Cladistic Methodology and West Germanic

Don Ringe, U. of Pennsylvania

For about a century and a half now one of the usual tasks of historical linguistics has been reconstructing the prehistory of linguistic communities. (That is, of course, the side of reconstruction that non-linguists are interested in.) A perennial question is what sort of model of linguistic diversification best fits the data. This paper will explore that question in the light of data supporting a West Germanic subfamily of languages.

The earliest model of language diversification, with obvious points of contact in evolutionary biology and manuscript stemmatics, was the Stammbaum or "family tree" model, and it is virtually the only model used by non-linguists (such as computer scientists or population geneticists) who try to make use of linguistic data. Linguists, by contrast, spent much of the 20th century downplaying its usefulness in favor of other models (such as the "wave" model) suggested by dialect geography or sociolinguistics. However, it's noticeable that the most rigorous theoretician of historical linguistics, Henry Hoenigswald, continued to work with the tree model throughout his career. That alone would be reason enough to keep it in our toolkit. Other reasons will presently become clear.

There seems to be a general belief that a model of linguistic diversification should try to capture what actually happens as speech communities expand, diversify, disintegrate, and split up. It's certainly true that one can't ignore that without violating the Uniformitarian Principle (A.1 on the handout). In a paper published about a decade ago Malcolm Ross worked out a model of diversification based on the tree model but incorporating additional features to handle network-like diversification of dialects and even reintegration into a single speech community of historically related communities that have not diverged too much. This is necessary because (as he points out) the tree, wave, and convergence models are not really alternatives: "one cannot take the same set of language data and, by the application of different models, reconstruct alternative sets of linguistic events" (p. 211). Rather, different models are appropriate for different datasets generated by different historical processes. Ross works on Austronesian languages in western Oceania, where a very large number of languages have diversified at various

time depths, including quite recently, making it possible to estimate how often different types of diversification events occur. It seems clear to Ross that "clean" splits of the kind that can be modelled straightforwardly by a "family tree" are uncommon, and I see no reason to doubt that.

But any working historical linguist knows that almost all cladistic trees of language families are full of clean splits, especially in the higher reaches of the tree corresponding to greater time depths. The Uralic family tree is typical: Baltic Finnic is (or until recently was) obviously a dialect continuum that diversified beyond the point of mutual intelligibility; Saami is another; there are several less thoroughly diversified continua, such as that of the Mansi dialects; but all the higher-level speciations are reconstructable as clean binary or ternary splits, and that's not because large amounts of mess have been swept under the rug. If clean splits are rare on the ground, why must we so often reconstruct them for prehistory?

So far as I'm aware, the only colleague who has discussed this problem in print is Johanna Nichols (using Uralic as an example, since it's well-known and clear). In her famous article of 1990 she observed that language death must be part of the picture: probably most episodes of diversification involve dialect networks yielding multiple closely related languages, but if the episode occurred in the distant past, typically only two or three products of the diversification survived long enough to be recorded (or to give rise to further daughters which would eventually be recorded), creating an appearance of clean splits in most cases (A.2). If we insist that our model should replicate only the initial process of diversification, then no model can be appropriate for cases like this, since the pattern of language death usually ensures that we actually can't tell what the process was. But why should we insist on that? Why not settle for a model that helps us make sense of the pattern we actually see in the data—so long as we're careful not to misinterpret it? In many cases that will in fact be the tree model.

A third factor that influences how much information about specific diversification episodes we wind up with is the gradual deterioration and loss of linguistic information over time. I don't recall ever seeing the impact of that process discussed, though it obviously must have an impact.

But to take this discussion any further it will be useful to have some actual data to

refer to. I therefore present, under B on the handout, a summary of the evidence that the West Germanic languages (English, Frisian, Netherlandic, Low and High German) constitute a valid clade of the Germanic family—that is, that they share a "critical mass" of significant innovations which demonstrate that they must have developed together as a single speech community for a significant period of time. Only a few examples of each shared change are given, but there are very many examples of some; those cases are marked with "etc."

B.1 lists a series of three phonological developments which must have occurred in the order listed if the first and third make any sense phonetically. What's important is that in PGmc the voiced dental obstruent had both fricative and stop allophones (roughly as in modern Spanish). In WGmc the allophony was levelled in favor of the stop allophone; we know that because our phonetic information about the earliest recorded stages of all the daughter languages is relatively good. But before the fricative allophone was replaced by the stop, a very peculiar sound change assimilating voiced coronal fricatives to a following *w occurred; and probably after the stop allophone was generalized, a restricted syncope of short *i between dental stops occurred. The first of the three changes is probably unusual enough to validate the WGmc clade by itself; the sequence of three changes certainly does so. (And this is item 1 of 13.)

B.2 lists another sequence of sound changes. The first, the loss of unstressed short low vowels in final syllables, could in principle have been a parallel innovation of the WGmc languages, since we can show that it also occurred (independently) in Gothic and Norse. But in WGmc the postconsonantal semivowels that were stranded by loss of the following vowels uniformly became syllabic, whereas in Gothic only the front semivowel did (and in Norse apparently neither did). Moreover, that was an early complex of changes, since we can show that it was followed by the uniform WGmc gemination of consonants by following *j (B.2.c). This is reasonable supporting evidence for a WGmc clade.

B.3 lists yet another sequence of sound changes. Again the first, the loss of nasalization on word-final vowels, could have been a parallel development; I list it here only because it makes some of the examples of the second change easier to understand. But the unrounding of inherited bimoric long *ō in final syllables, accompanied or followed

by the loss of the contrast between bimoric and trimoric vowels and the monophthong-ization of *au in final syllables, is further validation of the WGmc clade.

The sound change listed under B.4, the loss of word-final *z in unstressed syl-lables, is obviously less impressive (though it is uniform throughout WGmc and did not occur in Gothic or Norse). Whether the change listed under B.5 is significant depends on whether it violated the OCP; if it did, it might be a highly significant innovation. The lengthening of the vowel in two adverbs (only) listed under B.6 is significant precisely because of the lexical restriction: something that quirky is likely to have happened only once.

Beginning with B.7 I list morphological innovations exclusively shared by WGmc, and the first is so bizarre that it unarguably validates the clade by itself. The PGmc past tense of strong verbs and the PGmc present tense of the small class called "preterite-presents" reflect the PIE perfect stem. The indicative 2sg. of those categories exhibits a shape unproblematically inherited from PIE. In Gothic and Norse no further change occurred. In WGmc, however, the inherited indicative 2sg. form survived only in the presents of preterite-present verbs; in the past tenses of strong verbs it was replaced by the corresponding *subjunctive* form, which in a majority of cases was constructed from a different stem! In other words, in the past tenses of strong verbs (but not of weak verbs) the 2sg. indicative and the 2sg. subjunctive underwent syncretism *under the form of the subjunctive,* for reasons which are not now recoverable. I submit that this is weird —too weird to have happened more than once, so that it would force recognition of a WGmc clade even if it were the only innovation that the WGmc languages exclusively shared.

But the innovation listed under B.8 turns out to be almost as peculiar. The PGmc stem suffix of j-presents was long *-ī- alternating with *-ija- if the root syllable was heavy, and that ought to have been a stable situation, since long *-ī- is transparently contracted from underlying */-iji-/ and the alternation between short *-i- and *-a- at the end of the suffix complex was normal for PGmc present stems. Nevertheless, as Warren Cowgill pointed out in 1959, this long *-ī- was replaced by short *-i-, which was the corresponding suffix alternant if the root syllable was light, throughout WGmc; otherwise we can explain neither the usual syncope of the 2sg. and 3sg. endings in West Saxon OE

nor the consistently short vowel in those forms in OHG. But there is no plausible way to motivate the replacement working from surface forms (see the Excursus on p. 6). No old-fashioned proportion will work. It's true that the 2sg. of the imperative should have exhibited short *-i, since word-final long high vowels had been shortened already in the PNWGmc period; but there is no realistic way to motivate the spread of an alternant from one imperative form through the entire paradigm.

However, it is possible to explain the change as a result of native learner reanalysis. Take a look at the comparison between representative forms of PGmc simple thematic presents and j-presents in the middle of p. 8. [Explain "strong" vs. "weak" here.] Among forms with heavy root syllables, both *draibīþi and *draibijanþi contrast with *drībidi and *drībandi respectively; that is, long *-ī- contrasts with short *-i- and the sequence *-ija- contrasts with *-a-. Among forms with light root syllables the situation is different. True, the *-ja- of *habjanþi contrasts with the *-a- of *grabandi; but *habiþi and *grabidi have superficially identical stem vowels, namely short *-i-. The reason is transparent enough to the linguist—there was a rule dropping *j before high front vowels unless it was in absolute word-initial position—but that might not be so obvious to four-year-olds. This is the weak point in the system from which a reanalysis might be expected to start. Moreover, the distribution of j-presents is heavily skewed in directions that might prompt a reanalysis. Hundreds of simple thematic presents with strong pasts and hundreds of j-presents with weak pasts survived (or were created) in PWGmc; but there were also exactly eleven j-presents with strong pasts—and all but one of them had light root syllables. I suggest that native language learners inferred (mistakenly) that j-present forms of strong verbs like 3sg. *habiþi had no underlying *j in the stem vowel complex, and that in forms like 3pl. *habjanþi the *j was inserted before *a—a plausible "crazy rule" for a language learner. But this crazy rule worked well; not only could it be generalized to j-presents of weak verbs with light root syllables, it could also be generalized to j-presents with heavy root syllables (nearly all of which belonged to weak verbs) because Sievers' Law still operated before back vowels in PWGmc. But once that had happened, the long *-ī- of forms like *draibīþi became opaque: it should have no underlying *j, since *j was inserted by rule in forms like *draibijanþi, so why should the stem vowel be long? And in due course it was replaced by short *-i-. If anything like that is

what happened, it should only have happened once.

And we're still not done with weird and idiosyncratic morphological innovations. B.9 details an unusual loss of *-i- in the past tense stems of weak verbs with root syllables ending in *-al-. This is usually called an early syncope (when it is not incorrectly projected back into PGmc), but it is hard to see why syncope should have occurred after light syllables but not after heavy ones, and after *l but not after *r or *n. A much more plausible hypothesis is that the loss began as an instance of lexical analogy, which probably happened once, in PWGmc, and then spread from lexeme to lexeme (at least partly in the individual histories of the daughters).

Most of the remaining morphological innovations are less striking (though the fact that the inflected infinitive, listed under B.13, is a ja-stem is so hard to explain that the formation probably should have been created only once). But the fact that so many are exclusively shared by the WGmc languages *is* striking.

Taken all together, these changes amount to massive evidence for a valid WGmc clade. But the received opinion among specialists is that WGmc is a "problem". Why??

A number of reasons can be given, all of them bad. The most egregious is a failure to understand, or perhaps a failure to accept, that shared innovations are proof of shared history, whereas shared retentions of ancestral material prove nothing (since the default development for any item in any line of linguistic descent is that native learners acquire it correctly). One seldom encounters that confusion explicitly in print. But the same result proceeds from a rather different confusion, namely failure to distinguish innovations from retentions correctly. Rösel's book of 1962 is full of examples of this. One which seems to have been widespread in the middle of the 20th century involves the etymological identity of the 3rd-person pronoun. In OHG it is vowel-initial and clearly cognate with the functionally corresponding Gothic pronoun. In northern WGmc we find a stem which rhymes with the one just mentioned but has an initial *h-*; Norse also exhibits forms with initial *h-*, though the rest of the stem is quite different. It is easy to find statements (even in encyclopedia articles) that this pattern indicates some shared history between Gothic and OHG on the one hand and Norse and northern WGmc on the other. But two different arguments based on different pieces of evidence converge on an entirely different conclusion. On the one hand, other IE languages have vowel-initial 3rd-

person pronouns (or weak deictics) that are cognate with the Gothic and OHG pronouns and deictics meaning 'this' which begin with reflexes of palatal *ḱ and are cognate with the northern WGmc stem; it follows that Gothic and OHG have retained the inherited situation, while northern WGmc and Norse have innovated (differently, therefore not necessarily together) in shifting the old deictic meaning 'this' into 3rd-person pronoun function. That's classic outgroup analysis, which any evolutionary biologist would recognize. On the other hand, fossilized relics of the *h*-initial deictic do occur in Gothic and OHG, where they consistently mean 'this'; the same conclusion follows. In other words, both outgroup analysis and the routine interpretation of linguistic fossils force the conclusion that Gothic and OHG share retained material, not an innovation, in this case. (They also make the baroquely complex hypotheses of Klingenschmitt 1987 completely superfluous.) Not to recognize that is simply incompetent. The other instances of this confusion listed under C.2 are not nearly so bad, but all are untenable. The underlying reasons for this sort of failure are of course various; the most insidious is a kind of hypercaution which has become part of the culture of comparative philology and appears deeply uneasy with any clear and precise hypothesis.

Other scholars seem to have doubted the validity of the WGmc clade from a simple failure to consider all the evidence. Again, there are all sorts of reasons for this. Perhaps the most important is that too little attention has been paid to the relative chronology of changes and to changes in unstressed syllables, both admittedly more difficult to discover and exploit than obvious across-the-board sound changes, but both crucial in this case. There was a long, intensive, and manysided discussion of those problems about a hundred years ago, but thereafter several generations of specialists either assumed that all the details had been worked out or were interested in completely different sorts of problems. A quarter of a century ago Patrick Hollifield discussed a few points ably (though not definitively) in a long and important article. Dirk Boutkan's recent book ought to have updated the discussion, but it suffers from too many flaws to have done that effectively. [Will discuss if people are interested.]

But the most important reason why the WGmc clade has been impugned is a misunderstanding of what the tree model actually models. I've discovered by personal experience that many colleagues *assume* that each line of development in a tree model

—each "edge" of the tree, to use the technical term—*must* indicate the development of a *completely uniform* dialect if the model is to be valid at all. But that is another example of the confusion between modelling what actually happened in real time and modelling what we can recover after the fact from evidence that has significantly deteriorated in the meantime. We *know for a fact* that closely related, but not identical, dialects can develop together as a single "speech community" in the larger sense; we can't rule out that possibility for the past without violating uniformitarianism. We are also aware, as working historical linguists, that even in the written record we usually cannot distinguish between closely similar dialects or subdialects; the same should be true for prehistory *a fortiori* because the evidence is even poorer. From all this it ought to follow that, in the default case, each edge of a tree represents not a single uniform dialect but a tangled skein of closely related dialects developing together.

And the reason why I have used WGmc as my example in this paper is that we can probably demonstrate that that is what PWGmc was. The evidence is the distribution of variants of a rule lowering PGmc stressed short *u before a syllable containing a nonhigh vowel.

Throughout WGmc inherited short *u was *not* lowered when immediately followed by a nasal in the syllable coda; it also was *not* lowered when the next syllable contained *j, short *i, long *ī, or short *u. (There are no examples before long *ū; the only example before *w does show lowering.) On the other hand, throughout WGmc inherited short *u was normally lowered to *o when it was not immediately followed by a nasal *and* the following syllable contained a nonhigh vowel with no *j intervening. D.1 gives examples of uniform lowering throughout WGmc, and it can be seen that they are diverse and very numerous. But there are also cases in which lowering occurred only in some WGmc languages. Those under D.2 present a clear pattern: if the consonant intervening between the *u and the nonhigh vowel was a single nasal, lowering usually occurred in southern WGmc (that is, in the OHG dialects) but not in northern WGmc (the so-called Ingvaeonic dialects). That could be accounted for positing uniform lowering followed by a re-raising of the vowel before single nasals, and some handbooks do exactly that; it might be a little less probable than failure of lowering in the first place, but not so much so that it can be excluded. Of course there were also inflectional paradigms

and derivationally related families of lexemes in which lowering occurred in some forms and not in others, and the non-lowered vowel was later levelled through; as expected, those are found throughout the family, though non-lowered *u is suspiciously common in such cases in the northern languages. Finally, and importantly, there is a small handful of words in which the northern languages unexpectedly fail to lower *u. They always contain at least one labial consonant and an *l and no other coronal consonant in the root-syllable; but stating a coherent phonological environment for the failure of lowering is impossible, because there are about as many words of very similar shape in which lowering did occur after all. Once again we have lexical exceptions uniformly distributed (as in B.6 above), which argues a single historical change.

But when did lowering of *u, with its northern WGmc exceptions, occur? It turns out that it must have preceded the loss of short low vowels in final syllables (adduced in B.2 above), which resulted in stranded word-final semivowels becoming syllabic, which must have preceded the gemination of consonants by *j—and even the last of those changes was uniformly shared by all WGmc languages. We know that because of the data adduced in D.5: consonant-stem nouns, which never had a nonhigh vowel in any reconstructable case-ending except the genitive plural, never undergo lowering of short *u in any WGmc language, whereas a-stem nouns usually do (as the numerous examples in D.1 demonstrate). So we have a non-uniform change, demonstrating that the dialects of WGmc were already different, followed by a sequence of uniform changes demonstrating that they were still developing together as a single speech community.

That's actually what we ought to expect according to the UP. We must see it so seldom because the information we have about linguistic prehistory is relatively poor. We're able to show it in this case because the information we have for PWGmc is much better than usual. Even so, it's reasonable to suspect that we could find more points in which the dialects of PWGmc differed at an early date if our information had not deteriorated as much as it has.

I submit that these considerations render irrelevant much of the endless arguing about the appropriateness of the "family tree" model—and, of course, nearly all the ink that's been spilled about the "problem" of WGmc.