

Adaptive Word Prediction for AAC

Keith Trnka Kathleen F. McCoy
University of Delaware

problem

many inappropriate predictions

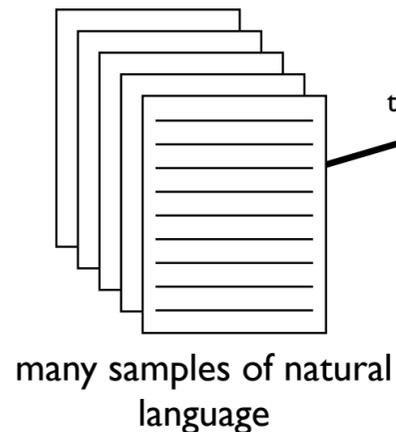
they bought a _____

house
skateboard
fashionable
carton
dilapidated

effects of context

- what if the user was already talking about milk?
- what if they were writing a novel vs. talking to family?

background



- basic ngram methods are reasonable at predicting grammatically appropriate words
- prediction quality depends on the samples
- training texts similar to actual usage are best
- but the topic and style of actual text varies!

solution

tune to training text

topic

style

- predict words similar to those being typed
- focus the training on texts that use the same words as the user
 - automatically boost words that are often go with the words already typed
- learn the grammatical patterns the user is typing (part of speech)
- alter the predictions to focus on similar grammatical patterns
- example: active vs. passive voice

integrate user text

- basic example: recency
- our solution: integrate user's words seamlessly using part of speech information
- directly addresses the need for appropriate training text
- using part of speech allows us to take advantage of the learned part of speech model

email excerpt

Switchboard is really low .
NNP VBZ RB JJ .

This could reflect that we chose
DT MD VB IN PRP VBD

a good corpus originally , maybe that
DT JJ NN RB , RB IN

the cleanup was more consistent
DT NN VBD RBR JJ

(I do n't think it 's any
-LRB- PRP VBP RB VB PRP VBZ DT

more advanced than the others ,
RBR JJ IN DT NNS ,

but I think I spent far more time on it
CC PRP VBP PRP VBD RB JJR NN IN PRP

paper excerpt

The self-test analysis is affected
DT JJS NN VBZ VBN

by both the size of the corpus
IN DT DT NN IN DT NN

as well as the diversity of the corpus
IN RB IN DT NN IN DT NN

, which explains the trend with Switchboard
, WDT VBZ DT NN IN NNP

: participants in the corpus collection
: NNS IN DT NN NN

were restricted to one of roughly 70 topics
VBD VBN TO CD IN RB CD NNS

, most of which are represented
, JJS IN WDT VBP VBN

in every set of Switchboard .
IN DT NN IN NNP

... I think this analysis of Switchboard shows
PRP VBP DT NN IN NNP VBZ

user's word model