Announcing the

Final Examination of

Nathan Sanders

for the

Degree of Doctor of Philosophy in Linguistics

June 11th, 2010, 1 PM

Room 317, Memorial Hall

Dissertation: A Statistical Measure for Syntactic Dialectometry

This dissertation investigates a statistical measure for syntactic dialectometry, using grammatical features of Swedish dialects to measure how much they differ. It analyzes variant feature sets and

measures to determine which provide the most accurate results. The results are compared to existing Swedish syntactic dialectology and phonological dialectometry.

The statistical measure requires two functions: the first extracts syntactic features from an interview corpus, and the second compares the features from two such corpora. It does this by comparing the number of each feature type between the two corpora; the sum of differences is the total distance. Several variants of both functions are defined: the feature extraction functions vary in complexity from surface-oriented parts of speech up to subtrees of the syntax tree. The comparison functions also vary in complexity; the simplest is described above, while the more complex are based on information theoretic measures.

The distances are compared to previous work in three ways: agreement with (1) overall Swedish dialect geography, (2) specific dialect regions and (3) specific dialect phenomena. For traditional dialectology, the overall geography agrees moderately well, the specific regions are very similar, and agreement with specific phenomena is inconclusive. For phonological dialectometry, the agreement is very good in all three areas.

This dissertation shows that a statistical measure of syntactic dialect distance is a useful addition to dialectometry, extending the field beyond phonology and the lexicon. It shows that the measure's results on Swedish agree with those of related fields. And it provides a path for future work by finding the best variations of the measure, given various practical conditions.

|  |  |
| --- | --- |
| *Outline of Current Studies*  Major: Linguistics  Minor: Computer Science | *Educational Career*  Ph D, Indiana University, 2010  MA, Indiana University, 2006  BA, College of the Ozarks, 2004 |

Assistant Professor Sandra Kübler, Chairperson

(812-855-3268), Linguistics

Assistant Professor Markus Dickinson

(812-856-2535), Linguistics

Associate Professor Michael Gasser

(812-855-7078), Computer Science

Professor Steven Franks

(812-855-8169), Linguistics and Slavic Languages and Literatures

Approved:

Sandra Kübler

(Any member of the Graduate Faculty may attend. As a courtesy, please notify the Committee Chairperson in advance.)