Studying low-frequency grammatical variables in dialects of English on Twitter

Researcher

Dr Chris Strelluf

Research focus

Some linguistically important grammatical variables occur infrequently or unpredictably in natural language, so it has not been possible to study productions of these variables.

Studies of these variables have reported responses to elicited grammaticality judgments. However, speakers are not always conscious of these variables occurring in speech, and can therefore not reliably answer questions about their grammaticality.

Methodology

I collected more than 4 million tweets containing low-frequency grammatical variables that had received attention in dialectological, semantic, and syntactic research, particularly in American Englishes. I mapped the locations of tweets’ authors to describe the regional distribution of these features. I analysed syntactic, semantic, and other features in these messages.

Key findings:

- Low-frequency grammatical features provide long-term evidence of historical settlement patterns (e.g., migration from Scotland to Ireland to the US Midland) long after more prominent features of the language varieties have changed.

- Low-frequency grammatical features adhere to boundaries for regional dialects that were previously determined on the basis of phonetic and lexical data, even after differences in phonetics and lexis have begun to disappear.

- Production results broadly confirm the conclusions of previous studies which were based on problematic methodologies. However, production data reveal fine-grained intra-regional differences. For instance, grammatical acceptability judgments show that the feature positive-

  anymore

is present through the US Midland. Production data show that, while the feature may be grammatical throughout the Midland, it is used much more in eastern Midland cities than western ones.

To find out more: