Теоретически информированные параметры

Существует множество суждений о параметрах, важных для репрезентативности выборки. Более- менее все согласны, что выборка должна быть сбалансирована ареально и генеалогически. Кроме того, нередко требуют различий по грамматическим параметрам (разумеется, не проверяемым в самом исследовании). Здесь и возникает проблема.

Грамматические параметры нередко оцениваются поверхностно и «нетрансформационно», что приводит к неструктурированным и нелогичным признакам. Наиболее ярким примером служит базовый порядок слов в предложении, который, по (Dryer & Haspelmath, 2013), имеет семь независимых значений: SOV, SVO, VSO, VOS, OSV, OVS и «нет базового порядка». Более того, ввиду поверхностной процедуры оценивания язык с синтаксисом вроде немецкого, различающимся для, например, главных и придаточных предложений, получает последнее значение.

В то же время, если подойти к этой же проблеме с точки зрения «трансформационных» теорий, можно на месте одного неструктурированного признака получить несколько структурированных. Например, на месте порядка слов можно получить четыре признака: «левое/правое ветвление (предсказывающее, в частности, OV/VO)», «вынос глагола влево (в главных/во всех/нет)», «вынос VP влево (+, например, у OVS и VOS; потенциально тоже в главных/во всех/нет)», «вынос чего-либо на левую периферию (+, например, у OSV и немецкого типа; ярлык условный)». Они уже предсказывают 20 типов, так как вынос глагола влево и вынос VP влево не сочетаются – это не теоретическое ограничение, а ограничение данных, где их сочетание едва ли будет отличимо от выноса объекта (т. е. + у последнего признака). Среди этих типов все старые не теряются и дробятся неслучайным образом, позволяя извлечь более глубокие закономерности. Например, у «левого ветвления» корреляция с «предлоги/послелоги» будет намного выше, чем у изначального порядка слов.

Разумеется, возникает вопрос о том, не оказывается ли определение значений признаков в этом случае зависимым от выбранной теории. Однако никакие признаки не являются внетеоретическими (ввиду имплицитных предположений лингвистов), так что речь идёт о выборе между теориями, а не о наличии или отсутствии оной. И теории, используемые для ряда признаков сейчас, довольно поверхностны – как в прямом, так и в переносном смысле.

Библиография

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. The world atlas of language structures online. Leipzig: Max-Planck-Institut für evolutionäre Anthropologie. http://wals.info/