", "arm", "hand", "eye", "ear", "tooth", "blood", "brother", "sun", "moon", "night", "water", "die", "drink", "see".

— Comparative method in syntax.

—No syntactic analogue to the regularity of sound change. Some sentences are actually stored, e.g. proverbs and idioms, and these often show syntactic archaisms. Also, earlier syntax often survives in fossilized form in later morphology, we have another rich source of data for diachronic syntax.

—Problems with non-independence of "features". Also, how many features is e.g. SVO — relative order of subject and verb; relative order of verb and object; relative order of subject and object where necessary to disambiguate.

—Lack of "tertium comparationis" ("basis for comparison"). Cf. phonology, where you can compare Greek pater, pod-, with English father, foot, because these pairs have the same

MEANING. Since "the sign is arbitrary", it's unlikely the p-f correspondence in initial position could be due to chance. Basic word order — only 6 possibilities. Sharing a rare syntactic trait (e.g. postposed articles in Romanian, Bulgarian, Scandinavian) no proof of genetic relationship.

— Relative stability of different parts of the grammar. Jacques Guy says "From my experience with languages of Vanuatu, morphological paradigms are the LEAST stable features, followed by phonology, then, most stable, lexical."

— Sally Thomason "Structures do get borrowed, sometimes. So, for instance, there is general agreement that the Tanzanian language Ma'a (also called Mbugu) was not originally a Bantu language ... dramatically mixed structure ... it has few structural features that are clearly of Cushitic origin, and it has an entire inflectional morphology (as well as other features) adopted wholesale from Bantu languages ... One of these features is the irregular negative + 1sg prefix, which (as in some Bantu languages) contrasts with other members of the negative paradigm, which have separate negative and person/number prefixes. This is just the same type of feature that Teeter cites as obvious evidence of the relationship between (say) German & Latin." cf. Nichol's "individual-identifying evidence".

BIBLIOGRAPHY

Anttila, Raimo & Sheila Embleton. 1988. Review of Vitalij V. Shevoroshkin & T. L. Markey, eds. *Typology, Relationship and Time*, 1986. *Canadian Journal of Linguistics* 33: 79-89.

Hamp, Eric P. 1992. "On Misusing Similarity". *Explanation in Historical Linguistics*, ed. by Garry W. Davis & Gregory K. Iverson. Amsterdam & Philadelphia: John Benjamins. pages 95-103.

Justeson, John S. & Laurence D. Stephens. 1980. "Chance Cognation: A probabilistic model and decision procedure for historical inference". *Proceedings of the Fourth International Conference on Historical Linguistics*, ed. by Elizabeth Closs Traugott, Rebecca Labrum & Susan Shepherd. Amsterdam & Philadelphia: John Benjamins. pages 37-46(?).

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago & London: Univ. of Chicago Press.

Ruhlen, Merritt. 1994. *On the Origin of Languages: Studies in linguistic taxonomy*. Stanford: Stanford Univ Press.

Salmons, Joe. 1992. "A Look at the Data for a Global Etymology: **tik* 'finger'". *Explanation in Historical Linguistics*, ed. by Garry W. Davis & Gregory K. Iverson. Amsterdam & Philadelphia: John Benjamins. pages 207-228.

FL024275 LSA 1995

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Comments at Symposium "Against Multilateral Comparisons"	
Author(s): Sheila Embleton	
Corporate Source: none	Publication Date: January 5/1995

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☑ Sample sticker to be affixed to document

Check here
Permitting microfiche (4" x 6" film), paper copy, electronic, and optical media reproduction.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Level 1

Sample sticker to be affixed to document ☐

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Level 2

or here
Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Sheila Embleton	Position: Associate Professor (Linguistics)
Printed Name: Sheila Embleton	Organization: York University
Address: South 929 Ross Bldg. York University, 4700 Keele St North York, Ontario, CANADA M3J 1P3	Telephone Number: (416) 736-5260
	Date: May 19, 1995

OVER

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information reguarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:	
Address:	
Price Per Copy:	Quantity Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name and address of current copyright/reproduction rights holder:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on
Languages & Linguistics
1118 22nd Street NW
Washington, D.C. 20037