Contact-induced Change, Genetic Relationship, and Scales of Comparison [1]

James A. Matisoff
University of California, Berkeley

1.0 Perceived similarities and scale of comparison

At relatively shallow time-depths, "microlinguistic" comparative reconstruction is possible, even in the absence of extensive written records, as long as one is dealing with a well-ramified family with surviving members in several branches. Regularity of sound correspondences can be insisted upon (even for vowels!), and exceptions to phonological rules or semantic discrepancies can be explained to everyone's satisfaction. This happy state is familiar to specialists in Tai, Loloish, or Bantu -- and a fortiori to Romance philologists.

Extensive written records and morphological complexity (as in Indo-European or Semitic) and/or a large number of highly diversified daughter languages (as in Austronesian, Tibeto-Burman, or Austroasiatic) permit "macrolinguistic" work, enabling us to push back the clock to perhaps 6000 years B.P. At this level there are many unsolved and perhaps insoluble problems, though the basic validity of the family grouping is not in serious question.

At remoter time-depths, the classic distinction between genetic and other types of relationship breaks down. Too many alternative explanations for perceived similarities are possible: chance, borrowing, areal typological convergence, universal tendencies, faulty analysis, wishful thinking. Virtually every possible genetic supergrouping of the major language families of East and Southeast Asia has been proposed. Partisans of the opposing theories talk past each other, fixating on different scraps of evidence. Usually the problem gets posed in a simplistic, "all-or-nothing" way. The long-despised concept of the Mischsprache needs to be revitalized. At the megalo-level, language superstocks are as nebulous a construct as biological "races" of humanity. Notational devices, semantic promiscuity, and analytical sleight-of-hand can make any two language families look related.

All scales of linguistic comparison are legitimate, as long as one realizes that the rules of the game are quite different at the megalo-level. In the "megalocomparative" realm, it ill behooves anyone to be dogmatic.

1.1 Types of relationship and degrees of perceived similarity

Degrees of organicity of relationship	***Degrees of surface similarity***
close genetic	close genetic
remote genetic	heavy borrowing
substratum	remote genetic
borrowing	substratum
chance	chance

[1] *This preliminary draft is merely an extended outline, not yet a finished paper, and is not suitable for citation in its present form.*

But this is very crude. Consider, e.g. 'eye' Latin **oculus** / Modern Greek **mati**, both from Proto-Indo-European ***okw-** (> Gk. **op-mn-ti-on** > **om-mati-on**). An original morphological element has become the root-initial in Mod.Gk., and all trace of the root-morpheme has disappeared.

Yet we wouldn't want *every* etymology to be like this, with Benedictine "split cognates"![2] Much independent evidence is required in extreme cases of dissimilarity of reflexes.

2.0 Theoretical issues in establishing genetic relationship

•Relative weighting of areas of structure
--The notion of "core vocabulary" (Diffloth's "scruff and smegma" critique [1990]; core vocabulary replaceable by taboo in Nicobarese and Aslian [Diffloth, p.c. 1985]); only cute and/or culturally interesting words preserved in "deep substratum" of children of shifting bilinguals (below 4.0).
--Idiosyncratic morphological features; special difficulties presented by languages without much morphology.

•Glottochronology
Why must we assume that languages replace their basic vocabulary at a universally constant rate? (see JAM 1978, esp. the discussion of "replacement tolerance quotients".)

•Semantic latitude
--Practicing what one preaches: Vovin 1993:1: "...the abundance of comparisons of the type 'water' - 'sap' over the type of 'water' - 'water' seriously diminishes the credibility of any hypothesis of genetic relationship'. Then proceeds to compare Proto-Japanese ***momo** 'peach' to Proto-Manchu-Tungus ***ñang-ta** 'nut'.
--PKB 1990: Indonesian **ikan** 'fish' / Jse **ika** 'squid'
--Sino-Austronesianists: Sagart (1990) compares PAN ***pusuq** 'heart; central leaf' with OC ***swia** (re-reconstructed ***s-j-wa?**) 'marrow' (since marrow is "the heart of a bone"). Yet there's no evidence that marrow has ever been conceived in a "heartlike" way by East Asian peoples. MARROW <--> BRAIN yes, MARROW <--> HEART no! The reality of "areal semantics." (See JAM 1978 ("VSTB"), passim.)
--Semantic associations are unpredictable, but unusual ones require ample evidence.
--It should not automatically be assumed that semantic associations attested in one linguistic area are universally valid.
--Once a semantic association has already been established on independent grounds within a linguistic area, similar associations found elsewhere have confirmatory force. (See JAM 1988)

•"Regularity" of correspondence (see JAM 1994)
--Making the megalo- look micro-:
abuse of notational devices; ad hoc explanations; pushing the FAMILY TREE METAPHOR too far back, trying to get to the original Greenbergian trunk;[3] failure to recognize that at a certain point of remoteness of genetic relationship the "laws of

[2]By the way that **-mn-** infix looks very Mon-Khmer! See below 7.0.
[3]A better visual image for the linguistic relationships of SEA might be a *thicket*; or even the tangled interconnected Moebius-band like horns of a Dr. Seuss animal.

nature" change, just as the "laws" of Newtonian physics no longer apply at the subatomic or intergalactic level.
--Not enough to set up "tables of correspondences" without presenting all the data *that either confirm or disconfirm* the fillers of the cells in the table.
--Chimera of proto-regularity, when every observable modern language is rife with irregularities. Every modern language is a proto-language with respect to the future" (M.R. Haas).

3.0 Borrowing

•Importance of **social factors** in contact-induced change. "Intimate" vs. "vertical" borrowing.
Is there such a thing as a "normal/average/normative" amount of borrowing between a random pair of contiguous languages? If there were such a normative range, one could then speak of abnormal (i.e. abnormally great or small) amount of interinfluence.
--too many variables:
••amount of physical and cultural contact (natural barriers, trade, marriage, relative prestige)
••closeness of relationship of the lgs, genetically and/or typologically

•typologically favored borrowing, e.g. borrowing between closely related languages or dialects
(exs. from Thomason/Kaufman (1988; henceforth "T/K"):
••Norse --> northern Old English
••Cakchiquel --> Quiché: rule that neutralizes final /m/ and /n/ as [≥].
••Mamean --> Quiché: dissimilation rule that palatalizes velar stops /k, k'/ before a vowel followed by /q/, /q'/, or /x/.
••Marathi --> Bombay Hindi: question particle switched from sentence-initial to sentence-final position

Types of interference
Widespread belief that syntax is in some way the "deepest" level of the grammar. "While there may be some aspects of a lg's syntax which, because of internal structural cohesion, are especially resistant to foreign interference...*syntactic interference is as common as phonological interference.*"[4] (T/K:118)

"Structural borrowing"
Question on my Ph.D. orals (1964): "Is structural borrowing possible?"
I gave ex. of Chinese syntactic influence on Japanese - e.g. compounding of nouns by simple juxtaposition; and of -able/-ible added to native English roots.

3.1 Borrowing scale / scale of intensity of contact (T/K:74-6)

(1) Casual contact: lexical borrowing only
content words; non-basic vocabulary borrowed before basic

[4]Exceptions: where educated speakers shift to, or use as a second lg, a major international literary lg such as English. In such cases we sometimes find (as in India or Africa) a local variety of the international lg that is phonologically, but not necessarily morphosyntactically, influenced by the indigenous substratum (e.g. unaspirated initial voiceless consonants in Indian English, trilled r, lack of v/w contrast, realization of Eng. interdental fricatives as dental stops, contrasting with realization of Eng. /t d/ as retroflex stops; characteristic intonation); monophthongization of Eng. vowels in African English, etc.) (T/K:119, 129)

•*Casual borrowing* donor language not well known by learner
(a) physical proximity
--Farangs resident in Thailand learning a few Thai words (e.g. fruit names, *tuk-tuk*, "samlor")
(b) at a distance
--U.S. intellectuals using French or German phrases (*déjà vu* all over again; *Zeitgeist*)
--How many Americans realize that *jihad* 'holy war' and *mujahadin* 'Afghan rebels' come from the same Semitic triliteral root? Or that Russian *glasnostj* 'openness' and Polish *solidarność* 'solidarity' reflect the same Slavic suffixes? (Cf. Khmer infixation processes faintly perceivable in Thai, e.g. **trùat** 'examine' / **tamrùat** 'police".)

Differences between British and American borrowings from French:
--stress shift to first syllable in Britain: *ballet, pâté* (T/K:349)
--[æ] in Britain / [a:] in US: *cantata* (JAM)
Sociolinguistic cause: the British are more secure in their own pronunciations, so they anglicize more?

•Special cases: borrowing from a high prestige written language (may be well known to an elite)
--Japanese < Chinese; Urdu < Arabic; Hindi < Sanskrit; Thai < Khmer; Yiddish < Hebrew

(2) Slightly more intense contact: slight structural borrowing
lexicon: function words - conjunctions and various adverbial particles
structure: minor phonological, syntactic, and lexical semantic features
maybe new phonemes, but only in loanwords
syntactic borrowings restricted to new functions or functional restrictions, or new orderings; little or no typological disruption
/e.g.Slavic influence on Yiddish reflexive pronouns (became invariant instead of inflected)/

(3) More intense contact: slightly more structural borrowing
lexicon: function words (adpositions); derivational affixes may be abstracted from borrowed words and added to native vocabulary; inflectional affixes may enter, but confined to loanwords; personal and demonstrative pronouns and low numerals more likely to be borrowed
structure: phonemicization, even in native vocabulary of previously allophonic alternations (esp. those that exploit distinctive features already present in the borrowing language); acquisition of salient prosodic features (e.g. stress rules) and syllable-structure features (e.g. addition of final consonants in loanwords); some aspects of word-order shift may be found, e.g. borrowed postpositions in an otherwise prepositional lg.

(4) Strong cultural pressure: moderate structural borrowing
structure: major structural features that cause relatively little typological change; introduction of new distinctive features into native vocabulary, and perhaps loss of some phonological contrasts; new syllable structure constraints, also in native words; acquisition of a few natural allophonic and mppc rules, such as palatalization or final obstruent devoicing
fairly extensive word order changes; other syntactic changes that cause little categorial alteration; borrowed inflectional affixes and categories (e.g. new cases)