Conceptual maps using multivariate statistics: Building bridges between typological linguistics and psychology

ASIFA MAJID

1. Why the gulf between typology and psychology?

What sorts of entities do people distinguish in the world? How do we conceptualize ''objects'' and ''events'', ''time'' and ''causation''? How do we think about people interacting with each other in social worlds? Typological linguists answer these questions by uncovering universals and variation in language structure, the assumption being that universals inform us of common conceptual repertoires while variation enlightens us about the potential of human mental faculties. These are the same properties that psychologists struggle to understand. Yet, most work by typologists goes unnoticed by psychologists. Why should this be so, when both groups are interested in the same underlying questions?

Part of the problem is the divergence of methodologies. Psychologists espouse the scientific method. Phenomena are operationalized so that they can be measured, and data are collected through a standard procedure. This ensures that any finding can be replicated by other researchers. Typologists have relied more (though not exclusively) on extracting patterns from grammatical descriptions, with little concern for how the language data are collected or whether another typologist might replicate any conclusions they reach. This has been true of work on classical semantic maps. A semantic map is a plot of related grammatical functions, where functions that are expressed by a single construction are plotted adjacent to each other (usually with a connecting line). Functions not encoded by a single construction are not adjacent (or connected). The functions themselves, as well as the extensions of specific language constructions, are typically derived by the analyst from secondary sources, such as grammars. Language is idealized away from observable, measurable,

empirical evidence. Croft and Poole's (this issue) paper is exciting because it suggests a possible scientific approach to the study of typological data, and as a consequence promises to build bridges between linguistic typology and psychology.

2.   Using multivariate statistics to construct conceptual maps

Croft and Poole show how typology can move away from inducing patterns based on idealized grammars to collecting real speaker data and using statistical models to uncover patterns. They analyze data from Östen Dahl's (1985) tense, mood, and aspect-questionnaire (TMA-Q), which is based primarily on native speaker responses from 64 languages for almost 200 sentence contexts. Croft and Poole examine how constructions from these different languages categorize the diverse sentential contexts, and thereby infer the most important underlying conceptual dimensions for this domain (past-future; perfective-imperfective), as well as functions (as determined by sentence-context clusters).

My colleagues and I have used a comparable method to study the linguistic encoding of events of cutting and breaking (Majid et al., 2007, forthc.)[1]. There are minor differences in how the data are elicited and how they are subsequently analyzed, but the approach is broadly similar. In terms of the data, Dahl's database is constructed from responses to linguistic materials. Speakers are presented with sentences in English and they have to provide TMA-marked translation equivalents in their native language. In contrast, my colleagues and I use non-linguistic stimuli, such as pictures or videoclips, to limit possible influences of a contact language. The stimuli are used to elicit descriptions, and these are then subjected to multivariate statistics. Here again there is a slight difference between our work and Croft and Poole's, namely in the precise statistical techniques employed. Overall, however, the logic is the same. When large numbers of languages are studied, the statistical tools reveal underlying structure common to all languages which can be interpreted as reflecting a universal conceptual structure (cf. Croft, 2003).

---

[1]   See also Levinson et al. (2003; cf. Bowerman & Pederson, 1992) on adpositions.

For the domain of cutting and breaking events, we found a small number of dimensions underlying lexical categorization in nearly 30 typologically, geographically and genetically different languages (Majid et al., 2007, forthc.). The first and most important dimension distinguishes events where the precise locus of separation in the object is predictable from those where the locus of separation is not predictable. This corresponds roughly to the distinction between *cut* and *break* in English. The second dimension pulls apart events in which flexible 2-dimensional objects are separated by hand (i.e. *tear*). Among events where the locus of separation is unpredictable, the third dimension further distinguishes separation of a 1-dimensional rigid object by pressure on both ends (i.e., *snap*) from separation caused by a sharp blow to a rigid object (i.e., *smash*). In our study, languages respect the overall conceptual space, regardless of whether they use just three verbs to cover all the videoclips (as Yélî Dnye does; Levinson, 2007) or more than fifty (as Tzeltal does; Brown, 2007). This can be demonstrated statistically by calculating correlations between the dimensions extracted by the general solution and those extracted for each language individually (Majid et al., forthc.).

In order to discover these regularities, we used correspondence analysis. Croft and Poole (this issue, p. 13) criticize this method, along with two other common multivariate statistical techniques – factor analysis and principle components analysis. They do this on the grounds that their technique, multidimensional scaling (MDS) captures all of the variance in the data in just a few dimensions, while correspondence analysis, factor analysis and principle components analysis capture variance in many more dimensions. They argue that MDS should be preferred over the other techniques because (1) their technique calculates the degree of fit between the data and the model and (2) MDS models are invariant with respect to translation and rotation of the final solution extracted.

MDS is a powerful tool for analyzing typological data but I disagree with the implication that a single model is appropriate for all language datasets. MDS, correspondence analysis, and the rest are just tools. Their selection should be based on the particulars of the job at hand. Poole's nonparametric binary unfolding algorithm is an excellent addition to the collection, particularly for its abilities to handle lopsided datasets, a property that seems to be rampant in language data. But other tools are

more suited to other datasets. For example, if Dahl's dataset had information from dozens of speakers for each language (instead of just one or two), we could use more powerful models to better capture the parametric nature of the data.

Croft and Poole's objections to other plotting techniques do not pick out fatal flaws. First, there are methods for calculating the goodness-of-fit of a factor-analytic solution to the data, for example, using maximum likelihood estimation (Fabrigar et al., 1999). This technique can be used to determine how many dimensions to retain in the analysis, just as for MDS. Second, for models that do not provide a measure of ''fit'', like correspondence analysis, there are other ways to confirm the goodness of the model. For the cutting and breaking study, we did this by correlating the general solution with individual language solutions. In this way, we could determine to what extent the individual languages were making the same distinctions as found in the general solution. Third, all of the techniques – MDS, factor analysis, correspondence analysis – provide a measure of how much variation in the data is accounted for by the model. Variation accounted for is a valid and informative measure for determining whether statistical models are revealing interesting structure from linguistic datasets, and are equally as important as goodness-of-fit measures. A dimension that only accounts for 1% of the data is less interesting than one that accounts for 50%. Croft and Poole assume that it is always preferable to have the smallest number of dimensions accounting for the majority of the data. But this can lead to distortions. If the underlying structure is actually three-dimensional, then a model with only a one-dimension solution is not a good representation. Finally, Croft and Poole worry that the solutions from factor analysis are not invariant under translation or rotation. Why is this so important? Sometimes the dimensions extracted by a model are not the most interpretable ones. In Croft & Poole's analysis of the TMA-Q data, the past-future and perfective-imperfective dimensions are orthogonal to dimensions 1 and 2 of the MDS solution. So statistical fit to the data is not the only criteria of importance; the fit has to be meaningful too. In factor analysis there are tools for rotating a matrix so that items load on interpretable dimensions. New eigenvalues are calculated, and the model's ability to account for the data can be measured. So, Croft and Poole's objection is not critical.

Croft and Pool lump together correspondence analysis with factor analysis and principle components analysis. But correspondence analysis differs interestingly from these techniques in how a similarity space is constructed and how the data are analyzed. In correspondence analysis, a matrix can be asymmetrical. For example, the rows could be stimulus events and the columns linguistic forms, such as the verbs or constructions used to describe them. The values in the matrix are counts, e.g., how many instances of each form were used to describe that stimulus. Poole's algorithm also operates on an asymmetrical matrix. But unlike Poole's algorithm, correspondence analysis performs a dual factoring on both rows and columns. This means that not only can stimuli be plotted on the basis of their similarity to each other, but verbs can also be plotted in the same space. This makes it a very powerful technique for some issues. For instance, we have used it in developmental work to study how children's use of cutting and breaking verbs compares to that of adults (Bowerman et al., 2004) – to establish the use of verbs by children of different ages, and how verbs are distributed over stimuli. This would not have been possible in a single analysis using these other techniques.

3.   The psychological consequences of conceptual maps

Croft and Poole suggest that their models uncover "a universal conceptual structure in the minds of human beings" (p. 5f.), and that "the conceptual space modeled by MDS is hypothesized to be the same for all speakers" (p. 11). This is an intriguing claim that promises fruitful exchange between typologists and psychologists. Conceptual maps extracted from typological data can provide the foundations for hypotheses-testing. For example, on the basis of the TAM-Q conceptual map we could predict that all people should distinguish between past and present and between perfective and imperfective. Another prediction could be that sentence contexts within a cluster of the MDS solution should be more confusable with each other than contexts in different clusters. Hypotheses derived from conceptual maps need not be restricted to linguistic materials. Croft's (2003) claims are about conceptual structure more generally, so hypotheses can be derived about the nature of non-linguistic representations too.
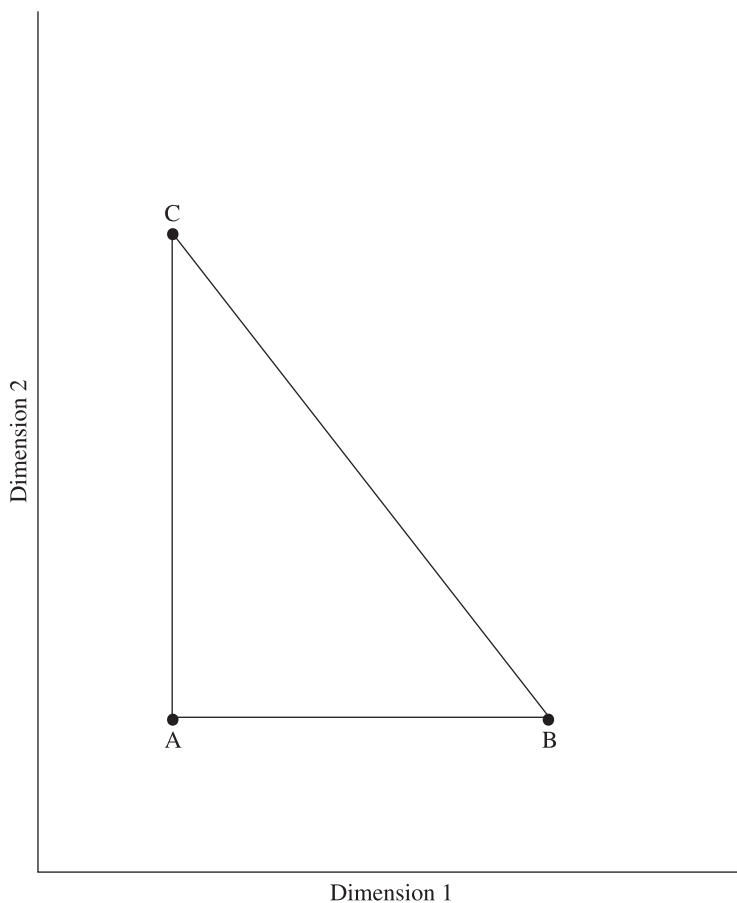
Figure 1.   *Stimuli A, B and C plotted in a two dimensional space.*

If the statistical models extracted from typological datasets are to be considered psychologically real, then it is worth examining the underlying assumptions of these techniques more closely. In MDS, correspondence analysis, etc. similarity between objects is depicted in Euclidean space. There are three core ideas underlying Euclidean space: (1) the distance between a point and itself is identical, (2) the distance between point A and B is the same as the distance between point B and A, and (3) the distance between point B and C is defined by the Pythagorean relation (see Figure 1).

There are reasons to think that psychological space does not always correspond to Euclidean space. In a classic paper, Tversky (1977) asked participants to make a range of similarity judgments and showed that all three of the above principles are defeasible. For example, in one study participants were asked to rate the similarity between China and North Korea. Participants always rated North Korea as being more similar to China than China is to North Korea. This is contrary to principle (2), which states that the distance between two objects be symmetrically identical. Principle (3) can also be challenged. In Euclidean space the distance between B and C is along the diagonal in Figure 1. But it is also possible to get from B to C via A. This is the city-block metric. In perceptual judgments, integral concepts (e.g., color) are modeled better with Euclidean space while separable concepts (e.g. geometric concepts varying in form and size) are modeled better with a city-block metric (Shepard, 1964).

Neither of these illustrations concern language data, but in all likelihood the general points carry over to linguistic judgments too, a discomfiting fact for those of us using metric models. Nevertheless, metric modeling has been highly useful in psychology for understanding perceptual and conceptual spaces and in this new guise – as a tool for linguistic typologists – it promises to uncover hitherto unknown structures.

Max Planck Institute for Psycholinguistics
Asifa.Majid@mpi.nl

Acknowledgments

References

Bowerman, M. and Pederson, E. (1992). Crosslinguistic perspectives on topological spatial relations. Paper presented at the annual meeting of the American Anthropological Association, San Francisco, CA, December.

Bowerman, M., Majid, A., Erkelens, M., Narasimhan, B. and Chen, J. (2004). Learning how to encode events of cutting and breaking: A crosslinguistic study of semantic devel-

opment. Poster presented at the Child Language Research Forum, Stanford University, CA, April.

Brown, P. (2007). 'She had just cut/broken off her head': Cutting and breaking verbs in Tzeltal. *Cognitive Linguistics*, 18, 319–330.

Croft, W. (2003). *Typology and Universals, 2nd ed.* Cambridge: Cambridge University Press.

Croft, W. and Poole, K. T. (in press). Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics*, this issue, 1–38.

Dahl, Ö. (1985). *Tense and Aspect Systems.* Oxford: Basil Blackwell.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.

Levinson, S. C. (2007). Cut and break verbs in Yélî Dnye, the Papuan language of Rossel Island. *Cognitive Linguistics*, 18, 207–218.

Levinson, S. C., Meira, S. and The Language and Cognition Group (2003). 'Natural concepts' in the spatial topological domain – Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language*, 79, 485–516.

Majid, A., Bowerman, M., van Staden, M. and Boster, J. S. (2007). The semantic categories of cutting and breaking: A crosslinguistic perspective. *Cognitive Linguistics*, 18, 133–152.

Majid, A., Boster, J. S. and Bowerman, M. (forthc.). The crosslinguistic categorization of everyday events: A study of cutting and breaking.

Shepard, R. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54–87.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.