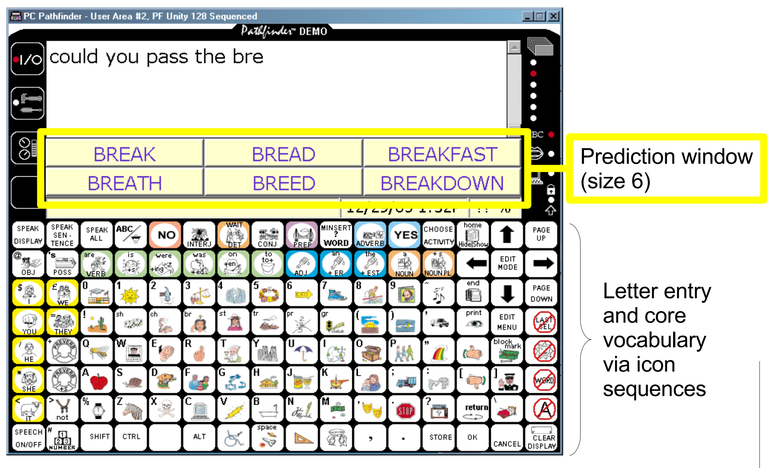
Topic Modeling in Fringe Word Prediction for AAC

Keith Trnka, Debra Yarrington, Kathleen McCoy
University of Delaware

Christopher Pennington
AgoraNet, Inc.

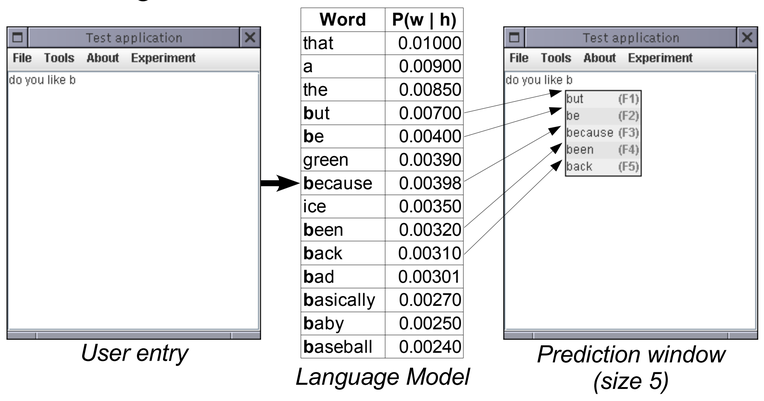
Motivation

Alternative and Augmentative Communication (AAC) is the field of research concerned with finding ways to help those with speech difficulties communicate more easily and completely. AAC devices such as PRC's Pathfinder (below) attempt to mitigate the lowered communication rate with unique user interfaces.



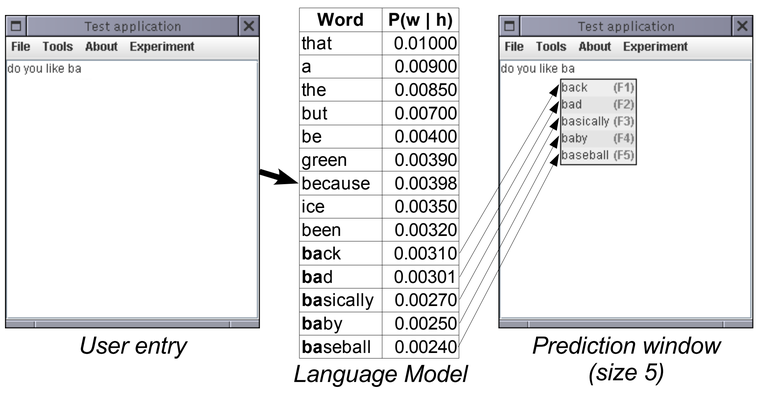
Word Prediction for AAC

Suppose a user is asking a friend “do you like baseball games?”



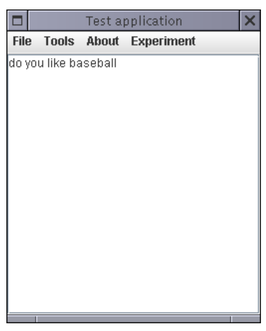
Word	P(w | h)
that	0.01000
a	0.00900
the	0.00850
but	0.00700
be	0.00400
green	0.00390
because	0.00398
ice	0.00350
been	0.00320
back	0.00310
bad	0.00301
basically	0.00270
baby	0.00250
baseball	0.00240

User presses 'a'



Word	P(w | h)
that	0.01000
a	0.00900
the	0.00850
but	0.00700
be	0.00400
green	0.00390
because	0.00398
ice	0.00350
been	0.00320
back	0.00310
bad	0.00301
basically	0.00270
baby	0.00250
baseball	0.00240

User presses F5



The user has finished typing “baseball ”

# of characters: 8 + 1 (space incl.)
# of keys pressed: 2 + 1

keystroke savings for “baseball ” (window size 5):

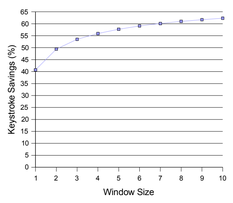
$$\frac{9-3}{9}\times 100\% = 66.7\% \text{ savings}$$

In general, the number of characters entered and keys pressed are accumulated over larger pieces of text:

$$KS = \frac{keys_{orig} - keys_{\text{with prediction}}}{keys_{orig}} \times 100\%$$

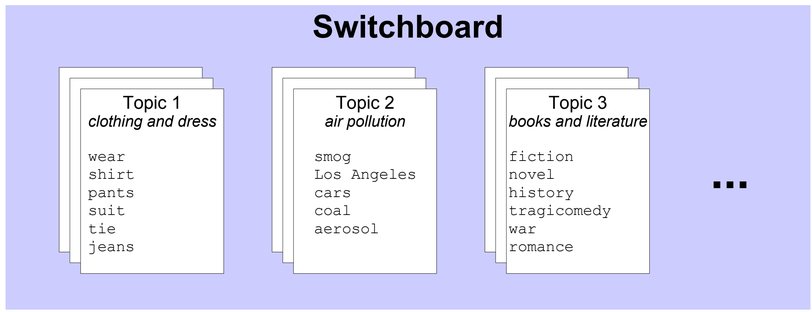
Trigram model: $P(w|h) = P(w|w_{-2}\ w_{-1})$

The keystroke savings of the trigram baseline is shown at several window sizes. Larger windows require fewer keystrokes.



Topic Modeling Approach

Topic Representation



Language models are constructed from conversations in each topic. Some example high-probability keywords are shown.

Topic Identification

Questions to consider for each word
- How often is it used?
- How recently was is it used?
- How good is it for topic identification?

Conversation 2001 – in progress

B: okay hi
A: hi yeah i'd like to talk about how you dress for work and what do you normally what type of outfit do you normally have to wear
B: well i work in corporate control so we have to dress kind of nice so i usually wear skirts and sweaters in the winter time slacks i guess and in the summer just dresses
A: um-hum

Important keywords underlined

Word	Weight
sweaters	2.59
slacks	2.34
skirts	2.33
dresses	2.30
dress	1.85
outfit	1.70
hi	1.69
wear	1.09
corporate	1.00
winter	0.54
normally	0.47
summer	0.22
control	0.17

Representation of the conversation

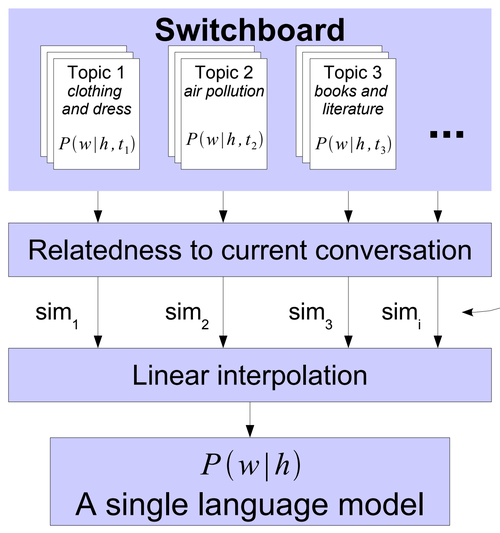
Word	Weight
wear	100
outfit	50
slacks	25
suit	23
sweaters	21
dresses	20
dress	18
shirt	17
tie	17

Topic unigram model

Cosine metric

Similarity score

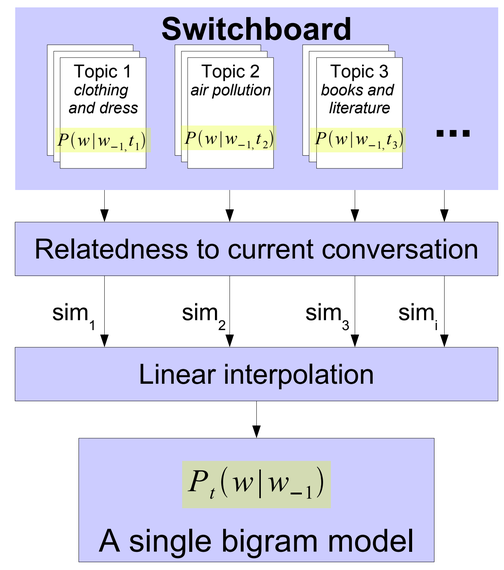
Topic Application



$$P(w|h) = \sum_{i \in topics} \frac{\text{sim}_i}{norm} \times P(w|h,t_i)$$

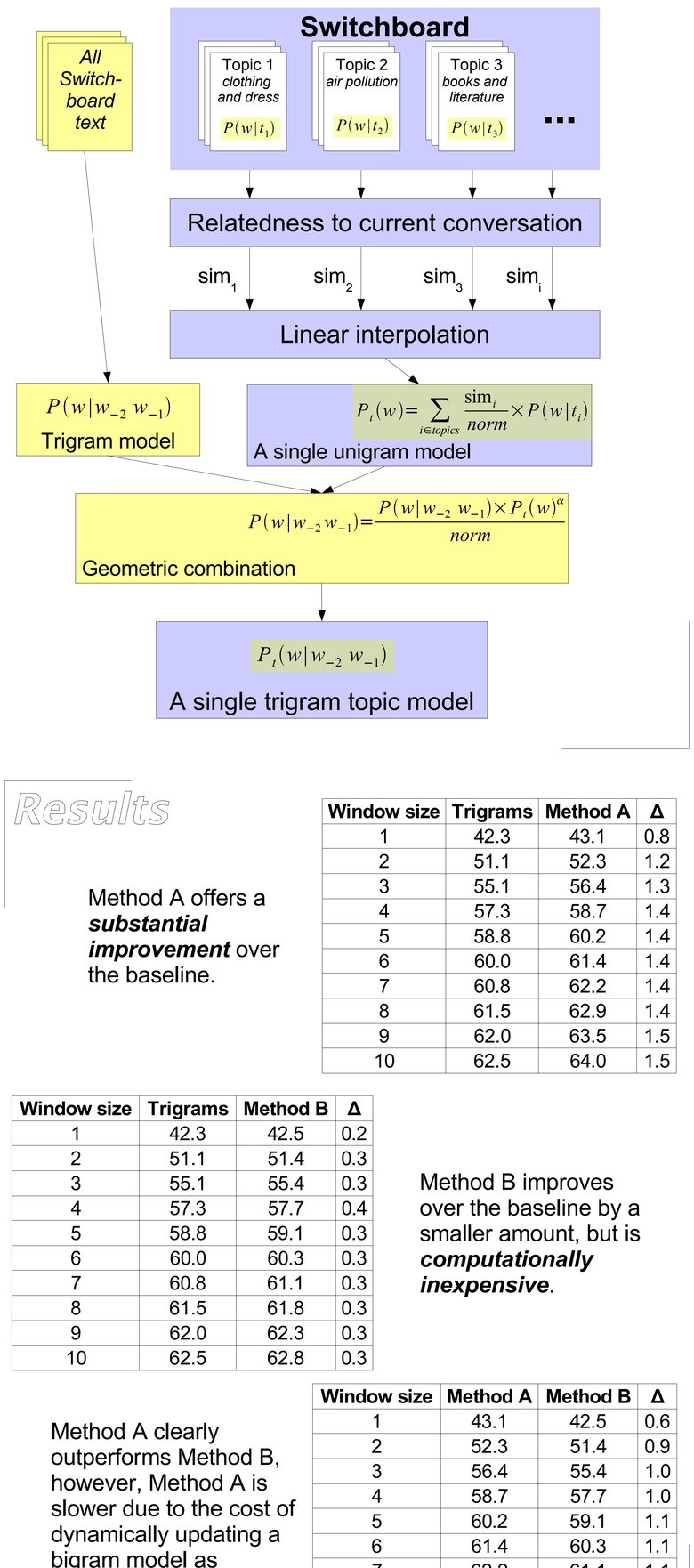
Topic Modeling in Detail

Method A



$$P(w|w_{-1}) = \sum_{i \in topics} \frac{\text{sim}_i}{norm} \times P(w|w_{-1},t_i)$$

Method B



Results

Method A offers a ***substantial improvement*** over the baseline.

Window size	Trigrams	Method A	Δ
1	42.3	43.1	0.8
2	51.1	52.3	1.2
3	55.1	56.4	1.3
4	57.3	58.7	1.4
5	58.8	60.2	1.4
6	60.0	61.4	1.4
7	60.8	62.2	1.4
8	61.5	62.9	1.4
9	62.0	63.5	1.5
10	62.5	64.0	1.5

Window size	Trigrams	Method B	Δ
1	42.3	42.5	0.2
2	51.1	51.4	0.3
3	55.1	55.4	0.3
4	57.3	57.7	0.4
5	58.8	59.1	0.3
6	60.0	60.3	0.3
7	60.8	61.1	0.3
8	61.5	61.8	0.3
9	62.0	62.3	0.3
10	62.5	62.8	0.3

Method B improves over the baseline by a smaller amount, but is ***computationally inexpensive***.

Method A clearly outperforms Method B, however, Method A is slower due to the cost of dynamically updating a bigram model as opposed to a unigram model.

Window size	Method A	Method B	Δ
1	43.1	42.5	0.6
2	52.3	51.4	0.9
3	56.4	55.4	1.0
4	58.7	57.7	1.0
5	60.2	59.1	1.1
6	61.4	60.3	1.1
7	62.2	61.1	1.1
8	62.9	61.8	1.1
9	63.5	62.3	1.2
10	64.0	62.8	1.2

Funded by US Department of Education grant H113G040051, under the National Institute on Disability and Rehabilitation Research program