Celtic and Afro-Asiatic

Graham R. Isaac
(National University of Ireland, Galway)

It is not remarkable that structural similarities between the Insular Celtic and some Afro-Asiatic[1] languages continue to exert a fascination on many people. Research into any language may be enlightening with regard to the understanding of all languages, and languages that show similar features are particularly likely to provide useful information. It *is* remarkable that the structural similarities between Insular Celtic and Afro-Asiatic languages continue to be interpreted as diagnostic of some sort of special relationship between them; some sort of affinity or mutual affiliation that goes beyond the fact that they are two groups of human languages. This paper investigates again the fallacious nature of the arguments for the Afro-Asiatic/Insular Celtic contact theory (henceforth AA/IC contact theory). It takes its point of departure from Gensler (1993). That work is as yet unpublished, but has had considerable resonance. Such statements as the following indicate the importance that has been attached to the work: "After the studies of Morris-Jones, Pokorny, Wagner[2] and Gensler it seems impossible to

---

[1] Many recent discussions of the matters in hand retain the older designation 'Hamito-Semitic'. I shall use the more recent term 'Afro-Asiatic' irrespective of what term is used by a particular author whose work I might at any time be discussing. The term is imprecise, insofar as Omotic, Cushitic and Chadic, other Afro-Asiatic subfamilies, are not generally implicated in the theory in question. But with that caveat borne in mind, there is no reason to insist pedantically on a narrowing of the terminology.

[2] I.e. Morris-Jones (1900), Pokorny (1927, 1928, 1930, 1949, 1962, 1964), Wagner (1959, 1967, 1977, 1981, 1987) (selected references). To these may be added some further work on the theory by Shisha-Halevy (1995, 2000 a, 2000 b) and Vennemann (2000, 2001, 2002, 2003 a, 2003 b, 2003 c). In the present paper, I shall not be discussing the history of the theory over more than a century. A historiographical summary, together with critique, will be provided by Hewitt (fc.; cf. further Zeidler 2004; McCone 2006: 20-40). Nor shall I be addressing the differences of detail between the various presentations, differences in the features chosen as significant, differences in the mechanisms of contact or of immediate affiliations of the substrate language, etc. The base of all versions of the theory is that there is a continuum of contact between the Insular Celtic languages and the Afro-Asiatic lan-

deny the special links between Insular Celtic and Afro-Asiatic" (Jongeling 2000: 64). And the ideas in question have been propagated in the popular scientific press,[3] with the usual corollary that it is these ideas that are perceived by the interested but non-specialist public as being at the cutting edge of sound new research, when in fact they may simply be recycled ideas of a discredited theory. For these reasons it is appropriate to subject Gensler's unpublished work to detailed critique.[4] In particular, with regard to the twenty features of affinity between Insular Celtic and Afro-Asiatic which Gensler investigated, it will be shown (yet again, in some cases):

1) that several of the features are merely implicational correlates of other features, not genuinely separate features at all;

2) that the triviality of some of the features positively disallows them as being diagnostic of a special relationship (and some consideration will be given to the claim, essential to the contact theory, that a group of individually non-diagnostic features can combine to make a diagnostic ensemble of features, or: how do twenty inconclusive arguments add up to one conclusive argument?);

3) most particularly, many of the features imply positively bizarre realist interpretations:
  a) either there is a causal link between the fact of the Celtic languages gaining certain grammatical features in the tenth century A.D. and the (supposed) fact that they had been in contact with an Afro-Asiatoid language five hundred, seven hundred or over a thousand years earlier (realist comparison: 21st-century English develops a certain syntactic feature because it was in contact with Proto-Norse in the tenth century); or
  b) the Celtic languages were still in contact with that Afro-Asiatoid language in the tenth century A.D.

---

guages which is seen evidentially in their shared typological features. Nowhere are the nature of the features involved and the structure of the arguments relating to them explicated in more detail than by Gensler (1993). All detailed discussion will therefore be restricted to the facts and arguments given in the latter work. I shall assume that any criticism of that treatment is transitively applicable to any other treatment relying on, or derived from, Gensler's or using the same data and argumentative direction. If anyone would challenge that assumption, i.e. if anyone would claim that the theory can be upheld without reliance on *any* of the features discussed by Gensler, or on *any* of the implications of his discussion for the linguistic history and geography of the ancient and medieval world, let the critique of that be the task of another day.

[3] Wuethrich (2000).

[4] The criticisms presented in this paper are harsh. The argumentative context requires this. Gensler has helped the field immensely by formulating the issues in the most explicit manner ever. This will have been in vain if the matter is not treated as deserving the most penetrating attack possible.

Neither of the latter realist positions is empirically defensible. Consequently, either the AA/IC contact theory is a metaphysical theory, in which case its usefulness is questionable, or it is wrong.

Here follow the features of affinity between Afro-Asiatic and Insular Celtic (henceforth AA and IC) which Gensler identifies and discusses. He first presents them as seventeen features,[5] but later separates 2 and 4 into subfeatures and argues for counting them separately,[6] giving the array of twenty features. These twenty features are taken by Gensler as defining a shared AA/IC linguistic-typological 'macrotype'. I have added to most of the features a value T, V and/or D, which are explained immediately following.

1.	Conjugated prepositions	T/V
2a.	Clause-level word order VSO	D^1
2b.	NP-level word order VO (Head-Dependent)	D^1
3.	Relative particle or zero (not relative pronoun)	T/D^2
4a.	Genitival relative clauses: pronoun copying not gapping	D^2
4b.	Prepositional relative clauses: pronoun copying not gapping	D^2/V
4c.	Prepositional relative clauses: preposition moves to verb	D^2
5.	Special relative form of verb	V
6.	Polypersonal verb	T/V
7.	Infixing/suffixing alternation	D^1
8.	Definite article in genitive embedding	V
9.	Non-concord of verb before full-NP subject	V
10.	Verbal noun, not infinitive	
11.	Predicative particle	V
12.	Prepositional periphrastic continuous	T/V
13.	'DO' periphrastic	V
14.	Adverbial clauses 'and …'	
15.	Verbal noun/Infinitive instead of finite main-clause verb	T
16.	Syntactically governed word-initial change (mutation)	V
17.	Non-literal kin-term use[7]	

T = Trivial:

The feature is one that, while unfamiliar in Standard Average European (SAE), is of such widespread occurrence throughout languages of the world that its chance occurrence in the two groups IC and AA is not diagnostic.

---

[5] Gensler (1993: 5-6).
[6] Gensler (1993: 293-4; 298-306). The arguments for separation appear more operational than analytical; it is possible and convenient to count 2a and 2b, 4a, 4b and 4c, separately.
[7] Irish *mac léinn* 'student' lit. 'son of learning', Arabic *ibn al-sabīl* 'traveller' lit. 'son of the road'. This feature will not be discussed descriptively or diachronically. In effect, I might thereby be said to be tacitly conceding its significance for the argument. Whether in the argumentative context its importance is therefore enhanced I doubt, but leave open.

1. Cf. in Europe, Hungarian, and in general, fusion of local markers and pronouns found throughout the world.[8]
3. Relative pronoun usual in Modern Europe, but particle or zero common throughout the world, including early Germanic (patterns with relative pronoun, English *who*, *which*, and German *der*, *die*, *das*, etc., are not inherited Germanic: cf. OE *þe*, ON *er*, *at*).[9]
6. Atypical for Indo-European and SAE (but cf. Basque), but otherwise common for verbal systems of inflexional, agglutinative and polysynthetic types throughout the world.[10]
12. Widespread throughout world, and hardly unique in Europe! (German, Dutch, Icelandic, Georgian, with varying degrees of grammaticalisation).[11]
15. In regard to the use of verbal nouns in main clauses conjoined with an initial clause with a finite verb, this is a simple case of conjunction reduction, a universal principle, and without diagnostic force. Use of the narrative verbal noun without conjunction is a stylistic feature of texts, easily paralleled by the Latin narrative infinitive, so giving no scope for placing it in a group of typological features diagnostic of prehistoric contact.

D = Dependency:

There is an implicational correlation between the features, two series of features thus related indicated by superscript numerals.[12]

a. Given clause level VSO, it is an implicational typological commonplace that at NP level the orders NAdj. and NGen. would be expected; so 2a and 2b cannot be allowed to stand as *two* features in the list: they are *one*

---

[8] See Appendix 1.

[9] E.g. Korowai (Papuan), particle (van Enk & de Vries 1997: 114-15); Lake Miwok (Utian), particle (Mithun 1999: 264-5); Bella Coola (Salish), zero with deictic congruence of verb (Davies & Saunders 1997: 97-105), Somali (Cushitic), zero (Kirk 1905: 125); Swahili (Bantu), particle (Ashton 1947: 110-14); Korean, zero (with special relative form of the verb) (Sohn 1999: 309-14). This sample is tiny and useless for statistical analysis. But I assume the ease with which these relativisation procedures can be found is indicative of their triviality for the AA/IC contact theory.

[10] It is so common that it appears to me futile to give even token examples. I wonder, in fact, if polypersonality in the verb is not, purely numerically, really the dominant pattern of verb-argument cross-referencing in world languages. I do not know. But it is clearly a typological commonplace.

[11] I assume that the minimal survey by Comrie (1976: 98-103), suffices to indicate the relative ease with which this feature can be found in languages throughout the world.

[12] Some of these dependencies are explicitly signalled by Gensler through the structure of the numeration, and even discussed, but he still ultimately insists on counting them as separate features (see note 6). I do not take issue with Gensler's observation (293) that the correlation between clause-level and NP-level word order is not deterministic. But it remains trivial for a clause-level VSO language to have NP-level noun-adjective, noun-genitive order, e.g. I refer to the 'Expanded Sample' of Hawkins (1983: 283), in this matter. See further Appendix 2.A.

word-order type. Also, given VSO type, the suffixing/infixing alternation is implicationally preprogrammed, once a system of clitic elements centred on the verb is posited, and this is congruent with the Indo-European background of Celtic (not exotic for Indo-European).

b. The structure of relative clauses forms a single subsystem type. The choice between copying *or* gapping (i.e. *not* relative pronoun) in 4a and 4b is already given by 3, relative particle or zero (i.e. *not* relative pronoun). 4a, copying, not gapping, in genitival relative clauses, is the rule for extant Brittonic and post-Old Irish. But Old Irish has genitival relative gapping *also* (Thurneysen 1946: 321-2, §507b, d), restricted to certain constructions where the genitival case role internal to the relative clause can be read from the construction unambiguously without copying; where this is not possible, copying is obligatory (Thurneysen 1946: 321-2, §507c, e), i.e. structurally inevitable, given 3, relative particle or zero and 5, special relative form of verb (i.e. *not* relative pronoun). Given the syntactic restrictions gapping must be subject to (unambiguous reading of case role from clause), it is trivial that if anything is generalised, then it is the functionally more versatile copying, because it is clearer in more contexts. Further, given copying in genitival relative clauses 4a, then copying, not gapping, in prepositional relative clauses 4b is trivial, though obviously not inevitable, as shown by 4c. Prepositional-pronoun copying 4b and prepositional movement 4c are structurally mutually exclusive; the *historically visible* shift from 4c to 4b in IC therefore represents a transition entirely internal to Celtic, and one cannot have *both* 4b (AA in Ar., He., Eg.) and 4c (AA in Be.) as diagnostic features of the 'macrotype'.[13]

V = Vacuous:

The feature is either of such limited occurrence within IC and/or AA as to be non-diagnostic for the respective groups as a whole (e.g. 12), or is of such demonstrably late emergence within one of the groups as to make it impossible to connect with any 'contact' between prehistoric IC and AA speakers (e.g. 4b, 16).

1. In IC emerging only in the period of apocope rules; the proto-forms of the conjugated prepositions were grammatically analysable syntagmata, e.g. OIr. *airib* < *are swes*; OIr. *intiu* < *ande sūs*; OIr. *duaib* < *do ēbis*;

---

[13] Gensler (1993: 439), is aware of the mutual exclusivity of 4b and 4c, but only expresses this in terms of the co-occurrence of the features in the AA languages. He misses the point in the diachrony of IC that one cannot self-consistently posit prepositional relative movement as being due to contact with the substrate language, only to posit the later shift to prepositional relative copying as being due to contact with the same language. In Gensler's defense, one can note that the priority of prepositional relative movement in Brittonic was not widely known when he was writing, see Isaac (2003).

W *gennyf* < **kanta mī*; OW *cennin*, OB *centen* < **kanta snī*; MW *yt*, MB *dit* < **do tī*; MW *ynn* < **do snī*, vel sim, etc.[14]

4b. In Irish only later in Old Irish period. So no justification for attribution to prehistoric IC.[15]

5. Emerges in IC only at apocope period; prior to that grammatically analysable syntagmata with enclitic relative particle **yo*: OIr. *bertae* < **beronti-yo* 'who carry' vs. *tuthēgot*[16] < **to-yo·tēgont* 'who come' vs. *aratobarr*[17] < **are-yo·toberor* 'for which is given'; Gaul. *dugiIontiIo*,[18] *toncsiIontIo*[19] vs. OW *emmi guollig*[20] < **ambi-yo·wollunget*, etc.[21]

6. Post-apocope reflex of pre-apocope.

7. Infixing/suffixing.[22]

8. Only definable from phase when definite article itself became grammaticalised, not very long period prior to historical languages (no trace in Old/Continental Celtic sources).

9. In IC only well into the periods of historical attestation of the languages. (Contact with AA in 10th century!)

11. For IC, in historical period, general in Welsh only, develops from use of particle *int* (OW/OB) with adjectives in adverbial use (Mod.Irish *Tá sé go maith*, etc., limited to few adjectives).

12. In AA only Egyptian! So no justification for attribution to the type of AA as a whole. In IC grammaticalisation postdates Brittonic unity (terminus

---

[14] The transparency of the syntagmata behind the IC conjugated prepositions seems not to have been noted by the AA/IC contact theorists. The univerbated treatment of these syntagmata by the rules of apocope indicates that the prepositions were tonic, the pronouns clitic. But that does not in itself imply fusion of preposition and pronoun, merely the usual relationship of tonic and clitic. While the IC conjugated prepositions are synchronically exotic for Indo-European languages of the same period, their proto-historical diachrony is fully in accord with inherited PIE grammatical patterns, not exotic at all in the context. Tonic PIE adverb (> adposition; cf. Beekes 1995: 219) + clitic pronoun is a PIE pattern the reflex of which is also clearly seen in Vedic syntax, e.g. **práti vā́ṃ** *sū́ra údite vidhema* 'when the sun has arisen, we pay homage to you two' (*Rgveda* 7.63.5.) This pattern is also implicated in the development of IC verb-initial syntax, compound verbs and infixed pronouns. Adpositions governing nouns and lexical preverbs compounding with verbs are two parallel developments of the same original elements, things like Vedic *práti* in the above example.

[15] See Isaac (2003) for the structure of prepositional relative clauses in early extant IC languages.

[16] Cambrai Homily (Stokes and Strachan 1903: 247).

[17] Wb. 12[d]29 (Stokes and Strachan 1901: 579). The nasalisation in the relative clause, in place of inherited lenition, is analogical (Isaac 2003: 89, n. 25, with references).

[18] An inscription of Alise-Sainte-Reine (Lejeune 1988: 149-50).

[19] Inscription of Chamalières (Lambert 2002: 270-1).

[20] Computus Fragment (Williams 1927: 256; Falileyev 2000: 54).

[21] These patterns too are discussed in Isaac (2003).

[22] The verb only becomes polypersonal by virtue of the syllable loss which transformed the transparently suffixed and infixed object pronouns (syntagmata) into more morphology-like affixes.

post quem (tpq.) c. 400 A.D.), since Cornish (*ow* < *writ*) and Breton (*o* < *writ*) use a different preposition to Welsh (*yn* < *onkus* = OIr. *oc*, Isaac 1994).[23]

13. Only Brittonic, and of uncertain priority there, and Egyptian. But cf. German. Trivial, as well?

16. Mutation systems only emerge as side-effect of rules of apocope (5th-6th centuries A.D.); prior to that, there was only trivial sandhi.[24]

I present the features again in a tabular arrangement, to illustrate their range in the languages in question. This data is merely taken from Gensler's work. I have silently added a couple of 'hits' where a language has a feature not, apparently, noted by Gensler (thus, incidentally, ostensibly bolstering his case). In no case have I removed a 'hit' recorded by Gensler. For AA I have restricted the presentation to the languages of Arabic, Hebrew, Berber and Egyptian. These are the languages consistently referred to by Gensler, others, such as Phoenician, Akkadian or Aramaic, being mentioned only occasionally. I am happy to concede for argument's sake that the representation of a feature by all four of the AA languages included here can be taken as equivalent to full representation by all relevant AA (excluding Omotic, Cushitic and Chadic), even though that may not actually be the case. I assume that in the cases where other AA languages do or do not have specific features, the full picture would not differ radically from that given here, based on Gensler's work. This array of features in common between AA and IC may indeed be thought to be 'amazing'.[25] But the present article,

---

[23] In the conclusion to the cited paper, I pointed out, with respect to the *relative* chronology of the construction in Brittonic and Irish, that it 'might have to be regarded as inherited from their respective, or even common, proto-languages, and of some considerable antiquity' (Isaac 1994: 380). I do not recall whether I had any specific *absolute* chronology in mind when I wrote that. I interpret the extant facts of the prepositions used in the construction as meaning that the development of the construction belongs to the Late British/Late Proto-Irish period, with strong influences of contact between the languages, but with South-West British (> Cornish / Breton) already genetically separate. Hence tpq. c. 400 A.D.

[24] There is no doubt that systems of initial mutation comparable to those of the IC languages are genuinely an extreme rarity in the languages of the world. But I am not persuaded that the comparison with Berber initial vocalic alternations is apposite. There are significant, universal differences between the phonological functions and distributions of vowels vs. consonants in the structures of languages; initial vowel mutation and initial consonant mutation are not really the same thing morphophonologically (the prefixing of *h-* to vowels in the IC languages is also not a 'vowel mutation'; it is a consonant mutation of the structure ø → *h*), and the Berber initial vowel 'mutation' is really far more appropriately described as a sort of ablaut, which is quite a different kettle of typological fish. For the initial *consonant* mutations of IC, Gilyak and Finnish (Jakobson 1971: 86-7, resp. Karlsson 1984: 23; both cases are 'multicategorial' in the term of Gensler 1993: 247) provide far more apposite objects of comparison. But they are useless for the AA/IC contact theory. The case of Gilyak, and some others, is mentioned by Gensler (id.: 247).

[25] Gensler (1993: 6), on the first presentation of the 17 undifferentiated features.

building on what has been observed above, will continue to question the argumentative strength of the array.

But (see Table 1 on the next page and Table 2 below):

1, 5, 6, 16 are contingent on 5th-6th-century apocope rules, so not present in 4th-century Celtic 'macrotype' (6 polypersonal verb, prior to apocope = 7 inf./suff.).

4b, 9, 11 are developments deep into historical period, i.e. not present in 6th-century Celtic 'macrotype' (11 generalised only Welsh; Irish sporadic, later).

Table 2. Diachronic tabulation:

tpq. c. 900	9^{IWCB}, 11^{IW}
c. 600-c. 900	4b?[26]
c. 400-c. 600	1, 5, 6, 12, 16, (13?[WCB])
taq. c. 400	2a, 2b, 3, 4a, 4c, 7, 8, 10, 15?

14. Adverbial clauses 'and …' omitted.[27]
17. Non-literal kin-term use omitted.

---

[26] There are very rare instances of prepositional relative copying in Old Irish (Thurneysen 1946: 322; McCone 1985: 96). There are no clear data for this for the Old Brittonic languages. The argumentation of Isaac (2003) points to the rise of prepositional relative copying within or immediately prior to the periods of attestation of the languages, so c. 600-900 as given here.

[27] Consensus has not been reached with regard to the antiquity of the IC 'and ...' clauses. The question has not been answered definitively whether OIr. *os mé...* etc. (ostensibly 'and I...,' etc.) really contains a contracted form of *ocus* 'and' (Thurneysen 1946: 548) or is actually originally a participial form of the substantive verb (as O'Brien 1923), or something else again. The prehistoric diachrony of the adverbial 'and ...' clauses is therefore quite unknown at this time. The study of its diachrony in the history of Old and Middle Irish has, however, recently been given a firmer foundation by Ronan (2002) whose data and arguments are not obviously consistent with the proposal that the presence of the construction in Celtic languages is explicable by contact with an Afro-Asiatic or Afro-Asiatoid language. Cf. further Tristram (1999: 271-3).

Summary Table 1.:

		1	2a	2b	3	4a	4b	4c	5	6	7	8	9	10	11	12	13	14	15	16	17
AA	Ar.	■	■	■	■	■	■	□	□	■	□	■	■	■	■	□	□	■	□	□	■
	He.	■	■	■	■	■	■	□	□	□	□	■	■	□	□	□	□	■	■	□	■
	Be.	■	■	■	■	■	□	■	■	■	■	□	□	□	■	□	□	□	□	■	□
	Eg.	■	■	■	■	■	■	□	■	■	□	□	□	■	■	■	■	□	■	□	□
IC	OIr.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■
	MW	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□
	Co.	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	□
	Br.	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	■	■	■	■	□

% hits:

- Overall   AA = 60   Ar. = 65   He. = 55   Be. = 55   Eg. = 65
- Overall   IC = 93   OIr. = 95   MW = 95   Co. = 90   Br. = 90

This injection of diachronic realism into the arguments forces the AA/IC contact theory into a very strong empirical prediction.[28] If the appearance of features in the sixth or seventh centuries, or even after 900 is to be attributed to contact with an Afro-Asiatic or Afro-Asiatoid language, then that language must be visible. We know that the IC languages were in contact with Old English, Proto-Norse, Latin and Romance, in various places at various times during the period in question, as well as with each other at all times. The period is traditionally called the 'Dark Ages,' but we are still dealing with the fully historical period. Hand-waving arguments about invisible languages of conquered pre-Celtic people will not do here for this period. If the contact language were influencing Welsh, Cornish, Breton and Irish at this time, we would have to be able to see it. The AA/IC contact theorists must be able to show it to us.[29]

The cluster of features comprised by prepositional pronouns, relative verbal forms, polypersonal verbs and initial mutations, which are all dependent on the fifth and sixth-century apocope rules, present a particular challenge to the AA/IC contact theory. Not only is their date problematic for the theory, but the mechanism of their development is not obviously consistent with the causal relations posited by the theory. For the mutations, it is true that the apocope rules, before which there was trivial external sandhi and after which there were grammaticalised mutations, are not in themselves sufficient to motivate that grammaticalisation itself. Other languages, including many in Europe, undergo various kinds of apocope, without developing mutations. On the other hand, as I pointed out many

---

[28] There is a curious instance of such diachronic realism in Gensler's work. At the end of a paragraph discussing the diachrony of Berber initial vowel 'mutations,' the following conclusion is stated: "If valid, such considerations argue that Berber word-initial change as we know it did not come into existence at least until the time of the Arab conquest – much too late to be implicated in any hypothetical pre-Celtic substratum on the British Isles" (Gensler 1993: 248). It has not become clear to me how this conclusion can be reconciled with the subsequent continued inclusion by Gensler of initial mutations in the discussion of the 'macrotype'. Clearly, the tacit assumption would have to be made that 'such considerations' are, after all, not 'valid'. Such seems to be the thrust of Gensler's argument when he returns to the point (455), to suggest that Berber itself is showing the influence of a further non-extant African contact language here, though what that language was is unspecified, and apparently unspecifiable. The question of multiple contact languages is addressed below in the main text. (On Berber 'mutations,' cf. also note 24).

[29] A handful of questionable etymologies would not suffice. It must not be thought that the difficulty of the chronology is ignored in the pleas for AA/IC contact. 'Chronological anomalies' are indeed addressed at length by Gensler (1993: 442-56). It appears to me that he has himself formulated cogent chronological arguments against the AA/IC contact theory. I have not been able to discern that the various dismissals of these arguments in the cited passage amount to a coherent argument. They reduce to the adoption of the license to assume that one can always posit influence from invisible languages, an indefinite period (i.e. as required) prior to the extant language under investigation. Insofar as this license amounts to the renunciation of the need to be consistent with any data at all (if the required language is not extant, then it can just be posited to have been present but invisible and belonging to an indefinitely earlier period), this appears to be a barren principle for an empirical discipline to adopt.

years ago in the context of a discussion of the origins of the IC word-order type, there is a typological correlation between word-initial morphological alternations (which the mutations are) and verb-initial word order.[30] Since it is not in dispute that IC had moved to VSO word-order typology several centuries before the apocope in question, there is a hint here of an additional typological correlation, which would again not need the input of a contact language for explanation. Furthermore, since the infixing/suffixing alternation, and the pre-apocope relative particles which are part of that subsystem, are themselves also connected to the verb-initial syntactic typology, and are also transformed by the apocope rules into the polypersonal verb and the special relative form of the verb, we are seeing here further implicational chains describable and explicable entirely in terms of the grammars and histories of the IC languages themselves. A contact language again adds nothing in the way of explanation to the analysis. It has been suggested, probably many times, that the two types of explanation are not mutually exclusive.[31] But this is obscurantism. The correct logical relations have been deliberately blurred. Given that we are seeking explanations, we find an explanation for a set of phenomena within their own structures and histories (the grammatical and typological links between apocope, conjugated prepositions, polypersonal verbs, relative verbal forms and verb-initial word order). The question is irrelevant how compatible these internal explanations are with an explanation in terms of a contact language. They *are* compatible with the hypothetical *presence* of a contact language. But since the phenomena in question are explained by their own structures and histories, that *presence* need not be hypothesised in the first place. It is not that the internal explanation contradicts the AA/ IC contact theory: it is just that it makes it redundant. There is no need to explain that for which there is already an explanation. As a result, there is no need for the hypothesis of an AA contact language in the British Isles in the fifth and sixth centuries.[32] One may of course hypothesise anything. But then, in the case

---

[30] Isaac (1993: 12-13). See further Appendix 2.B.

[31] Gensler (1993: 436-7).

[32] Why does the 'internal' explanation take precedence over the explanation by contact? This is dictated by parsimony. Parsimony cannot tell us what languages were spoken in the British Isles in the fourth, fifth and sixth centuries A.D. We must use other approaches to establish that (analysis of data). But the grammatical features in question in this paragraph all have histories and prehistories, and those histories and prehistories are linked with each other. One would say more precisely that the ontogeny and the ontology of the grammatical features are bound together, e.g. it is in the very nature of the Insular Celtic polypersonal verbs and relative verbal forms that they exist (ontology) because they arose (ontogeny) from the interaction of rules of apocope and syncope with VSO word-order typology and patterns of enclisis: the state of the grammar reflects the history of the grammar, ontology and ontogeny are intertwined. Thus, the explanation for the phenomena in question is derived from the phenomena themselves, and nothing further need be posited to explain their existence and nature. And since nothing further *need* be posited, nothing further *should* be posited. This is parsimony. To posit an unknown contact language is a hypothesis posited 'in vain' ('frustra'), as William of Ockham would have put it.

of hypothesising the existence of a language in Dark-Age Britain and Ireland, that language must be presented.

We may revise the array of features to take into account the diachronic details that have been discussed, and present a typological comparison of AA and IC based on the type of the latter in the fourth century.[33] I also remove 2b, 4a, 4b, 4c and 7 in accordance with points which have been discussed previously:

    2a, 2b are a single word-order-typological complex.

    4a, 4b, 4c are contingent on 3 (if there is a non-referential relative particle, or zero, then copying or prepositional movement is trivial; also restricted gapping in OIr.).

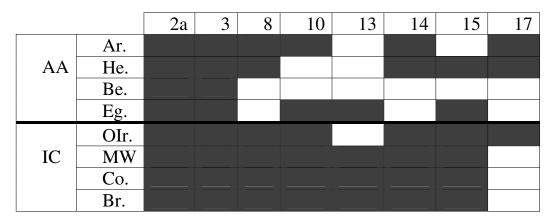   7 (infixing/suffixing) is contingent on 2a (VSO).

Table 3. Resultant 4th-century picture:

		2a	3	8	10	13	14	15	17
AA	Ar.	■	■	■	■	□	■	□	■
	He.	■	■	■	□	□	■	■	■
	Be.	■	■	□	□	□	□	□	□
	Eg.	■	■	□	■	■	□	■	□
IC	OIr.	■	■	■	■	□	■	■	■
	MW	■	■	■	■	■	■	■	□
	Co.	■	■	■	■	■	■	■	□
	Br.	■	■	■	■	■	■	■	□

% hits:

- Overall AA = 59   Ar. = 75   He. = 75   Be. = 25   Eg. = 63
- Overall IC = 88   OIr. = 88   MW = 88   Co. = 88   Br. = 88

Whereby:

2a   (VSO): while this is superficially synchronically exotic for Indo-European, it remains the case that it is unproblematically derivative of the principles of Wackernagel's Law[34] and Vendryes' Restriction.[35] The exot-

---

[33] These diachronic details apply *only* to IC. I am not at this point discussing AA diachrony.

[34] Enclitics come second in their syntactic domain (Wackernagel 1892).

[35] Enclitic objects are restricted to following part of the verbal predicate (Vendryes 1911-12). Coupled with a universal dynamic tendency (not a rule) to keep the morphosyntactic elements of the same semantic lexeme together, the combination of Wackernagel's Law and Vendryes' Restruction naturally caused the verb to gravitate towards the beginning of its clause. I have argued for recognition of a greater role of communicative function in the developments associated with Celtic word order than is generally posited (Isaac 1996: 146-7).

icism of IC VSO is exaggerated by the AA/IC contact theorists.[36] With the addition of Vendryes' Restriction, the syntax of IC reflects the same basic PIE sentence structure as Vedic or Hittite. The Restriction makes a considerable typological difference, to be sure, but is not, in itself, a startling or puzzling phenomenon. And however 'exotic' a feature may appear when considered diachronically context-free, if that feature is shown to be derivative of – for Indo-European – non-exotic diachronic principles, then its diagnosticity for the presence of an 'exotic' contact language is compromised. Such is the case for IC VSO word-order typology. There is much research still to be done to understand the mechanisms and motivations for diachronic IC word-order changes, but that is research which is being done, and showing successes, and does not point to the presence of any non-extant contact language, AA or otherwise.[37]

3   (non-referential relative particle): trivial and widespread.

13   'DO'-periphrasis is included in the 'macrotype' on the strength of its presence in AA in Egyptian alone, but the grammaticalisation of 'DO'-periphrasis in Late Egyptian (tpq. c. 1500 B.C.)[38] coincides with and is symptomatic of the movement of Egyptian *away* from the 'macrotype'. The outcome of this movement, Coptic, is not as good a representative of the 'macrotype' as Old and Middle Egyptian. It has moved away from VSO (2a), partially through generalisation of the 'DO'-periphrastic, and furthermore cannot be said to have 'DO'-periphrasis synchronically, since the conjugation patterns resulting from Late Egyptian 'DO'-periphrasis are fully grammaticalised and lexically opaque in Coptic. Coptic has also lost the prepositional periphrastic continuous (12). Coptic therefore has only 50% hits for the twenty-feature array constituting the full 'macrotype'.[39] How do we justify including both VSO and 'DO'-periphrasis in the AA 'macrotype,' when the latter is only Egyptian and it is part of the

---

This position opposes both the more mechanistic Indo-Europeanist approaches and the equally function-neutral AA contact approach.

[36] See Appendix 3.

[37] As a postulate, Vendryes' Restriction is relevant to more than the prehistory of the medieval Insular Celtic languages alone, as witnessed by the existence of the inscriptions in Cisalpine Gaulish of Voltino *tomedeclai obalda natina* 'Obalda Natina placed me here' (Thurneysen 1923: 8-10; Meid 1989: 17-26) and Vercelli *akisios arkatoko*[.]*materekos tośokote atom teuoxtonion eu* 'Acisios Argantocommaterecos, he has set it up, the boundary of gods and of men, *ex uoto*' (Lejeune 1988: 26-37; Koch 1983: 187-9; Eska, 1990; there are grounds in the analysis of the latter to posit a left-dislocated topicalisation, Eska, 1990; Isaac 1996: 120-1; I emphasise that none of the insights mentioned in this footnote are due to me). The structures of the verbal segments here, *to-me·deklai* (segmentation as by Eska and Weiss 1996) and *to-so*[*n*]·*ko*[*n*]*de*, and their positions, are in exact accord both with the principles of Indo-European syntax (with the addition of Vendryes's Restriction) *and* the well-attested structures of medieval Insular Celtic word order. I return to this point below in the main text.
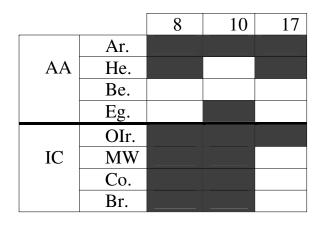
[38] Loprieno (1995: 7, 220, 225).

[39] For details of Coptic grammar I rely on Layton (2000).

drift away from VSO there? A further point of argument: the presence of 'DO'-periphrasis in Germanic (NB *not* just English)[40] makes it look typologically rather trivial.

14   It is not at all clear that adverbial 'and …' clauses are that old.[41]

15   is a trivial case of conjunction reduction and/or literary stylistics, and cannot be made part of a diagnostic 'macrotype'.

Taking further account of these considerations gives us the full array of features that can reasonably be posited as being significant points of resemblance between fourth-century IC and AA (with no restrictions on the historical state of the latter) as in Table 4:

Table 4:

		8	10	17
AA	Ar.	■		■
	He.	■		■
	Be.			
	Eg.		■	
IC	OIr.	■	■	■
	MW	■	■	
	Co.	■	■	
	Br.	■	■	

8.   Definite article in genitive embedding
10. Verbal noun not infinitive[42]
17. Non-literal kin-term use

---

[40]   Lockwood (1968: 157).

[41]   See note 27. Exclusion of the feature due to this uncertainty may be arbitrary. I will not labour the point.

[42]   With regard to the place of verbal nouns in the arguments for AA/IC contact, there is a double irony. 1) The view is widespread enough to be regarded as consensual that the verbal nouns of IC actually represent the archaic preservation of the original pattern of non-finite verbal abstracts in Proto-Indo-European (cf. Disterheft 1980: 197; McCone 1994: 175; Russell 1995: 275-6, amongst others), not an 'exotic' divergence from a 'standard Indo-European' pattern. 2) It is only a minority view that this interpretation of the state of affairs is incorrect, and that the IC verbal nouns are actually *relatively* recent innovations (cf. Jeffers 1978; Lehmann 1994: 105-6; Ziegler 1997; and Isaac 1996: 431-6; Isaac fc.). Clearly, only the latter view is compatible with theories of AA/IC contact, though it would be mistaken to conclude that the representatives of the minority view were thereby automatically favourably disposed to the theory. I leave the feature in question in the array for argument's sake, without prejudice to its validity, however this may ultimately be judged.

% hits:
- Overall AA = 50   Ar. = 100   He. = 67   Be. = 0   Eg. = 33
- Overall IC = 75   OIr. = 100   MW = 67   Co. = 67   Br. = 67

These three features are all that can really be said to be left of the AA/IC contact theory from the perspective of fourth-century IC. It is obviously not particularly 'amazing,' in either the number and quality of features involved or in the numbers of hits the languages score.

In this paper, I have started by looking at the finest details of the arguments in question. I have examined the feature array proposed as the 'evidence' for some sort of contact between AA and IC, I have attempted to clarify why this array of features does not stand up to scrutiny, and how it forces us to posit realist interpretations in chronology that go beyond what may properly be defended. In the latter parts of this paper, I shall look at some broader implications of the AA/IC contact theory.

Firstly, there is the question of the argumentative construction of the 'macrotype' itself. The point has been made explicitly that even if the features of the 'macrotype' are individually trivial or susceptible to alternative explanation, it is the whole combination, the 'ensemble,' that is diagnostic.[43] This is a crucial step

---

[43] "The point is not any individual feature, but the cumulative weight of the ensemble" (Gensler 1993: 439). Gensler's work contains frequent comparisons between the 'typological method' of language comparison which he attempts to develop and 'traditional' comparative-historical method. The 'Assessment' with which the argumentative section of his work concludes largely consists of the assessment of just that comparison (Gensler 1993: 456-63). This ethos of comparison of the two 'methods' can be adopted here, in respect of the argumentative role of 'ensembles' of features in comparative-historical linguistics. One would be mistaken in thinking that the theory of Indo-European (or any other reconstructable genetic language family) is dependent on such arguments. Gensler writes, "It is not *typically* the case that the historical linguist, when dealing with a pair of resemblant word forms in two languages, must make it his or her first task to defend the resemblance against charges of coincidence" (Gensler 1993: 10). Allowing for a certain ambiguity in the term 'word forms,' this observation seems to be based on a common misapprehension of the substance of arguments for Proto-Indo-European. It is not *words*, or *word forms*, that are the core of that substance, but grammar, mostly in the form of paradigms. The mere presence in, say, Latin, Greek and Sanskrit of similar *words* does not, could not, and never has been thought to, suffice for the postulate of a common genetic ancestor of these languages, any more than the presence of many French words in English is indicative of the derivation of English from Latin. The identification in the twentieth century of the Anatolian languages and Tocharian as Indo-European did not follow from the discovery of Indo-European *words* in these languages. If there had just been Indo-European *words* in these languages (be they ever so many, an 'ensemble'), embedded in quite different grammatical systems, then Anatolian and Tocharian would never have been classified as Indo-European (the presence of Indo-European words in these languages in those locations at those periods would have been a matter of considerable interest in itself, of course). They are classified as Indo-European languages because they have Indo-European grammars. It is grammars that form the basis of comparison, and the paradigms of which they consist, a

in the structure of the argument, and I have not seen it explicated just how this step is achieved. If it is conceded that the features may indeed be individually trivial, then it must be a matter of urgency that the nature of this 'cumulative' argument be clarified.

The 'problem' under investigation, the *explicandum* posited, is the appearance of apparently similar grammatical features in the IC languages and the AA languages. The 'solution' proposed, the *explicans*, is that the two sets of languages are linked by influence in some sort of contact situation (further implications of which will be discussed below). But if, for instance, 'conjugated adpositions,' relative particles (not pronouns), polypersonal verbs and prepositional periphrastic continuous tenses are, in fact, *relatively* common features of languages throughout the world (some more than others to be sure), how does this '*explicans*' actually function as such? The contact theory can only explain the commonality in the sets of features if it explains the presence of those features in the languages in question. The argument has the structure of *modus ponens*:

(Major Premiss)	If *A* and *B* are two genetically unrelated languages[44] and have the features *w*, *x*, *y*, *z*,[45] then the presence of those features in the respective languages must be due to a causal link through contact.
(Minor Premiss)	*A* and *B* are two genetically unrelated languages and have the features *w*, *x*, *y*, *z*.
(Conclusion)	Therefore the presence of those features in the respective languages must be due to a causal link through contact.[46]

point that was made already in the classic passage of 1786 by William Jones, which specifically postulates that the evidence for genetic commonality is to be found 'both in the roots of verbs and in the forms of grammar'. The mere presence, in languages from Western Europe to India, of verbal paradigms of systematically relatable structures is sufficient as an indicator of that genetic link. The verbal paradigm is a single feature, a feature with an internal structure, certainly, but nevertheless a single feature, and it is itself diagnostic of the genetic link. The fact that the languages show other similarities in many grammatical subsystems corroborates the postulate of such linkage. But this is not a cumulative argument. This is not an argument from a 'diagnostic ensemble' of individually non-diagnostic features. The paradigmatic features are themselves, individually, diagnostic of genetic linkage. This is therefore a difference *in principle* between the comparative-historical method and the so-called (by Gensler) 'typological method,' not just a difference in degree of some parameter. As a final point of clarification: the typological comparison of languages and the insights to be gained therefrom into the nature and structure of human language, are not in question here, only the application of typological comparison to the argument for prehistoric contact. It is the latter application which is the meaning of Gensler's term 'typological method'.

[44] Or language families. This synthetic distinction has no bearing on the argument structure.

[45] Or up to twenty features. Again, the number has no immediate bearing on the argument structure.

[46] I cannot see how hedging this, as "... are *probably* due to a causal link," could enhance the validity of the argument. We want to know what is the case. Even if a rigorous metric were

But the Major Premiss can only be a valid implication if what is said about the features *w*, *x*, *y*, *z* together is also valid for each individual feature. For it to be true that there must be a causal link to explain the presence of features *w*, *x*, *y*, *z* in languages *A* and *B*, it must also be true that there must be a causal link to explain the presence of feature *z* in languages *A* and *B*. If it is not true that there must be a causal link to explain the presence of feature *z* in languages *A* and *B*, then it cannot be true that there must be a causal link to explain the presence of the set of features *w*, *x*, *y*, *z* in languages *A* and *B*. Thus, if *w*, *x*, *y*, *z* are all individually more or less trivial features of languages, whether in Europe, or in Indo-European languages or taking the languages of the world as a whole, then the mere combination of those features into a set, an 'ensemble,' cannot make them diagnostic of some prehistoric connection between languages *A* and *B* (a more formal proof is given in Appendix 4). I assume it can be seen that the same is true in the case of features that may not exactly be trivial, but which can be shown to arise in the diachrony of the extant structures themselves; an 'ensemble' of such features also cannot attain a degree of diagnosticity that the individual features do not have.

I recognise that Gensler's statistical analyses are an attempt to demonstrate exactly the required diagnosticity. But their adequacy is in question. Firstly, the mere statistical frequencies of the individual features in his survey of 70 languages do not seem very encouraging for the theory. With one exception, the frequencies of the individual features do not, by Gensler's analysis, drop below a possible proportion of 1:8.[47] Granted that, on a scale, 1:8 is 'rarer' than 1:2, nevertheless, 1:8 does not really seem absolutely very rare at all. Given a total of world languages at the estimate of c. 5000, that would give c. 556 languages with a feature of frequency 1:8. The ideal, and probably unattainable, complete survey of all world languages would presumably give a different figure to this, but it is equally presumably Gensler's assumption that the statistical method he adopts guarantees that it would not be radically different.[48] Is this really 'rare,'

---

being applied by means of the term 'probably' (which I doubt), we would then be lacking an argument that such and such is the case. If the AA/IC contact theorists insist that they are merely presenting a 'probability,' which can *by definition* not be shown not to be the case, i.e. refuted (for what data could refute the assertion that something is 'probably' the case?), then they are *by definition* not presenting an argument about the real world as it is. Hence my earlier assertion that one interpretation of the AA/IC contact theory is that it is a metaphysical theory, and therefore not interesting as a theory within an empirical discipline. I do not myself accept this interpretation. I do not believe that the AA/IC contact theory is a metaphysical theory. I believe it is an empirical theory, i.e. a theory about how the world is, and as such, wrong.

[47] Gensler (1993: 374). The exception is initial mutations, with a proposed frequency of 1:18. I have no reservations conceding that this indeed constitutes a rarity. But it can only be included in the AA/IC 'macrotype' on the basis of Berber initial vocalic alternations, and on these, see above, notes 24 and 28.

[48] Even my formulation here is not entirely accurate, since the full count of approximately 5000 'world languages' (whatever the exact figure) is itself merely a contingent sample. It

rather being just less common than some other features? This is only my impression, but I have not been able to discern how Gensler's method demonstrates that my 'impression' in this matter is incorrect, his, that 1:8 somehow shows something significant for the theory, correct. There does not seem to be any clear statement derivable from the analysis that a feature with a frequency of 1:8, or an 'ensemble' of features with frequencies in the languages of the world ranging from 1:2 to 1:8 must be diagnostic of something.

Secondly, though it must be conceded that Gensler's statistical argument goes beyond mere frequencies of the individual features, it is still not clear that the full analysis reaches a conclusion that can go beyond the impression that there are noteworthy typological similarities between IC and AA (not in dispute). Gensler examines the frequencies of various pairings of features, concluding:

> Our examination of interface correlations provides new confirmation that the exotic features are indeed exotic: they may sporadically recur outside [IC and AA] and Africa, but only minimally do they ever occur together in the same language.[49]

This statement is not in question. But it does not, nor do the analyses it summarises, provide any argumentative, implicational link between the occurrence of the individual features in IC and AA and the postulate of a continuum of contact between those language families. The AA/IC contact theorists conclude on such a basis that those occurrences are 'probably' not coincidental, and therefore that there was 'probably' a contact continuum linking AA and IC in prehistoric Europe. But the latter 'probability' is not derived from the statistical analysis, which only gives us a very rough 'probability,' or rather 'improbability,' of the occurrences and co-occurrences of the features. No deductive path leads from that 'improbability' to a 'probability' that there was AA/IC contact. The latter is a hypothesis posited to explain the observed distribution of the features in the languages. The 'probability' of the occurrence of the features is numerical, based

---

does not include any non-extant past languages (e.g. the languages spoken around 30,000 B.C.), nor does it include any future languages (of, say, 30,000 A.D.). I assume that it can be agreed that the full number of all 'human languages' is, and always will be, *in principle* inaccessible, any accessible statistics of the sort in question here therefore being very broad approximations. One can make descriptive statistical statements based on the tiny sample of languages extant over the approximately 5,000-year period which is available to us. But to justify the conclusion that the occurrences of features or groups of features in AA and IC are not coincidental, or 'probably' not coincidental, one must make the untestable assumption that the tiny sample of languages available to us is representative not only of the tens or hundreds of thousands of years of languages inaccessible to us in the past, but also of the indefinite (if not infinite) number of potential future languages. In truth, however, there is no justification for the assumptions that either the inaccessible linguistic past or the inaccessible linguistic future of human beings followed the same trends as are visible in the extant sample. Consequently, the 'conclusion' of non-coincidence for the appearance of common features in AA and IC is, and will always be, in fact, an unsupportable assumption.

[49] Gensler (1993: 414).

on the statistical analysis of the sample of languages:[50] the 'probability,' so-called, of there having been AA/IC contact is not. The term 'probability' is thus being used in two different senses in the two parts of the argument. And the use of the term 'probably' in the second sense has been discussed, see note 46. To assert that there was 'probably' AA/IC contact does not tell us anything about how the world is, because no conceivable facts could contradict such an assertion.[51]

So, after all the statistical analysis, we still have no criterion for the answer to the question, why *these* features, why *this* 'ensemble,' are 'diagnostic,' i.e. imply, that there must have been contact between the languages that show them. And if the diagnostic nature of this 'ensemble' cannot be read off the logic or numbers of the argument, it must be derived from elsewhere. The only place it can be derived from is the nature of the features themselves. So we are brought back after all to the dependence of the diagnosticity of the 'ensemble' on the diagnosticity of the individual features. And, as has already been discussed, the latter is wanting.

It seems intuitively 'obvious' that a set of twenty features in common between two language families *must* be indicative of some sort of connection between them. It is just such obviousness which keeps the AA/IC contact theory in the realm of debate. The argument might perhaps be said to reduce thereby to 'common sense'. But it is the task of the scholar to test common sense, not to accept it uncritically. The AA/IC contact theory should consist of arguments that *demonstrate* the diagnosticity for that contact of the 'ensemble' of features, not of arguments that are formulated on the *assumption* of that diagnosticity. The rhetoric of the large 'ensemble' of features is what sustains belief in the theory.[52] But 'obviousness' is only so from certain perspectives, and the choice of perspectives is largely an aesthetic matter. It is not sufficient as an argumentative basis.

I have discussed the nature of the individual features themselves, arguing that they are not such as could support the hypothesis of an AA contact language as a substratum to the IC languages, and I have discussed how these individual features are putatively combined into an organic 'macrotype,' arguing that that pro-

---

[50] For argument's sake, I do not question the validity of the statistical analysis itself, though some may wish to do this.

[51] The statement 'x is not so' does not contradict the statement 'x is probably so'. So even if it were shown without any room for doubt that there was no AA/IC contact, one could still logically maintain that there 'probably' was. Of course, it can be highly improbable that there was no contact, but nevertheless the case. This is the difficulty of all non-metric 'probabilistic' arguments. It is just because the statement 'x is probably so' implies the statement 'either x is so or x is not so,' which is a tautology, analytically true, therefore saying nothing whatsoever about the world.

[52] It is on similar rhetorical turns that statements about 'solid nonimpressionistic method' (Gensler 1993: 458) or the 'specialness' of the AA/IC resemblances (Gensler 1993: 460) depend. But 'solidity,' 'specialness' or the 'ensemble' cannot replace engagement with data and argumentative structure. There have been plenty of data and arguments presented in all expositions of the AA/IC contact theory. The present paper takes issue with the quality of engagement with these things that has been seen.

cess of combination is merely a rhetorical device with no logical basis which could uphold the AA/IC contact theory. Finally, wider issues of the chronology and linguistic geography implied by the AA/IC contact theory are to be addressed.

Old and Middle Egyptian are good representatives of the 'macrotype'. But that means that the type is fully established there c. 3000 B.C. If the contact language in Western Europe was itself AA (remaining invisible the whole time), then it must have been making its presence 'felt' since c. 3000 B.C. But IC only 'joins' the 'macrotype' in the early centuries A.D. In what geolinguistic and sociolinguistic context was a language or language group influencing Egyptian by contact prior to 3000 B.C. *and* IC *after* 100 A.D.?

The combined AA/IC 'macrotype' is better represented by IC languages than by AA languages, a point already made by Hewitt (2003). For all twenty of Gensler's features over the four languages each, AA scores 60%, IC 93%.

This must be interpreted in the light of the following facts:
  i) IC's representation of the 'macrotype' is the result of developments in the historical period, from c. 1st century A.D. on; and
  ii) the 'classic' AA languages (except Berber) are attested *much* earlier than IC, and must therefore be closer to the 'pure' 'macrotype,' which the IC languages only secondarily reflect.

Therefore:

A. If the language in contact with IC was itself genetically AA, then either

  i) it influenced them to be more like the 'macrotype' than it was itself (a paradoxical suggestion?),[53] or
  ii) it was itself more like the 'macrotype' than any of the *actually extant* AA languages (a counter-empirical suggestion?).[54]

B. If the language in contact with IC was not itself genetically AA (so 'Afro-Asiatoid'), then either

  i) it was itself influenced by AA by contact, thus transitively passing the features on to IC, in which case the features of the 'macrotype' became stronger the more distant they were from the causal source (another paradoxical suggestion?),[55] or

---

[53] Contact with a non-extant AA language cannot 'explain' the typology of IC if that language did not itself fully represent the type to be explained.

[54] Contact with a non-extant AA language cannot 'explain' the typology of IC if not only must that language be conjectured, but its type too must be conjecturally 'enhanced'.

[55] The suggestion compounds the difficulty in A.(i) by transferring it to the hypothetical intermediary. The logic of this variant of the explanation also requires us to posit that the hypothetical intermediary acted not only as a channel of diffusion, but also as a typological

ii)   it was itself the source of the 'macrotype' features in AA, in which case it had already done its work in North-East Africa in AA by 3000 B.C. (Old Egyptian), and yet was present, 'potent,' and *typologically unchanged* over 3000 years later in the British Isles (another counter-empirical suggestion?).[56]

It does not appear to me that these difficulties can be reduced by introducing multiple intermediary languages.[57]

'lens,' 'magnifying' the feature complex involved: the (non-AA) intermediary made IC typologically 'more Afro-Asiatic' than Afro-Asiatic itself, an obviously self-contradictory assertion.

[56] The core assertion of B.(ii) is *in principle* irrefutable, because it is an affirmative existential statement: "There was [such] a language." No conceivable data could refute this assertion (this is probably true of the other assertions in A and B, but in those cases, different problems arise in the arguments leading to them). But we can at least observe that its implications can be shown to be contrary to (though not contradicted by) real data. There are a few languages that have extant histories extending over a period of 3000 years or more. Egyptian itself is one. Greek is another, also Iranian and Indo-Aryan languages. Semitic as a family has a history of up to 4500 years, but no individual Semitic language approaches this length of continuous history. None of the languages just listed show typological stability over the period of the order required by the AA/IC contact theory. More languages are extant for a period of between 2000 and 3000 years. The Romance languages < Latin are an obvious European instance. The Celtic languages themselves fall into this category. There is also Chinese, and possibly a number of others. The typological stability over such a large array of features as implied by the AA/IC contact theory is not given for any of these languages, and typological flux is rather the rule. One or other feature, can, of course, remain stable for very long periods, e.g. noun-phrase-level word order in Semitic languages. But it is surely crucial to the AA/IC contact theory that if it is the whole array of twenty features of the 'macrotype' which is diagnostic of the contact situation in question, then it must be the whole array that remains stable over the period in excess of 3000 years. If it is conceded that not all of the features need remain stable for that time, then it is in effect conceded that no causal link need be posited between the appearances of individual features in the individual languages. But then that unravels the argument for there being any causal link in the first place. I assume it can further be seen that this difficulty besets all variants of the AA/IC contact theory in general to a greater or lesser extant, not just in the extreme formulation of B.(ii): the gap of over 3000 years in attestation of the features between Old Egyptian and IC remains the same whatever variant explanation is adopted.

[57] It should be noted additionally that the arguments have an inherent circularity. Contact with a non-extant language (the *explicans*) is posited to explain the typological similarity of the IC and AA languages (the *explicandum*). The only empirical effect of ('evidence for') this language is the typological similarity of the IC and AA languages. The *explicandum* is the only evidence for the *explicans*: circularity. Contrast the structure of the theory of Proto-Indo-European. In that case, what is posited as *explicans*, to the *explicandum* of the grammatical similarity of the languages involved, is a theory of phonological and morphological correspondences, largely in the form of reconstructive and predictive algorithms, which lead deductively to a hypothetical 'proto-grammar'. The 'Proto-Indo-European language' itself is the realist interpretation of this hypothetical grammar (where there was a grammar, there must have been a language). But the 'Proto-Indo-European language' itself does not have the status of *explicans* within the theory: it is an interpretative result, not an explanation. It should not be necessary to state that the fact that it was *his-*

The comparison with the Balkans is instructive, negatively so for the AA/IC contact theory. It has been suggested that the feature complex for AA/IC compares favourably with that for the Balkan area.[58] But if we are identifying contact areas solely on the basis of feature complexes, then we must do so consistently. Looking *only* at the features and their distributions, without noting their geographical and chronological contexts, then for AA/IC, we would have to conclude that it is IC, the better representative of the feature complex, which is the causal trigger, the 'donor,' for the development of the feature complex in AA. It is only the geographical and chronological details that forbid us from positing this. We are therefore bound to include the geographical and chronological information in our analyses. And that vitiates the comparison with the Balkan area. The feature complex itself is *never* sufficient to posit a contact area. The information on geographical and chronological context is essential input for the establishment of a contact area. The Balkan languages form a contact area *because* they are in geographical and chronological proximity. It matters little whether the chosen language families are on neighbouring continents or at opposite ends of the earth. If they are separated by several thousand miles and, with regard to the appearance of the features, several centuries, if not millennia, then there is no area.

The geographical and chronological contexts of the AA/IC contact theory impose further restrictions on the proposals. There are more data available for these matters than appears to have been taken into account by the theory's advocates. It is not sufficient to wave a hand vaguely over the map of Europe and decree that first there was the contact language and then there was Indo-European. For a realist interpretation of the theory, there are more details that need to be clarified. We have seen that the contact language or languages must link the Egyptian of 3000 B.C. and IC of the early centuries A.D. This cannot be done without taking account of the where and the when. The link must by definition be a continuum, geographically and chronologically defined, with a direction, as a field of vectors of contact. There must be places and times of contact, and there must be paths of contact.

The first path of contact that can be ruled out is the one that would lead from North Africa and the Middle East, through Anatolia and then through Europe in a westerly direction. For the period 2000 B.C. - 100 B.C. we can see sufficiently clearly what is going on in Anatolia to be able to rule this out as one of the paths of the continuum. During this period we see the presence of various Indo-European languages ('Anatolian' Hittite, Luwian, Palaic, later Lycian, Lydian, Carian; otherwise also, Phrygian, Greek, Celtic) and non-Indo-European Hurrian and Hattic. None of these languages remotely show traces of the 'macrotype' in ques-

---

torically the language that was posited first is irrelevant for the logical structure of the theory which was developed.

[58] Gensler (1993: 6).

tion, and there is no trace of any Afro-Asiatic or Afro-Asiatoid continuum through Anatolia during this period.

The documentation for non-Mediterranean Europe in the period 3000 B.C. - 100 B.C. is obviously sparse, and restricted to the later end of the period. But we do know that during this period Indo-European must have been spreading somehow through the continent. It is a matter of sufficient controversy just how it was doing so to force that issue to be avoided here. But that this is exactly the period during which Europe was becoming Indo-European, or, by some models, those parts of it that were not already so, is not in dispute. If the Afro-Asiatoid contact continuum linking Egyptian with IC were to have been in the westerly direction through Europe, then it must have been in place already prior to the Indo-Europeanisation of Europe along the same path. This must therefore have been so before 3000 B.C. It is therefore doubly suspicious not only that IC only develops the features of the 'macrotype' over 3000 years later than they are seen in Egyptian (with the purely chronological problems discussed above), but also that no other Indo-European language so much as approaches the 'macrotype,' even though, in this model, by definition *every* Indo-European language of Europe must have come into contact with the same language or language family.

The points of the last two paragraphs appear to me to be decisive in excluding the east to west direction through Anatolia and Europe as the vector field of the AA/IC contact continuum. I shall go out on a limb and simply assume agreement that the vector of contact cannot have entered Europe directly through the Balkan or Italian peninsulas. To the north of these peninsulas, the problems of this model are identical with those of the preceding paragraph, and there seems to be sufficient material from c. 1400 B.C. in the Greek context and c. 800 B.C. in the Italian context to give us confidence that the reason we do not see an Afro-Asiatoid language in these places is because it was not there. The Balkans in the period 1400 B.C. - 100 B.C. give us Greek, Illyrian, Thracian, Dacian, Pannonian, Celtic (all IE), and Lemnian (non-IE, related to Etruscan).[59] By comparison the more restricted territory of the Italian peninsula in the first millennium BC is also rich in linguistic remains, which, however, include nothing remotely answering the genetic or typological demands of the AA/IC contact theory: very roughly south to north, Sicel (IE?), Greek (IE), Messapic (IE), Oscan-Umbrian (IE), Latin-Faliscan (IE), Etruscan (non-IE), South Picenian (IE), North Picenian (non-IE), Ligurian (IE), Celtic (IE; on the Celtic of northern Italy, see below), Venetic (IE), Raetic (non-IE, related to Etruscan).

The last available possibility, as far as I can see, is to posit the continuum as running primarily through North Africa, and establishing a European vector through the Iberian Peninsula. Since the case of IC is a phenomenon of the European far west, the Iberian vector would at least have the virtue of keeping the

---

[59] Undeciphered as it is, the language of the Minoan Linear A script could be Afro-Asiatic, and there have been attempts to decipher it on that assumption (Gordon 1982: 131-52). The presence of Afro-Asiatic in the eastern Mediterranean is not in dispute, however, and that does not constitute the evidence needed for a contact continuum through Europe.

language in the right general region. It would avoid all the difficulties mentioned in the preceding paragraphs of having the contact continuum run through the whole of Europe. It can be kept as an exclusive phenomenon of Western Europe.[60] The Iberian Peninsula is nearly as well covered for prehistoric data on linguistic geography as Italy. In the first millennium B.C., the following languages are clearly seen: 'Tartessian' (probably non-IE) in the south-west, Iberian (non-IE) in the east, Proto-Basque (non-IE) in the north, Lusitanian (IE) in the west and Celtic (IE), possibly of more than one variety, in central and north-western areas.[61] None of the non-Indo-European languages in question here are candidates for the contact language that is sought: the peninsula seems rather dominated by SOV typology. For the contact vector through the Iberian Peninsula to be plausible, therefore, we must assume that the contact language must have been present there at a time significantly earlier than the presence of the extant languages. This necessarily implies not only a 'coming' of the Indo-European Celts and Lusitanians,[62] but a 'coming' of, presumably, at least the Proto-Basques and Iberians also. There can have been no continuum of contact between North Africa and the British Isles if these languages, radically incompatible with the AA/IC 'macrotype,' had been blocking the way through the Iberian Peninsula. If there was a continuum of the required Afro-Asiatic or Afro-Asiatoid languages running through prehistoric Iberia in, say, the second millennium B.C., then we must say that all the languages that we actually see in Iberia in the first millennium B.C. got there shortly before their attestation and wiped out the AA contact language without a trace, both genetically and typologically. So where did they all 'come from'? As has already been mentioned, the notion of the 'arrival' of Indo-European, about which we have much information, is difficult enough to model. What do we then do with Iberian and Proto-Basque, to get them 'arriving' in Iberia in such a way as to obliterate all trace of another language, which we have only posited to save another theory? It is important to observe that Proto-Basque and Iberian are extant on both sides of the Pyrenees ('Proto-Basque' = 'Aquitanian').[63] They are still 'in the way' of the continuum north as well as south of the Pyrenees.

---

[60] Obviously, at no point do I address the question of the plausibility of this North-African continuum in its geographical, chronological and linguistic details. This I must leave to others.

[61] The historically visible presence of the Phoenicians in the Iberian Peninsula ensures the existence of Afro-Asiatic linguistic elements there also, but I assume that this is not what is meant by the advocates of the AA/IC contact theory.

[62] In line with certain fashions in archaeology and prehistory, some researchers would probably be uncomfortable with the terminology of 'comings' even for speakers of Celtic languages in the Iberian Peninsula. I would not be one of them, but in the interests of consensus, one must only realise that there must have been a time before which Celtic was not spoken there. It is in that earlier period that the AA/IC contact theorists must posit the presence of the contact language.

[63] Villar (2005) has proposed that the Basque presence in northern Spain is due to linguistic spread from north of the Pyrenees in historical times. This is not the place to examine his arguments critically (a desideratum, certainly, relying heavily as he does on data from ge-

The fact that the Gaulish of northern Italy of the first century B.C. already shows constructions with clause-initial compound verbs in accordance with the combination of Wackernagel's Law and Vendryes' Restriction[64] should give further pause to the AA/IC contact theorists, as VSO typology is one of the major cornerstones of the whole theory. We are not justified in asserting that the language of the inscriptions in question was characterised by VSO basic-word-order typology, nor that it was not. On the other hand, even if those instances are not realisations of full-blown VSO-typology, they do clearly show at least the protohistorical prerequisites for Insular Celtic VSO. While the observed existence of verb-initial syntax in Cisalpine Gaulish of the first century B.C., and Transalpine Gaulish in the first century A.D.,[65] is no certain evidence of basic order, nevertheless, in light of such instances, it hardly stands to reason either that Celtic verb-initial syntax is a phenomenon exclusively restricted to the far north-west, the British Isles. Note that Gaulish is also showing us genitive following head-noun (feature 2b),[66] non-referential relative particle (feature 3)[67] and infixing-suffixing alternation (feature 7).[68]

Gensler is aware of some of these features of Gaulish, though not verb-initial clause structure. He also interprets the instances at notes 18-19 as special relative verbal forms and those of note 37 as cases of polypersonality in the verb. That gives six out of the twenty features already present in Gaulish. It is worth emphasising also that the fragmentary nature of the attestation of Gaulish imposes limits on what could be seen there. We have no data in Gaulish for the structures of genitival or prepositional relative clauses (features 4a, 4b and 4c). There is no evidence that Gaulish had a definite article, so feature 8, the distribution of the article in genitive phrases, is irrelevant. There is no evidence either way for feature 10, verbal noun or infinitive. Features 12, prepositional periphrastic continuous, 13, 'DO'-periphrasis, 14, adverbial 'and ...' clauses, and 15, verbal noun/infinitive as main clause verb, are such that, without an extensive corpus, their absence from the fragmentary Gaulish corpus is not indicative of anything. Features 1, conjugated prepositions, and 16, initial consonant mutations, have been seen to be dependent in IC on the apocope rules of the fifth and sixth centuries. And features 9, non-concord of verb with plural subject, and 11, predicative particle, are recognised as very late developments in IC anyway. Making allowance for the limited nature of the corpus of Gaulish, and consider-

---

netics), but it can be mentioned that Villar's model of the prehistoric linguistic geography of Western Europe is maximally incompatible with the notion of a north-African substratum language there.

[64]   See above, notes 35 and 37.

[65]   As seen at La Graufesenque, *sioxti albanos panna extratuθ .xxx.* 'Albanos supplied additional vessels, 300' (Marichal 1988: 136). The instance is not unique as a case of Gaulish verb-initial construction (cf. Isaac 1996: 113-23).

[66]   E.g. *doiros segomari* 'Doiros [son] of Segomaros' (Lambert 2002: 352-3).

[67]   E.g. *dugiIonti-Io*, *toncsiIont-Io*, see notes 18-19.

[68]   In the contrasting positions of the suffixed enclitic relative particle in, e.g. *dugiIonti-Io* vs. the infixed pronoun of, e.g. *to-so[n]·ko[n]de*, see note 37.

ing therefore only the features that could reasonably be expected to be visible at
all in such circumstances, it is not six out of twenty that Gaulish scores, but six
out of ten, a higher percentage than Hebrew and Berber score over the whole
array of features (55% each).

In dismissing the difficulties for the theory that the presence of such features
in Gaulish creates, it is the argument from the mutual compatibility of the ge-
netic-internal and contact-external explanations on which Gensler relies,[69] on
which see note 32, and the main text at that place. I will not push the argument
from the figures of the last paragraph any further. The point is just that Gaulish,
with the features it *does* have, shows that they were already in place in a Celtic
language spread over a large territory of Western Europe in the last centuries
B.C., including northern Italy.[70] So either the substrate language was spread
throughout that territory around that time also – in which case, where is the in-
dependent, non-circular evidence for it? – or, as has been noted before, the hy-
pothesis of a substrate language as causal trigger for the presence of these struc-
tures in Celtic is a hypothesis too many, and the language was just not there,
anywhere, in the first place. This remains true even if the distinction is made
between the origins of the patterns in question and their generalisation. If it is
conceded by the AA/IC contact theorists that the structures were present in an-
cient Celtic anyway, independent of any substratum, then from what is the ar-
gumentative *need* derived that their generalisation *must* be due to contact? There
is none.

There are, presumably, many ways one might want to tweak the argument in
the light of the Gaulish data, and it is indicative of at least the hope of a realist
position that Gensler himself does not go very far in doing so. The most obvious
to me would be to argue that while the substrate language was originally spread
throughout Western Europe, passing on its features to Gaulish before dying out
in that territory, it survived long enough in the British Isles to continue influenc-
ing IC. The more or less extensive evidence for other languages in southern
France and northern Italy has already been noted, with no trace of the required
substrate language, which must therefore have been extinct in those places ear-
lier than the first century B.C. But this argument cannot get round the complete
invisibility of the language in Roman and early medieval Britain. We have al-
ready seen the difficulty of having the contact continuum linking Egyptian of
3000 B.C. and IC of 500 A.D. The difficulty is hardly less in the microcosm of
Western Europe itself, whereby we would have to have the invisible contact lan-
guage influencing Gaulish in the second century B.C., say, and IC between 500
and 1000 A.D. The oft-cited 'principle' that features can remain hidden in 'sub-
standard' forms of language 'for generations' must presumably be invoked at
some point here. But how many generations? To fit the theory into the chronol-

---

[69] Gensler (1993: 453).
[70] Lepontic, the non-Gaulish Celtic of northern Italy, extant several centuries earlier than Gaulish
      there, gives us three verbal clauses with non-initial verbs (cf. Eska and Evans 1993: 45).

ogy of the real extant data, we seem to have to be thinking in terms of enough generations to fill many hundreds of years. I have not seen a demonstration of the plausibility of this. Apparently, we must just take this on faith.

It seems clear that the AA/IC contact theory fails to provide the possibility of a realist interpretation, by which I mean an interpretation that locates the languages in question in a geographical and chronological context which can be consistently confronted with extant data, of which there are more than the proponents of the theory seem to have taken into account, for the histories and pre-histories of the languages in question and for the linguistic geography of prehistoric Europe. This is in addition to the unparalleled typological stability of the contact language over the extraordinary gap of more than 3000 years between the appearance of the features in extant Egyptian and their development in IC, and the remarkable fact that the contact language was able to make its presence felt in the IC languages over an extended period in their *extant* histories, well into the Middle Ages, without, over all this time, there having been the slightest trace in uncontroversial data[71] of the presence of this language anywhere in ancient Europe, let alone in medieval Britain or Ireland.[72]

I do not regard any of the arguments I have stated in this paper as being intrinsically new. Many of them may not have been explicitly stated before, but I believe that in such cases I have merely stated what has been clear as implicit

---

[71] It may be necessary to elaborate on the point about the quality of data. Whether or not such and such a toponym, or even group of toponyms, say, could be shown, with much imagination, to be similar to a Semitic lexeme, would be beside the point in the face of contextualised use of languages in considerable corpora in, say, Anatolia, Italy and the Iberian Peninsula. If no plausible path through the extant languages can be posited, then imaginitive etymologies of West-European toponyms would be neither here nor there. This is a point of principle, independent of the specific arguments for any such etymologies of European toponyms that may have been proposed. An etymology can never be used as specific evidence for a theory of a non-extant language, since that theory, including the assumptions which it incorporates, is the argumentative background upon which the etymology is predicated in the first place. E.g. the Indo-European etymology of OIr. *athair*, Eng. *father*, Lat. *pater*, Gk. πατήρ, Skt. *pitā́*, expressed as the reconstruction PIE *$ph_2t\acute{e}r$*, is not 'evidence for' Proto-Indo-European; rather it is the theory of Proto-Indo-European which provides the means for giving these words the etymology, a kind of historical explanation. Similarly, whether the etymology of the name of Ireland in Semitic *'iy-weri'um* 'Isle of Copper' (Vennemann 1998) is right or wrong has no bearing as 'evidence for' the AA/IC contact theory, since it is the theory of the presence of a Semitic language in ancient Western Europe which provides the theoretical basis for the etymology in the first place. In the context it is apposite to add that the etymology is wrong, because such a proto-form could only result in OIr. **Íriu*, not correct *Ériu*.

[72] Once again, the following statement should not be necessary, but I make it just in case. The question is not whether or not there were non-Indo-European languages in Britain and Ireland before Celtic languages were spoken there. Of course there were. The question is not even whether or not these influenced the typology of the Insular Celtic languages. Maybe they did, maybe they did not. The question is only and specifically whether these prior languages can be placed in a geographical and chronological continuum with the Afro-Asiatic languages. I have not seen any data or arguments which cause me to think that they can.

counter-arguments to the AA/IC contact theory for many decades. The need to state these things now arises from the fact that the theory has gained new advocates in some quarters in recent years, not from any genuinely new insights gained, on either side of the debate.

A point about which much seems to be made in the arguments for AA/IC contact is that it is unreasonable to demand 'solid proof' of some theory or other in order to accept it as a valid piece of theorising. Such meta-discursive observations in principle invite detailed discussion in their own right, but, at the end of this paper, in which I have tried to show the relationship of the AA/IC contact theory with real data, I believe I have almost come as far as need be in discussion of the theory itself. But I will add a few words on this meta-issue. The first sentence of the present paragraph was in my own formulation. Here is the idea in the words of an AA/IC contact theorist:

> It may seem good sober linguistic practice to say of a problem, this remains unproven and even unprovable. But to say that something is unproven and therefore to dismiss it out of hand is not good science.[73]

I do not recognise in this formulation the careful attitude of countless linguists who have worked diligently on the analysis of real data relating to the diachrony of the Insular Celtic languages, both the extant diachrony and the prehistoric diachrony accessible by the comparative method, and who have accepted that the rigour of the discipline, 'scientific' or however one wishes to characterise it, imposes constraints on what can or cannot be accepted as valid argumentation, postulates or results. I unreservedly include John Morris-Jones amongst those linguists, well aware that much of his own comparative-historical work[74] was flawed from the start and has been duly rejected in the meantime (but by no means all of it). As nominal originator of the AA/IC contact theory,[75] Morris-Jones presented a sound and reasonable hypothesis to explain a puzzling phenomenon. But Morris-Jones was writing at a time when there was no working theory of the diachrony of Insular Celtic phonology, or of the diachrony of initial mutations, or of the diachrony of Insular Celtic verbal syntax. These are all products of the twentieth century (work in progress, to be sure). At the time Morris-Jones put forward his hypothesis, the inscriptions in Iberian writing were undeciphered, and Celtiberian was not recognised as such. For all Morris-Jones could have known, if he had considered the possibility, the undeciphered inscriptions in question could have been instances of the very contact language that he implicitly postulated. Morris-Jones's observations were undoubtedly 'good

---

[73] Gensler (1993: 463).

[74] As embodied primarily in his *Welsh Grammar* (Morris-Jones 1913).

[75] In Morris-Jones (1900). The notion of connections between the Celtic and the Semitic languages are much older than that, but it was Morris-Jones's observations that initiated modern research into the matter in a way compatible with notions of typology and prehistoric linguistic contact.

science'. But it is essential also to recognise the way new data and new theories impinge on the old. We have theories of the diachrony of Insular Celtic phonology, mutations and verbal syntax, and we know that the inscriptions in the Iberian writing are definitely not in a language remotely compatible with the postulate of an Afro-Asiatoid contact language.

The position I have taken in this paper is certainly not that the AA/IC contact theory is unproven and unprovable, therefore to be rejected. On the contrary: it has been my intention to demonstrate as clearly as possible that it is simply wrong. But perhaps I am being too rigorous. Perhaps, in my insistence on details of argumentative logic and objective testing of the theory against real-world data, I am making unreasonable demands of a 'probabilistic' theory, by its nature unprovable, but no less valid for that. In case that is so, I may myself be allowed to end with an unprovable argument of plausibilty. I can hardly be criticised for doing so by those who insist that their own theories are of the same nature. The central empirical postulate of the AA/IC contact theory is that ancient and early medieval Europe, whether as a whole or in the west alone, was permeated by a language or group of languages, which, throughout a period of significantly more than 3000 years, much of it in the full light of extant linguistic history and geography, remained completely invisible and typologically, unwaveringly stable. How rigorous a 'proof' do we need of the implausibility of that?

Appendix 1

Fusion of adpositions and pronouns ('conjugated adpositions')

As a way of doing a spot check on the rarity of conjugated adpositions in languages of the world, I took six grammars off my shelves. In the case of Hungarian (Bánhidi and Jókay 1962), there was no randomness, since I already knew that Hungarian had pronominal suffixation to its adpositions. But in the case of the other five, I had not previously taken note of what structures the languages employed in this domain. The five languages were Amele (Roberts 1987), Itelmen (Georg and Volodin 1999), Tariana (Aikhenvald 2003), Udi (Schulze 1982), Yimas (Foley 1991). I then noted *post facto* that Hungarian and Yimas, both with the feature, were included in Gensler's own sample. So I checked two other languages, not in his sample, Kashmiri (Wali and Koul 1997) and Bella Coola (Davis and Saunders 1997).

It would be futile to pretend to any sort of statistical randomness in this sample. I clearly aimed at a spread over several continents, and I knew in advance that I could have taken 30 grammars of European languages, say, without finding any conjugated adpositions. And there would be my motivation for acquiring grammars of those particular languages in the first place (beyond interest). Still, if they were as rare as they would need to be to constitute 'amazing' evidence for AA/IC contact, I would be surprised that I can find four languages with them out of a sample of eight taken from my own bookshelves.

Hungarian (Finno-Ugric) (Bánhidi and Jókay 1962: 346):

alattam 'under me'	*alattunk*
alattad 'etc.'	*alattatok*
alatta	*alattuk*
etc.	

Tariana (Arawak) (Aikhenvald 2003: 228-9), e.g.:

[*di-wika-se*]	*di-wasa*	*di-swa*	*diha*	*ñaki-ne*
3sgfnf-on.top-LOC	3sgnf-jump	3sgnf-stay	ART	evil.spirit-FOC.A/S

'The evil spirit jumped on top of him.'

[*nu-dalipa*]	*pi-nu*	*pi-ema*
1sg-near	2sg come	2sg-stand

'Come and stay near me.'

Cf. cross-referencing prefixes and pronouns (id.: 122).

Yimas (Papuan) (Foley 1991: 109):

'*nampan*, *naŋkun*, and *nampayn* contrast with *kantk* in that their objects, when pronouns, must be in the form of the bound possessive prefixes.'

mpu-nampan	*mpu-naŋkun*	*mpu-nampayn*	*pun kantk*
3PL-for	3PL-toward	3PL-like	3PL-with
'for them'	'toward their house'	'like those'	'with them'

Cf. 179, bound possessive prefixes, (id.: 179); verbal cross-reference prefixes, (id.: 200).

Bella Coola (Salish) (Davis and Saunders 1997: 118):

ʔałps-ił	ʔał-cx
eat-we	Prep-her

'We're eating with her' / 'We ate with her'

ʔapsuł-Ø	ti-staltmx-tx	ʔał-tχʷ
reside-he	-chief-	Prep-them

'The chief lived with them'

Cf. deictic suffixes (id.: 86).

Appendix 2

The publication of Haspelmath, et al. (2005) has made available the largest yet sample of typological profiles of languages. The data can be used in connection with some of the typological dependencies referred to in this paper.

A. Correlations of Clause-Level and Noun-Phrase Level Word-Order Types

The data point in the same direction as indicated in footnote 12. Figures have been compiled using the 'Composer' programme on the CD accompanying the *Atlas*; the samples for each binary combination are not identical, languages showing no dominant order in any parameter have been ignored. 'V^+' = 'verb medial/final,' 'V^1' = 'verb-initial'; neither here nor in any analysis below have artificial means been used to ensure that rounded percentages for each feature or combination give totals of 100%; not all do:

Position of verb & order of noun and genitive (781 languages)

[V^+ & GN]	[V^+ & NG]	[V^1 & GN]	[V^1 & NG]
445 languages	243 languages	6 languages	87 languages
57% of sample	31% of sample	1% of sample	11% of sample
65% of V^+	35% of V^+	6% of V^1	94% of V^1
99% of GN	74% of NG	1% of GN	26% of NG

Position of verb & order of noun and adjective (850 languages)

[V^+ & AN]	[V^+ & NA]	[V^1 & AN]	[V^1 & NA]
224 languages	533 languages	23 languages	70 languages
26% of sample	63% of sample	3% of sample	8% of sample
30% of V^+	70% of V^+	25% of V^1	75% of V^1
91% of AN	88% of NA	9% of AN	12% of NA

Order of noun and genitive & order of noun and adjective (862 languages)

[GN & AN]	[GN & NA]	[NG & AN]	[NG & NA]
212 languages	289 languages	55 languages	306 languages
25% of sample	34% of sample	6% of sample	35% of sample
42% of GN	58% of GN	15% of NG	85% of NG
79% of AN	49% of NA	21% of AN	51% of NA

In all three combinations, it is the pattern which matches the AA/IC 'macro-type' ($[V^1$ & NG] as a percentage of V^1 languages, $[V^1$ & NA] as a percentage of V^1 languages, [NG & NA] as a percentage of NG languages) which scores the highest percentage points (except [NG & NA] as a percentage of NA languages). The data of Haspelmath, et al. (2005) can be combined further for a survey of the ternary correlation of position of verb & order of noun and genitive & order of noun and adjective. This gives a sample of 682 languages with recorded values for all three features, excluding cases, as above, where no dominant order is discernible in one or more feature (more detail is given here than above for clarity in the ternary comparison): V^+ 599 languages (88% of sample), V^1 83 languages (12% of sample), GN 382 languages (56% of sample), NG 300 languages (44% of sample), AN 199 languages (29% of sample), NA 483 languages (71% of sample). First of all, the binary comparisons are extracted from the new sample, for comparison with the percentages already given, as a control for the congruence of the now smaller sample with the larger and different, separate samples used above.

Position of verb & order of noun and genitive

$[V^+$ & GN]	$[V^+$ & NG]	$[V^1$ & GN]	$[V^1$ & NG]
377 languages	222 languages	5 languages	78 languages
55% of sample	33% of sample	1% of sample	11% of sample
63% of V^+	37% of V^+	6% of V^1	94% of V^1
99% of GN	74% of NG	1% of GN	26% of NG

Position of verb & order of noun and adjective

$[V^+$ & AN]	$[V^+$ & NA]	$[V^1$ & AN]	$[V^1$ & NA]
180 languages	419 languages	19 languages	64 languages
26% of sample	61% of sample	3% of sample	9% of sample
30% of V^+	70% of V^+	23% of V^1	77% of V^1
90% of AN	87% of NA	10% of AN	13% of NA

Order of noun and genitive & order of noun and adjective

[GN & AN]	[GN & NA]	[NG & AN]	[NG & NA]
159 languages	223 languages	40 languages	260 languages
23% of sample	33% of sample	6% of sample	38% of sample
42% of GN	58% of GN	13% of NG	87% of NG
80% of AN	46% of NA	20% of AN	54% of NA

So the different sample makes no significant difference to the proportions of binary types extant. The ternary combinations themselves are now given, with percentages of total languages in the new sample for each of the individual features and the binary combinations.

[V$^+$ & GN & AN]	[V$^+$ & GN & NA]	[V$^+$ & NG & AN]	[V$^+$ & NG & NA]
157 languages	220 languages	23 languages	199 languages
23% of sample	32% of sample	3% of sample	29% of sample
26% of V$^+$	37% of V$^+$	4% of V$^+$	33% of V$^+$
41% of GN	58% of GN	8% of NG	66% of NG
79% of AN	46% of NA	12% of AN	41% of NA
42% of [V$^+$ & GN]	58% of [V$^+$ & GN]	10% of [V$^+$ & NG]	90% of [V$^+$ & NG]
87% of [V$^+$ & AN]	53% of [V$^+$ & NA]	13% of [V$^+$ & AN]	47% of [V$^+$ & NA]
99% of [GN & AN]	99% of [GN & NA]	58% of [NG & AN]	77% of [NG & NA]

[V^1 & GN & AN]	[V^1 & GN & NA]	[V^1 & NG & AN]	[V^1 & NG & NA]
2 languages	3 languages	17 languages	61 languages
Negligible	Negligible	2% of sample	9% of sample
2% of V^1	4% of V^1	20% of V^1	73% of V^1
1% of GN	1% of GN	6% of NG	20% of NG
1% of AN	1% of NA	9% of AN	13% of NA
40% of [V^1 & GN]	60% of [V^1 & GN]	22% of [V^1 & NG]	78% of [V^1 & NG]
11% of [V^1 & AN]	5% of [V^1 & NA]	89% of [V^1 & AN]	95% of [V^1 & NA]
1% of [GN & AN]	1% of [GN & NA]	43% of [NG & AN]	23% of [NG & NA]

In the matter of correlations of clause-level word order and NP-level order, the AA/IC 'macrotype' is consistent with obvious general typological trends (verb-initial correlating with **[NG & NA]**), and therefore does not have any significant implications for contact between the two families. It would have been more noteworthy if the 'macrotype' were characterised by some sort of mismatch here (e.g. **[V^1 & GN & NA]**). But given the contrary extant facts, I leave open the question as to *how* noteworthy that would have been. None of the insights of this section of the appendix are original, but previously available observations have been shown to be corroborated by newly available data-sets.

B.   *Correlation of Verb-Initial Clause Order and Initial Mutations*

(Ad footnote 30 and main text thereto). I believe the diffidence with which I suggested this correlation (Isaac 1993: 12-13) is no longer appropriate. The reference that I gave at the time (Lehmann 1978: 23) was not to broad or detailed quantified data. But the data of Haspelmath, et al. (2005) allow a more detailed assessment, by means of the correlation of the feature of order of verb, subject and object with the feature of suffixing vs. prefixing in grammatical morphology. The sample available for this correlation includes 793 languages. Of these, 220 languages can be ignored as not showing a dominant word order and/or significant grammatical affixing. From the remaining sample of 573 languages the following statistical statements are derived, in which 'suffixing' means 'suffixing index > 60% of affixing index,' 'prefixing' means 'prefixing index ≥ 40% of affixing index' (terms as defined by Dryer 2005: 110-11, in Haspelmath, et al. 2005: figures compiled using the 'Composer' programme on the accompanying CD):

V^+ & suffixing	338 languages
V^+ & prefixing	170 languages
V^1 & suffixing	24 languages
V^1 & prefixing	41 languages

Any structurally further unspecified language is more likely to be suffixing (63%) than prefixing (37%).

Any structurally further unspecified language is much more likely to be verb-medial/final (89%) than verb-initial (11%).

However:

A verb-initial language is nearly twice as likely to be prefixing (63%) as suffixing (37%).

A verb-initial language is nearly twice as likely to be prefixing (63%) as a verb-medial/final language (33%).

A prefixing language is nearly three times as likely to be verb-initial (19%) as a suffixing language (7%).

No explanation for the correlation is derivable from these statistics, but a large-scale quantitative survey corroborates a significant correlation between verb-initial word order and prefixing. The initial mutations of IC are not prefixes, but, like prefixes, they are a subtype of word-initial morphological oppositions. When initial consonant mutations were grammaticalised in the fifth to sixth centuries in the already VSO IC languages, it seems reasonable to conjecture that a real linguistic trend was thereby being followed.

Appendix 3

(Ad 'The exoticism of IC VSO is exaggerated by the AA/IC contact theo-
rists,' main text at fn. 36). It has not frequently been noted in the context of his-
torical word-order studies of Indo-European, and never to my knowledge in the
Celtic context, that Old Russian and Old Serbian are strongly verb-initial in their
extant texts.[76] Berneker (1900: 1-16) surveys several Old Russian texts, and finds,
for example, "nicht weniger als 130" instances of verb-initial construction out of
160 verbal sentences in 25 pages of one Old Russian text.[77] Berneker's study
jumps from Old Russian to the modern language, so giving no view of the details
of word-order practice in the intervening centuries. But it seems reasonable to
consider texts from the interim in the same light. The following example has no
probative value, but is offered purely illustratively, in view of the general lack of
recognition of the phenomenon of Slavic verb-initial constructuion in the rele-
vant contexts. (From the *bylina* of *Volx Vseslav'evič*, edition and translation by
Jakobson and Szeftel (1966); verbs = v, subjects = s, direct objects = o; these are
shown without prejudice to any actual constituency relations in the syntax):

[*Obvernétsja*]$_v$ [*Vol'x*]$_s$ *jasnym sókolom* –
'Volx now turned into a bright falcon,

[*zvilsjá*]$_v$ [*on*]$_s$ *vysóko po podnébes'ju,*
'Soared high up beneath the vault of heaven,

[*poletél*]$_v$ [*on*]$_s$ *daléče vo čistó pole*
'He flew far away into open plains,

140   *ko svóej ko družíne xoróbryja.*
'To his brave retinue.

[*družína*]$_s$ [*spít*]$_v$*, tak* [*Vol'x*]$_s$ *ne* [*spít*]$_v$*,*
'The retinue sleeps, but Volx does not sleep.

[*razbudíl*]$_v$ [*on*]$_s$ [*udályx dobryx mólodcov*]$_o$*:*
'He awakened his bold and hardy valiants:

"Goj esí vy, družína xoróbraja.
'"Hail to you, my valiant retinue!

Ne vrémja spat', porá vstavát', –
'No time to sleep, 'tis time to rise,

145   [*pojdém*]$_v$ [*my*]$_s$ *ko cárstvu Indéjskomy."*
'Let's march to the Indian Realm."

I [*prišlí*]$_v$ [*one*]$_s$ *ko stené belokámennoj, –*

---

[76] A notable exception is Friedrich (1975: 61-3), the work which has drawn my attention to
     the phenomenon.
[77] For Old Serbian, see Berneker (1900: 38-9).

'And they reached the white-stone wall –

krepká stená belokámenna,
'That white-stone wall was strong,

voróty y góroda želéznyju,
'The city had gates of iron made,

krjukí, zasóvy vse médnyja,
'The hinges, bars all of copper,

150   [*stoját*]ᵥ [*karaúly*]ₛ *denný-noščný,*
'Sentinels stand on watch day and night,

[*stoít*]ᵥ [*podvorótnja*]ₛ – *dorog rýbej zub,*
'The gate has an undersill – costly walrus bone,

mydrény výrezy výrezeno,
'Ingenious slots cut out,

a i tól'ko v výrezy [*myrašý*]ₛ [*projtí*]ᵥ.
'And through the slot only an ant could pass.'

Apart from ll. 141 and 153, where there is transparent motivation for subject focus, this passage illustrates the narrative use of verb-initial construction. But it will correctly be wondered to what extent such a passage is representative. Given the apparent unfamiliarity of the Slavic VSO constructions outside of the field of Slavic linguistics, it seems legitimate to present a more detailed analysis as a case study. I have analysed the constituent orders for all finite clauses in the *bylina* of *Volx Vseslav´evič*. The accessability of the edition of Jakobson and Szeftel should make reproduction of the entire text unnecessary here. I have identified 120 finite clauses in this poem of 204 lines (in this edition). In the discussion that follows within this note, 'initial' means only with respect to the core constituents of verb, subject and object; use of the term does not imply that no material at all (e.g. particles, conjunctions, prepositional modifiers, etc.) precedes the 'initial' constituent of the clause in question. Insofar as it is specifically the relative order of subject and verb that is of primary interest, I ignore for convenience here 31 clauses which do not contain an explicitly expressed subject (18 with just **V**, 6 **VO** and 7 **OV**). I also ignore 9 clauses with imperative or hortative verbs (all initial), as this pragmatic mode is characteristically connected with verb-initial order in many otherwise non-verb-initial languages. Of the 80 clauses left, 43 are subject-initial, 37 verb-initial, in the lines as follows (listed in full for transparency and replicability): **SV** 4, 6-7, 10, 17, 20, 21, 22, 27 (×2), 40, 42, 60, 65, 69 (×2), 70, 76, 81 (×2), 82, 98, 99, 104, 109, 116, 117, 126, 128, 129, 136, 141 (×2), 153, 154, 159, 179, 195, 200; **SVO** 77, 188; **SOV** 167, 199, 202; **VS** 2, 12, 13, 15, 16, 39, 44, 58, 83, 100, 106, 110, 112, 118, 123, 124, 137, 138, 139, 146, 150, 151, 163, 165, 184; **VSO** 53, 73, 79, 84, 89, 142, 156, 162, 191; **OVS** 45, 47, 49 (these three lines in close sequence have three mutually

contrasting objects of the same verb with the same subject). Purely numerically then, subject-initial orders are marginally dominant. Objective criteria for assessing the motivation for this distribution would be welcome. These are provided by a number of easily accessible factors. In the following clauses, the parallelism of construction with different subjects gives transparent motivation for contrastive subject focus:

> 20   *rýba pošlá v morskúju glubinú,*
> 21   *ptíca poletéla vysóko v nebesá,*
> 22   *túry da oléni zá gory pošlí*
>
> 'Fishes went into the depth of the sea,
> 'Birds flew high heavenward,
> 'The aurochs and deer went o'er the mountains'

> 27   *Vól´x govorít, kak gróm gremít*
>
> 'Volx speaks as the thunder roars'

> 69   *Družína spít, tak Vól´x ne spít:*
>
> 'The retinue sleeps, but Volx sleeps not'   (+ 81, 141)

A second factor that can easily be examined is the recoverability of the referents of explicit subject pronouns. I tabulate all clauses with such pronouns in Table 5. Columns: A = line number; B = clause (those marked * have the subject both as a pronoun and as an explicit nominal, in left-dislocation or a related construction); C = number of lines previously the referent of the pronoun was named as an explicit nominal; D = number of lines previously a coreferent pronoun was used, if lower than the number in C. The subsections of the table are ordered by column C (ascending).

Initial subject pronouns (including the constructions with a dislocated subject of some sort) tendentially correlate with low recoverability of the referent (explicitly mentioned up to 23 lines previously, referred to pronominally up to 17 lines previously), subject pronouns following their verbs correlating with high recoverability (explicitly mentioned up to nine lines previously, but only with a pronominal mention in the immediately previous line; previous pronominal mentions otherwise no more than four lines back).

However, in view of the apparent lack of differentiation between subject-initial and verb-initial amongst the instances with the lower figures, some additional detail is called for on the observed patterns of distribution recorded in the table. In ll. 70 and 82, the subject pronoun takes up an explicit nominal referred to only in the previously line, implying high recoverability, contrary to the proposed trend. The passages are given below.

Table 5

A	B	C	D
Subject-Initial			
4	**Ona** s kámenju **skočíla** [na ljutá zmeja],	1	-
70	**ón obvernétsja** serým volkóm,	1	-
82	**ón obvernétsja** jasnym sókolom,	1	-
128	**on** té-to de réči **povýsplušal**,	2	-
202	**on** zláta-sérebra **výkatil**,	2	-
129	**ón obvernúlsja** gornóstalem,	3	1
188	**On berét** carjá za belý ruki,	4	-
76	**on** zájcam, lisícam ne **brézgival**.	7	3
65	A vtápory **Vól´x on** dogádliv **býl**:*	9	-
136	a vsé **on** v zémlju **zakápival**.	10	7
103 -4	A tút takovój **Vsesláv´evič**, **ón obvernétsja** gnedým tyróm,*	14	3
159	**Mólody Vól´x**, **on** dogádliv **býl**:*	18	17
109	**Ón obvernétsja** jasnym sókolom,	20	2
179	A sám **on Vol´x** vo paláty **pošél**,*	20	12
126	A **vtápory Vól´x**, **on** dogádliv **býl**:*	23	14
Verb-Initial			
53	**stál** sebe **Vól´x on** družínu pribirat´:*	1	-
79	**nosíli one** šúby sobolínyja,	1	-
138	**zvilsjá on** vysóko po podnébes´ju,	1	-
142	**razbudíl on** události dobryx mólodcov:	1	-
162	**Prošlí one** sténu belokámennu,	1	-
156	**Govorját** [**oné**] takovó slovo:	2	-
83	**poletél on** daléče na síne more,	2	1
139	**poletél on** daléče vo čistó pole	2	1
84	a **b´ét on** guséj, belyx lébedej,	3	1
106	**pobežál on** ko cárstvu Indéjskomu:	3	2
73	A **b´ét on** zvéri soxátyja,	4	3
146	I **prišlí one** ko stené belokámennoj,	4/5	4
110	**poletél on** ko cárstvu Indéjskomu.	7	1
89	A **stál on Vólx** vražbú činít´:*	8	2
112	i **sél on** na paláty belokámenny,	9	1

69	*Družína spít, tak Vól´x ne spít:*	$S^1 V$	$S^2_i V$
70	*ón obvernétsja serým volkóm*	$^{pro}S_i V$	

'The retinue sleeps, but Volx sleeps not:
'He turned into a grey-haired wolf ...'

81	*Družína spít, tak Vól´x ne spít:*	$S^1 V$	$S^2_i V$
82	*ón obvernétsja jasnym sókolom*	$^{pro}S_i V$	

'The retinue sleeps, but Volx does not sleep:
'He turns into a bright falcon ...'

In both cases, there are two nominal subject referents in the preceding line. The pronoun therefore picks out just one of them. The gender of *on* 'he' is in principle adequate for disambiguation, but the selecting focus of the pronoun is enhanced and therefore clarified by initial placement. This contrasts with the roughly parallel lines (quoted above):

141	*Družína spit, tak Vol´x ne spit,*	$S^1_i V$	$S^2_j V$
142	*razbudíl on udályx dobryx mólodcov*	$V {}^{pro}S_j O_i$	

'The retinue sleeps, but Volx does not sleep,
'He awakened his bold and hardy valiants ...'

The pronoun *on* again picks 'Volx' as the subject of the second line, but in this case, the other nominal of the first line is also taken up in the second: the 'bold and hardy valiants' are the same as the 'retinue,' and their role in the clause of the second line is syntactically explicit. Consequently, there is no need for contrastive focus on the subject pronoun, which is thus post-verbal. So, although the figures for reference recoverability recorded in the above table come out the same for 70/82 vs. 142, the cases are not identical.

The clauses of ll. 126-9 also have high recoverability values apparently in conflict with the trend for subject-initial construction to be used for low recoverability. Comparison of the lines in their contexts, 126-9, with 81-4, with similar recoverability values, is instructive:

81	*Družína spít, tak Vól´x ne spít:*	$S^1 V$	$S^2_i V$
82	*ón obvernétsja jasnym sókolom,*	$^{pro}S_i V$	
83	*poletél on daléče na síne more,*	$V {}^{pro}S_i$	
84	*a b´ét on guséj, belyx lébedej*	$V {}^{pro}S_i O$	

'The retinue sleeps, but Volx does not sleep:
'He turns into a bright falcon,
'He flew afar to the blue sea,
'And he strikes the geese, the white swans ...'

126	*A vtápory Vólx, on dogádliv býl:*	S_i $^{pro}S_i$ V
127	*sídjuči na okóške kosjáščatom,*	–
128	*on té-to de réči povýsplušal,*	$^{pro}S_i$ V
129	*ón obvernúlsja gornóstalem*	$^{pro}S_i$ V

'And at that time Volx, he was resourceful:
'While perching on the small window in a wooden frame,
'He listed to these very speeches,
'He turned into an ermine ...'

The verb-initial clauses of 83-4 follow the subject-initial clause of 82, with its pronoun in focus for the reasons already discussed above. With such prominence given to the referent over the two lines 81-2, by 83, it is highly recoverable, and the text reverts to verb-initial order, with no special reason to front the pronoun. By contrast, ll. 128-9 are separated from the last nominal and pronominal mention of the referent of the pronominal subjects by a line without mention of that referent at all. There is therefore a small but significant difference in the recoverability of the referent of the pronouns in ll. 128-9 as compared with ll. 83-4, a difference registered in the subject-initial construction of ll. 128-9.

Two other lines in the table have the referent of the pronoun explicitly mentioned two lines previously with no pronominal mention in between, 202 (subject-initial) and 156 (verb-initial). The three-line passages in question are as follows:

154	*I vsé molodcý zakručínilisja,*	S_i V
155	*zakručínilisja i zapečálilisja.*	V V
156	*Govorját [oné] takovó slovo*	V $^{pro}S_i$ O

'And all the valiants became worried,
'Became worried and grieved.
'They voice this sort of speech …'

200	*A i mólody Vól´x tut carém nasél,*	S_i V
201	*a to stáli – ljúdi posádskija:*	– (different subject)
202	*on zláta-sérebra výkatil*	$^{pro}S_i$ O V

'And now young Volx enthroned himself as Tsar,
'And his valiants became townfolk:
'He rolled out many kegs of gold and silver for them ...'

Between ll. 156 and 154, there is a line of two finite verbs still with the same subject as explicitly mentioned in l. 154, the subject taken up pronominally in l. 156, which is therefore maximally recoverable, so verb-initial construction (the pronoun of l. 156 of this edition is admittedly an editorial insertion; it is consistent with the analysis proposed here, but no significant change in the analysis would

be forced by the omission of the pronoun). By contrast, between ll. 202 and 200, a line with completely different subject intervenes, interrupting the recoverability of the referent in l. 202 and motivating fronting focus on the pronominal subject there (the fronting of the object in l. 202 is motivated by parallel contrast with the object, also fronted, in the subsequent l. 203).

All but one of the subject-initial clauses in the table above which have apparently high recoverability values for their pronominal subjects (ll. 70, 82, 128, 129, 202) are therefore seen to have their construction conditioned by additional factors interfering with the recoverability in terms of distance from the referent in lines alone, in contrast to clauses with similar recoverability values for their pronominal subjects in verb-initial construction. The exception in l. 4, *Ona s kámenju skočíla [na ljutá zmeja]* 'She leaped off a stone onto a serpent fierce,' is not obviously explicable in one of the ways discussed so far, but the depiction of the princess leaping onto the snake is, after all, the act which, by leading to the birth of Volx, initiates the events of the whole narrative. 'She,' as subject of this clause, therefore carries a considerable burden of semantic and pragmatic implications, which may well justify sharp foregrounding, and so initial position. I shall assume this is so until it may be demonstrated otherwise.

One final passage will be chosen for special comment with regard to referent recoverability, ll. 141-6, which have already been quoted above. In l. 146, verb-initial, but apparently with quite a distance between it and the nearest previous explicit mention of the referent(s), the pronoun *one* 'they' actually refers to both 'Volx' and his 'retinue' of 'hardy valiants'. Though 'they' have not been mentioned in exactly that way in the immediately preceding lines, 'they' are semantically present in the discourse as 'we' within Volx's speech, directly quoted, and no other argument has intervened as subject in these lines. So recoverability is high. So verb-initial construction in l. 146.

A third criterion which is easily accessible in the analysis of the motivation for different clause types is that of the meanings of the verbs involved. I note that the verbs in subject initial construction in this *bylina* have the following meanings (including negation):

berét 'seized' (188); *ne brézgival* 'disdained not' (76); *govorít* 'speaks' (27, 117); *gremít* 'roars' (27); *narjažáetsja* 'starts outfitting' (60); *nasél* 'enthroned' (195, 200); *obvernétsja* 'turns into' (70, 82, 104, 109); *obvernúlsja* 'turned into' (129); *obviváetsja* 'winds himself' (6-7); *obyvál-onevál* 'shod and clothed' (78); *otdaét* 'gives (order)' (167); *pereženílisja* 'took' (199); *poíl-kormil* 'regaled' (77); *poletéla* 'flew' (21); *ponós poneslá* 'conceived' (10); *pošlá/pošlí/pošél* 'went' (20, 22, 40, 42, 179); *povýsplušal* 'listened' (128); *priklonjáetsja* 'bows down' (99); *pristiláetsja* 'flatten out' (98); *projtí* 'passes' (153); *skočíla* 'leaped' (4); *skolybálosja* 'billowed' (17); *spít* 'sleeps' (69, 81, 141); *ne spít* 'sleeps not' (69, 81, 141); *tjanut* 'blow' (116); *výkatil* 'rolled out' (202); *zakápival* 'buried' (136); *zakručínilisja* 'became worried' (154).

(Three instances of the past-tense copula *býl* in 65, 126, 159, present a different case, insofar as the actual predicates of the sentences in question are the adjectives dependent on the copula, not the copula itself; these cases will not be considered further).

All these verbs denote an action or process prototypically entailing volition on the part of the subject; not necessarily that the subject wills the action or process, as in 'seized, speaks,' etc., but that volition is implied as belonging to the subject, even if that volition is not itself directed to the initiation of the action or process: 'sleeps' can hardly be said to be an action prototypically resulting from the will of the subject, but it does assume that the subject is in possession of will, which, by definition, is suspended in sleep. This analysis includes several figurative usages, either by grammaticalisation, as in *gramota Vólxu v naúk pošlá* (40) lit. 'reading went into knowledge for Volx,' i.e. 'Volx mastered reading,' or metaphorically, as in *gróm grémit* (27) 'thunder roars'.

The verbs in verb-initial construction present a slightly different picture:

b´ét 'strikes' (73, 84); *govoríla* 'spoke' (118); *govorít* 'spoke' (184, 191); *govorját* 'spoke' (156); *nosíli* 'wore' (79); *obernul* 'turned' (165); *obvernétsja* 'turned' (137); *otdavála* 'sent' (39); *otvečájut* 'reply' (100); *pobežál* 'scampered' (106); *podrožála* 'trembled' (15); *poletél* 'flew' (83, 110, 139); *poučílsja* 'was instructed' (44); *prišlí* 'reached' (146); *prošlá* 'travelled' (58); *prošlí* 'crawled' (162); *prosvetjá* 'shone forth' (12, 123); *razbudíl* 'wakened' (142); *rodílsja* 'was born' (13, 124); *sél* 'alighted' (112); *stál* 'started' (53, 89); *stáli* 'stood' (163); *stoít* 'stands' (151); *stoját* 'stand' (150); *strjaslósja* 'shook' (16); *učílsja* 'studied' (45, 47, 49); *xodíla-guljála* 'walked-meandered' (2); *zvilsjá* 'soared' (138).

There is no need to note in detail the verbs which here again entail volition in the subject, but in contrast to the verbs in subject-initial construction, the verbs in verb-initial construction include the following:

prosvetjá 'shone forth' (12, 123)
rodílsja 'was born' (13, 124)
podrožála 'trembled' (15)
strjaslósja 'shook' (intransitive) (16)
stáli/*stoít*/*stoját* 'stood/stands/stand' (150, 163, 151)

These nine verbs (out of the 37 in verb-initial construction in total) do not entail volition on the part of the subject, but denote processes or positional states quite independent of the fact of the subject's possession of volition or not (including the reflexive-as-passive *rodílsja* 'was born'; he who was born is, ultimately, possessed of will, but that fact has little if anything to do with the process of his being born itself). This appears to be another objectively accessible difference between subject-initial and verb-initial construction in the text in question: only verbs entailing volition on the part of the subject (regardless of

whether that volition is a prerequisite of the action or process of the verb itself) can take initial subjects. It is possible, and should probably be assumed for argument's sake, that over a suitably larger sample, one would find a wider variety of verb-types in both constructions, with non-subject-volition-entailing verbs in subject-initial construction also. But the distribution of the verb-types in the *bylina* of *Volx Vseslav´evič* remains a notable fact.

There is clearly justification for pursuing such analyses further. Since I shall not do so here, 'basic' VSO clause structure in this Russian text, as in the rest of the relevant corpus, may legitimately be regarded as an incompletely tested hypothesis, rather than as an uncontroversial fact. Be that as it may, my investigation, such as it has been, justifies the statement of that hypothesis in the following form. Of the 43 subject-initial constructions in the *bylina* of *Volx Vseslav´evič*, twelve are motivated by transparent contrastive subject focus (20, 21, 22, 27 (×2), 69 (×2), 81 (×2), 141 (×2) and, with subject pronoun (4), and fifteen, with subject pronouns, by the relatively low recoverability of the referent of those pronouns (4, 65, 70, 76, 82, 103-4, 109, 126, 128, 129, 136, 159, 179, 188, 202). Furthermore, initial position is reserved for those subjects with referents characterised by volition (whether literally or figuratively). Granted that I have not provided any analysis of the motivation for the remaining sixteen instances of subject-initial construction, nevertheless, on the evidence of the twenty-seven that I have analysed, a pattern of subject focus emerges. All these data indicate that the pre-verbal position is a position of marked focus. The proper domain for verb-initial order in some types of Russian (texts) is thus negatively defined: the verb comes first when there is no pressing reason for anything else to come first. This is as close to a definition of a 'basic' order as may be practical or necessary. Compare the following formulation: "The Basic Sentence Type in a language is that in which the Subject is old or topical, and the focus of new information falls on the Predicate" (Hopper 1986: 124; capitals *sic*). 'Topical' here corresponds with 'maximally recoverable,' as I have put it. And in connection with the textual approach I have taken, note Hopper's warning: "It is only from discourse that we can tell what is 'new,' what is 'old,' what is 'contrastive,' and so on; without textual analysis, the data base for a typology is suspect" (*ibid.*, 125).

As stated at the outset, the discussion of the *bylina* of *Volx Vseslav´evič* is purely illustrative. I have no information at present in what detail the text-grammatical motivation for the word-order patterns of Russian texts of this type and of earlier periods has been analysed. But the references already given show that the phenomenon of verb-initial construction in some phases of Russian is a matter of record, not something for which I have to plead.

There is obviously some difference between the relative freedom of Slavic sentence-structure in general, with verb-initial order favoured in some periods in some traditions, and the 'rigid' verb-initial syntax of IC. It is neither implicit in, nor a prerequisite of, my argument that some varieties of Slavic are to be analysed as 'VSO languages'. But it *is* implicit in, and a prerequisite of, the arguments for AA/IC contact that IC VSO as a structural feature is exotic and iso-

lated within the range of Indo-European languages, something that cannot be explained in Indo-European terms, and to explain which a non-Indo-European substrate language must therefore be invoked. With a substantial corpus of early Russian literature formulated in a 'basic' VSO order (regardless of the overall 'type' of the language), IC VSO order looks considerably less 'exotic' and 'isolated' than is sometimes assumed or claimed.

Appendix 4

The 'cumulative' argument for AA/IC contact from the feature 'ensemble' was formulated above as follows:

(Major Premiss)   If *A* and *B* are two genetically unrelated languages and have the features *w*, *x*, *y*, *z*, then the presence of those features in the respective languages must be due to a causal link through contact.

(Minor Premiss)   *A* and *B* are two genetically unrelated languages and have the features *w*, *x*, *y*, *z*.

(Conclusion)   Therefore the presence of those features in the respective languages must be due to a causal link through contact.

This formulation can be simplified. We assume that '*A*' and '*B*' stand for two specific, genetically unrelated languages. We are talking about the possession of features *w*, *x*, *y*, *z* by both *A* and *B*.[78] And we are assuming that the possession of features *w*, *x*, *y*, *z* by both *A* and *B* implies that *A* and *B* are linked by a continuum of language contact.[79] Using '$\rightarrow$' as usual to symbolise the conditional 'if ... then ...,' we may put the whole argument, fairly informally still, thus:

(Major Premiss)   *A* and *B* have (*w*, *x*, *y*, *z*) $\rightarrow$ *A* and *B* must be linked by contact.

(Minor Premiss)   *A* and *B* have (*w*, *x*, *y*, *z*).

(Conclusion)   Therefore: *A* and *B* must be linked by contact.

This is just *modus ponens*.

Using '$\wedge$' to symbolise the conjunction, obviously, the proposition [*A* and *B* have (*w*, *x*, *y*, *z*)] is compounded of

[(*A* and *B* have *w*) $\wedge$ (*A* and *B* have *x*) $\wedge$ (*A* and *B* have *y*) $\wedge$ (*A* and *B* have *z*)]

[*A* and *B* have (*w*, *x*, *y*, *z*)] cannot be true unless (*A* and *B* have *w*), (*A* and *B* have *x*), (*A* and *B* have *y*) and (*A* and *B* have *z*) are all true also.

Since the four propositions of the conjunction all have the same arguments, *A* and *B*, it will be convenient to symbolise them more simply as '***W***,' '***X***,' '***Y***,' '***Z***': i.e. let '***W***' stand for '*A* and *B* have *w*', etc. And let '***C***' stand for '*A* and *B* must be linked by contact'.

---

[78] I am not using any established notational convention for assigning symbols to proposition or variable types: the notation is arbitrary and irrelevant.

[79] It is irrelevant to the formalisation which of the models of this continuum discussed in the body of this paper is thought to be correct.

So the argument has the form:

(Major Premiss) $(W \wedge X \wedge Y \wedge Z) \rightarrow C$

(Minor Premiss) $W \wedge X \wedge Y \wedge Z$

(Conclusion)   Therefore: C

The form of this argument is undoubtedly valid.

The form of the conditional $[(W \wedge X \wedge Y \wedge Z) \rightarrow C]$ says that the truth of C is guaranteed by the truth of $(W \wedge X \wedge Y \wedge Z)$ (the 'ensemble' of statements that languages A and B have the features w, x, y, z). It seems from this form of the argument that one could say that the truth of W alone cannot guarantee the truth of C, because the truth of W does not guarantee the truth of X, Y and Z. So $(W \rightarrow C)$ does not follow from $[(W \wedge X \wedge Y \wedge Z) \rightarrow C]$. So possession by the languages A and B of the individual feature w does not guarantee the truth of C. Only the 'ensemble' of features can do that. Only the 'ensemble' is diagnostic, not the individual features. This is the apparent logical form of the cumulative argument.[80] It makes no difference whether the argument is demonstrated on the basis of four features, as in this illustration, or of twenty, as in Gensler's twenty features.

If fact, the same logical basis is represented if we formulate the argument with only two terms in the conjunction. We are therefore formulating the argument as follows:

(Major Premiss) $(X \wedge Y) \rightarrow C$

(Minor Premiss) $X \wedge Y$

(Conclusion)   Therefore: C

The form of the conditional $[(X \wedge Y) \rightarrow C]$ says that the truth of C is guaranteed by the truth of $(X \wedge Y)$, but not by the truth of X alone. The cases of $(X \wedge Y)$ and $(W \wedge X \wedge Y \wedge Z)$, and any number of conjoined propositions, can be generalised using the 'product' symbol applied to generalised propositions p:

$$\prod_{i=1}^{n} p_i$$

This is defined as $(p_1 \wedge p_2 \wedge \ldots p_{n-1} \wedge p_n)$, where n can be defined as any number. Thus, $(W \wedge X \wedge Y \wedge Z)$, in effect $(p_1 \wedge p_2 \wedge p_3 \wedge p_4)$, has $n = 4$, whereas $(X \wedge Y)$, in effect $(p_1 \wedge p_2)$, has $n = 2$. The twenty-feature array with which

---

[80] For argument's sake I shall assume that this is the case in what follows. But this assumption will ultimately have to fall.

Gensler works would take its place in the generalised argument with $n = 20$.[81] The *modus ponens* argument we have been looking at is therefore generalised to conjunctions of any size as (with q arbitrarily as the consequent of the conditional):

(Major Premiss) $(\prod_{i=1}^{n} p_i) \rightarrow q$

(Minor Premiss) $\prod_{i=1}^{n} p_i$

(Conclusion) Therefore: q

This is the generalised form of the cumulative argument. It is only the full conjunction of n propositions that guarantees the truth of C. No individual proposition p_i is sufficient to do this, or so we would have to believe, in order to accept the argument of the AA/IC contact theorists. But, although the form of *modus ponens* just given is undoubtedly a valid argument form, there is a problem in this for the AA/IC contact theory.

Although we cannot infer the individual propositions $(W \rightarrow C)$, $(X \rightarrow C)$, $(Y \rightarrow C)$ or $(Z \rightarrow C)$ from $[(W \wedge X \wedge Y \wedge Z) \rightarrow C]$, we can infer the disjunction $[(W \rightarrow C) \vee (X \rightarrow C) \vee (Y \rightarrow C) \vee (Z \rightarrow C)]$ from it.[82] This means that, in order for $[(W \wedge X \wedge Y \wedge Z) \rightarrow C]$ to be true, at least one of $(W \rightarrow C)$, $(X \rightarrow C)$, $(Y \rightarrow C)$ or $(Z \rightarrow C)$ must be true.[83] In terms of the AA/IC contact theory, at least one

---

[81] The actual order of the propositions is irrelevant: 1, 2, 3, etc., are a purely numerical convention with no argumentative import.

[82] Or a disjunction of as many conditionals as correspond to the propositions in the original conjunction, four in the illustration here, twenty in the case of Gensler's argument from the twenty features. Since all disjunctions and conjunctions can be reduced to two terms anyway, the proof can be simplified by doing so. '$\vee$' = 'either ... or ... or both'; '$\equiv$' = 'is semantically equivalent to,' '$\sim$' = negation:

$[(X \wedge Y) \rightarrow C]$ $\equiv$ $[\sim(X \wedge Y) \vee C]$
         Material Implication, $(p \rightarrow q) \equiv (\sim p \vee q)$.
       $\equiv$ $[(\sim X \vee \sim Y) \vee C]$
         One of *De Morgan's Theorems*, $\sim(p \wedge q) \equiv (\sim p \vee \sim q)$.
       $\equiv$ $[(\sim X \vee C) \vee (\sim Y \vee C)]$
         Association, $[p \vee (q \vee r)] \equiv [(p \vee q) \vee r]$ (recursively), *Idempotent*, $p \equiv (p \vee p)$, and *Commutation*, $(p \vee q) \equiv (q \vee p)$.
       $\equiv$ $[(X \rightarrow C) \vee (Y \rightarrow C)]$
         Material Implication, $(p \rightarrow q) \equiv (\sim p \vee q)$.

[83] Since the general form of the 'cumulative argument' has been given, the general form of the equivalence in question here will be given for completeness. Using the summation sign for the generalised disjunction, with the same definitory conventions as for the use of the product sign for generalised conjunction, the equivalence in question is as follows:

feature of the array must be individually diagnostic of contact. The following formulation was adopted above: 'Possession by the languages *A* and *B* of the individual feature *w* does not guarantee the truth of ***C***. Only the "ensemble" of features [*w*, *x*, *y*, *z*] can do that.' We now see that this formulation is incomplete. While possession by the languages *A* and *B* of the feature *w* does not necessarily guarantee the truth of ***C***, we see now that, in fact, *at least one* of the individual features *w*, *x*, *y*, *z* must do that. If not even at least one of those features is individually diagnostic of contact, then all the conditionals of the disjunction $[(W \rightarrow C) \vee (X \rightarrow C) \vee (Y \rightarrow C) \vee (Z \rightarrow C)]$ are false, therefore the disjunction itself is false, in which case, the conditional of the 'cumulative' argument, $[(W \wedge X \wedge Y \wedge Z) \rightarrow C]$, is also false. We see then that, though the 'cumulative' argument, as shown above, undoubtedly has a valid logical form, it is, nevertheless, in fact redundant. In order for any 'cumulative' argument of the general form,

(Major Premiss) $\qquad (\prod_{i=1}^{n} p_i) \rightarrow q$

(Minor Premiss) $\qquad \prod_{i=1}^{n} p_i$

(Conclusion) $\qquad$ Therefore: q

to be valid, at least one of the propositions p_i must imply q individually anyway. If not, then the disjunction corresponding by the proof of note 82 is false, therefore, the Major Premiss of the 'cumulative' argument is false, and we cannot deduce q. But if one of the propositions p_i individually implies q anyway, then there is no need for the argument from the conjunction in the first place. The 'cumulative' argument would thus be valid, but irrelevant.

It is apparent that the reason for discussing the nature and structure of the 'cumulative' argument from the 'ensemble' of features is because it is recognised by the AA/IC contact theorists that no single one of the features in question *is* sufficient to imply that there was contact, no single one of those features is individually diagnostic. But if no single feature is individually diagnostic, then no diagnostic 'ensemble' can be derived, and the argument collapses. In the formulation used towards the beginning of this paper: 'How do twenty inconclusive arguments add up to one conclusive argument?' They do not.

There is a simpler way out of the impasse as to how to motivate the diagnosticity of exactly *this* 'ensemble' in the cumulative argument. All discussion in this ap-

---

$$(\prod_{i=1}^{n} p_i) \rightarrow q \quad \equiv \quad \sum_{i=1}^{n} (p_i \rightarrow q)$$

The left-hand side is true if, and only if, the right-hand side is true. The right hand side is true if, and only if, *at least one* of the implications $(p_i \rightarrow q)$ is true. This is the general form of the statement that the 'ensemble' of features of the AA/IC 'macrotype' is diagnostic of AA/IC contact if, and only if, at least one of the individual features of the 'macrotype' is diagnostic of that contact on its own, which, it seems to be agreed, none of them are.

pendix has been of the question of how to establish the criteria for diagnosticity of the common features of the languages for contact between those languages. And this question is predicated on the *assumption* that the following is a valid conditional: 'If *A* and *B* are two genetically unrelated languages and have the features *w*, *x*, *y*, *z*, then the presence of those features in the respective languages must be due to a causal link through contact'. The difficulties in establishing the criteria for diagnosticity can be removed simply by rejecting this assumption. In that case, the presence of some set of common features in two genetically unrelated languages is just not indicative of a causal link through contact.[84]

It is noteworthy that so much effort should have been invested in a theory based on an argument which is ultimately, by its very nature, invalid. A possible source of uncertainty in this matter may have been the confusion of definition with discovery. For instance, it is, of course, the case that the 'macrotype' in question is *defined* by the full array of twenty features together, not by the individual features. Therefore, it is the fact that the AA and IC languages display the full 'ensemble' of features which is 'diagnostic' of their representing the 'macrotype'. This is diagnosticity of their conforming to a certain *definition*. Neither this conformity, nor the definition of the 'macrotype' itself are in question (even if the problematic nature of the respective degrees of conformity of the languages is). The nature of the 'macrotype' as a *defined* term (not a *discovered* fact) is signalled throughout this paper by giving the word in inverted commas. However, the diagnosticity which is sought in the argument for the AA/IC contact language is of discovery, not of definition. The question is not whether we can *define* the contact language from the array of AA/IC shared features,[85] but whether we can validly deduce its very existence from them, a matter of *discovery*, not of *definition*. Consequently, the deductive arguments involved must themselves conform to the usual rules of inference. And that leads to inconsistencies, as above.

It is unusual for such a detailed formal exposition of the structure of an argument to be given in the discipline(s) with which this paper is concerned. That is probably a good thing. But there are cases where a certain 'rhetorical orientation' of an argument (whether consistent with some 'intuition' or not) can obscure the underlying logical structure. In such a case, a more formal exposition to clarify the arguments cannot be objectionable.

References

Aikhenvald, A.Y., 2003, *A Grammar of Tariana, from Northwest Amazonia*, Cambridge: Cambridge University Press.

---

[84] This is not to say that they cannot arise through contact, only that their mere presence, in any combination, is not sufficient to indicate that.

[85] We can indeed thereby define it typologically to a certain extent. But that is irrelevant.

Ashton, E.O., 1947, *Swahili Grammar (Including Intonation)*, London: Longman. (2nd ed.).

Bánhidi, Z. & Z. Jókay, 1962, *Ungarisch über Ungarn. Lesebuch und Grammatik für Fortgeschrittene*, Budapest: Verlag für Lehrbücher.

Beekes, R.S.P., 1995, *Comparative Indo-European Linguistics. An Introduction*, Amsterdam, Philadelphia: John Benjamins Publishing Company.

Berneker, E., 1900, *Die Wortfolge in den slavischen Sprachen*, Berlin: B. Behr.

Comrie, B., 1976, *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*, Cambridge: Cambridge University Press.

Davies, P.W. & R. Saunders, 1997, *A Grammar of Bella Coola*, Missoula, MT.

Dryer, M.S., 2005, "Prefixing versus Suffixing in Inflectional Morphology," in: Haspelmath, M. et al., 2005, 110-13.

Eska, J.F., 1990, "Gaulish *To-so·KoTe* (Vercelli)," in: *Études Celtiques* 27: 193-5.

Eska, J.F. & D.E. Evans, 1993, "Continental Celtic," in: Ball, M.J. & J. Fife, eds., 1993, in: *The Celtic Languages*, London, New York: Routledge, 26-63.

Eska. J.F. & M. Weiss, 1996, "Segmenting Gaul. *tomedeklai*," in: *Studia Celtica* 30: 289-92.

Falileyev, A., 2000, *Etymological Glossary of Old Welsh*, Tübingen: Niemeyer.

Foley, W.A., 1991, *The Yimas Language of New Guinea*, Stanford: Stanford University Press California.

Friedrich, P., 1975, *Proto-Indo-European Syntax*, Bute, Montana: Montana College of Mineral Science and Technology.

Gensler, O., 1993, "A Typological Evaluation of Celtic/Hamito-Semitic Syntactic Parallels," unpubl. PhD dissertation, University of California, Berkeley.

Georg, S. & A.P. Volodin, 1999, *Die itelmenische Sprache. Grammatik und Texte*, Wiesbaden: Harrassowitz.

Gordon, C.H., ca. 1982, *Forgotten Scripts. Their Ongoing Discovery and Decipherment*, New York: Basic Books. (Revised and enlarged edition).

Haspelmath, M., et al., eds., 2005, *The World Atlas of Linguistic Structures*, Oxford, New York: Oxford University Press.

Hawkins, J.A., 1983, *Word Order Universals*, New York, London: Academic Press.

Hewitt, S., 2003, "The Hamito-Semitic Connection: Fact or Fiction?," paper read at the Twelfth International Congress of Celtic Studies, Aberystwyth, 24-30 August, 2003.

Hewitt, S., fc., "Remarks on the Insular Celtic/Hamito-Semitic Question," in: Karl, R. & D. Stifter, eds., fc., in: *The Celtic World*, London, New York: Routledge.

Hopper, P.J., 1986, "Discourse Function and Word Order Shift: a Typological Study of the VS/SV Alternation," in: Lehmann, W.P., ed., 1986, *Language Typology 1985. Papers from the Linguistic Typology Symposium, Moscow, 9-13 December 1985*, Amsterdam, Philadelphia: John Benjamins Publishing Company, 123-40.

Isaac, G.R., 1993, "Issues in the Reconstruction and Analysis of Insular Celtic Syntax and Phonology," in: *Ériu* 44: 1-32.

Isaac, G.R., 1994, "The Progressive Aspect Marker: W *yn* / OIr. *oc*," in: *Journal of Celtic Linguistics* 3: 33-9.

Isaac, G.R., 1996, *The Verb in the Book of Aneirin. Studies in Syntax, Morphology and Etymology*, Tübingen.

Isaac, G.R., 2003, "The Structure and Typology of Prepositional Relative Clauses in Early Welsh," in: Russell, P., ed., 2003, *Yr Hen Iaith. Studies in Early Welsh*, Aberystwyth, Wales: Celtic Studies Publications, 75-93.

Isaac, G.R., fc., "The Verb in the Book of Aneirin: addenda and corrigenda," in: *Zeitschrift für celtische Philologie* 56.

Jakobson, R., 1971, "Notes on Gilyak," in: Jakobson, R., 1971, *Selected Writings II*, The Hague, Paris: Mouton, 72-97 (First published in: *The Bulletin of the Institute of History and Philology. Studies Presented to Yuen Ren Chao on his 65th Birthday*, Academica Sinica 29, Taipei, (November, 1957), 355-81).

Jakobson, R. & M. Szeftel, 1966, "The Vseslav Epos," in: Jakobson, R., 1966, *Selected Writings IV*, The Hague, Paris: Mouton, 301-68 (First published in: Jakobson, R & E.J. Simmons, eds., 1949, *Russian Epic Studies*, Philadelphia: American Folklore Society, 13-86).

Jongeling, K., 2000, *Comparing Welsh and Hebrew*, Leiden: Research School of Asian, African, and Amerindian Studies, University of Leiden.

Karlsson, F., 1984, *Finnische Grammatik*, Hamburg: Buske.

Kirk, J.W.C., 1905, *A Grammar of the Somali Language with Examples in Prose and Verse and an Account of the Yibir and Midgan Dialects*, Cambridge: Cambridge University Press.

Koch, J.T., 1983, "The Sentence in Gaulish," in: Koch, J.T. & J. Rittmueller, eds., *Proceedings of the Harvard Celtic Colloquium* 3, Harvard: Harvard University Press, 169-216.

Lambert, P.-Y., 2002, *Recueil des inscriptions gauloises*, vol. II, fascicule 2: *Textes Gallo-Latins sur instrumentum*, Paris: CNRS Éditions.

Layton, B.R., 2000, *A Coptic Grammar. With Chrestomathy and Glossary: Sahidic Dialect*, Wiesbaden: Harrassowitz.

Lehmann, W.P., 1978, "The Great Underlying Ground-Plans," in: Lehmann, W.P., ed., 1978, *Syntactic Typology. Studies in the Phenomenology of Language*, Austin: University of Texas Press, 3-55.

Lehmann, W.P., 1994, "Modifications in Syntactic Change from SOV to VSO," in: Bielmeier, R. & R. Stempel, eds., 1994, *Indogermanica et Caucasica. Festschrift für Karl Horst Schmidt zum 65. Geburtstag*, Berlin, New York: de Gruyter, 97-109.

Lejeune, M., 1988, *Recueil des inscriptions gauloises*, vol. II, fascicule 1: *Textes Gallo-Étrusques, textes Gallo-Latins sur pierre*, Paris: CNRS Éditions.

Lockwood, W.B., 1968, *Historical German Syntax*, Oxford: Clarendon Press.

Loprieno, A., 1995, *Ancient Egyptian. A Linguistic Introduction*, Cambridge: Cambridge University Press.

McCone, K., 1985, "The Würzburg and Milan Glosses: our Earliest Sources of 'Middle Irish'," in: *Ériu* 36: 85-106.

McCone, K., 1994, "An tSean-Ghaeilge agus a Réamhstair," in: McCone, K., et al., eds, 1994, *Stair na Gaeilge in Ómós do P[h]ádraig Ó Fiannachta*, Maynooth: St. Patrick's College, 61-219.

McCone, K., 2006, *The Origins and Development of the Insular Celtic Verbal Complex*, Maynooth: National University of Ireland.

Marichal, R., 1988, *Les graffites de La Graufesenque*, Paris: CNRS Éditions.

Meid, W., 1989, *Zur Lesung und Deutung gallischer Inschriften*, Innsbruck: Institut für Sprachwissenschaft.

Morris-Jones, J., 1900, "Pre-Aryan Syntax in Insular Celtic," Appendix B in: Rhŷs, J. & D. Brynmor-Jones, *The Welsh People*, London: Fisher Unwin.

Morris-Jones, J., 1913, *A Welsh Grammar. Historical and Comparative. Phonology and Accidence*, Oxford: Clarendon Press.

Mithun, M., 1999, *The Languages of Native North America*, Cambridge: Cambridge University Press.

O'Brien, M.A., 1923, "Hibernica. 2. Old Irish *os me*, *os se*, *ote*, etc.," in: *Zeitschrift für celtische Philologie* 14: 311-15.

Pokorny, J., 1927, "Das nicht-indogermanische Substrat im Irischen," in: *Zeitschrift für celtische Philologie* 16: 95-144; 231-66; 363-94.

Pokorny, J., 1928, "Das nicht-indogermanische Substrat im Irischen," in: *Zeitschrift für celtische Philologie* 17: 373-88.

Pokorny, J., 1930, "Das nicht-indogermanische Substrat im Irischen," in: *Zeitschrift für celtische Philologie* 18: 231-48.

Pokorny, J., 1949, "Zum nicht-indogermanischen Substrat im Inselkeltischen," in: *Die Sprache* 1: 235-45.

Pokorny, J., 1962, "Die Sprachen der vorkeltischen Bewohner Nordwest-europas," in: *II. Fachtagung für indogermanische und allgemeine Sprachwissenschaft*, Innsbruck), 129-38.

Pokorny, J., 1964, "Zur Anfangsstellung des inselkeltischen Verbums," in: *Münchener Studien zur Sprachwissenschaft* 16: 75-80.

Roberts, J.R., 1987, *Amele*, London, New York, Sidney: Croom Helm.

Ronan, P., 2002, "Subordinating *ocus* 'and' in Old Irish," in: Filppula, M., J., Klemola & H. Pitkänen, eds., 2002, *The Celtic Roots of English*, Joensuu: University of Joensuu, 213-36.

Russell, P., 1995, *An Introduction to the Celtic Languages*, London: Longman.

Schulze, W., 1982, *Die Sprachen der Uden in Nord-Azerbajdžan. Studien zur Synchronie und Diachronie einer süd-ostkaukasischen Sprache*, Wiesbaden: Harrassowitz.

Shisha-Halevy, A., 1995, "Structural Sketches of Middle Welsh Syntax I: the Converter," in: *Studia Celtica* 29: 127-223.

Shisha-Halevy, A., 2000 a, "Stability in Clausal/Phrasal Pattern Constituent Sequencing: 4000 Years of Egyptian (with some Theoretical Reflections, also on Celtic)," in: Sornicola, R., E., Poppe & A. Shisha-Halevy, eds., 2000, *Stability, Variation and Change of Word-Order Patterns over Time*, Amsterdam, Philadelphia: John Benjamins Publishing Company, 71-100.

Shisha-Halevy, A., 2000 b, "Structural Sketches of Middle Welsh Syntax II: Nominal Predications and Presentatives," in: *Studia Celtica* 33: 155-234.

Sohn, H.-M., 1999, *The Korean Language*, Cambridge: Cambridge University Press.

Stokes, W. & J. Strachan, 1901, *Thesaurus Palaeohibernicus: a collection of Old-Irish glosses, scholia, prose and verse*, vol. 1, Cambridge: Cambridge University Press.

Stokes, W. & J. Strachan, 1903, *Thesaurus Palaeohibernicus: a collection of Old-Irish glosses, scholia, prose and verse*, vol. 2, Cambridge: Cambridge University Press.

Thurneysen, R., 1923, "Gallisches," in: *Zeitschrift für celtische Philologie* 14: 8-12.

Thurneysen, R., 1946, *A Grammar of Old Irish*, Dublin: Dublin Institute for Advanced Studies.

Tristram, H.L.C., 1999, "'The Celtic Englishes' – Zwei grammatische Beispiele zum Problem des Sprachkontaktes zwischen dem Englischen und den

keltischen Sprachen," in: Zimmer, S., R. Ködderitzsch & A. Wigger, eds., 1999, *Akten des zweiten deutschen Keltologensymposiums (Bonn, 2.-4. April 1997)*, Tübingen: Niemeyer, 254-76.

van Enk, G.J. & L. de Vries, 1997, *The Korowai of Irian Jaya. Their Language in its Culture*, Oxford, New York: Oxford University Press.

Vendryes, J., 1911-1912 "La place du verbe en celtique," in: *Mémoires de la Société de Linguistique de Paris* 17: 337-51.

Vennemann, T., 1998, "Zur Etymologie von *Éire*, dem Namen Irlands," in: *Sprachwissenschaft* 23/4: 461-9.

Vennemann, T., 2000, "English as a 'Celtic' Language," in: Tristram, H.L.C., ed., 2000, *The Celtic Englishes II*, Heidelberg: Winter, 399-406.

Vennemann, T., 2001, "Atlantis Semitica: Structural Contact Features in Celtic and English," in: Brinton, L., ed., 2001, *Historical Linguistics 1999. Selected Papers from the 14th International Conference on Historical Linguistics*, Amsterdam, Philadelphia: John Benjamins Publishing Company, 351-69.

Vennemann, T., 2002, "Semitic ⇒ Celtic ⇒ English: the Transitivity of Language Contact," in: Filppula, M., J. Klemola & H. Pitkänen, eds., 2002, *The Celtic Roots of English*, Joensuu: University of Joensuu, 295-330.

Vennemann, T., 2003 a, *Europa Vasconica – Europa Semitica*, Berlin: Mouton de Gruyter.

Vennemann, T., 2003 b, "Languages in Prehistoric Europe North of the Alps," in: Bammesberger, A. & T. Vennemann, eds., *Languages in Prehistoric Europe*, Heidelberg: Winter, 319-32.

Vennemann, T., 2003 c, "Syntax und Sprachkontakt: Mit besonderer Berücksichtigung der indogermanischen Sprachen des Nordwestens," in: Bammesberger, A. & T. Vennemann, eds., *Languages in Prehistoric Europe*, Heidelberg: Winter, 333-64.

Villar, F. & B.M. Prósper, 2005, *Vascos, Celtas e Indoeuropeos. Genes y lenguas*, Salamanca: Ediciones Universidad de Salamanca.

Wackernagel, J., 1892, "Über ein Gesetz der indogermanischen Wortstellung," in: *Indogermanische Forschungen* 1: 333-436.

Wagner, H., 1959, *Das Verbum in den Sprachen der britischen Inseln*, Tübingen: Niemeyer.

Wagner, H., 1967, "Zur unregelmässigen Wortstellung in der altirischen Alliterationsdichtung," in: Meid, W., ed., 1967, *Beiträge zur Indogermanistik und Keltologie. Julius Pokorny zum 80. Geburtstag gewidmet*, Innsbruck: Institut für Sprachwissenschaft, 289-314.

Wagner, H., 1977, "Wortstellung im Keltischen und Indogermanischen," in: Schmidt, K.H., ed., 1977, *Indogermanisch und Keltisch: Kolloquium der Indo-*

germanischen Gesellschaft am 16. und 17. Februar 1976 in Bonn: Vorträge, Wiesbaden: Reichert, 204-35.

Wagner, H., 1981, "Near Eastern and African connections with the Celtic world," in: O'Driscoll, R., ed., 1981, *The Celtic Consciousness*, Toronto: McClelland & Stewart; Dublin: The Dolmen Press, 51-67.

Wagner, H., 1987, "The Celtic invasions of Ireland and Great Britain: Facts and Theories," in: *Zeitschrift für celtische Philologie* 42: 1-40.

Wali, K. & O.N. Koul, 1997, *Kashmiri. A Cognitive-Descriptive Grammar*, London, New York: Routledge.

Williams, I., 1927, "The Computus Fragment," in: *Bulletin of the Board of Celtic Studies* 3: 245-72.

Wuethrich, B., 2000, "Linguistics: Peering into the Past, with Words," in: *Science* 288/5469 (May 19, 2000), 1158 (consulted at URL: <http://www.sciencemag.org/cgi/content/full/288/5469/1158>).

Zeidler, J., 2004, "Ägyptisch – Hethitisch – Indogermanisch. Methodische und komparative Ansätze," in: Schneider, T., ed., 2004, *Das Ägyptische und die Sprachen Vorderasiens, Nordafrikas und der Ägäis: Akten des Basler Kolloquiums zum Ägyptisch-Nichtsemitischen Sprachkontakt, Basel 9. – 11. Juli 2003*, Münster: Ugarit, 271-300.

Ziegler, S., 1997, "Zur Syntax und Entwicklung des Verbalnomens in den keltischen Sprachen," in: Crespo, E. & J.L. García Ramón, eds., 1997, *Berthold Delbrück y la syntaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft (Madrid 1994)*, Madrid: Ediciones de la UAM, 631-44.