The Question of a Hamito-Semitic Substratum in Insular Celtic

Article *in* Language and Linguistics Compass · July 2009

DOI: 10.1111/j.1749-818X.2009.00141.x · Source: DBLP

1 author:

Steve Hewitt
United Nations Educational, Scientific and Cultural Organization
16 PUBLICATIONS **58** CITATIONS

The Question of a Hamito-Semitic Substratum in Insular Celtic

Steve Hewitt*
UNESCO

Abstract

The hypothesis of a Hamito-Semitic (or Afro-Asiatic) substratum in the Insular Celtic languages elaborated successively by Morris Jones, Pokorny and Wagner to explain striking structural resemblances between Insular Celtic and Hamito-Semitic is enjoying a revival. Linguists have generally assumed that the parallels between Insular Celtic and Hamito-Semitic are to be explained in terms of Greenbergian typology (all languages of the VSO type). However, recent work by Gensler, and also Jongeling and Vennemann, compels us to revisit the substratum hypothesis. This article presents the main contributions on the question, provides a table showing the principal points of similarity by author and language, briefly comments on each of these points, and, regretting the reluctance of substratalists to consider typological explanations, sounds a note of caution against what might be termed "substratum frenzy".

1. Introduction

Striking structural resemblances have long been noted between the Insular Celtic languages (divided into Goedelic – Irish Gaelic, Scottish Gaelic, Manx; and Brythonic – Welsh, Breton, Cornish) and various Hamito-Semitic (more broadly, Afro-Asiatic) languages. Few of these traits appear to have been noteworthy of the now extinct Continental Celtic languages (Gaulish, Celtiberic, Lepontic, Galatian), and they are not generally found in other Indo-European languages. This fact has led some to raise the possibility of a Hamito-Semitic substratum in the British Isles to explain their origin. Among the most prominent shared features are: VSO order; singular (apersonal) verb-marking with plural post-verbal lexical subjects; the Semiticists' "construct state" – a [HEAD [*the*-DEPENDENT]] genitive construction; "conjugated" prepostions; and oblique relatives with pronoun copies, each illustrated below in Breton and Arabic.

Breton					*Arabic*			
VSO order								
…e	*skrivas*	*Yann*	*ul*	*lizher*	*kataba*	*Yaḥyā*	*risāla*	
…AFF	wrote.PRET°	John	a	letter	wrote.PRET.M°	John	letter	

"John wrote a letter"

Singular (apersonal) verb-marking with plural post-verbal lexical subjects

…e	*skrivas*	*ar*	*merc'hed*	*katabat*	*al-banāt*	
…AFF	wrote.PRET°	the	girls	wrote.PRET.F°	the-girls	

"The girls wrote"

*"Construct state" – [HEAD [*the-*DEPENDENT]] genitive construction*

ti	*ar*	*roue*	*bait*	*al-malik*
house	the	king	house	the-king

"The house of the king, the king's house"

"Conjugated prepositions"

gant,	*ma',*
ganin, ganit,	*ma'ī, ma'ak, ma'ik,*
gantañ, ganti,	*ma'uh, ma'hā*
ganeomp, ganeoc'h, gante	*ma'nā, ma'kum, ma'hum*
"with,	"with,
with me, with you.SG,	with me, with you.SG.M, with you.SG.F,
with him, with her	with him, with her
with us, with you.PL, with them"	with us, with you.PL, with them"

Oblique relatives with pronoun copies

ar	*gwele*	*a*	*meus*	*kousked*	*ennañ*	*as-sarīr*	*al-ladhī*	*nimt*	*fīh*
the	bed.M	AFF	I.have	slept.PP	in.it.M	the-bed.M	the-REL.M	I.slept	in.it.M

"The bed I slept in, the bed in which I slept"

2. Main authors

Both Gensler, in his thesis (1993:57-191), and Jongeling (2000:6-64) provide extensive surveys of earlier authors on the Insular Celtic / Hamito-Semitic question. The first mention of structural similarities between Insular Celtic and Hamito-Semitic was in Davies (1621), where a number of resemblances between Welsh and Hebrew were noted. John Rhŷs (1877:189f) and subsequent works raises the possibility that pre-Aryan languages may have exerted structural influence on the Insular Celtic languages.

Morris Jones (1900) reviews startling similarities between Welsh and Egyptian, including (1900:625-6) the periphrastic conjugation *be* + preposition + verbal noun:

"In Welsh and Irish, although these languages retain many of the Aryan tenses, this construction is extremely common ... The three prepositions commonly used for this purpose in Egyptian are *em* 'in', *er* 'to, for', *ḥer* 'above' [= Loprieno (1995:80) *m* "in", *r* "towards", *ḥr* "on"] indicating the present, future, and perfect respectively. These correspond in use with the Welsh prepositions *yn* 'in', *am* 'for', *wedi* 'after'." He also notes the surprising parallels between Welsh *yn* and Egyptian *em*: (1) preposition "in", (2) "in" + verbal noun = progressive, and (3) predicative and adverbializing "in". Other similarities are noted with Berber, but no Semitic language is examined. He concludes (1900:639) that the resemblance

> seems to involve an intimate connection of some kind between the two families of speech in the prehistoric period, though they are probably not actually cognate. It is with Hamitic, however, rather than Semitic, that Celtic syntax is in agreement; for, as we have seen, it agrees with Berber where the latter differs markedly from Arabic, as, for instance, in the shifting of the pronominal suffix from the verb to a preceding particle ... Is the influence of a Hamitic substratum to be discovered in the simultaneous development on the same analytic lines of French, Spanish, and Italian in their use of infixed and postfixed pronouns?

Pokorny's *magnum opus* on the subject (1927-1930), has been conveniently reduced to 64 features by Vennemann (2002:324-6). Pokorny's discursive text is often impressionistic, with numerous examples (never glossed, at best paraphrased) from Hamito-Semitic languages, as well as Cushitic, Bantu (including unseemly references (1927:137) to "*Negersprachen*" deemed "*ungemein primitiv*"), Basque, Finno-Ugric, Caucasian, etc., all grist to his substratal mill. He proposes (1927:100ff) that in a language with a strong, aristocratic literary tradition such as Irish, substratal influence may take a long time to become apparent in the written language. Pokorny continued to write on the subject throughout his life; his most concise statement of the linguistic features shared by Insular Celtic and Hamito-Semitic is (1959), where he identifies a more managageable list of 20 shared features, most of which are included in the table of shared features below. See also Pokorny (1960).

Wagner's main work on the Insular Celtic / Hamito-Semitic question (usefully summarized by Gagnepain (1961)) is (1959), especially the third part, entitled "The linguistic geography position of the Brittonic verb", with sections on "Celtic, Berber, Basque, English and French as representatives of a North African – Western European linguistic stratum, exemplified by the structure of the verb", "The Berber verb system" and "Remarks on the Semitic verb system". He describes his basic position (see also 1981) as follows (1987:19-20):

> Between the fourth and sixth centuries A.D. Insular Celtic suffered revolutionary changes … As a result of it Insular Celtic developed features and grammatical categories hardly found in other Indo-European languages. They have, however, close parallels in Berber and Egyptian, the Hamitic languages of Northern Africa, as well as in Basque … The linguistic structure of Insular Celtic compels me to assume that, long before the arrival of Celtic or Belgic tribes, these islands were populated by people, who spoke languages or dialects which, from the point of view of E. Lewy's typology could be described as Hamito-Semitic, languages not necessarily connected with but of a similar type as Berber and Egyptian and, somewhat more distantly Hebrew and Arabic. For my latest position on this subject, cf. my articles of 1976 and 1982. When Celtic was

> adopted by pre-Celtic populations, the structure of their original language(s) began to impose itself on the language of the Celtic invaders. The result was a linguistic revolution which led to the making of the mediaeval and modern Celtic languages.

Hewitt (1985) was written before I became aware of the Pokorny-Wagner tradition; I assumed that the resemblances could only be typological. The similar features reviewed in that article include head-dependent typology, VSO~SVO word order (main and subordinate clauses), verb-subject agreement/non-agreement, collective/singulative, conjugated prepositions, expression of "have", the construct state genitive, compound "construct-state" adjectives, double "topic ≠ subject" sentences, relatives (restrictive, non-restrictive, on prepositional objects, on possessives), the dummy sentential pronoun, and circumstantial subordinating *and*. The chief shared feature omitted is the verbal noun, most certainly historically speaking a verbal noun in Breton, but in modern Breton it behaves like an infinitive (accusative rather than possessive object pronouns).

Gensler (1993) examines 12 Insular Celtic or Hamito-Semitic languages plus a random sample of 58 other languages from all over the world. He identifies a set of 17 "exotic" structural features shared by Insular Celtic and Hamito-Semitic, but not common among languages worldwide. Assigning scores for each feature in each of the 70 languages, he concludes (1993:426):

> On the basis of the sample used in this study, nothing remotely close to the full-blown Celtic/Hamito-Semitic [CHS] linguistic type recurs anywhere else in the world. The relatively few languages which are 'best matches' – actually rather poor matches – are scattered all over the globe, from the West Coast of North America to the Caucasus and New Guinea. However, the continental average score for Africa is higher than for any other continent, and drops only slightly when the CHS languages Egyptian and Berber are omitted; West Africa scores especially well, and appears especially hospitable to several of the CHS features (adpositional periphrastic, word-initial change, kin terms, inter alia). Conversely, Europe has one of the lowest average scores, and when Welsh and Irish are excluded its score drops far below that of any other continent. Celtic is thus radically out of place in a European landscape, whereas the Hamito-Semitic languages simply intensify a structural trend seen over much of Africa. A weak form of the CHS type, then, would appear to have a natural home in Africa, in particular Northwest Africa. Within Afroasiatic, the highest-scoring languages are on the Mediterranean; scores fall away in every direction, but the Chadic language Hausa (in West Africa) scores much higher than Cushitic Afar (in East Africa). The diachronic evidence, too, argues that the (weak) CHS type is something quite old in Africa: the African and Arabian case studies all show stronger CHS-ness further back in time. All this, in conjunction with the blood-type agreement between the British Isles and Northwest Africa, argues for some sort of prehistoric scenario specifically linking these two regions.

While Gensler does not claim to have proved the Hamito-Semitic substratum hypothesis for Insular Celtic, he does appear to be saying: "in the face of such statistical results, what else can it be?" He takes great care with the weighting (0, ±½, ±1) of scores for individual features, but no account is taken of the centrality or frequency of each feature within a particular language, such that pervasive features such as the genitive construction (feature 8) have the same weight as more marginal ones like the idiomatic genitive kinship constructions (feature 17). Furthermore, the

languages that score highest in Gensler's table, thus most strongly exemplifying what he calls the "Celtic/Hamito-Semitic type", are the Insular Celtic languages rather than the Hamito-Semitic languages, i.e. the languages that are said to have been influenced by the Hamito-Semitic substratum which purportedly gave rise to the type. Another drawback is that possible typological explanations for some of the features (cf. remarks on the genitive construction below) are not envisaged or examined. Nevertheless, with his thorough analysis and the sheer wealth of linguistic evidence he has marshalled, Gensler has certainly put the Hamito-Semitic substratum hypothesis back on the map; all authors on the subject will henceforth need to take due account of his arguments.

Jongeling (2000) provides an excellent, lengthy introduction to the history of the subject. The features he surveys include VSO, head-dependent order, numerals, nominal clauses, circumstantial subordinating *and*, relatives, the verbal noun, conjugated prepositions, and the lack of a verb "*have*". He proposes (2000:149-50) an interesting variant of the substratum hypothesis:

> Supposing that the explanation of certain peculiarities of Insular Celtic are due to substratum influence, one might suppose that the same or a similar substratum has influenced some subgroupings of Afro-Asiatic [Hamito-Semitic] … In short, this scenario would mean that we should consider Western Europe and North Africa as an old coherent area of VSO-character. The influence on the three northern Afro-Asiatic groups, Semitic, Egyptian and Berber is comparable to the influence on the Celtic sub-grouping of Indo-European ... one might suppose that Western Europe and Northern Africa once formed one great contiguous VSO area. This area was split by the incoming Indo-Europeans. The proportion of Indo-Europeans on the continent was so great that any influence of a pre-existing language was blotted out, while the number of pre-Indo-Europeans inhabitants on the British Isles was such that their influence there was felt long after they were gone from memory.
>
> This scenario not only explains the congruity in syntax of Welsh and Hebrew but at the same time gives a reason for the lack of lexical correspondences not only between Welsh and Hebrew, but in general between Afro-Asiatic and Insular Celtic.

Vennemann sees a Hamito-Semitic substratum as having influenced Insular Celtic, and through Celtic, English. In (2001:351) he claims: "The European Atlantic Littoral was, at the dawn of history, explored and colonized by Mediterranean, probably Palaeo-Phoenician seafarers." The main features examined in this article include the verbal noun and the related progressive construction, the English "Northern Subject Rule" reminiscent of Semitic and Celtic verb-subject agreement, and the "replacement of the sympathetic dative by the internal possessor construction" (*Jean s'est cassé le bras* vs *John broke his arm*, see feature 22 below). Explaining why what he calls the "Atlantic type" arose only in Middle English, he reiterates (2001:364) an argument of Pokorny's and Wagner's: "substratal influence originates in the lower strata of a society and usually takes centuries to reach the written language, and regularly only after a period of social upheaval." In (2002), in which he draws attention to circumstantial subordinating *and*, and the prevalence of tensed verb/auxiliary responsives in Celtic and English, Vennemann states:

> In my view the case is closed, the thesis of a Hamito-Semitic substratum underlying Insular Celtic being one of the most reliably established pieces of scientific knowledge there is in any empirical discipline. As Gensler has shown, the substratum really was not simply Hamito-Semitic, which is a huge family including hundreds of languages in Africa and Asia (which is why it is also called Afro-Asiatic or Afrasian), but more specifically Hamito-Semitic of the Mediterranean type, which includes Libyco-Berber, Ancient Egyptian, and Semitic. In order to stress the similarity of the substratum to this particular manifestation of Hamito-Semitic, I sometimes refer to it as Semitidic or simply Semitic.

Celtic influence on English. The idea that certain features of English may be attributable to Insular Celtic is particularly in vogue among Finnish and German scholars, cf. articles by Filppula, Klemola, Vennemann, and especially Filppula, Klemola, Pitkänen (eds) (2002) and Tristram (ed.) *Celtic Englishes I, II, III*, *IV* (1997-2006). This theory presupposes that the Anglo-Saxons were (thinly) superimposed on a British-speaking population which eventually shifted to Anglo-Saxon, leaving subsequent substratal structural traces in English. However, a recent study by Capelli et al. (2003) has found genetic evidence to support the more traditional picture of massive population shifts (2003:979): "By analyzing 1772 Y chromosomes from 25 predominantly small urban locations, we found that different parts of the British Isles have sharply different paternal histories." Coates (2004) casts doubt on the influence of British Celtic on Anglo-Saxon.

Typological approaches. Not all authors dealing with similarities between Insular Celtic and Hamito-Semitic languages subscribe to the substratum hypothesis. Recent papers adopting a purely typological stance include Borsley (1995), Isaac (2004), and Roberts (2004). An important methodological critique of the substratum approach is Isaac (2008).

3. Shared features

The following table shows the main features shared by Insular Celtic and Hamito-Semitic, according to author and language. The first 17 features correspond to Gensler's (1993) shared features. These are followed by additional features identified by Morris Jones (1900), Pokorny (1927-1930, 1959), Wagner (1959), Hewitt (1985), and Jongeling (2000). Features marked with (✓) in the column for Gensler are mentioned by him, but are not central to his thesis; features 35-39 from Hewitt (1985) are given by no other author, and are purely typological in nature.

Similar features in Insular Celtic and Hamito-Semitic according to author

Feature	Morris Jones (1900) W, (Ir.); Eg., Bb.	Pokorny (1927-30, 1959) Irish, HS	Wagner (1959) IC, HS	Hewitt (1985) Breton, Arabic	Gensler (1993) IC, HS	Jongeling (2000) Welsh, Hebrew
1. Conjugated prepositions	✓	✓	✓	✓	✓	✓
2. Word order: VSO, head-dependent, prepositions	✓	✓	✓	✓	✓	✓
3. Invariable relative clause linker, not relative pronoun		✓	~		✓	✓
4. Relative clause copying, not gapping: *the bed that I slept* ***in it***	✓	✓	✓	✓	✓	✓
5. Special relative tensed verb form		✓	✓		✓	
6. Subject and object marking in verb	✓	✓	✓		✓	
7. Object marker: preverb-infix-V/V-suffix	✓	✓	✓		✓	
8. Genitive construction: def. art. on dependent only: *house* ***the-****man*		✓		✓	✓	✓
9. Non-agreement of verb with plural noun subject	✓		✓	✓	✓	✓
10. Verbal noun, not infinitive (object in genitive, not accusative)		✓	✓		✓	✓
11. Predicative particle: *he is* ***in*** *a farmer*	✓	✓	✓		✓	✓
12. Prepositional periphrastic: *he is* ***at*** *singing*	✓	✓	✓		✓	
13. Periphrastic DO: *he* ***does*** *singing*	✓	✓	✓		✓	
14. Circumstantial clause *and S PRED* (subordinating *and*)		✓	✓	✓	✓	✓
15. Nonfinite possible instead of finite main-clause verb			✓		✓	✓
16. Word-initial phonetic changes (mutations), various syn. functions			✓		✓	
17. Idiomatic genitive kinship constructions: *son of X*		✓			✓	
18. Nominal clause (absence of copula)	✓	✓	✓		(✓)	✓
19. Amplification of negative by noun after verb: French *pas*	✓				(✓)	
20. Numerals followed by singular	✓					✓
21. Prepositional expression of *have*		✓	✓	✓	(✓)	✓
22. Possessive *he broke his arm* rather than dative *il s'est cassé le bras*		✓			(✓)	
23. Preference for parataxis (Pokorny: *anreihend* "stringing along")		✓	✓		(✓)	
24. Basic unit word group rather than single word		✓	✓		(✓)	
25. Subjectless sentences (impersonal constructions)		✓			(✓)	
26. No present/active participle		✓			(✓)	
27. Distinction between essential and contingent BE (*is/tá*)		✓	✓		(✓)	
28. States/relations expressed with *N (PREP-O) PREP-S Tá scilling agam ort*		✓	✓		(✓)	
29. Welsh *yn*, Egyptian *m* "in": predicative, locative, progressive		✓			(✓)	
30. Old Irish infixed pronoun *–d-* identical to Berber *–d-*		✓	✓		(✓)	
31. Comparatives (and superlatives) predicative only, not attributive		✓			(✓)	
32. Initial focus clefts		✓	✓		(✓)	
33. Yes/no responsives with auxiliary/verb/entire clause		✓	✓		(✓)	
34. Prepositional relative: fronting of bare preposition *the bed* ***in*** *I slept*		✓	✓		(✓)	
35. Unmarked collective, derived singulative		✓		✓		
36. Virtual complement clause VSO; factual complement clause SVO~VSO				✓		
37. Construct state *bahuvrīhi* adjectives (Ar. "improper annexation")				✓		
38. Topic ≠ subject ("double subject" sentences)				✓		
39. Yes/no dummy sentential pronoun: *I don't know and* ***he*** *they came*				✓		

Similar features in Insular Celtic and Hamito-Semitic according to language

Feature	Irish	Welsh	Breton	Arabic	Hebrew	Egyptian	Berber
1.	✓	✓	✓	✓	✓	✓	✓
2.	✓	✓, AUX S V	✓PSO~SPO	✓+SVO	✓>SVO	✓+SVO	✓
3.	✓	✓	✓	~✓gen.num.	✓	✓agreem.	×
4.	✓	✓	✓	✓	✓	✓	×
5.	✓	~ *sydd, sy*	~ *so (zo)*	×	×	✓	✓
6.	✓ OI	~	~	~✓	~✓	×	✓
7.	✓ OI	~	~ ✓trad. B	×	×	×	✓
8.	✓	✓	✓	✓	✓	×	×
9.	~	✓	✓VS, SV aff.	✓ VSO	~ VSO	✓	×
10.	✓	✓ ~	✓ trad.	✓	✓	? ~ ×	✓
11.	✓ *in his*	✓	~× ✓adv.	~×	~×	✓	
12.	✓	✓	✓	×	×	✓	×
13.	~×	~✓	✓	×	×	✓	~×, ?
14.	✓	✓	✓	✓	✓	✓	✓ < Ar.
15.	×	✓	~✓	×	✓		×
16.	✓	✓	✓	×	~× ✓phon.	×	~ vow. red.
17.	~✓	~✓	~✓	✓	✓		✓ < Ar.
18.	~✓	~✓gnomic	~✓gnomic	✓	✓	✓	✓
19.	~✓	✓	✓	~✓ (dial.)	×	✓ (Coptic)	✓
20.	~✓ (20, x)	✓	✓	✓ (11+)	✓ (11+)	✓ ME>	✓ (11+)
21.	✓	✓	✓	✓	✓	✓	✓
22.	✓	✓	✓	✓	✓		
23.	✓	~✓	~✓	?	?	??	✓
24.	✓	✓	✓	~✓	✓	?	✓
25.	✓	~✓	✓	✓	✓	✓	✓
26.	✓	✓	✓	×	×	×	✓
27.	✓	×	✓	×	×	✓	✓
28.	✓	✓	✓	✓	✓	✓	✓
29.	~✓	✓	~✓	~✓	?	✓	×
30.	✓	×	×	×	×	×	✓
31.	✓	×	×	×	×	✓	✓
32.	✓	✓	✓	✓	✓	✓	✓
33.	✓	✓	✓	×	?	✓	✓
34.	✓	×	×	×	×	×	✓
35.	?	✓	✓	✓	?		
36.	×	×	~✓	✓	?		
37.	✓ + poss.	✓ + poss.	✓ + poss.	✓	✓		
38.	×	×	✓	✓	?	✓	
39.	×	×	✓	✓ Eg. Ar.	×		

1. Conjugated prepositions. In both Insular Celtic and Hamito-Semitic, this looks historically very much like incorporation of a pronominal in the preposition, a commonplace process; in Hamito-Semitic there is a single set of endings for prepositions, possessives and objects of verbs.

2. Word order: VSO, head-dependent, prepositions. The word order typology of both families is basically verb-subject-object (VSO), with head-dependent order and prepositions. However, a number of qualifications are in order. As suggested in Hewitt (2002b), Breton simultaneous VSO and V2 is better described as PSO (predicate-subject-object) and T2 (tense-second).

Irish and Welsh are generally considered to be classic examples of VSO order, and (apart from a strong V2 period in Middle Welsh, see Willis (1998)) show no signs of a tendency towards SVO. However, Jones & Thomas (1977), analysed Welsh as TSPO, T attaching either to an auxiliary or, if there is none, to the main verb. So which is Welsh in the case of *Mae Mair yn dysgu Cymraeg* [isº Mair in learn.VN Welsh] "Mair is learning Welsh", SVO or VSO? It depends on whether for V you see as primary T the tense-bearing function (> VSO), or P the predicative function (> SVO).

Most Hamito-Semitic languages are reckoned to be VSO, with the exception of Amharic and Akkadian (both SOV). Hebrew shows a steady progression from clear VSO in the Biblical period to SVO in the Massoretic and modern periods. Similarly, Arabic is normally considered to be VSO, although, as in Breton, SVO is a common alternative order, even from the Koranic period; SVO has gained in prominence in modern times; certain styles of journalistic Arabic are reckoned to be more SVO than VSO, and some dialects, in particular Egyptian, are thought to be basically SVO, with only residual VSO effects. While the *World Atlas of Language Structures Online* (Feature 81: Order of Subject, Object and Verb) does show VSO with the Celtic and Semitic languages (Modern Hebrew and Syrian Arabic being "neutral", and other modern dialects of Arabic being predominantly SVO), the other concentrations of VSO are in the Rift Valley in Africa, Sumatra – Philippines, Central America, and the Northwest Coast of North America; central sub-Saharan Africa, where pronoun copies with oblique relatives, as in the Celtic languages, are so heavily concentrated (see below, feature 4), is strongly SVO rather than VSO.

However, in the case of Arabic VSO, another analysis is possible, with potentially far-reaching consequences for word order typology in general (cf. Hewitt 2002a, 2006): in verb-initial clauses, there are numerous and regular violations of the canonical VSO order, such as VOS, VoS (o=pronominal object), VPREPOS (PREP=preposition), VPREPoSO, VPREPoOS, and VoOS. A principle of increasing "information salience" of post-verbal nominal constituents (given, known information > new information) appears to provide a unitary account of all the observable orders, *including VSO*. Strict SO order is thus called into question for Arabic, and replaced by a strict GN (given-new) order. Languages traditionally described as VSO (however, the principle of information salience-governed word order does not appear to apply to any of the Celtic languages) or SOV (e.g. Turkish, Hindi/Urdu, Tibetan) might need to be revisited in order to see whether VGN or GNV does not provide a better account of their functioning than VSO or SOV. The *World Atlas of Language Structures Online* assumes that predicate-and-arguments

word order can only be based on syntactic function – subject, object, despite the questionable universality of those concepts – and does not envisage the possibility that for some languages the primary factor may instead be information salience. The whole question of VSO is thus rather more complex than it might appear at first sight.

3. Invariable relative clause linker, not relative pronoun. The precise syntactic status of the Celtic relativizer, e.g. Breton direct (subject, object) *a*, indirect (oblique – other elements) *e* is debatable, being variously analysed as either an affirmative tense particle or a relative pronoun of sorts. The invariable Hebrew *ăšɛr* has been analysed both as a relativizer and a relative pronoun. The Arabic relative pronoun *al-ladhī* M, *al-latī* F, *al-ladhīna* M.PL, etc. agrees in gender and number with its antecedent.

4. Relative clause copying, not gapping: *the bed that I slept in it*. Yes, Breton *ar gwele a meus kousked ennañ* [the bed.M AFF I.have slept in.it.M]; Arabic *as-sarīr al-ladhī nimt fī-h* [the-bed.M REL.M I.slept in-it.M] "the bed I slept in/in which I slept". The *World Atlas of Language Structures Online* (Feature/Chapter 123: Relativization on Obliques) does show a heavy concentration of this strategy, apart from in Irish and Scottish Gaelic, in Semitic (Hebrew and Arabic) and in numerous languages of central sub-Saharan Africa, but also in Persian, Eastern Kayah Li (Thailand, Myanmar), Paamese (Vanuatu, South Pacific), and Guaraní (Paraguay, Brazil). Berber never has resumptive pronouns, only movement of the bare preposition; Old Irish is similar, but the order of relator and preposition is the reverse: Berber REL+PREP; Old Irish PREP+REL.

5. Special relative tensed verb form. This is present in Irish (*-as* vs. *-aidh*, *-ann*, etc.) and apparently in Egyptian and Berber; in Brythonic, the only modern trace is in the present relative form of the copula Welsh *sy(dd)*, Breton *so* (Modern Breton *zo*); apart from Akkadian, this is unknown in Semitic; the Berber "relative form" is commonly called a "participle".

6. Subject and object marking in verb. Yes, for both Insular Celtic and Semitic; this concerns most strongly Old Irish, Berber, and Egyptian. Object pronouns are traditionally proclitic in Celtic and postclitic in Semitic; this is not the same, however, as the true subject-and-object-marking verbal morphology of Georgian. The cliticization of object pronouns on the verb is hardly a rare trait.

7. Object marker: preverb-infix-V/V-suffix. This concerns especially Old Irish and Berber. It should be noted more generaly that pre- versus post-cliticization of object pronouns concerns many languages, cf. Romance, Serbo-Croat, etc.

8. Genitive construction: def. art. on dependent only: *house the-man*. Breton *ti ar roue* [house the king], Arabic *bait al-malik* [house the-king] "the king's house". Known as the "construct state" CS among Semiticists, this is not necessarily due to substratal influence. In the typology of genitive constructions, there is a limited number of parameters: (1) the order of head and dependent: H D or D H (VSO normally implies H D order); (2) the presence or absence of an article (on D only; on both D and H; no examples of the article on H only; all Insular Celtic and most, but not all, Hamito-Semitic languages have a definite article); (3) the relation marking may be on either H or D; and finally (4) a limited number of relator mechanisms (one or more are possible): (a) simple adjacency (H D, as in all Insular Celtic and

Hamito-Semitic languages except Amharic, or D H); (b) phonetic modification of either H or D: phonetic CS marking of H in Hebrew: *bayiṯ* "house", but *bēṯ ham-mɛlɛk* [house.CS the-king]; *dābbār* "word", but *dǎbar ham-mɛlɛk* "the king's word"; or phonetic modification of D in Berber: *agellid* "king", but *axxam (n) ugellid* [house (of) king.CS]; (c) case: GEN, DAT, OBL, etc.; (d) possessive POSS: Turkish D-GEN + H-POSS: *kral-ın ev-i* [king-GEN house-his]; (e) link particle LNK: in both Hindi-Urdu and Swahili the link agrees with H: Hindi-Urdu *laṛkā* "boy.NOM", *laṛkē kā ghar* [boy.OBL LNK_x $house_x$] "the boy's house"; Swahili: *nyumba ya mfalme* [house.CL9 LNK.CL9 king.CL1] "the king's house"; (f) adposition: preposition, postposition, cf. English *the door of the house*.

In Insular Celtic and Hamito-Semitic languages which no longer have case, the construct state is defined solely by the adjacency of the head and dependent, and the restriction of the article to the dependent. However, in both families this is probably the result of independent evolution. While Breton, Welsh, Hebrew and colloquial Arabic have no cases: Breton *dor an ti*, Welsh *drŵs y tŷ*, Hebrew *dɛlɛṯ hab-bayiṯ*, colloquial Arabic *bāb al-bait* [door the-house] "the door of the house", Classical and formal Modern Standard Arabic and Irish (for some items at least) conserve case endings: Arabic *bāb-u l-bait-i*; Irish *doras an tí* [door.NOM the-house.GEN], and these help to define the genitive relation. It is only with the loss of the case endings that the [H [*the*-D]] structure becomes crucial to defining the genitive construction.

Germanic has both a compact genitive construction *the king's house*, with genitive case and only one article possible, on D, and a periphrastic construction *the house of the king* with two articles and the genitive relation expressed by the preposition *of*. While it is not obvious to ordinary English-speakers which element the article *the* applies to in *the king's house*, other Germanic languages provide a clue: German *des Königs Haus* [the.GEN king.GEN house.NOM] or Swedish: *kungens hus* [king.the.GEN house]. It therefore seems logical to bracket the phrase as follows: [[*the king's*] *house*], which is simply the reverse of the order of the two main constituents H and D in Breton [*ti* [*ar roue*]] or Arabic [*bait* [*al-malik*]]. Indeed, in one Germanic language which has lost all genitive case-marking, the [[*the*-D] H] order actually defines the genitive relation: the highly evolved form of English found in Jamaican Creole: [[*di king*] *hoos*]. Seen in this light, the Insular Celtic genitive structure is rather less exotic than it might appear at first sight; there is little need to appeal to Hamito-Semitic for a source.

Insular Celtic and Hamito-Semitic differ with regard to adjective placement in genitive constructions. While neither Celtic nor Germanic has any problem in attaching adjectives to either or both H and D: Breton *ti bihan ar roue bras* [house little the king big] "the big king's little house", Semitic cannot do this; any adjectives go obligatorily after the genitive construct, which is more akin to a compound noun "king-house", so *bait al-malik al-kabīr* [house the-king the-big] can in principle mean either "the king's big house" or "the big king's house". Only in formal Arabic is it possible to tell which the adjective applies to, from the case-marking. The usual way of applying adjectives to both H and D is to use a longer construction with two articles and a preposition, structurally similar to the periphrastic construction *the house of the king*: *al-bait aṣ-ṣaghīr li-l-malik al-kabīr* [the-house the-little to-the-king the-big] "the big king's little house". If the genitive construction in Insular Celtic

really had its origins in a Hamito-Semitic substratum, it is difficult to understand why this major structural constraint prohibiting the insertion of adjectives between H and D would not also apply in Celtic.

9. Non-agreement of verb with plural noun subject. This is a striking parallel, strongest in Welsh, Breton, Egyptian, Classical Arabic and to some extent Biblical Hebrew, in the latter two with VSO order only; in Breton also with SVO order in the affirmative, but not in the negative; not in Berber. Non-agreement is fairly common with VS order worldwide, cf. Greenberg (1966), Universal 33: "When number agreement between the noun and verb is suspended and the rule is based on order, the case is always one in which the verb precedes and the verb is in the singular." Number non-agreement with plural post-verbal subjects is lost in spoken Arabic, and was lost in Hebrew from the Mishnaic period, so those two languages have moved away from non-agreement. There was usually *agreement* with post-verbal plural subjects in Old Welsh and Old Breton, so non-agreement appears to have come in since those periods. Non-agreement appears to be even more recent in Gaelic; indeed, in some dialects, such as Munster, there is often still agreement. These are exceedingly long times for some putative substratal non-agreement to have filtered through.

10. Verbal noun, not infinitive (object in genitive, not accusative). There appears to be more of a cline than a sharp distinction between the abstract verbal noun (Arabic, Georgian *masdar*) and the infinitive. The criterion for distinguishing between the two is whether objects are in the genitive (verbal noun) or accusative (infinitive). With the development in Breton since the 18th century (with the exception of the SE Gwened/Vannes dialect) of true "accusative" object pronouns (etymologically "*of* + pronoun"): *ma gweled* [my seeing] > *gweled ahanon* [see.INF/VN of.me] "to see me", little now distinguishes the Breton verbal noun from the French infinitive. In Insular Celtic, only the Irish verbal noun seems truly *masdar*-like.

Verbal noun or infinitive?

Nominal features	Irish	Welsh	Breton	Berber	Egyptian	Arabic	Hebrew
Gender	✓	×	×	✓	–	✓	?
Article possible	✓	✓	✓	–	–	✓	?
Pron. object: possessive	✓	✓	✓>×	?	✓	✓	✓
Lexical object: genitive	✓	(✓)	(✓)>×	?	(✓)	✓	(✓)

11. Predicative particle: *he is <u>in</u> a farmer*. Especially in Welsh and Egyptian, if Welsh predicative *yn* really is "in", and this has recently been challenged by Gensler (2002). The construction is "*In* ***his*** *farmer*" in Irish. This feature is very limited in Breton, and is marginal in Hebrew and Arabic.

12. Prepositional periphrastic: *he is <u>at</u> singing*. As Comrie points out (1976:100-102), apart from Insular Celtic languages, which all have prepositional periphrastic constructions, copular locative phrases expressing the progressive are found in numerous other languages: Chinese, Georgian, Yoruba, Shona, Igbo, Kpelle, other Nigero-Congolese languages, Hindi/Urdu, Punjabi, North American Indian, etc., not to mention Icelandic, various German dialects, and Continental

Scandinavian. However, no Semitic languages do this, although there is an increasing use of active participles. In Egyptian *ḥr* "on", *r* "towards", and *m* "in" are all used with verbal nouns to express a progressive. Paradoxically, the Breton progressive is much closer in force to the English progressive than the Welsh or Scottish Gaelic periphrastic constructions, which have become a general cursive (imperfective) which freely allows statives (cf. Hewitt (1986, 1990)). Indeed, there appears to be a general tendency in many languages for a parallel evolution of simple tense > specialized uses, and progressive > general imperfective.

13. Periphrastic DO: *he does singing*. This conflates at least three distinct uses: (1) activity DO with dynamic (non-stative) VPs: (Middle) Welsh, Breton; (2) DO with NPs: numerous languages; (3) DO as an empty auxiliary: North Welsh; Breton to avert V-1 in the affirmative (also English *do* with negative, interrogative). This is not typical of Semitic.

14. Circumstantial clause *and S PRED* (subordinating *and*). This feature is typical of both Insular Celtic and Hamito-Semitic (in Berber it is possibly borrowed from Arabic): Breton *gweled neus ahanon ha me o tond er-maes* [seen he.has me.OBJ and I PROG come.INF out], Arabic *laqad ra'ā-nī wa-'anā ṭāli'* [PFV he.saw-me and I coming.out] "he saw me as I was coming out". The construction is syntactically coordinate (in both Celtic and Semitic, the order after "*and*" is always SVO), but semantically subordinate; an adversive "*although*" connotation is possible.

15. Nonfinite possible instead of finite main-clause verb. This is particularly prevalent in Welsh, followed by Hebrew (infinitive absolute), but not in Arabic. There are sporadic examples in Irish and Breton.

16. Word-initial phonetic changes (mutations), various syn. functions. The highly grammaticalized Insular Celtic initial consonant mutations are hardly comparable to the Berber "construct state" initial changes *argaz* > *urgaz (wərgaz)* "man", *tamγart* > *tmγart* "town", which appear to be more akin to vowel contraction, cf. the Hebrew construct state forms described under feature 8. As formulated, this is rather too abstract a feature to be confidently attributed to substratal influence, and there are numerous instances worldwide of results of phonetic changes acquiring a grammatical function.

17. Idiomatic genitive kinship constructions: *son of X*. This is very productive in Semitic, cf. Iraqi Arabic *abu chegāyir* [father.CS cigarettes] "(street) cigarette seller". It is not typical of Brythonic; the few examples in Insular Celtic are in Irish: *mac tíre* [son land.GEN] "wolf".

18. Nominal clause (absence of copula). There is no copula in the present tense in Semitic, only in non-present (future, past) tenses. This is not the same as ellipsis of the copula in Insular Celtic, especially Welsh, and to a lesser extent Breton, in gnomic expressions.

19. Amplification of negative by noun after verb: French *pas*. Arabic dialects (Palestine and westwards) have developed a French-like circumfix *mā V-sh* (<*shi* "thing"); Welsh *ni V S ddim*, Breton *ne V S ked*. This is surely part of a general linguistic tendency to amplify function words that otherwise risk being lost altogether.

20. Numerals followed by singular. Yes, in Brythonic; in Irish, originally nouns after 20 and higher multiples of 10 stood in the GEN.PL; due its identity with

the nom. sg. in some declensional classes, this gave rise to its reinterpretation as singular. In Semitic, yes for 11 and higher; 3-10 are followed by nouns in the plural. Numerals are followed by singular nouns in many languages, for example Persian, Basque, Hungarian, Georgian, Tibetan...

21. Prepositional expression of *have*. Yes, Breton is the only Celtic language to have developed a verb "have": *m-eus*, etymologically [to.me-there.is°] "I have", used as an auxiliary with perfect tense and as a lexical verb "possess"; possession of definites is usually expressed with prepositional periphrasis *an arc'hant so ganin* [the money is° with.me] "I have the money", as it is in the other Celtic languages, Hamito-Semitic, and many other languages worldwide.

22. Possessive *he broke his arm* rather than dative *il s'est cassé le bras*. Yes, this is true of both Insular Celtic and Hamito-Semitic; "internal" possession with a possessive, rather than "external" possession with a dative is rare in European languages, with the exception of Celtic and English, cf. Payne & Barshi (1999). It is unclear which is more common worldwide. Note that Breton requires a combination of both types (possessive + dative) where the possessor of the object is not the same as the subject (possible, but not obligatory in French):

Mae Ieuan wedi torri 'i fraich	*Mae Ieuan wedi torri braich Pedr*	(Welsh)
is° Ieuan after break.VN his arm	is° Ieuan after break.VN arm Pedr	
"Ieuan broke his arm"	"Ieuan broke Pedr's arm"	
Jean s' est cassé le bras	*Jean (lui) a cassé le bras à Pierre*	(French)
Jean RFL is broken the arm	Jean (to.him) has broken the arm to Pierre	
"Jean broke his arm"	"Jean broke Pierre's arm"	
Yann neus torred e vrec'h	*Yann neus torred e vrec'h da Ber*	(Breton)
Yann has.M broken his arm	Yann has.M broken his arm to Per	
"Yann broke his arm"	"Yann broke Per's arm"	

Deus 'ta heol benniged da dommañ o revrioù d'ar ffiliped (Breton)
come then sun blessed to warm.INF their backsides to the sparrows
"Come on, dear sun, and warm the sparrows' backsides"

23. Preference for parataxis (Pokorny: *anreihend* "stringing along"). It is unclear how such a feature, identified by Pokorny, could be measured, and if it could be demonstrated, whether it is really unique to Insular Celtic and Hamito-Semitic.

24. Basic unit word group rather than single word. Again, it is unclear how such a feature might be measured, and whether it is unique to Insular Celtic and Hamito-Semitic.

25. Subjectless sentences (impersonal constructions). Both Pokorny and Wagner list this as a common feature. This probably covers a number of distinct phenomena which need closer definition, and in any case, it is hardly unique to Insular Celtic and Hamito-Semitic.

26. No present/active participle. Yes, for Insular Celtic, but not true of Semitic at all.

27. Distinction between essential and contingent BE (*is/tá*). Yes, in Irish (*is/tá*) and Breton (*eo/emañ*), but not in Welsh or in Semitic. Again, such a distinction is quite common worldwide.

28. States/relations expressed with *N (PREP-O) PREP-S*. Yes, cf. Irish *Tá scilling agam ort* [is.SIT° shilling with.me on.you] "you owe me a shilling", *tá tart orm* [is.SIT° thirst on.me] "I am thirsty". Felt to be very typical of IC; not particularly typical of Semitic.

29. Welsh *yn*, Egyptian *m* "in": predicative, locative, progressive. Attention was first drawn to this amazing coincidence by Morris Jones, and it is tantalizing if true. However, there is some question as to the identity of the three *yn*'s in Welsh, cf. Isaac (1994), where he proposes that progressive *yn* derives from *wnc* "close", and Gensler (2002), who claims that Welsh predicative *yn* is derived from a deictic *int*.

30. Old Irish infixed pronoun *-d-* identical to Berber *-d-*. Pokorny drew attention to this, for instance in (1959:157). Its significance is unclear.

31. Comparatives (and superlatives) predicative only, not attributive. This is especially true of Irish, but not of Brythonic or Semitic.

32. Initial focus clefts. Such structures are common across many languages.

33. Yes/no responsives with auxiliary/verb/entire clause. This is considered to be very typical of Insular Celtic (which is held by some to have given rise to English "*yes, it is*", "*no, I don't*", etc.), but is not particularly characteristic of Semitic.

34. Prepositional relative: fronting of bare preposition *the bed <u>in</u> I slept*. Only Irish and Berber have this.

35. Unmarked collective, derived singulative. This is particularly productive in Breton *blew/blewenn* and Arabic *sha'r/sha'ra* "hair/strand of hair", cf. Hewitt (1985), but less so in Welsh and Hebrew, and it is quite marginal in Irish. It is also found in other languages, such as Swahili *nywele/unywele* "hair/strand of hair".

36. Virtual complement clause VSO; factual complement clause SVO ~ VSO. "Virtual" complement clauses of the type "I want John to come" are obligatorily VSO in both Arabic and Breton, whereas "factual" complement clauses such as "I think John will come" are obligatorily SVO in Arabic; traditionally they have been VSO in Breton, but since the 18th century, an alternative SVO order has become increasingly frequent. It is unlikely that this should be French influence because Breton never has SVO order in the "I want John to come" type, cf. Hewitt (1985). This similarity is far more likely to be typological rather than substratal in origin.

Arabic			*Breton*				
'urīd	*'an yajī'*	*Zaid*	*Me*	*meus*	*c'hwåant e*	*teuffe*	*Yann*
I.want	that come.SUBJ.M°	Zaid	I	I.have	desire AFF	come.COND°	Yann
"I want Zaid to come"			"I want Yann to come"				

'aẓunn	*'anna*	*Zaid*	*sa-yajī'*	*Me*	*a*	*soñj*	*din*	*e*	*teuo*	*Yann*
I.think	that	Zaid	will-come.3.SG	I	AFF	think°	to.me	AFF	will.come°	Yann

Me a soñj din (penaos) Yann a deuo
I AFF think° to.me (how) Yann AFF will.come°
"I think that Zaid will come" "I think that Yann will come"

37. Construct state *bahuvrīhi* adjectives (Arabic "improper annexation"). Many languages have such adjective-noun compounds, including English, cf. "*pure-hearted*", "*great-winged*" below. The construct is formed with a possessive in Celtic, but has the form of a normal construct state in Semitic; note, however, the difference in treatment of the article between Hebrew and Arabic: in Arabic, the dependent noun always has the definite article; when the compound is definite, the whole construct state has, quite exceptionally, a definite article prefixed to it; in Hebrew, it is the article on the dependent noun that determines, in rather more orthodox fashion, whether the compound is definite or not.

un den ledan e chouk (Breton)
a man broad his nape
"a well-to-do man" (i.e. who can bear a heavy [financial] load)

rajul ṭāhir al-qalb (Arabic)
man pure the-heart
"a pure-hearted man"

ar-rajul aṭ-ṭāhir al-qalb (Arabic)
the-man the-pure the-heart
"the pure-hearted man"

nɛšɛr gădōl kănāpayim (Hebrew, Ezek. 17:7)
eagle great wings
"a great-winged eagle"

han-nɛšɛr hag-gāḏōl, gădōl hak-kănāpayim (Hebrew, Ezek. 17:3)
the-eagle the-great great the-wings
"the large, great-winged eagle"

38. Topic ≠ subject ("double subject" sentences). This is very common in both Breton *Per eo klañv e vab* [Per is° ill his son] and Arabic *Zaid marīḍ ibn-uh* [Zaid ill son-his] "Per/Zaid's son is ill"; Breton *Chirac a dalc'h e fri da voanâd: c'hwessa a ra partoud!* [Chirac AFF keeps° his nose to narrow.INF: sniff.INF AFF he.does everywhere] "Chirac's nose keeps getting narrower / keeps narrowing: he's sniffing everywhere [for votes]". Again, this is surely typological; it is found in numerous languages.

39. Yes/no dummy sentential pronoun: *I don't know and he they came*. This concerns a partial resemblance between Breton and Egyptian Arabic, cf. Hewitt (1985); the main difference is that in Breton the dummy pronoun is invariable, whereas in Egyptian Arabic it must agree in number with the subject: Breton *n-onn ked hag-eñv e oa aed ar baotred* [NEG-I.know not and-**he** AFF was gone the boys] "I don't know whether the boys went"; Egyptian Arabic *sa'al-ni humma r-riggāla mishyu* [he.asked-me **they** the-men they.went] "he asked me whether the men had gone". This is unlikely to be substratal.

4. Concluding remarks

The most prominent and high-frequency features shared by the Insular Celtic and the Hamito-Semitic languages are probably the following:

VSO order (feature 2). This apparently concerns both families, but there are numerous other VSO languages worldwide, even if the proportion of VSO languages is low (around 15%). Furthermore, there has been drift of varying degrees towards SVO in both Hebrew and Arabic, and also in Breton. Welsh (and to some extent even Irish) could be analysed as SVO if the predicative function is taken as being more important than the tense-bearing function for V. In the *World Atlas of Language Structures Online*, VSO does not cluster well with feature 4 (pronoun copies with oblique relatives), where many other languages partial to that strategy are strongly SVO. Finally, analysis of Arabic suggests that for that language at least, and possibly for other Hamito-Semitic languages, but clearly not for any Insular Celtic language, the basic order of the verb and its arguments is not based on the syntactic status of the arguments (subject, object: VSO), but rather on their information salience (given, new: VGN), the VGN principle economically accounting for *all* observable verb-initial orders, *including* the most frequent verb-initial order VSO.

Construct state (feature 8). The [*house* [*the-king*]] "the king's house" [HEAD [*the*-DEPENDENT]] structure is certainly one of the most striking features shared by the two families. This involves: head-dependent order; availability of a definite article; relation-marking on the dependent only; relator mechanism of simple adjacency. Historically in both families, the relation was marked on the dependent with the genitive, and to some extent continues to be so in Arabic and Irish and Scottish Gaelic. The [HEAD [*the*-DEPENDENT]] structure becomes crucial in defining the genitive construction only with the lost of case (Colloquial Arabic, Hebrew, Welsh, Breton). Finally, Hamito-Semitic allows absolutely no adjectives to be placed between the head and the dependent, whereas all Celtic languages do so freely. The two families thus appear to have arrived quite independently at this construction, which is not so exotic as might be thought: Jamaican Creole, which has lost all case-marking, has [[*di-king*] *hoos*], which is merely the dependent-head inverse of [*house* [*the-king*]].

Non-agreement with plural post-verbal subjects (feature 9). Greenberg's (1966) Universal 33 states that: "When number agreement between the noun and verb is suspended and the rule is based on order, the case is always one in which the verb precedes and the verb is in the singular." The tendency in Semitic (Arabic and Hebrew) has been to relax or lose non-agreement with plural post-verbal subjects; in the older stages of the Celtic languages (Old Welsh, Old Breton, Old Irish) number *agreement* was general; *non-agreement* has come in since the old periods, and in Irish, for instance, is still not complete, agreement being common in the southwestern dialects of Munster. Thus the trends appear to be going in opposite directions in the two families. In the negative, Breton functions exactly like Standard Arabic in both affirmative and negative: number *agreement* with preverbal subjects; *non-agreement* with post-verbal subjects.

"Conjugated prepositions" (feature 1). The apparent "conjugations" are probably historically, in both families, simply the result of a morphological process of incorporation of post-prepositional pronominals.

Pronoun copies with oblique relatives (feature 4). In addition to Insular Celtic and Hamito-Semitic, the *World Atlas of Language Structures Online* shows this strategy to be common in numerous languages of central sub-Saharan Africa, but also in Persian, Eastern Kayah Li (Thailand, Myanmar), Paamese (Vanuatu, South Pacific) and Guaraní (Paraguay, Brazil).

Verbal noun rather than infinitive (feature 10). The distinction between the two is a matter of degree, the chief polar features being: gender: yes/no; definite article possible: yes/no; pronominal object: possessive/accusative; lexical object: genitive/accusative. Only Irish and Arabic have the first option (verbal noun) for all four features; other languages are less clear-cut. Breton (mainstream KLT dialects) and Modern Spoken Welsh have reached a point at which little distinguishes the untensed citation form of the verb from the infinitive of French and English. Arabic is not moving in this direction at all; as in Georgian, for instance, its *maṣdar* is fully nominal.

Circumstantial subordinating *and* (feature 14). I am unaware of any other languages apart from Hamito-Semitic and Insular Celtic which have this feature.

The apparent correspondence of **Egyptian *m* "in", *r* "towards", *ḥr* "on" and Welsh *yn* "in", *am* "for", *wedi* "after"** expressing, respectively, **predicative/locative/progressive, future, perfect** (Morris-Jones and **feature 29**) is quite tantalizing, but to this level of detail concerns these two languages only, and is not widely shared by the rest of the Insular Celtic and Hamito-Semitic languages (*tar éis* "after" with a verbal noun may express the perfect in Irish, and *war* "on" with the verbal noun/infinitive may be used for a prospective in Breton).

For none of these prominent shared features is a substratal explanation demonstrably more plausible than a typological explanation or mere coincidence.

ཐི་

The existence of striking structural similarities between the Insular Celtic and the Hamito-Semitic languages is beyond question. However, the matter of whether this is to be attributed to substratal influence through prehistoric contact or to typological tendencies and correlations remains unresolved. Gensler's statistical approach (the low likelihood of such clustering of "exotic" features in two genetically unrelated families) is in itself skewed: by focusing on the shared features, he loses sight of the bigger picture, including all the features that are not shared by the two families. Furthermore, he has no way of accounting for the relative frequency or centrality of his various features in the languages concerned.

Authors inclined to a substratal explanation for the shared features appear to be prey to a kind of "substratum frenzy", as if prehistoric contact must be the only possible explanation for "un-Indo-European" traits in an Indo-European language. They pay scant attention to the possibility of typological explanations, even though, as we have seen, such explanations are perfectly plausible for many of the shared traits. With lexical items, the number of possible phonetic sequences is so vast that

any significant accumulation of lexical similarities between two genetically unrelated languages can hardly be anything other than a sure sign of contact and borrowing. With structures, however, the range of possibilities across languages is far more limited – there are, for instance, only so many ways of expressing a genitive relation –, so it is less surprising for unrelated languages to possess analogous structures. It is therefore important always to bear in mind and investigate thoroughly the possibility of a typological explanation.

A major problem with the substratal explanation is the precise identity of the substratum. A subsidiary puzzle is the special affinities noted between Welsh and Hebrew (several authors, most recently Jongeling), Welsh and Egyptian (Morris Jones), Irish and Berber (Morris Jones, Pokorny and Wagner), and Breton and Arabic (Hewitt – probably typological: both SVO~VSO topic-prominent languages); in each case, however, the number of features concerned is low enough for the "special affinity" to be coincidental.

Surprisingly, Gensler's scores suggest that it is the Insular Celtic languages which are most typical of the "Celtic/Hamito-Semitic type" rather than the Hamito-Semitic languages, and this is borne out by our table of shared features by author and language, where the various features are more consistently present in Insular Celtic than in Hamito-Semitic. This is the reverse of what one would expect if the shared features really had their origin in Hamito-Semitic.

Rather than positing some Berber *Urvolk*, or Phoenician settler ghosts (who have somehow managed to leave no archaeological traces), substratalists (and it should by now be clear that I am sceptical) might take a cue from Jongeling (2000), who moots a single prehistoric substratum to both Hamito-Semitic and Insular Celtic. Such a substratum might have been centred on north-western Europe or even the British Isles, where it might have affected the incoming Celtic languages strongly, but the more distant Hamito-Semitic and North African languages less so. The identity of such a substratum would, however, perforce be so shrouded in the mists of prehistory as to be quite unknowable.

Clearly, more work is needed on both the substratal and the typological approaches to this fascinating question.

Tibetan for "dog", pronounced /ˉkʰi/, cf. Welsh *ci*, Breton *ki*, Irish *cú*.

Abbreviations

adv.	adverb(ial)	Ir.	Irish	RFL	reflexive
AFF	affirmative tense particle	LNK	link particle	S	subject
		M	masculine	SG	singular
aff.	affirmative	ME	Middle English	SIT	situative
agreem.	agreement	N	new	SUBJ	subjunctive
Ar.	Arabic	NEG	negative tense particle	T	tense
AUX	auxiliary	NOM	nominative	trad.	traditional(ly)
Bb.	Berber	num.	number	V	verb
Br.	Breton	O	object	VN	verbal noun
CL	class	OI	Old Irish	vow.	vowel
COND	conditional	o	pronominal object	W.	Welsh
CS	construct state	OBJ	object	X	some initial element: P, S, O, Adv, etc.
D	dependent	OBL	oblique	º	apersonal verb form: no person-marking
DAT	dative	P	predicate		
dial.	dialect(al)	PFV	perfective		
Eg.	Egyptian	phon.	phonetic	x	agreement, marked on both terms
F	feminine	PL	plural		
G	given	POSS	possessive	~	partially
GEN	genitive	PP	past participle	✓	present
gen.	gender	PRET	preterite	✗	not present
H	head	PREP	preposition	?	questionable; uncertain
HS	Hamito-Semitic	PROG	progressive		
IC	Insular Celtic	red.	reduction		
INF	infinitive	REL	relator		

Short biography

Steve Hewitt is Editor of Records at UNESCO, and as English Reviser is responsible for Arabic-English translation. Familiar with a wide range of languages, he is particularly interested in the Celtic languages (especially Breton, on which he has published extensively), Arabic, and more recently Georgian; his theoretical interests include verb systems (tense-aspect-mode, split auxiliarity, valency, case arrays of arguments), information structure, and word order typology. From 2000 to 2008, he held a weekly class on "La linguistique du breton" at the École Pratique des Hautes Études of the Sorbonne.

Endnote

* Correspondence address: Steve Hewitt, ADM/CLD/T, UNESCO, 7 place de Fontenoy, 75352 PARIS 07 SP, France. s.hewitt@unesco.org.

References and short bibliography

BORSLEY Robert D. 1995. On some similarities and differences between Welsh and Syrian Arabic. *Linguistics* 33:99-122.

CAPELLI Cristian *et al*. 2003. A Y chromosome census of the British Isles. *Current Biology* 13:979-84, (available on-line from www.sciencedirect.com).

COATES Richard. 2004. Invisible Britons: The view from linguistics. Paper read at the conference "Britons and Angles in Anglo-Saxon England", Manchester, 14-16 April 2004.

COMRIE Bernard. 1976. *Aspect*. Cambridge University Press.

DAVIES John. 1621. *Antiquae Linguae Britannicae, nunc communiter dictae Cambro-Britannicae, a suis Cymraecae, vel Cambricae, ab aliis Wallicae, Rudimenta*. London; reprinted Oxford, 1809.

FILPPULA Markku, Juhani KLEMOLA, Heli PITKÄNEN. 2002. Early contacts between English and the Celtic languages. FILPPULA, KLEMOLA, PITKÄNEN (eds). *The Celtic Roots of English*. 1-26.

FILPPULA Markku, Juhani KLEMOLA, Heli PITKÄNEN (eds). 2002. *The Celtic Roots of English*. Proceedings of the International Colloquium on Early Contacts between English and the Celtic Languages, University of Joensuu, Finland, 24-26 August, 2001. Joensuu, Finland: University Press.

FILPPULA Markku and Juhani KLEMOLA. *English and Celtic in contact: Bibliography*. On-line searchable bibliography, www.joensuu.fi/fld/ecc/bibliography.

GAGNEPAIN Jean. 1961. A propos du 'verbe celtique' [Review of WAGNER, *Das Verbum...*], *Etudes celtiques* 9:309-26.

GENSLER Orin David. 1993. A typological evaluation of Celtic/Hamito-Semitic syntactic parallels, unpublished Ph.D. dissertation, University of California, Berkeley, available from University Microfilms International, Ann Arbor, Michigan, No. 9407967. Extracts printed in KARL & STIFTER (eds). 2007. *The Celtic World*, Vol. IV, *Celtic Linguistics*. 151-229.

GENSLER Orin. 2002. Why should a demonstrative turn into a preposition: The evolution of Welsh predicative *yn*, *Language* 78/4:710-64.

GREENBERG Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. Joseph H. GREENBERG (ed.), *Universals of Grammar*, Cambridge, Mass. / London: MIT Press. 77-113

HEWITT Steve. 1985. Quelques ressemblances structurales entre le breton et l'arabe: Conséquence d'une typologie ordinale commune? *La Bretagne Linguistique* 1:223-262. Brest: Centre de recherche bretonne et celtique (CRBC).

HEWITT Steve. 1986. Le progressif en breton à la lumière du progressif anglais. *La Bretagne Linguistique* 2:132-48. Brest: Centre de recherche bretonne et celtique (CRBC).

HEWITT Steve. 1990. The progressive in Breton in the light of the English progressive. Martin BALL, James FIFE, Erich POPPE, Jenny ROWLAND (eds), *Celtic Linguistics: Readings in the Brythonic languages*. Festschrift for T. Arwyn Watkins. Amsterdam: John Benjamins. 167-88.

HEWITT Steve. 2002a. L'arabe – VSO ou VDN (verbe-sujet-objet ou verbe-donné-nouveau)? *Cahiers de linguistique de l'INALCO* 3/2000:211-21, *Ordre des mots et typologie linguistique*. Paris: INALCO.

HEWITT Steve. 2002b. The impersonal in Breton. *Journal of Celtic Linguistics* 7:1-39.

HEWITT Steve. 2006. Arabic: verb-subject-object or verb-given-new? Implications for word order typology. Expanded English version of HEWITT (2002a) presented to the Conference on Communication and Information Structure in Spoken Arabic, University of Maryland, 16-18 June 2006.

HEWITT Steve. 2007. Remarks on the Insular Celtic / Hamito-Semitic question. KARL & STIFTER (eds), *The Celtic World*, Vol. IV, *Celtic Linguistics*. 230-68. (Contains an extensive bibliography).

ISAAC Graham R. 1994. The progressive aspect marker: W. *yn*/OIr. *oc*. *Journal of Celtic Linguistics* 3:33-9.

ISAAC Graham R. 2004. The nature and origins of the Celtic languages: Atlantic seaways, Italo-Celtic and other paralinguistic misapprehensions. *Studia Celtica* 38:49-58.

ISAAC Graham R. 2008. Celtic and Afro-Asiatic. Hildegard TRISTRAM (ed.), *The Celtic Languages in Contact: Papers from the workshop within the framework of the XIII International Congress of Celtic Studies, Bonn, 26-27 July 2007*. Universität Potsdam. 25-80.

JONES Morris and Alan R. THOMAS. 1977. *The Welsh Language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.

JONGELING Karel. 1995. Afro-Asiatic and Insular Celtic. *DS-NELL* [Near Eastern Languages and Literatures Foundation] 1:135-65.

JONGELING Karel. 1996. The nominal clause in Hebrew and Welsh: Some remarks. *Zeitschrift für celtische Philologie* 68:259-86.

JONGELING Karel. 2000. *Comparing Welsh and Hebrew*. Leiden: CNWS Publications.

KARL Raimund and David STIFTER (eds). 2007. *The Celtic World: Critical concepts in historical studies*, Vol. IV, *Celtic Linguistics*. London & New York: Routledge.

LOPRIENO Antonio. 1995. *Ancient Egyptian: A linguistic introduction*. Cambridge University Press.
MORRIS JONES John. 1900. Pre-Aryan syntax in Insular Celtic. Appendix B. J. RHYS and D. BRYNMOR-JONES, *The Welsh people*, London: T. Fisher Unwin. 617-41. Reprinted in KARL & STIFTER (eds). 2007. *The Celtic World*, Vol. IV, *Celtic Linguistics*. 103-21.
PAYNE Doris and Immanuel BARSHI (eds). 1999. *External Possession*. Amsterdam: John Benjamins.
POKORNY Julius. 1927-30. Das nicht-indogermanische Substrat im Irischen. *Zeitschrift für celtische Philologie* 16:95-144, 231-66, 363-94; 17:373-88; 18:233-48.
POKORNY Julius. 1949. Zum nichtindogermanischen Substrat im Inselkeltischen. *Die Sprache* 1:235-45.
POKORNY Julius. 1951. Sprachliche Beziehungen zwischen dem alten Orient und den britischen Inseln, *ArOr* 19:268-70.
POKORNY Julius. 1955. Vorkeltische indogermanische Spuren im Kymrischen. *Zeitschrift für celtische Philologie* 25:87.
POKORNY Julius. 1959. Keltische Urgeschichte und Sprachwissenschaft. *Die Sprache* 5:152-164.
POKORNY Julius. 1960. The pre-Celtic inhabitants of Ireland, *Celtica* 5:229-40. Reprinted in KARL & STIFTER (eds). 2007. *The Celtic World*, Vol. IV, *Celtic Linguistics*. 122-32.
POKORNY Julius. 1964. Zur Anfangsstellung des inselkeltischen Verbums. *Münchener Studien zur Sprachwissenschaft* 16:75-80.
RHŶS John. 1877. *Lectures on Welsh Philology*. London.
RHŶS John. 1882. *Early Britain – Celtic Britain*. London.
RHŶS John. 1890. Traces of a non-Aryan element in the Celtic family, *Scottish Review* 16:30-47.
ROBERTS Ian. 2004. Parametric comparison: Celtic, Semitic and the anti-Babelic principle. Paper read at the Gregynog Welsh Syntax Seminar, Powys, Wales, 12 July 2004, and at a Hans Rausch Endangered Languages Project Seminar, School of Oriental and African Studies (SOAS), London, 25 January 2005.
TRISTRAM Hildegard L.C (ed.). 1997. *Celtic Englishes I*. Proceedings of the First Colloquium on Celtic Englishes, Potsdam, 28-30 September 1995. Heidelberg: Carl Winter.
TRISTRAM Hildegard L.C (ed.). 2000. *Celtic Englishes II*. Proceedings of the Second Colloquium on Celtic Englishes, Potsdam, 23-27 September 1998. Heidelberg: Carl Winter.
TRISTRAM Hildegard L.C (ed.). 2003. *Celtic Englishes III*. Proceedings of the Third Colloquium on Celtic Englishes, Potsdam, 20-23 September 2001. Heidelberg: Carl Winter.
TRISTRAM Hildegard L.C (ed.). 2006. *Celtic Englishes IV*. Proceedings of the Fourth Colloquium on Celtic Englishes, Potsdam, 20-26 September 2004. Potsdam: Universitätsverlag.
VENNEMANN Theo. 2001. Atlantis Semitica: Structural contact features in Celtic and English. Laurel BRINTON (ed.). *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics*. Amsterdam: John Benjamins. 351-69.
VENNEMANN Theo. 2002. Semitic → Celtic → English: The transivity of language contact. FILPPULA, KLEMOLA, PITKÄNEN (eds), *The Celtic Roots of English*. 295-330.
WAGNER Heinrich. 1959. *Das Verbum in den Sprachen der Britischen Inseln*. Tübingen: Max Niemeyer.
WAGNER Heinrich. 1969. The origin of the Celts in the light of linguistic geography. *Transactions of the Philological Society*. Oxford. 203-50.
WAGNER Heinrich. 1970. Studies in the origins of early Celtic civilization. *Zeitschrift für celtische Philologie* 31:1-58, 1970.
WAGNER Heinrich. 1971a. Nekrolog Julius Pokorny (1887-1970). *Zeitschrift für celtische Philologie* 32:313-19.
WAGNER Heinrich. 1971b. *Studies in the Origins of the Celts and of Early Celtic Civilization* [= The origin of the Celts in the light of linguistic geography + Studies in the origins of early Celtic civilization]. Tübingen: Max Niemeyer.
WAGNER Heinrich. 1976. Common problems concerning early languages of the British Isles and the Iberian peninsula. F. Jordá, J. de Hoz & L. Michelena (eds), *Actas del I coloquio sobre lenguas y culturas prerromanas de le peninsula ibérica (Salamanca, 17-31 mayo 1974)*, Salamanca: Ediciones Universidad de Salamanca. 388-407.
WAGNER Heinrich. 1977. Worstellung im Keltischen und Indogermanischen. Karl Horst SCHMIDT (ed.), *Indogermansich und Keltisch*. Wiesbaden: Ludwig Reichert. 204-35.

WAGNER Heinrich. 1981. Near Eastern and African connections with the Celtic world. Robert O'DRISCOLL (ed.), *The Celtic Consciousness*. New York: George Braziller. 51-76. Reprinted in KARL & STIFTER (eds). 2007. *The Celtic World*, Vol. IV, *Celtic Linguistics*. 133-50.

WAGNER Heinrich. 1982. Near Eastern and African connections with the Celtic world. R. O'Driscoll (ed.), *The Celtic Consciousness*, Portlaoise: Dolmen Press / Edinburgh: Cannongate Publishing, 51-67.

WAGNER Heinrich. 1987. The Celtic invasions of Ireland and Great Britain: Facts and theories. *Zeitschrift für celtische Philologie* 42:1-40.

WILLIS David. 1998. *Syntactic Change in Welsh: A study of the loss of verb-second*. Oxford University Press.

World Atlas of Language Structures Online. http://wals.info/ Online version of Martin HASPELMATH, Matthew S. DRYER, David GIL & Bernard COMRIE (eds). 2005. *World Atlas of Language Structures*. Oxford: Oxford University Press.