Introduction to Historical Linguistics
B. Joseph

The Seeds of a Solution to the Gbe Comparative Method Problem

Parts of this problem are very straightforward, but when you add all the pieces together, it has some "wrinkles" that make it rather interesting.

In this case, you are presented with 5 sets of correspondences across 5 languages of West Africa. As always with a Comparative Method problem, the task at hand is to look at the data, determine the correspondences among the languages (in this case, involving the initial consonant in the word), state those correspondences in a clear and concise manner, and then propose a reconstruction for each correspondence, that is a Proto-Gbe segment that could have given rise in a plausible way to the correspondence sets you identify. You have to pay attention to the distribution of the correspondences and decide if any of the correspondences can be grouped together as conditioned variants of one another (conditioned by phonetic environment, per the Neogrammarian view of sound change).

You are told that there are some natural groupings of the data and that you should work on them in the order items 1-5, then items 6-9, then items 18-22, then items 10-13, and finally items 14-17.

Items 1-5 give you a correspondence set of *t* ~ *t* ~ *t* ~ *t* ~ *t*, for which the obvious reconstruction (one consistent with Ockham's Razor) is *t. Items 6-9 give you a correspondence set of *s* ~ *s* ~ *s* ~ *s* ~ *s*, for which the obvious reconstruction (one again consistent with Ockham's Razor) is *s. But a further consideration is to see whether the sets occur in complementary distribution (i.e. completely separate and distinct phonetic environments) or in overlapping distribution (i.e. at least one phonetic environment that is the same for the two sets). In this case, looking at items 2 and 7 provides a clear answer, in that the *t* ~ *t* ~ *t* ~ *t* ~ *t* set and the *s* ~ *s* ~ *s* ~ *s* ~ *s* set occur in exactly the same environment, namely before a high-toned open *o* (i.e., *ɔ́*). This tells us that the two sets cannot be resolved into a single proto-language segment and our reconstructions *t and *s as separate and distinct proto-segments is the right way to go. Recall, by the way, that the reason that the environment matters is precisely the Regularity of Sound Change Axiom – if we started with the same sound for both groups of words, then items 2 and 7 should not show different outcomes, given Regularity of Sound Change.

Moving on to items 18-22, we see the set *t* ~ *t* ~ *t* ~ *s* ~ *s*, which is similar to the set for *t and the set for *s, but is not identical to it. So we have to consider the phonetic environments and the distribution, and item 19 becomes important, because it has the high-toned open *o* (*ɔ́*), and thus the *t* ~ *t* ~ *t* ~ *s* ~ *s* set overlaps with the *t* ~ *t* ~ *t* ~ *t* ~ *t* set and also with the *s* ~ *s* ~ *s* ~ *s* ~ *s* set so all three sets must come from separate and distinct proto-language segments. Since *t has been used for the first set and *s has been used for the second set, the proto-segment for this last set can be anything other than *t or *s (due to Regularity of Sound Change), unless we want to revise our reconstructions for the first two sets (but doing so would run afoul of Ockham's Razor so it is not advisable here). Thus we might entertain here a sound that has properties of *t* and properties of *s*, that is, a dental affricate, *t^{s}. You might think of other possibilities, e.g. *θ*, which

has a reasonable phonetic connection to *t* and to *s*, so that the phonetic changes needed would not be so drastic or so numerous (thus offering a tip of the cap to Ockham!). Let's say *t^s at this point but as our reconstructions are really nothing more than hypotheses about the proto-language and not hard-and-fast truths, we should be open to the option of revising our hypotheses if needed.

Looking next at items 10-13, with the correspondences $t^s \sim t^ʂ \sim s \sim s \sim s$ (where $t^ʂ$ stands for a voiceless prepalatal affricate). As with the other sets we have looked at, we need to examine the distribution because this is similar to the other sets, though not identical, and the similarity means that it could in principle be a conditioned variant of one of the other sets. Here again, the high-toned open *o* (ɔ́) environment appears, in item 11, so this set too overlaps with the other sets, and thus must be a different proto-segment from the other three that we have reconstructed.

As for what to reconstruct for 10-13, all we know for certain is that it has to be something different from *t, *s, and *t^s, that is something different from the proto-segments reconstructed for items 1-5, 6-9, and 18-22, respectively, unless we adjust the reconstructed proto-segment for 18-22 to, say, *θ (at which point we could not use *θ for 10-13). If we use *t^s for 18-22, then a reasonable possibility for 10-13 would be *$t^ʂ$ (the prepalatal affricate) but another way to proceed would be to posit *θ for 18-22 and use the dental affricate *t^s for 10-13. There is really no easy way to choose between these two possibilities (neither Regularity of Sound Change, nor Ockham, nor phonetic naturalness of the changes needed to go from the posited proto-segment to the attested outcomes in the individual languages (the "reflexes" of the proto-sound)). For what it is worth, my preference is for *$t^ʂ$ for 10-13, but I cannot give good arguments for this preference other than a sense that *t^s > t / s is somehow more "natural" than *θ > t / s.

This leaves items 14-17. Here the correspondence set is $t^s \sim t^ʂ \sim t^ʂ \sim t^ʂ \sim ʂ$, a new set that is different from all of the other four sets. As with the other sets we have looked at, we need to examine the distribution because this is similar to the other sets, though not identical, and the similarity means that it could in principle be a conditioned variant of one of the other sets. Here again, one could point to the high-toned open *o* (ɔ́) environment in item 14, and treat this as yet another proto-language sound. That would be a reasonable account, though the choice of proto-segment becomes complicated (one could think, for instance, of using the prepalatal affricate *$t^ʂ$, which would mean using something different for 10-13).

Another way of viewing the task of dealing with 14-17 is to notice that two of the forms, items 14 and 15 have a *–y-* in Alada and 15 has a *–y-* in Fon. This means that we could posit a *y in our reconstruction of all of the forms in 14-17 and further posit that that *y conditioned the outcomes in these forms that are different from the other four sets. If we do that, it means that we could use *t^s here and posit sound changes of *t^sy > t^s, > $t^ʂ$ etc. in the respective languages. This also means that we need a further sound change to eliminate the *-y- in those forms where it does not appear. While it might seem like hocus-pocus to posit a segment and then delete it, remember that the proto-language can be quite different from the attested individual languages, and as long as we have a principled basis for doing so.