

# Topic Modeling in Fringe Word Prediction for AAC

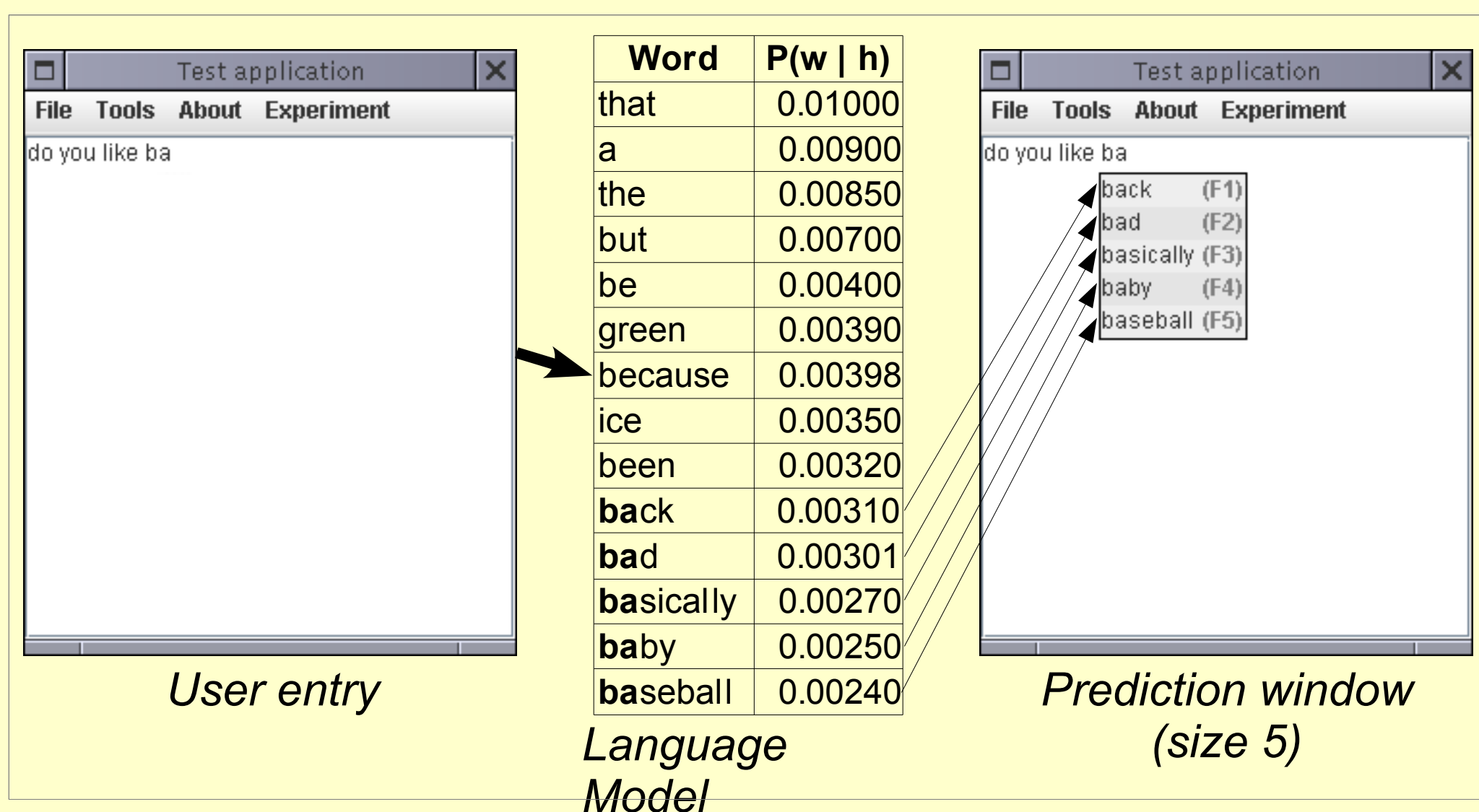
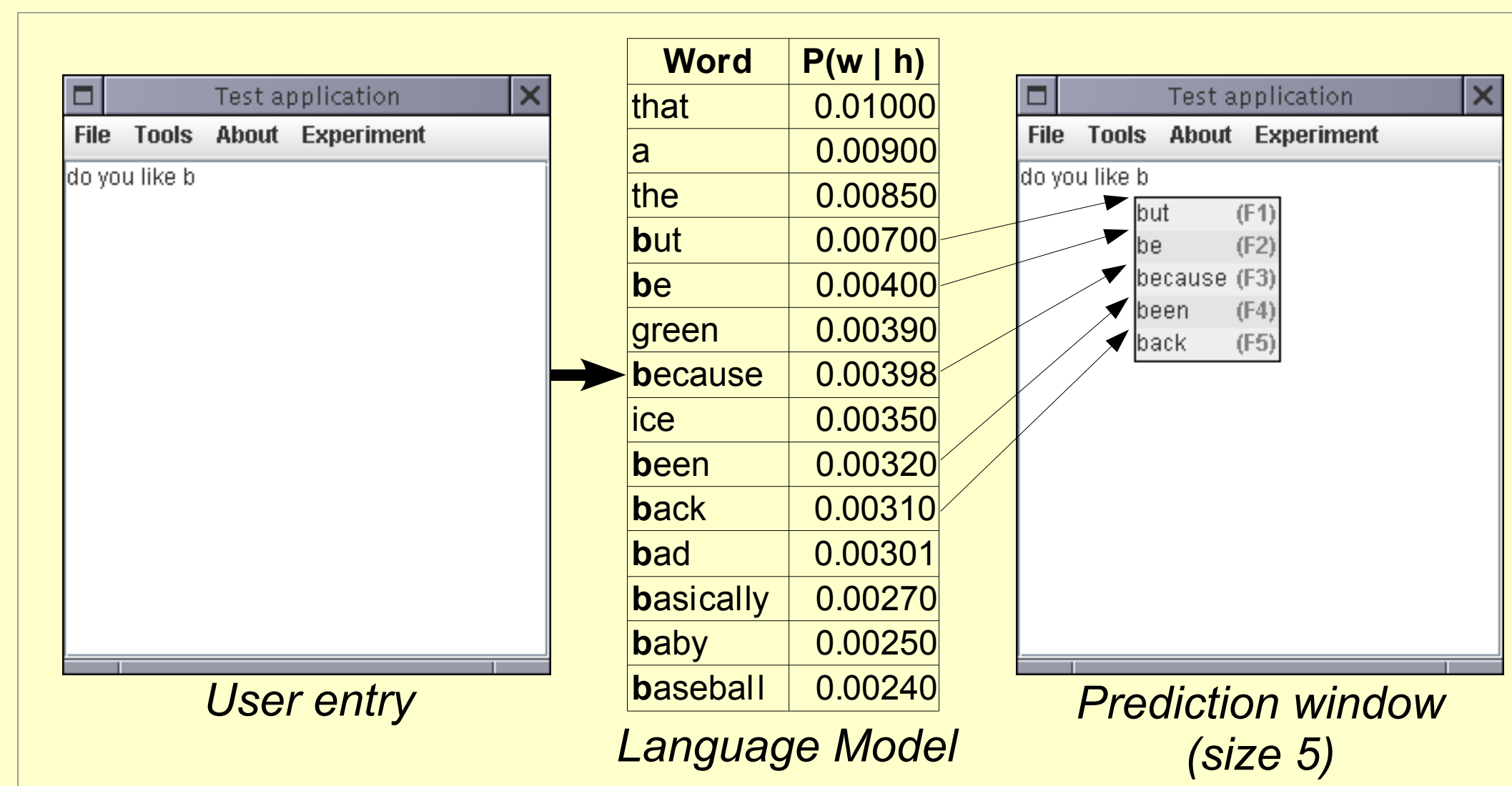
Keith Trnka, Debra Yarrington, Kathleen McCoy  
University of Delaware

Christopher Pennington  
AgoraNet, Inc.

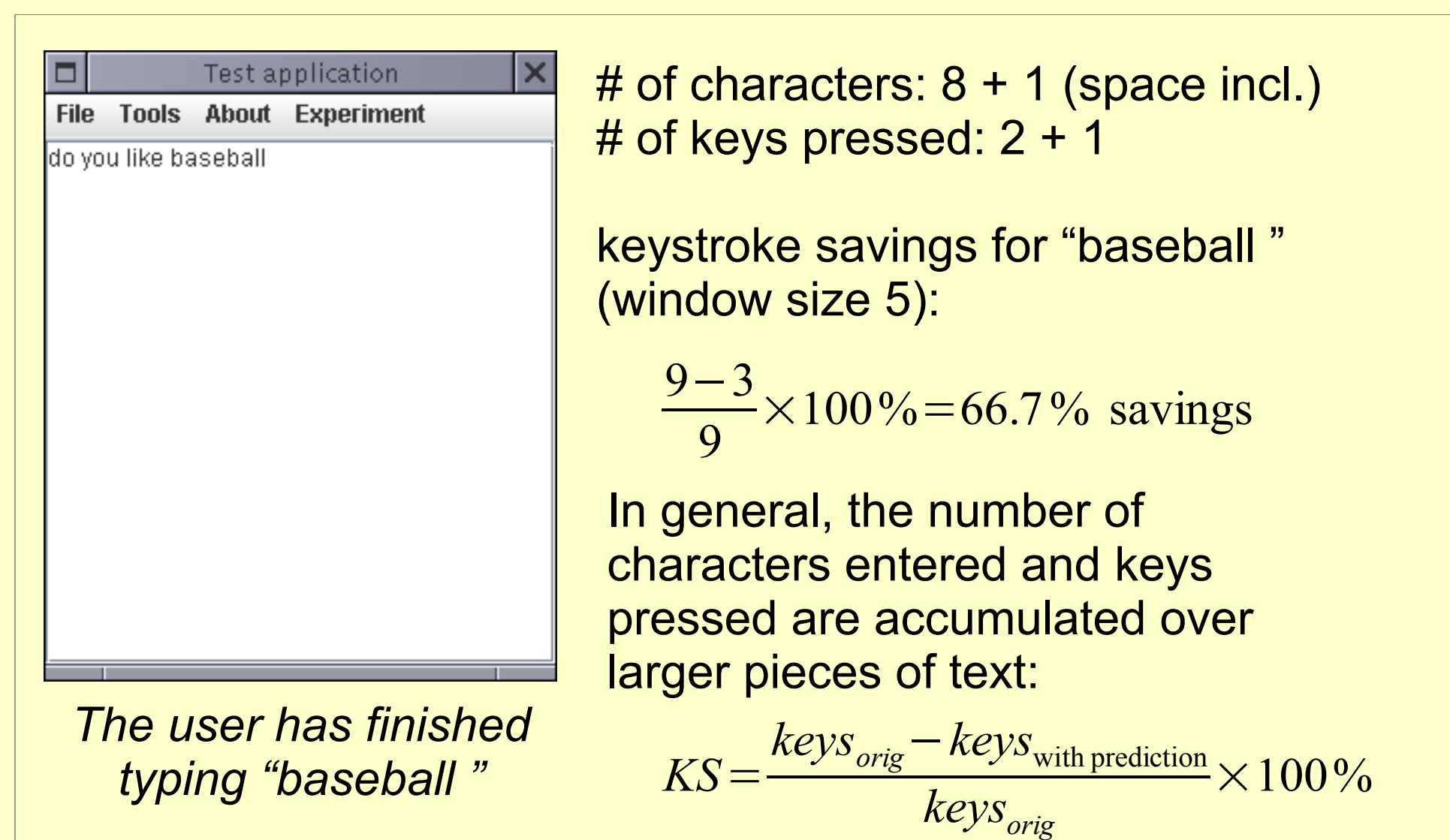
## Motivation

Many AAC devices try to optimize text entry by focusing on quick access to common words (*core vocabulary*). On the other hand, fringe words are often spelled out with the aid of word prediction. This work improves upon ngram-based fringe word prediction using topic modeling.

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

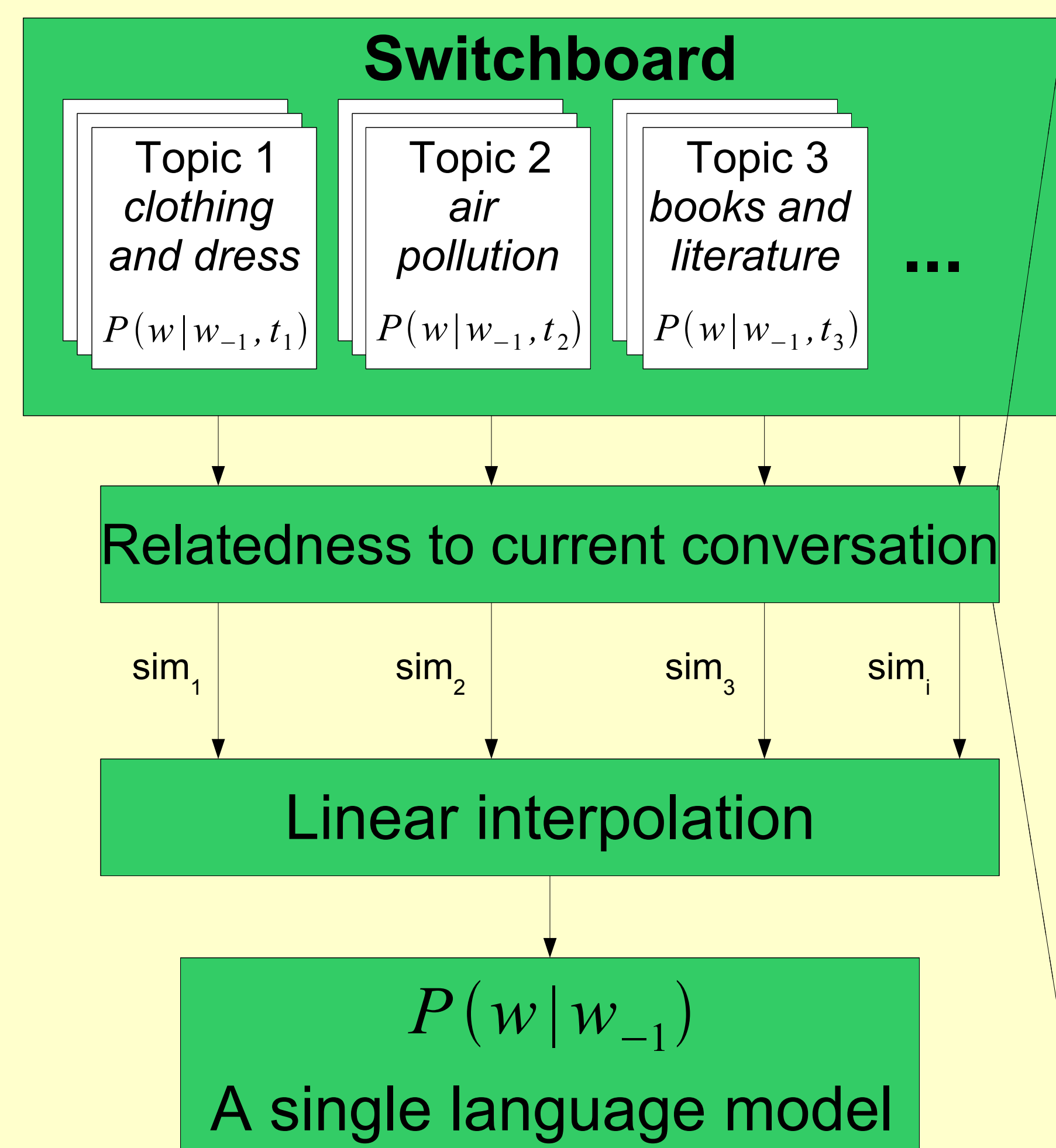


← User presses 'a'



## Methods

A language model is constructed from each topic in a corpus. Throughout a conversation, what has been said is compared to each topic to determine the contribution of each topic to the final language model. This language model is then used for word prediction.



Conversation 3682 (in progress)

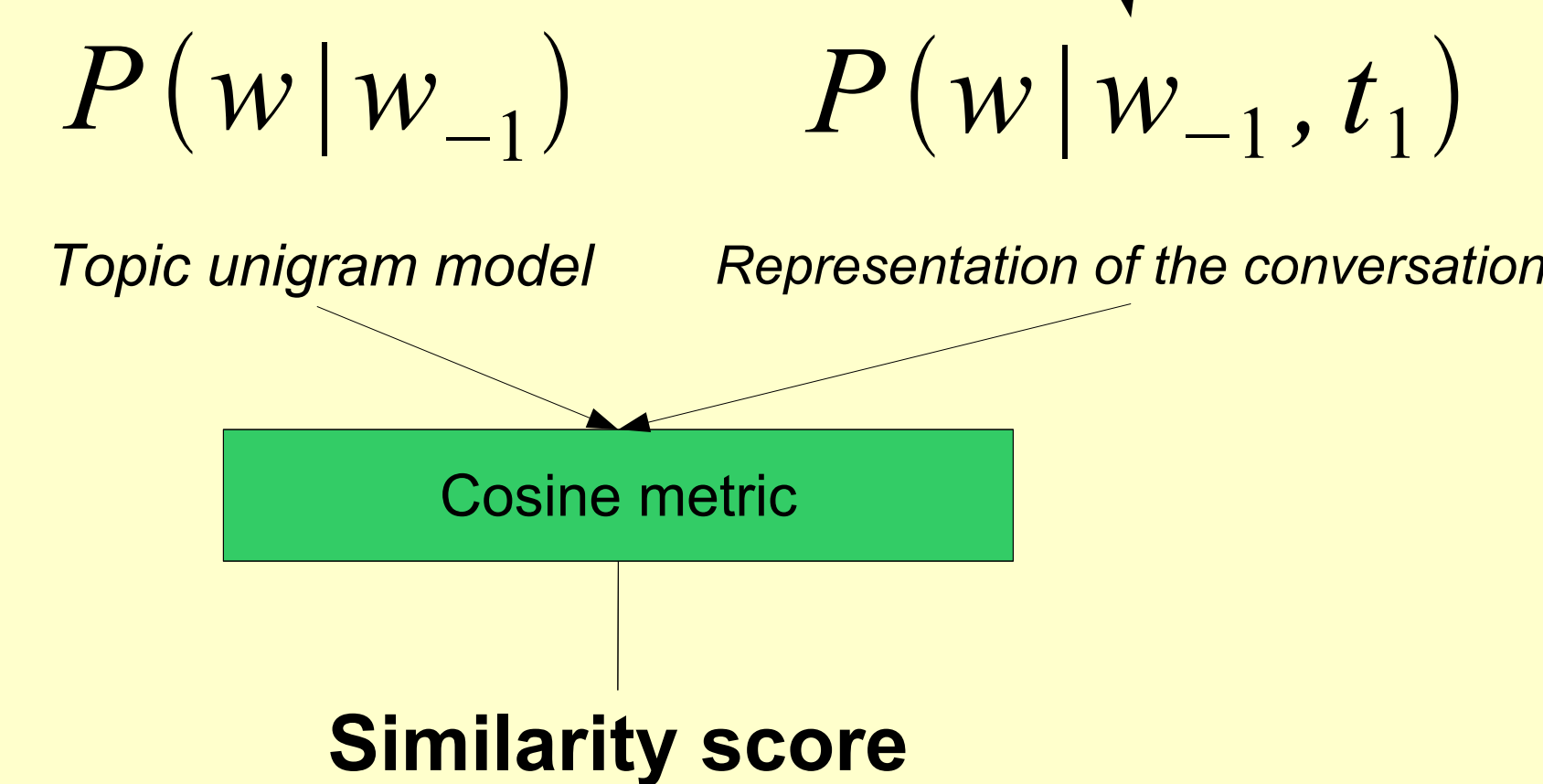
...  
**Shirley:** hm now that would be good how long do you have to bake it

**Jane:** yeah  
**Jane:** uh-huh let's see i think it was an hour i've done it in the microwave so often lately because it's really easy yeah

**Shirley:** i do microwave too a

**Jane:** but i think it's about an hour or maybe it's only half hour hm

Important keywords underlined



Example similarity scores for above conversation

cooking	restaurants	gardening	movies	taxes
23.4	4.42	1.43	0.47	0.33

## Results

Window size	Trigrams	Topic	Diff	t-score	p-value
1	42.3	43.1	0.8	9.45	<0.005
3	55.1	56.4	1.3	16.97	<0.005
5	58.8	60.2	1.4	21.29	<0.005
6	60.0	61.4	1.4	21.37	<0.005
7	60.8	62.2	1.4	22.84	<0.005

## Topic Modeling in Practice

### Example 1 - cooking

#### Conversation in progress

**Shirley:** hm now that would be good how long do you have to bake it

**Jane:** yeah  
**Jane:** uh-huh let's see i think it was an hour i've done it in the microwave so often lately because it's really easy yeah

**Shirley:** i do microwave too a

**Jane:** but i think it's about an hour or maybe it's only half hour hm

**Shirley:** how long would it be in the ...

#### Basic ngram word prediction

in the world winter city area United future mountains

in the mi military mid **microwave** midwest midst minority minimum

in the m mountains military mail mornings movie market mood

**Shirley says:** how long would it be in the microwave

#### Topic modeling

in the oven kitchen **microwave** mountains summertime world winter

**Shirley says:** how long would it be in the microwave

### Example 2 – public education

#### Conversation in progress

**Sally:** partitions or just to face the children different directions or to do something

**Linda:** just yeah

**Sally:** i know last year my son had to eat in the classroom because there was no cafeteria at his school so he spent all day in this one large open area

**Linda:** oh goodness

**Sally:** so i didn't really care for that it seems like the ...

#### Basic ngram word prediction

seems like the i kids news Cowboys older guy Dallas

seems like the sc **schools** science scene scout score scribes schooling

seems like the s second sport stuff system state spice sports

**Sally:** so i didn't really care for that it seems like the schools

#### Topic modeling

seems like the kids **schools** computer nation administration police person

**Sally:** so i didn't really care for that it seems like the schools

### Example 3 - recycling

#### Conversation in progress

**Sherry:** we're setting up curbside recycling bins for the city home owners and several of our local businesses also have recycling stations set up in their parking lots

**Kitty:** um-hum

**Sherry:** and how about in your area

**Kitty:** well in our i live in a very small community and we're just getting started as a matter of fact this last week we got a notice in the mail from City Hall and they've informed us that we'd be getting more instructions in the mail in a week or two but we're going to start a separation of our refuse and separating the ...

#### Basic ngram word prediction

separating the kids things problem news world government ones

separating the ca cat cars **cans** cats cable carpet cash

separating the c city computer crime cost credit cat community

**Kitty:** ... we're going to start a separation of our refuse and separating the cans

#### Topic modeling

separating the city recycling plastic trash garbage **cans** newspaper

**Kitty:** ... we're going to start a separation of our refuse and separating the cans