Contact-induced Change, Genetic Relationship, and Scales of Comparison

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 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

<table>
<thead>
<tr>
<th>Degrees of organicity of relationship</th>
<th>Degrees of surface similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>close genetic</td>
<td>close genetic</td>
</tr>
<tr>
<td>remote genetic</td>
<td>heavy borrowing</td>
</tr>
<tr>
<td>substratum</td>
<td>remote genetic</td>
</tr>
<tr>
<td>borrowing</td>
<td>substratum</td>
</tr>
<tr>
<td>chance</td>
<td>chance</td>
</tr>
</tbody>
</table>

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"\(^{12}\). 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 *nang-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 [ŋ].
• 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." (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

4Exceptions: 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 glasnost '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 / [aː] 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)
added to native words, esp. if there is a good typological fit in both category and ordering

(5) **Very strong cultural pressure: heavy structural borrowing**
**structure**: major structural features that cause significant typological disruption; added mppc rules; subphonemic changes in habits of articulation (change in basis of articulation), incl. allophonic alternations; loss of phonemic contrasts and of mppc rules; changes in word structure rules (e.g. adding prefixes in a lg that had been exclusively suffixing; or change from flexional toward agglutinative morphology); categorial as well as more extensive ordering changes in morphosyntax (e.g. development of ergativity; added concord rules, including bound pronominal elements)**

3.11 *Grading SEA'n contact situations in terms of these criteria*

**Type (5) cases:**
Extremely intense contact resulting in typological change. No aspect of structure is immune to borrowing under these conditions.

- Karen (< Mon, Tai)
- Vietnamese (< Chinese)
- Munda (< Indo-Aryan)
- Bai (< Chinese)
- Cham (Hainan: < Chinese; Cambodia/Vietnam: < Mon-Khmer)
- Tai (---< Chinese)
- Hmong-Mien (--- Chinese)
- Japanese (Mischsprache; off the scale)

**Type (4) cases:**
- Thai ---< Khmer
- Hayu (< Nepali)
- Kelantan Chinese (< Malay) [Teo Kok Seong 1993]
- Newari (< Indo-Aryan)

**Type (3) cases:**
- Mon --> Burmese
- Malay --> Aslian

**Type (2) cases:**
- Shan --> Lahu
- Chinese --> Lahu
- Thai --> Vietnamese
- Tai --> Khmuic
- Thai ---< Malay
- Thai --> Kelantan Chinese

**Type (1) cases:**
- MK --> Lepcha [substratal?]
- Loloish --> Palaung-Wa [adstratal]

---

5Cf. the ongoing controversy about the genetic antiquity of TB "verb pronominalization".
6For a somewhat analogous case cf. Chinook Jargon (Silverstein 1972), where the tense/aspect systems of different dialects have diverged under the influence of different outside languages. (T/K 17)
4.0 Borrowing vs. shift

Shift: target language imperfectly known by learner
(a) shift caused by influx from outside; i.e. no displacement of population that must shift
--Southern Min speakers learning Guoyu in Taiwan (older population has lower prestige)
(b) shift by a displaced population
--Early 20th c. Yiddish-speaking immigrants to US (older population has higher prestige)

Degrees of interference in bilinguals' languages (T/K:40)

<table>
<thead>
<tr>
<th></th>
<th>English -&gt; Yiddish</th>
<th>Yiddish -&gt; English</th>
</tr>
</thead>
<tbody>
<tr>
<td>borrowing</td>
<td>lexicon</td>
<td>moderate</td>
</tr>
<tr>
<td></td>
<td>phonology</td>
<td>strong</td>
</tr>
<tr>
<td></td>
<td>morphosyntax</td>
<td>strong</td>
</tr>
</tbody>
</table>

Viewed vertically, intergenerationally:

Yiddish adstratum /shift to English/ Yiddish adstratum

Dad b. 1910, Zembin, Belarus

JAM b. 1937, Boston

substratum

Nadja b. 1964, Berkeley

deep substratum

4.1 The notion of substratum

• Mistrust of unfettered substratal proposals
--Indo-Europeanists have a traditional prejudice against external explanations for change...especially when all the proposed interference takes place without actual transfer of gcl morphemes. (T/K:139-43)
• Abuse of this concept has given substratal advocates a bad name: "Substratomania" (Derek Bickerton).
• The worst cases are those where we have no structural info either about the abandoned lg or about the target lg before the shift. In order to make educated guesses we must:
--be able to identify a substratum lg (or lg group) some of whose speakers shifted to the target lg at the relevant time period;
--have information about structure of substratum lg;
--have information about structure of target lg before the shift.
"These methodological prerequisites have frequently been ignored by substratum enthusiasts, and this fact prob. accounts in large part for the widespread suspicion
with which historical linguists tend to view substratum explanations of lg change." (T/K:111)

• Emeneau 1980, quoted in T/K 350: "The only really valid evidence [of structural interference] is that derived from bilingual situations in which the lgs on both sides are well known. It will not do to deal in substrata that have long vanished entirely from our control. That way lie uncontrolled hypotheses, alias guesses, and the mysticism of so many substratum theories." T/K agree, but stress that the two languages may be well known indirectly (by comparative reconstruction) rather than directly.

• JAM 1972: "The Mon-Khmer substratum in TB." Much too grandiose a title: dealt with the *k- animal prefix in Lolo-Burmese, which has a convincing MK etymology. But this is merely a case of Type (3) contact.

• Superstratum and substratum

There seem to be two aspects to these concepts:

--relative prestige (synchronic)

the higher prestige one is the superstratum; the lower one is the "basolect", which later gets attenuated into a substratum.

--relative chronology (diachronic)

the older element is the substratum

• Superstratum, substratum, adstratum

Superstratum lgs are typically those of victorious invaders (who might then shift to the lg of the conquered people (e.g. Norman French in England).

Substratum lgs are those of conquered, or at least sociopolitically subordinate, indigenous populations and immigrants.

--Afrikaans is historically a descendant of Dutch, but its development into a separate lg was heavily conditioned by nonwhites who learned Dutch imperfectly as a second language (T/K:256).

--No Mongolian superstrate in Xixia (> Qiangic). The nomadic Mongols didn’t stick around long enough.

Adstratum lgs are those of invaded or invader groups that are neither dominant nor subordinate in the contact situation.

"Only one significant lgc consequence seems to depend on the super-/sub-/ad-
distinction: superstratum interference is more likely to include lexical items than substratum interference is." (T/K:116) (E.g. masses of [superstratal] French loanwords in English,7 vs. paucity of [substratal] Dravidian loanwords in Indo-Aryan or of [substratal] old Uralic loanwords in Slavic.)

"The traditional super-/ad-/substratum distinction is of limited usefulness for the interpretation of most past shift situations." (T/K:118) "We will...avoid the term 'adstratum', since its linguistic effects are not obviously distinguishable from those of superstrate and substrata."

• Cases where multiple explanations are possible, even in Indo-European:

Maybe a pre-IE substratum "accounts for" the strange Celtic lenition, as well as Celtic incorporated pronominal objects. On the other hand, maybe there is an ancient contact relationship with Hamito-Semitic (see e.g. Adams 1975) [T/K:111].

4.11 Convincing substratum/shift cases

7These are different from French loanwords in, e.g. Russian, Danish, or Dutch, which are strictly due to borrowing and not to shift.
•Dravidian ---> Indo-Aryan. A very strong case can be made for this. Dravidian speakers were once much more widespread in northern India than they are now, and the lgc map of India -- with a solid Dravidian-speaking south (except for intrusive Sinhalese) and Dravidian islands surrounded by Indic in the north -- supports this hypothesis.

Dravidian substratal interference on IA: the Indic quotative iti is a native morpheme, though its use to mark direct quotations is attributable to Dravidian substratal interference. (T/K 131).

•Uralic ---> Balto-Slavic.
Relic speech islands of Uralic lgs surrounded by Slavic.
substratal interference --> new cases in Russian and Lithuanian (T/K:114)
"Overall we believe that the case for old Uralic interference in Slavic is fairly strong, though fewer interference features have been identified here than in Indic from Dravidian." (T/K:145)

•Burushaski substratal influence on Shina (Dardic) (T/K:135-9).

4.12 Borrowing-cum-shift
•Cushitic influence on Ethiopian Semitic (Leslau 1945). But T/K feel that there has been so much adoption of words from the source lgs (Cushitic) that borrowing and shift interference must have been occurring simultaneously. (115) /Cushitic and Semitic are both descended from Proto-Afro-Asiatic/
•Borrowing by English speakers from Norman French, simultaneously with the shifting of those French speakers to English.

•Uralic/Balto-Slavic
Arose partly through shift by Uralic speakers to Slavic and Baltic, and partly through borrowing by Uralic from Slavic and Baltic. (I.e. for some speakers, Uralic substratum interfering with acquisition of B-S (old language interferes with new language); for other speakers, incorporation of B-S (superstratal or adstratal) elements into Uralic itself (new language interferes with retained old language)

4.13 Borrowing from superstratum: donor language well known by learner
--English speakers learning Norman French, but never shifting totally; massive lexical influx.
--Indonesian speakers learning Dutch; seems to have contributed to the growing reinterpretation of the di- focus marker as a Western style "passive voice".

4.14 Continuum of interference from the shifting language onto the target language

The relative sizes of the shifting and target language (TL) speaker groups offer a fairly reliable criterion for predicting whether or not structural interference will occur. "We know of no clear case in which a relatively large invader group has shifted to the language of the indigenous population." (T/K 122)

4.141 Shift without interference (T/K:119-21)
•Typical shift sans interference: urban immigrant groups of European origin in U.S.
•Shift of small superstratal immigrant group - little structural influence on local lg, but likely to leave numerous loanwords (Gothic and Burgundian [Gmc] infl. on medieval Romance dialects of southern France and Spain)..[Contrast to Frankish influence in N. France, where the Germanic presence was larger and lasted longer.]
• North American Indian languages -- shift sans interference -- overwhelming numerical and military preponderance of majority lg.

4.142 Shift with slight interference (T/K:121-9)
• Only one phonological innovation in English that is both non-simplificatory and not in line with internal developmental tendencies can be ascribed to French influence: the "Romance stress rule" (Chomsky and Halle 1968) - but even here confined to words of Latin and French origin.
• Frankish influence in N. France, where the Germanic presence was larger and lasted longer -- thus at least some minor structural interference from Frankish on northern Gallo-Romance. Franks arrived in northern France ca. 486, and did not shift to Gallo-Romance until ca. 850. Few undisputed Frankish structural interference features in French, though numerous Frankish loanwords. Romance lgs in general lack /h/ and initial /w/. But Modern French dialects of Picardy, Wallonia, and Normandy still have initial /h/ in Gmc loanwords. "h aspiré" in Gmc loans blocks liaison. /w/ appears word-initially in Picardy, Wallonia, and Lorraine. In other French dialects, this [w] was strengthened to [gw], then simplified to [g]. A few native Romance words acquired /h/ and /w/ thru Frankish influence: haut > haut 'high' (VL altus, influenced by Frankish h»oh); OF waste 'waste' (VL vastus, infl. by Frankish wöstî); OF waine 'sheath' (VL vagina) [mod. Fr. gaine].
• Two French syntactic features stimes claimed to be due to Gmc influence:
  --OF had an obligatory verb-second rule in prose, unknown elsewhere in Romance, but characteristic of Germanic.
  --Use of BE instead of HAVE as the perfect auxiliary with certain intransitive verbs of motion and change of state. (Rumanian, Spanish, and Portuguese use only HAVE; Italian follows the French pattern, but this feature was undoubtedly borrowed from French, probably during the 12th century.)

4.143 Shift with moderate to heavy interference (T/K 129-46)
Three different categories:
(1) Simplificatory changes resulting from learners' failure to learn TL structures that are more complicated than functionally corresponding structures in the shifting speakers' original lg.
  --development of fixed initial stress in Latvian under Livonian influence
  --JAM: what about devoicing of final stops in African American English?
(2) Changes which neither simplify nor complicate the system.
  --replacement of Semitic pharyngealized emphatic consonants by Cushitic-influenced glottalized emphatic consonants in Ethiopic Semitic. Both are highly marked types of sounds.
(3) Changes which seem to complicate the TL grammar.
  --Russian and Lithuanian added cases thru Uralic influence.
  --Irish English - development of a recent past construction using 'after' (He's after singin' = He has just sung).
  --Introduction of clicks into Zulu and other Nguni dialects of Bantu from Khoisan (Hottentot).

4.2 Austric: substratal or remote genetic relationship?

Shift vs. borrowing situations:
--If two lgs are known to have been in contact, the presence of noninherited universally marked features in one of them ...may provide good evidence of interference from the other. If the interfered-with lg has adopted few or no words
from the source language, then the mechanism of the structural interference must have been shift; if many words have been adopted, the mechanism was probably either borrowing alone or a combination of shift and borrowing... *Marked features are more likely to be transferred in borrowing than in shift.* It is easier to find evidence of interference if the mechanism was borrowing; light to moderate interference through shift, since it more often results in gc1 simplification, is harder to distinguish from internally motivated simplificatory changes.

"...though claims of structural borrowing are also hard to establish sometimes, they are in general easier to argue for than claims of interference through shift, for two main reasons: (a) in a borrowing situation both the source lg and the receiving lg are maintained, at least for a considerable period of time; (b) structural borrowing is invariably preceded by lexical borrowing, so the source lg (or lg group) will be identifiable from the loanwords." (T/K113-4)

**Application to Austric:** over 200 AA/AN comparabilia have been suggested in the literature, of which only a few are prime candidates for cognacy (WOOD, BONE, LIZARD...), but deep grammatical similarities (-um- [agent focus/nominalization] and -in- [object focus/nominalization] infixes; -p- causative; see esp. Reid 1993). So a borrowing relationship between them can be excluded. This combination (few cognates, impressive idiosyncratic morphological similarities) does not contradict a substratal interpretation, but this is disfavored by lack of evidence of hegemony of one group by the other. *Remote genetic relationship* seems like the best hypothesis.

**Genetic relationship:**
Schmidt 1906; Shorto 1976; Reid 1993; Blust 1993; Diffloth 1977, 1990, 1993
earlier Benedict: 1942

**Substratal relationship**
later Benedict: 1973, 1976

**Borrowing relationship**
still later Benedict: 1985bis, 1993

Agnostic: *not borrowing but not sure whether genetic*


5.0 **Sprachbund: directionality of multilateral interference not clear**
T/K don't discuss very much. "The reason for these omissions is that our interest in unraveling the causes, effects, and mechanisms of contact-induced language change has led us to focus on two-language contact situations in which the direction of interference can be definitely established. Sprachbund situations are notoriously messy." (T/K:95)

**Difference between "Sprachbund" and "linguistic area"**
--T/K distinguish between multilateral Sprachbünde, where directionality is unclear, and "non-multilateral well-established linguistic areas", where one can "establish

---

8"There is no longer any need, therefore, to posit a ‘substratum’ of any kind in this area (contra Benedict 1973) since ‘border phenomena’ (loans/influences) will do nicely as a general line of explanation."
both the source of interference features and the direction and mechanism of diffusion", e.g. the Dravidian > Indic interference (a case of SHIFT).
--"What a long-term multilateral Sprachbund seems to promote...is the gradual development of isomorphism (JAM: **intertranslatability**) in all areas of structure except the phonological shapes of morphemes." (T/K:96)

**New Guinea central highlands**
Capell (cited in Wurm 1956): the adjoining languages have very different vocabularies, but their grammatical features "recur with almost monotonous regularity from language to language."

**Pacific Northwest**
Jacobs 1954: areal phonological features such as glottalized phonemes, labialized dorsals, uvulars, and lateralized obstruents
Thomason 1983 ("Chinook jargon"): areal morphosyntactic features such as yes/no question markers, possessive affixes, and periphrastic imperative constructions

**Balkan Sprachbund**
"...the long period of widespread back-and-forth migrations of small groups, caused by the Turkish invasions, resulted in mutual bilingualism and multilingualism rather than the one-way bilingualism that is common...in two-language situations." (T/K:95)
(Turkish, Greek, Albanian, Bulgarian, Macedonian, Romanian, Hungarian, Serbo-Croatian, Czech-Slovak, Romany, Yiddish, German)
--periphrastic future
--replacement of morphological infinitives by a specific clausal construction
--postposed articles

**Mutatis mutandis, similar to SEA area** ("widespread back-and-forth migrations of small groups" -- caused by Chinese pressure, Mongol campaigns, local confrontations).
A feature of Sprachbünde is **possible change of directionality according to the waxing and waning of extralinguistic factors**

<table>
<thead>
<tr>
<th>Earlier</th>
<th>Later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tai &gt; Chinese</td>
<td>Chinese &gt; Tai</td>
</tr>
<tr>
<td>Mon &gt; Burmese</td>
<td>Burmese &gt; Mon</td>
</tr>
<tr>
<td>Khmer &gt; Thai</td>
<td>Thai &gt; Khmer</td>
</tr>
</tbody>
</table>

**Local diffusion vs. Sprachbund**
These are to be carefully distinguished. "...Multilateral Sprachbund cases usually have only a few area-wide features, together with many instances of localized diffusion. Sometimes we can establish directionality for local diffusion, but often we cannot..." (T/K:96-7).
JAM: but every Sprachbund begins as local diffusion - an accumulation of individual cases; 'multi-local diffusion'.

**6.0 Mixed languages; Mischsprachen⁹; creoles**

"Es gibt keine Mischsprache." Max Müller 1871.1:86.
"Es gibt keine völlig ungemischte Sprache." Hugo Schuchardt 1884:5 (T/K:1)

---
⁹This controversial concept is also susceptible to abuse by what one might call "misch-ugene-linguists".
• Shift situations where the shifting speakers acquire so few of the target language's gcl structures that transmission must be considered abnormal.
• The TL as a whole is not available to the shifting group (e.g. Haitian French-based creole). But subsequently decroelization is possible, when the TL becomes accessible.

Abrupt creolization (shift without normal transmission): (T/K:147-66)
• The emerging contact language at once becomes the primary lg of the community and is learned as a first lg by children. Talmy Givón's unfunded project: artificially transporting speakers of different languages to an uninhabited island, to observe the kind of creole they would develop.

Overwhelming cultural pressure: replacement of large portions of the inherited grammar
Intense cultural pressure (so that all the pressured speakers must learn the dominant language of the community) usually leads to one of three outcomes. (Can't predict which outcome will eventuate in any given case.) (T/K 100-9)
1. rapid shift to dominant language; abandoned language dies sudden death;
2. shift may take place over many generations; subordinate language undergoes slow attrition; "language death" in the usual sense;
• "semi-speaker phenomenon", when a person knows no language well. Bloomfield's Menomini speaker named White Thunder.
• Lg death is a process involving the simplification of language form along with the restriction of linguistic function; it is the loss of domains of usage that leads to loss of stylistic resources and, ultimately, to loss of grammatical structures, as new generations of speakers fail to learn forms their elders never or rarely use. Flathead speakers can no longer produce verb forms with 1st pl. subj. + 2nd pl. obj. or vice versa. [These forms are exceptionally complex in Salishan.]
• All Flathead speakers have near-native fluency in English. But Flathead has "borrowed nothing from English except a few words, as we would expect, given that the shift to English will have been completed within a few generations from the onset of intense cultural pressure." (T/K:102)

• Where lg death results from a SLOW shift process, by contrast, structural borrowing from the dominant language is likely to occur in addition to the loss of structure characteristic of language death. (Welsh <= English; (Hayu <=--- Nepali) from Hodgson to Michaelovsky
• All English gypsies speak English natively, but maintain part of Romani lexicon as a secret code (but used with English grammar).
• Ma'a (Mbugu), a Tanzanian language with perhaps 50 percent Cushitic vocabulary.
• Mednyj Aleut: Aleut elaborate noun morphology remains intact, as well as the nonfinite verb morphology; but the entire Aleut finite verb morphology has been

---

10T/K say that this second type "Though little studied until recently, may also be common". Seems quite common to me.
replaced by Russian finite verb morphology. Russian affixes, particles, and pronouns are used with native Aleut verb roots.

*Kormakiti Arabic* (Cyprus): Home language (unwritten) is Arabic of Maronite Christians who immigrated to Cyprus in 1191; everyone completely fluent also in Cypriot Greek. About 40% Greek loanwords. But Arabic words retain full inflections, and Greek words inflect exactly as in the standard lg. Numerous syntactic patterns also borrowed from Greek.

Michif and Mednyj Aleut are odd, in that level of borrowing is not uniform for nouns and verbs. But there are good reasons: Russians who learned Aleut found the even more elaborate Aleut verb morphology beyond them; borrowed noun phrases in Michif because foreign NP's were fairly easy to fit into the verb-centered Cree sentences (T/K:107).

All these languages arose through a process of *abnormal transmission*. Their grammars and lexicons (esp. basic vocabulary), as whole entities, are not borrowed from a single historical source. Their potential value for comparative reconstruction varies, however.

The development of these lgs was non-genetic (as with *abrupt creoles* and *nativized pidgins*). But abrupt creoles and pidgins arise in multilingual settings without any group achieving full bilingualism. The above languages all arose in two-lg contact situations with full one-way bilingualism. (Unlike pidgins and creoles, the above lgs have elaborate morphological systems - though this is not essential: *doesn't occur in a mixture of two isolating languages.*) (T/K:108-9)

**Grammar/Lexicon bifurcative cases**

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ma'a (Tanzania)</td>
<td>Cushitic Bantu</td>
</tr>
<tr>
<td>Asia Minor Greek</td>
<td>Greek Turkish</td>
</tr>
<tr>
<td>Mednyj Aleut (morphology)</td>
<td>Aleut Russian (finite verb</td>
</tr>
<tr>
<td>AngloRomani</td>
<td>Romani English</td>
</tr>
<tr>
<td>Japanese</td>
<td>Austro-Tai Altaic Altaic</td>
</tr>
</tbody>
</table>

Qst: If it's a grammar/vocabulary split, how can you tell which is older? Must one always assume that it is the grammatical similarities that reflect the older stratum?

Michif (ex-dialect of Cree): grammatical split case
Retains elaborate Algonkian verb morphology, but has borrowed French noun morphology and NP syntax along with French nouns:

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns</td>
<td>Verbs</td>
</tr>
<tr>
<td>French</td>
<td>Cree</td>
</tr>
<tr>
<td>Nouns</td>
<td>Verbs</td>
</tr>
<tr>
<td>French</td>
<td>Cree</td>
</tr>
</tbody>
</table>

*The speech of shifting bilinguals may show code-switching from lg to lg within a sentence, with homogenization of lexicon, phonology, and grammar. The "semi-speaker phenomenon", where neither lg is spoken perfectly.

---

11This is an extreme version of what goes on, e.g. with English Latinate plurals (alumni, alumnae) (T/K 106).
Giving the mixed languages concept a bad name: phony "creolization"
Some authors use this term loosely -- e.g. those who consider English to have been
creolized by French or Norse (Domingue 1975, Bailey and Maroldt 1977). "Bailey
views every case of significant structural interference as creolization, whereas we
draw a line between ordinary interference and interference so extreme as to disrupt
genetic continuity." (T/K:123)
--Schmidt 1906 called Cham a Mischsprache (langue mixte in the Fr. transl.).
Haudricourt 1954 called Hmong-Mien "intermédiaire" between Austroasiatic and
Tibeto-Burman.
--Probably never more than 50,000 native French speakers in England (whose total
population was 1.5-2 million). Though French lexical influence on English was
very heavy, the lexical influence was rather light in the basic vocabulary, only about
7 pct. French speakers had all shifted to English by about 1265. (T/K:126) The
relative sizes of the shifting and target lg speaker groups offer a fairly reliable
criterion for predicting whether structural interference will occur. "We know of no
clear case in which a relatively large invader group has shifted to the language of
the indigenous population." (T/K:122) "English speakers in England never lacked a
common means of communication, so they had no reason to develop a pidgin or
creole." (T/K:125)
JAM: Perfectly possible to write English without French loanwords (Tolkien), but
absolutely impossible to write an extended passage in natural English containing
ONLY French loans.

6.1 The affiliations of Japanese

Grammar:
Here there are striking similarities with Korean/Altai:
--- verb-final/postpositional; complex but agglutinative verbal morphology;
traces of vowel harmony12; extensive use of onomatopoeia (giseigo/gitaigo);
same overall typology; high degree of "intertranslatability" between Japanese and
Korean

Lexicon:
Detailed comparison necessary of suggested AN vs. Altaic etymologies for a given
Presupposes a solid reconstruction on the Altaic side. Many difficult cases will
doubtless remain a matter of opinion, esp. as concerns semantic latitude.
--Perhaps AN and Altaic words of accidentally similar sound/meaning
contaminated and reinforced each other. Many (often hilarious) examples can be
cited from the daily life of bilingual households.

--Korean/Altai: Martin 1966; Miller 1971; Russian scholars (Starostin, Peyros,
Vovin, Serafim.)
--Mixture: Murayama Shichirō 1976, 1978, etc.
--And what about Ainu?

Historical scenario:
--Initial plausibility of both an AN and an Altaic population movement into Japan is
high. The distance from Taiwan to the Ryukyus and Kyushu offered little challenge

---12 Even such an apparently profound feature as vowel harmony may be borrowed, as in Asia
Minor Greek, which is under massive Turkish influence (T/K:53).
to the navigational skills of the AN peoples, who settled the world from Hawaii to Madagascar; and it's also not very far from Pusan to Fukuoka.

--Murayama (and apparently other proponents of the Mischsprache theory) conceive of the earlier or substratal layer as AN, with a later overlay of Altaic (Murayama 1978:91-2, quoted in Miller 1987:647).

--This seems to contradict the received idea that grammatical similarities reflect a chronologically prior stratum.

--Yet this AN/Altaic melding is not necessarily to be conceived of being due to discrete waves of migration by invaders, but rather as a slow trickling of immigrants.

7.0 Long Rangers

Biggest problem at this level: judging among contradictory proposals. (One solution is to combine them, like Greenberg, or the latest version of Sagart's Sino-Austronesian, which now includes Austroasiatic -- i.e. in a kind of "Sino-Austric" grouping.)

Nostratic; Sino-Caucasian; Dene-Sino-Caucasian; Sino-Mayan
Pulleyblank 1990: Chinese/Indo-European

Sapir 1920, 1933: ST/Amerindian
Ohno 1980: Japanese-Tamil
Yoshiwara 1991: Japanese-Sumerian

The role of chance at the megalо- or hyper-megalо-level. Some linguists are now trying to quantify the probability of chance resemblances between any two languages or proto-languages taken at random (see e.g. Ringe 1992, 1994, 1994?).

Attempts at reductiones ad absurdum:
Callaghan and Miller 1962
Campbell 1988
Ffositam 1990: Sino-Tibetan/Amerind

Mistrust megalо-proposals with TOO MANY comparata/comparanda/comparabilia. A few good examples are worth hundreds of invalid ones.

The semantic latitude question becomes crucial here.

• Putative cognate identifications between monosyllabic and sesqui- or di-syllabic languages vary greatly in their persuasiveness, ranging from the obvious (Vietnamese and Muong; Cham of Hainan vs. Cham of Vietnam/Cambodia; Central Thai/Southern Thai); down through the often much less convincing Tai and Hmong-Mien comparisons with Austronesian; down to the often totally unconvincing (Chinese/Austronesian).

• Greenberg's "Indo-Pacific hypothesis" (1971) linking Andamanese, "Negrito", Papuan -- beyond our reach!

8.0 Linguistic stocks of Southeast Asia
Primitive dendrograms for Macro-Austric and Sino-Tibetan (*plura ex uno*), and a "retro-dendrogram" (*ex pluribus unum*) for Japanese:

```
Macro-Austric

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Austroasiatic</td>
<td>Austronesian</td>
<td>Tai-Kadai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hmong-Mien</td>
</tr>
</tbody>
</table>
```

```
Sino-Tibetan

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>Tibeto-Burman</td>
</tr>
</tbody>
</table>
```

- Any crossing of the Macro-Austric/ST line is taken as not due to normal genetic transmission.
- Two differences of interpretation: for PKB, the AA/AN relationship is not genetic, but at most substratal; and the AN/Japanese relationship is not substratal, but genetic.

### 9.0 Desiderata for the future

- Use of the computer. Instead of tracking down contradictory opinions (often by the same author who repeatedly changes his mind) in obscure or unpublished sources, the ideal is to have everything "sensibly" proposed concerning a particular etymon accessible in an instant for evaluation.
- "Sound law databases"

Hmong-Mien: Wang Fushi, Theraphan L. Thongkum, Ratliff, Strecker
Austronesian: Blust

---

13 After the fact I noticed that Diffloth has already used this term (1990, note 1)! Sjors van Driem has just emailed me that in an unpublished paper he is using the name "Greater Austric" to cover the same grouping here proposed. One term which will probably never catch on in Mon-Austronesian.
Mon-Khmer: Diffloth
Sino-Tibetan: the STEDT project

But blind faith in the computer is no good either! "Garbage in, garbage out!"

Selected References


Sapir, Edward. 1920. [Letters to Berthold Laufer re ST/Amerindian connections.]


THE KINSHIP TERMS USED BY
THE AHOMS OF NORTH EAST INDIA

Prof. (Mrs) RAJU BORUAH
GAUHATI UNIVERSITY
(JAGIROAD COLLEGE)
ASSAM, (INDIA)

The Ahoms of the North East India have contributed a lot in the field of language and culture of the land. The Tai people who migrated to Assam about seven hundred years back are known as Ahom/ahom/ after whose name perhaps their kingdom was named as Assam/assam/. Recently a few scholars termed them as Tai-Ahom. But in this paper the original term Ahom is used. The language of this part is also named after them as Asamiya/asamia/ i.e. Assamese, an englisigized word of Assam/assam/. This language is the lingua franca of this area of the country. Of course, Asamiya is an Indo-Aryan language.

The modern word Ahom by which the Tai-people, living in this eastern most corner of India, are known, evidently goes back to early Assamese Āsām: Āsām > Āsam > Āham, Āhom as elaborated by Dr. Banikanta Kakati, the pioneer linguist of Assam, in his scholarly work entitled, "Assamese, its Formation and Development". It is evident here that starting from the very name of the State of Assam, Tai-Ahom contribution, in the formation of Assamese language is conspicuous by its age-old affiliations.

Although Tai-Ahom language is not a major source which enriched the Assamese language with sufficient linguistic elements, it has distinctive features not commonly observed among the other communities of Assam. Such prominent terms exclusively