



# John Nerbonne

(Rijksuniversiteit Groningen / FRIAS Freiburg)

## „Der Traum des Strukturalismus“

Vortrag am 30.01.13 um 18 Uhr c.t.  
Raum S 005, Schellingstraße 3 (Vordergebäude)

Vom Strukturalismus kennt man den berühmten Spruch, die Sprache sei „un système où tout se tient“ (Meillet, 1903:407), mit Schwerpunktsetzung auf den *Beziehungen* von sprachlichen Elementen untereinander. Den Strukturalisten zufolge ist das System dieser Beziehungen wichtiger als der konkrete Inhalt der Elemente selbst. Das Phoneminventar wäre demnach wichtiger als die Details der Aussprache (oder der auditiven Verarbeitung). Das moderne Studium von Sprachlauten gibt den Strukturalisten allerdings Unrecht durch seine Schwerpunktsetzung auf die konkrete Realisierung, und es hält genaue Messungen von Sprachproduktion und -wahrnehmung immer im wissenschaftlichen Fokus. Instrumentelle Phonetik und „Laboratoriumphonologie“ sind in dieser Richtung führend.

Computerlinguistik und Korpuslinguistik betonen ihrerseits, dass allerlei Informationen in Verteilungen latent anwesend sind und doch zum Vorschein kommen können, wenn man sie richtig betrachtet. In diesem Vortrag leiten wir inhaltliche Information aus Beziehungen ab, insbesondere akustische Informationen aus den Verteilungen alternativer Aussprachen, wie sie in verschiedenen Sprachvarietäten vorkommen. Wir setzen voraus, dass man Daten über variante Aussprachen zur Verfügung hat und gebrauchen in konkreten Experimenten die Dialektatlasdaten von sechs verschiedenen Atlanten. Jeder Atlas enthält die phonetischen Transkriptionen der Aussprache derselben Wörter an Hunderten von Erhebungsorten. Wir gebrauchen einen einfachen ALIGNMENT-Algorithmus (entwickelt von Levenshtein), um die entsprechenden Segmente bei allen Wortpaaren zu finden.

Wir organisieren alle Segmententsprechungen in einer großen Kontingenztabelle, um die Assoziationsstärke mittels POINTWISE MUTUAL INFORMATION (PMI) aus der Informationstheorie zu messen. In diesem Schritt schließen wir auf die phonetische Nähe von zwei Segmenten aus der Frequenz von deren Entsprechung. Auf diese Weise erhalten wir ein Maß des phonetischen Abstands zwischen allen Segmentpaaren. Segmenten, die oft in Dialekten variieren, z.B. [s] und [z], werden kürzere Abstände zugewiesen. Ein zweiter einfacher Algorithmus iteriert Alignment und die Feststellung der Abstände so lange, bis beide stabil sind. Wir evaluieren den Prozess dann aufgrund der Korrelation zwischen den induzierten phonetischen Abständen und bekannten akustischen Messungen. Bei den Vokalen in allen Atlanten (Deutsch, Niederländisch, amerikanischem English, Gabon Bantu, Toskanisch und Bulgarisch) finden wir starke, signifikante Korrelationen zwischen den induzierten, verteilungsbasierten Abständen und den akustischen ( $0,6 < r < 0,8$ ). Der akustische Inhalt wird der Verteilung der Segmente entnommen.