

# TSHRDLU Modifications – Project Phase # 3

**Jim Evans**

Department of Linguistics  
University of Texas at Austin  
j.s.evans@utexas.edu

**Jason Mielens**

Department of Linguistics  
University of Texas at Austin  
jmielens@utexas.edu

## 1 Behaviors

We modified the behavior of the automated Twitter user "tshrdlu." First we give a brief description of the four behaviors, with more details in the next section. The first is a listening behavior. Every thirty minutes the bot searches for tweets that contain both the word "vaccine" and a link. The bot uses a corpus-trained Naive Bayes classifier to identify the linked webpage as either "pseudoscience" or "science." The bot tweets about classification of the webpage and provides the link to it in the tweet. The bot also has four possible behaviors for responding to tweets directed at it (two of those are the original tshrdlu behaviors):

- paraphrase: it responds by taking the original tweet and replacing longer words with randomly-selected synonyms of those words.
- original standard: searches Twitter using a query made up of a subset of the original tweet's words and tweets the most relevant result (original tshrdlu behavior),
- enhanced standard: the same as the previous behavior, except the query is made up of synonyms of the words from the original tweet, rather than the exact words from the tweet.
- follow: follows users that people ask it to follow (original tshrdlu behavior),

### 1.1 Listening

Periodically the bot searches for tweets that contain both the word "vaccine" and a link, and the bot classifies the linked page as "science" or "pseudoscience". This Naive Bayes classifier was trained on the corpus created by Beth and Evans (2012), in which documents (blog posts) are tagged as either "science" or "pseudoscience". The test accuracy was approximately 88 percent, and the training accuracy was about 92 percent.

### 1.2 Behaviors for mentions

Two novel mention behaviors were created for the bot, while retaining tshrdlu's original "follow" behavior, and "backing off" to tshrdlu's original "standard" behavior in some situations. The bot randomly chooses, with some probability, whether to use its "paraphrase" behavior