

FROM: MICHAEL DUNN

Thanks Martin for the kind words!

We've put together some materials for non-experts to help with understanding the paper: <http://language.psy.auckland.ac.nz/wordorder/>
I'm sorry that these are not currently at the appropriate level for typologists, but we'll add to them as necessary.

I don't quite get your parenthesis at the end: one of the points of the paper is that we use a particularly stringent and statistically powerful control for genealogical relatedness, and nevertheless (i) don't find many of the expected correlations, (ii) find many correlation which were unexpected, and (iii) find that even where dependencies are found between the same pairs of features in two lineages, the evolutionary models underlying these dependencies are different.

Best, Michael

On Wed, Apr 13, 2011 at 10:06:41PM +0200, Martin Haspelmath wrote:

```
> In my recollection, this is the first typology article to be
> published in Nature:
> http://www.nature.com/nature/journal/vaop/ncurrent/full/nature09923.html.
> Congratulations to our colleagues in Auckland and Nijmegen!
>
> There is also a popularized account in Nature News
> (http://www.nature.com/news/2011/110413/full/news.2011.231.html),
> and a Nature editorial about "Universal truths"
> (http://www.nature.com/nature/journal/v472/n7342/full/472136a.html).
> How wonderful to see that typology has become so important!
>
> (Unfortunately, I don't see what is new in the paper -- maybe
> someone can explain this? Didn't we know all along that we are not
> likely to get correlations if we don't control for genealogical
> relatedness?)
>
> Martin
```

FROM: MATTHEW DRYER

I am not sure whether this list is the appropriate venue for commenting on the paper in Nature by Michael Dunn et al. But since it is, as Martin said, unusual for a typological paper to appear in Nature, and since the paper was brought up on this list, I think some very brief comments are in order. My apologies to people on the list who have not had an opportunity to read the paper.

Put briefly, the paper is based on major misunderstandings of the claims of word order typology. All of the results of the paper are already familiar to me and are entirely consistent with claims that have been made in the word order literature. If I can take the liberty of quoting Michael's own words from his email,

"(i) don't find many of the expected correlations"

What the paper shows is that we often don't find the expected correlations WITHIN language families. But there are many reasons why we should expect this, and nothing in the word order literature would lead us to expect otherwise. We only expect to find the expected correlations ACROSS families.

"(ii) find many correlations which were unexpected"

In fact, the correlations of this sort mentioned in the paper are well-known, such as the correlation between the order of adjective and noun and the order of relative clause and noun. There is nothing unexpected about these correlations.

"(iii) find that even where dependencies are found between the same pairs of features in two lineages, the evolutionary models underlying these dependencies are different"

There are potentially novel results here, but I see no reason to think that these differences are due to anything other than random variation.

Matthew

FROM: MICHAEL DUNN

Thanks for your comments, Matthew. And thanks Jess Tauber for putting up a link. Here's another link that should give free access to both the paper and the supplementary materials (also important, given the brevity of Nature papers): <http://www.mpi.nl/publications/escidoc-95245/>

I'd like to respond to Matthew's comments on the list, because I think they touch on some issues which are important to all linguistic typologists.

The claim that some word order dependencies should only exist ACROSS families but not WITHIN them needs unpacking. Unrelated languages are products of historical processes too, it's just that these historical processes are unknown or (if you use sampling methods) ignored. The Phylogenetic Comparative Method test for dependency used in our paper (described in the Supplementary Materials) controls for genealogical relatedness by looking within a lineage and inferring the extent to which changes in two features are 'coupled'. In some cases, changes in one feature are regularly associated with changes in another, in other cases they're not. For the well-known correlation between verb-object order and Adposition order, phylogenetic comparative methods do detect the correlation within the Indo-European and Austronesian families. But the correlation is absent in Uto-Aztecan, despite that fact that verb-object orders and adposition orders do change within the history of the family.

It could be argued I suppose that this is a statistical universal. But the usual understanding of statistical universal (e.g. Dryer 1998) is that statistical universals are universal tendencies: the tendency itself should be present universally. If the tendency is absent in some lineage then it's not a statistical universal (unless you wish to admit statistical statistical universals!), it's a lineage specific process.

As to the unexpected correlations, Dryer 2007 ("Word Order", in the Shopen trilogy) makes strong statements about what word order features show intercorrelation and what don't. Our results (Figure 2) show that this differs from lineage to lineage. A correlation between SV and OV exists in Uto-Aztecan, but not any of the other families. There is no correlation between order of adjective and noun and relative clause order in Indo-European (although there is in Austronesian and Uto-Aztecan), but Indo-European alone does have an evolutionary correlation between adjective-noun and genitive-noun orders.

So far I've just been talking about tests for the existence of dependencies between particular pairs of features. An additional thing we get for free from the comparative phylogenetic approach is an explicit model of evolutionary change for each feature. Where a pair of features are correlated, the method infers which changes between states are more or less probable. Figure 3 in the paper shows the different models of transition probabilities inferred between V0 and adposition states in Austronesian and Indo-European. The patterns of evolutionary change inferred in the two families are different from one another: these features are dependent in both families, but it's not the same dependency. We can do this for any grammatical dependency detected in a family, something I am quite excited about for future work.

I hope this somewhat clarifies things for LingTyp readers. I would be happy to expand on this discussion if there's interest.

Best, Michael

FROM: S. WICHMANN

The model of coevolution was developed in biology, where it is highly meaningful to test whether features that seem to be correlated change in tandem along phylogenetic lineages, since biological traits normally don't spread laterally (bacteria providing a major exception). But for linguistic typological features we know that there is a strong areal effect, so if this effect is not somehow built into the evolutionary model one can only expect results that are difficult to interpret. It is hard to expect correct conclusions from the wrong premises. If the premise is that correlated linguistic feature should coevolve historically and the conclusion is that this evolution is lineage-specific then we should pause before accepting this conclusion. If we are to adopt methods from biology when studying linguistic typology we should be careful about which models to adopt.

Søren.

FROM: MICHAEL DUNN

Hi Søren,

Phylogenetic comparative methods are agnostic to the factor that prompts the change. The existence of a functional dependency would predict that the dependent elements would be regularly borrowed together, or if not, that they would regularly harmonize; conversely, lack of functional

dependency would predict that the features could be borrowed independently. I'm not sure what the systematic bias would be in that case: you could argue that contact would exaggerate the appearance of functional dependency where it doesn't really exist (so, the link between VO and ADP order in AN might be a replicate of an accidental association of these features in some non-AN languages), but you could equally argue that if these features were independent then they shouldn't be borrowed together so regularly. Basically, the source of innovation is not relevant at this level of analysis.

Best, Michael

FROM: M. DRYER

Regarding

"But the usual understanding of statistical universal (e.g. Dryer 1998) is that statistical universals are universal tendencies: the tendency itself should be present universally. If the tendency is absent in some lineage then it's not a statistical universal (unless you wish to admit statistical statistical universals!), it's a lineage specific process."

That is most emphatically not what anyone has ever meant by a statistical universal. I doubt that anyone has ever proposed a statistical universal that was not already known not to manifest itself in all language families. If something is found in more language families than one expects due to chance, then it is a statistical universal. That's all.

To make this more concrete, consider instances of competing motivations, since this a particularly clear case where we don't expect statistical universals to manifest themselves in all language families. Jack Hawkins has argued that there are in effect two competing motivations underlying the order of relative clause and noun. One of these is a principle whose effect can be expressed by a statistical universal "A language is RelN if and only if it is OV". The other is a principle whose effect can be expressed by a statistical universal "Languages tend to be NRel rather than RelN". Now these two principles are in competition in OV languages, one principle favouring OV&RelN, the other favouring OV&NRel. The claim is that these two principles are universal in the sense that they exert functional pressure on all languages. However, they certainly are not universal in the sense that we will see their effects in all language families. While there may be some families in which the effects of both principles manifest themselves, it is far more common for one principle or the other to "win out" in a given language family so that each of the two statistical universals that result from these two principles will manifest themselves only in some language families. Many of the differences between families discussed in the paper probably reflect nothing more than competing motivations.

There are two issues here. One is the detailed nature of the differences between language families discussed in the paper. These are certainly worth examining and potentially revealing and the software Dunn et al use may indeed be useful to bringing out these patterns. But in my opinion what they might reveal is universal principles. Where we find differences between families, they may reflect different options which are universally available. However, I admit that there is certainly room for various alternative interpretations of these patterns.

The second issue is whether the data provide any evidence against existing claims in word order typology. Here there is less room for debate: it is clear that they don't. While there is room for different interpretations of the data presented in their paper, this data does not provide any argument against existing claims.

Matthew

FROM: CROFT

We should not lose sight of what is good in Dunn et al.'s Nature article. Many typologists, starting with Greenberg, have argued that synchronic language universals are really just manifestations of diachronic universals. The state-process model used by Dunn et al. (see p. 5 of their supplementary materials) has been used by Greenberg (1978) and Maslova (2000). Dunn et al.'s method (and also Maslova's) allows us to take a major problem with sampling - historical dependence - and exploit it to uncover valid language universals. Using quantitative and statistical techniques will allow us to make more precise generalizations and assign a degree of goodness of fit to our theoretical models of typological universals.

Nevertheless, like others here, I am unconvinced of the results due to problems with the way they apply the method. Dunn et al.'s analysis (pp. 5-6 of the supplementary materials) treats two models as mutually exclusive: an "independent" model, in which a word order switches independently of other orders, possibly with a weighted preference for one order; and a "dependent" model, in which two word orders are linked and change together. These competing models are tested against each phylogeny (family tree), with one or the other winning out (*very* crudely, if branches with linked changes outnumber branches with single order changes, then the test takes the word-order correlation as justified for the tree). But this does not test the model that most typologists assume; at best it tests the oversimplified model of Theo Vennemann and Winfrid Lehmann from the 1970s, which reduced the diversity of word order in the world's languages to just two types, VO and OV. (This is also basically the generative head-ordering parameter.) But from Greenberg's original paper (Greenberg 1966) onwards, most typologists have adopted a model in which both single-order preferences (Greenberg's 'dominance') and linkages (Greenberg's 'harmony') compete with each other and jointly determine patterns of word order variation, as Matthew noted in his last post. This can be demonstrated statistically: Justeson and Stephens (1990) did a log-linear analysis on a large synchronic language sample and showed that the best-fit model included both dominance and harmony factors. I presume there is a dynamic equivalent of log-linear analysis that could be used to test the model that most typologists accept.

Regarding the issue of lineage-specific vs. universal patterns: Keith Poole and I had to deal with a similar question in using multidimensional scaling to find universals of grammatical categories (Croft and Poole 2008). There, the contrast was between language-specific MDS models (as used, for example, by Barbara Malt and her colleagues) and crosslinguistic MDS models. We argued that if the regularities were culture-specific, then mixing in languages with culturally-specific category structures would reduce the goodness of fit of the MDS model; but if the regularities were crosslinguistic (i.e. universal), then mixing together languages would improve the goodness of fit of the model. (We found the latter.) The same presumably would apply for word order patterns, *mutatis mutandis*. But here the problem with the Dunn et al. result is that they test only four lineages, representing only 7.5% of language genera (low-level language families) in the world; the other 92.5% of language genera occur

in other lineages. Also, for many of the pairwise correlations they test, including the two illustrated in their article, Bantu is too shallow a family to exhibit any variation; and for the two illustrated in their article, even in the other families very few branches undergo a word order shift. So the empirical sample, though well distributed geographically, is very small and has few independent changes to evaluate. And here we hit a problem that all researchers on language universals hit: our uncertainty about language phylogeny in most parts of the world, especially for deeper families, but even for subgrouping in accepted families. In this latter area, quantitative methods are also being applied; but that is another story.

I think typologists should welcome this effort to marry phylogeny and typology, even if we remain unconvinced of the particular result in this paper.

Bill