Exploring the Contribution of Conversational Context in Word Prediction

Haesik Min[1], Ph. D., Jeff Higginbotham[1], Ph. D., Greg Lesher[2], Ph. D., T. Fen Yik[1], M. A.

[1]State University of New York at Buffalo

[2]Dynavox Technologies

Background

Dynamic Word Prediction

- Many contemporary AAC systems tabulate word and inter-word frequencies to perform statistical n-gram word prediction.
- Recent research and development has explored sophisticated word prediction schemes to improve performance by including contextual factors such as topic priming, providing access to fringe vocabulary, utilizing web-based information, and exploiting geographic and local linguistic context.[2, 4, 5, 7]

Local Linguistic Context and Word Prediction

- N-gram based prediction databases in AAC systems provide a degree of local linguistic context sensitivity. When an intended word is infrequently used, it is less likely to appear on the prediction list despite its contextual relevance.
- **Speaker contribution**: word recency is available in many AAC applications to promote context-relevant words with low frequency such that they appear in the word prediction list.[5, 7]
  - With word recency, word prediction databases utilize messages formulated by the device users to selectively bias recently used words to show up in the word prediction list.
- **Partner contribution**: Partner talk can also be used to bias a word predictor, with the potential of improving keystroke savings.[1]
  - In the future, commercial speech recognition programs (e.g., Dragon Dictation, Google Voice, GoVivace and Siri Personal Assistant) and/or operator mediated talk with text services (e.g., CapTel[3]) may provide an opportunity for partner's previous utterances to be used by the word predictor.
- Previous research demonstrates that partner recency can provide additional word prediction improvement.[5, 7, 8]

Word Characteristics, Local Linguistic Context and Word Prediction

- **Open class words:** nouns, verbs, adjectives and adverbs
- **Closed class words:** articles, pronouns, conjunctions and prepositions
- **Word frequency**: Closed class words occur frequently despite having a small inventory. The repertoire of open class words is immense and used less repeatedly. Therefore, closed class words are more likely to appear on the prediction list than open class words.
- **Word length**: Short words are harder to predict than long words.
- **Word-recency** only biases words that are used verbatim in the previous utterances. Thus, open class words are at a disadvantage with respect to closed class words in benefiting from word recency. Word recency is most useful for placing less frequently used words (i.e., open class) into the prediction list,[5] whereas closed class words (usually short and frequently used) may not benefit from word recency.

Research Objectives

- The primary goal of the study was to investigate whether the contribution of local linguistic context (i.e., speaker recency and partner recency) provides additional improvement over simple word prediction performance (i.e. no recency involved) .
- Secondary goals were: (a) to investigate whether open and closed class words behaved differently in word prediction and (b) to examine how word length and word frequency of open and closed class words influenced word prediction.

Selected References

1. Arnott, J., Hannan, J., & Woodburn, R. (1993, May). *Linguistic prediction for disabled users of computer-mediated communication*. Paper presented at the Second European Conference on the Advancement of Rehabilitation Technology (ECART), Stockholm, Sweden.
2. Higginbotham, D. J., Bisantz, A., Sunm, M., Adams, K., & Yik, F. (2009). The effect of context priming and task type on augmentative communication performance. *AAC: Augmentative & Alternative Communication*, 25(1), 19-31.
3. Higginbotham, D. J., & Engelke, C. R. (2012, January). *InTra: Technologies for using partner speak in AAC*. Paper presented at the Assistive Technology Industry Association Conference, Orlando, FL.
4. Higginbotham, D. J., Lesher, G., & Luo, F. (2008, August). *Using a web crawler to enhance AAC*. Paper presented at the International Society for Augmentative and Alternative Communication, Montreal, Canada.
5. Lesher, G., & Rinkus, G. (2001, June). Domain-specific word prediction for augmentative communications. Paper presented at the RESNA 2001 Annual Conference, Reno, NV.
6. Patel, R., & Radharkrishnan, R. (2007). Enhancing access to situational vocabulary by leveraging geographic context. *Assistive Technology Outcomes and Benefits*, 4(1), 99-114.
7. Trnka, K., Yarrington, D., McCaw, J., McCoy, K., & Pennington, C. (2007, April). *The effects of word prediction on communication rate for AAC*. Paper presented at the Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers, Rochester, NY.
8. Wisenburn, B., & Higginbotham, D. J. (2008). An AAC application using speaking partner speech recognition to automatically produce contextually relevant utterances: Objective results. *AAC: Augmentative and Alternative Communication*, 24(2), 100-109.

Acknowledgments

UB's Signature Center for Excellence in Augmented Communication

The Rehabilitation Engineering Research Consortium on Communication Enhancement (AAC-RERC) is funded under grant #H133E080011 from the National Institute on Disability and Rehabilitation Research (NIDRR) in the U.S. Department of Education's Office of Special Education and Rehabilitative Services (OSERS). The contents do not necessarily represent the policy of NIDRR.

Methodology

Source Materials

Texts of two-party conversations were taken from the following corpora:

- The Santa Barbara Corpus (SBC) of Spoken American English: unscripted face-to-face conversations
- The Call Home American English Corpus (Call Home): unscripted phone conversations
- HCRC Map Task Corpus (HCRC): unscripted conversations while performing the map-completing task
- ALS narratives (ALS): face-to-face conversations between AAC users diagnosed with ALS and their familiar partners

	# of conversations	# of utterances	# of words
SBC	16	9,439	58,764
Call Home	120	25,374	205,149
HCRC	124	24,285	144,241
ALS	6	435	2,266

Data Conversion

Raw data

10.27 10.85	PAMELA:	Well you're right,
10.85 12.97		I think they're probably flip sides.
12.97 13.84	DARRYL:	(TSK) I mean who [are you].
13.53 13.87	PAMELA:	[But I'm] --
13.87 14.56	DARRYL:	Pollyanna?
14.56 16.95	PAMELA:	... (SWALLOW) (TSK) Ha=rdly=.
16.95 17.58	DARRYL:	@@@
17.58 18.56	PAMELA:	Hardly=,

Filtered & Tagged: utterances

$$10.27-10.85$$ @@ON-sbc0005-PAMELA@@ Well you're right, I think they're probably flip sides. @@DUMP@@
$$12.97-13.84$$ @@ON-sbc0005-DARRYL@@ I mean who are you. @@DUMP@@
$$13.53-13.87$$ @@ON-sbc0005-PAMELA@@ But I'm - @@DUMP@@
$$13.87-14.56$$ @@ON-sbc0005-DARRYL@@ Pollyanna? @@DUMP@@
$$14.56-16.95$$ @@ON-sbc0005-PAMELA@@ Hardly. @@DUMP@@

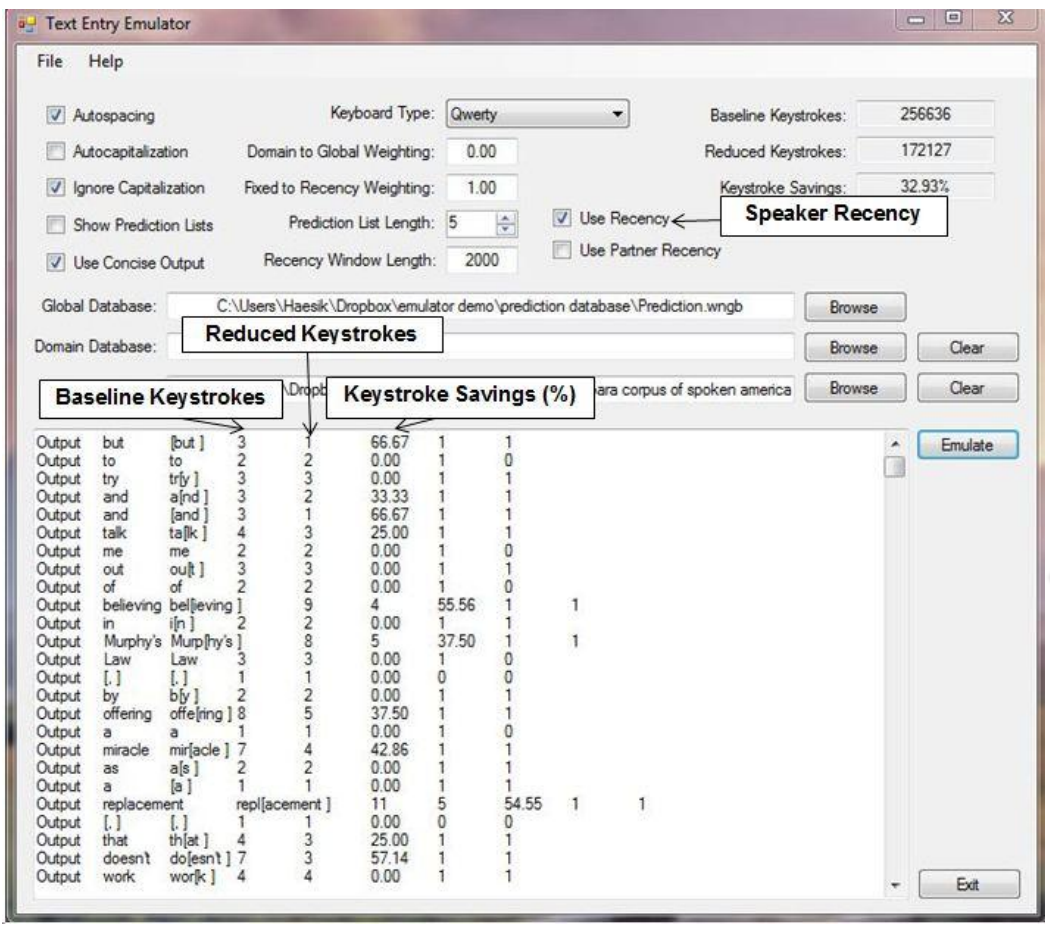
Various marking, typos and verbal fillers were cleaned up by using Perl.

Filtered & Tagged: words

$$10.27-10.85$$ @@ON-sbc0005-PAMELA@@ Well @@DUMP@@ you're @@DUMP@@ right @@DUMP@@ , @@DUMP@@ I @@DUMP@@ think @@DUMP@@ they're @@DUMP@@ probably @@DUMP@@ flip @@DUMP@@ sides @@DUMP@@ . @@DUMP@@
$$12.97-13.84$$ @@ON-sbc0005-DARRYL@@ I @@DUMP@@ mean @@DUMP@@ who @@DUMP@@ are @@DUMP@@ you @@DUMP@@ . @@DUMP@@
$$13.53-13.87$$ @@ON-sbc0005-PAMELA@@ But @@DUMP@@ I'm @@DUMP@@ - @@DUMP@@
$$13.87-14.56$$ @@ON-sbc0005-DARRYL@@ Pollyanna @@DUMP@@ ? @@DUMP@@
$$14.56-16.95$$ @@ON-sbc0005-PAMELA@@ Hardly @@DUMP@@ . @@DUMP@@

Another data conversion took place to examine the prediction performance of individual words.

Words emulated by Text Entry Emulator



The Text Entry Emulator is a testing interface that can simulate a person formulating messages by selecting letters/characters and calculate baseline keystrokes, reduced keystrokes, and keystroke savings (%).

Processed data

Word	Length	Frequency	NO	S	S&P
I	1	2325	0.0	0.0	0.0
a	1	1133	0.0	0.0	0.0
to	2	1233	0.0	0.9	0.5
it	2	1044	0.0	23.3	25.4
of	2	750	0.0	0.1	0.1
so	2	646	0.0	1.2	0.0
in	2	607	50.0	42.1	42.4
is	2	518	0.0	0.0	0.0
we	2	494	0.0	1.4	1.2
he	2	413	0.0	7.6	7.8
do	2	389	0.0	0.0	0.0
oh	2	383	0.0	1.0	0.0

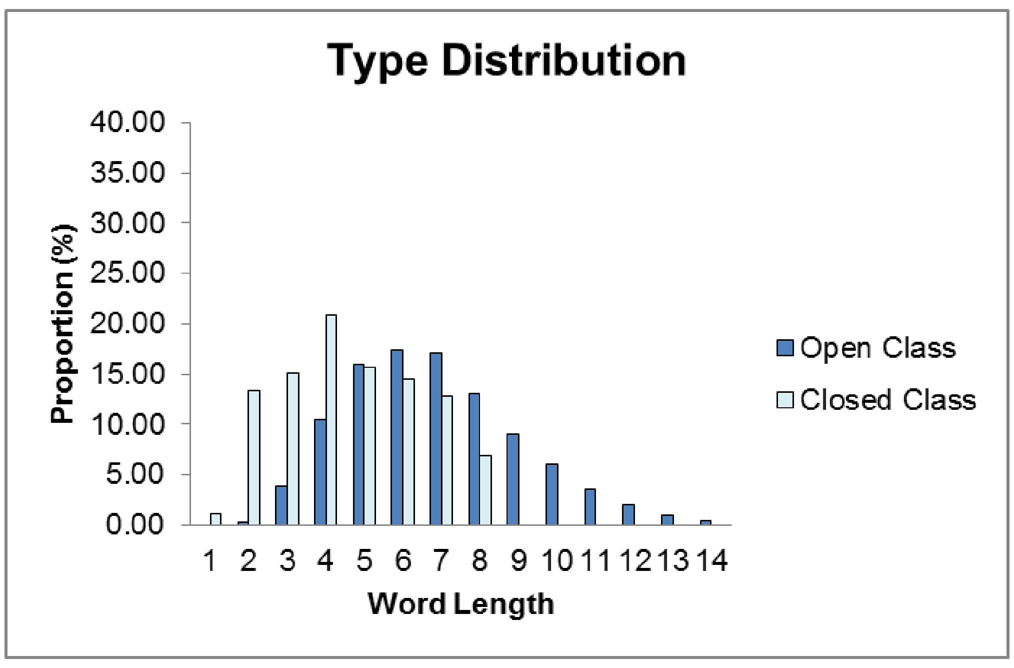
The outputs of the Text Entry Emulator were compiled by words used in the corpora with their average keystroke savings ACROSS three recency (no, speaker, speaker/partner) conditions and word characteristics (length & frequency).

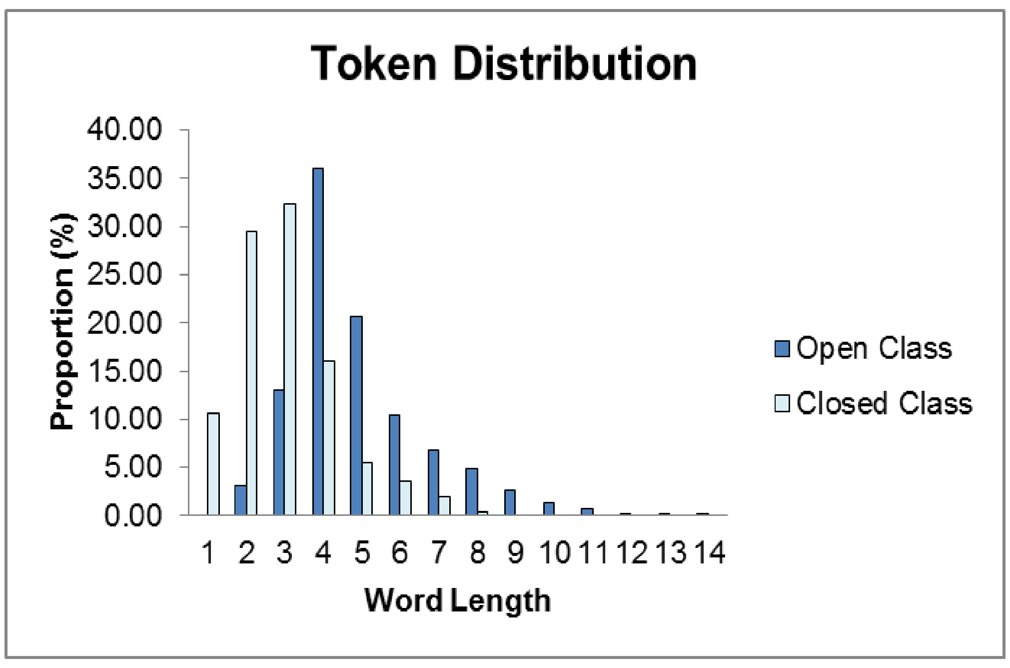
The number of words compiled from the texts of two-party conversations was 10,367. They were coded as open (N=10,194) or closed (N=173) classes and analyzed.

Research Design

- **Hierarchical Linear Modeling (HLM) regression analyses** was used to access the overall contribution of local linguistic context across recency techniques (no recency (NO), speaker recency (S), and speaker/partner recency (S/P)).
  - Dependent variable: keystroke savings
  - Independent variable: Level 1 – recency, Level 2 – word length & frequency
- **Repeated measures analysis of covariance** was used to assess the contribution S and S/P recency at each word length for open class words:
  - Dependent variable: keystroke savings, Within-group factor: recency, Covariate: word frequency
- **Friedman one-way analysis of variance** was used to measure the contribution of S and S/P recency at each word length for closed class words.
  - Dependent variable: keystroke savings, Independent variable: recency
- **Analysis of covariance** was used to explore the performance differences between open and closed classes at each word length.
  - Dependent variable: keystroke savings, Independent variable: syntactic class (open vs. closed), Covariate: word frequency
- Statistical significance was set at .05 of alpha. A family-wise alpha was adopted by making the Bonferroni adjustment for follow-up contrasts.

Results & Discussion





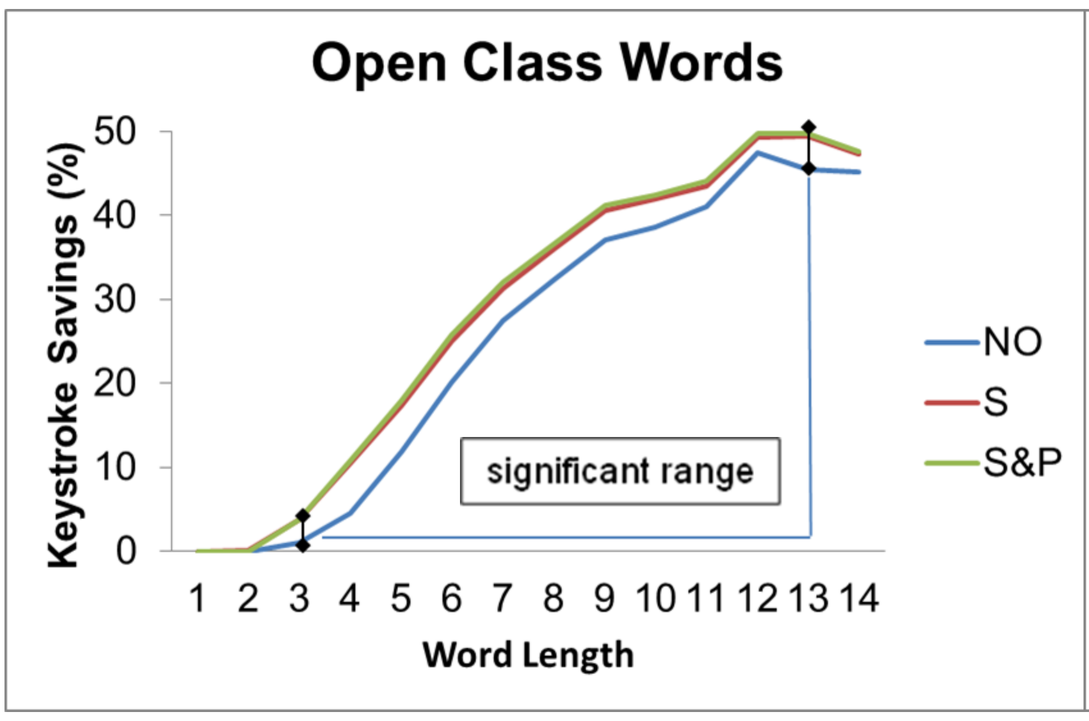
Descriptive results: Mean (SD)

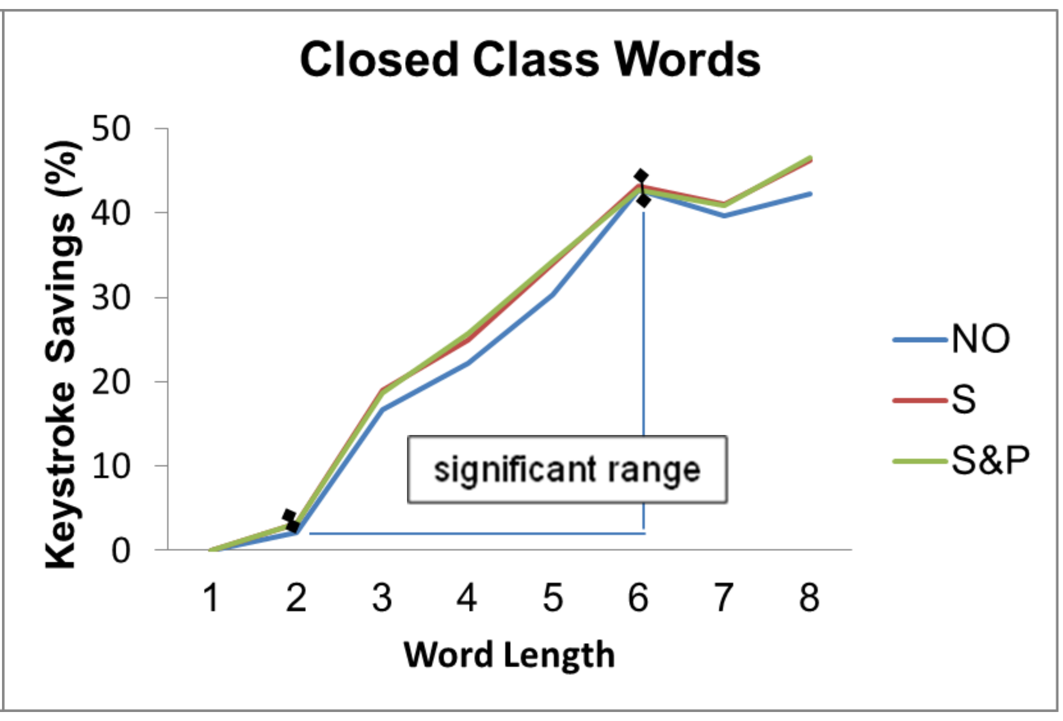
	Open Class (*N* = 10,194)	Closed Class (*N* = 173)
Word Length	6.9 (2.2)	4.7 (1.8)
Word Frequency (token count)	19.8 (169.2)	1206.5 (2,659.3)
Keystroke Savings (%): NO recency	23.5 (19.9)	26.3 (22.6)
Keystroke Savings (%): S recency	27.8 (19.9)	28.5 (21.2)
Keystroke Savings (%): S/P recency	28.4 (20.0)	28.5 (21.3)

The overall contribution of local linguistic context with recency techniques (p<.001)

- Open class: S/P > S > NO
- Closed class: (S/P = S) > NO
- Increases in word length > increases in word frequency : improving word prediction performance

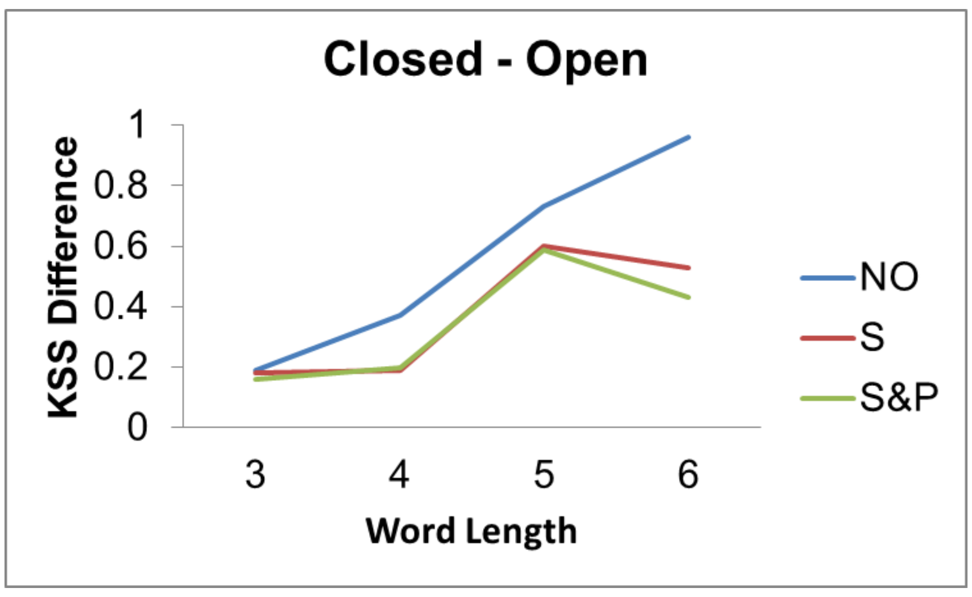
The contribution of local linguistic context with recency techniques at each word length





- Open class: omnibus test showed statistical significance in 3- to 13-letter words (p<.05).
  - S > NO (3- to 13-letter words)
  - S/P > S (4- to 11-letter words)
- Closed class: omnibus test showed statistical significance in 2- to 6-letter words (p<.05).
  - S > NO (2-letter words)
  - S/P = S (2-to 6-letter words)

Performance difference between closed and open class words at each word length



- Closed class > Open class: in all recency conditions (p<.05)
- S recency and S/P recency facilitated closing the keystroke savings gap between the two classes.
- The efficacy for open class words increased starting at 4-letter words and becoming more noticeable at 6-letter words.

Summary & Future Research

- Speaker recency contributes to the predictability of both open and closed class words, bolstering its utility for dynamic word prediction in current AAC systems.
  - 4- to 8-letter open class words benefited most from speaker recency.
  - 2- to 7-letter closed class words benefited most from speaker recency
- Statistically significant, but small effect sizes were found for Speaker/Partner versus Speaker recency. Thus, for current n-gram based prediction models, Speaker/Partner recency may not noticeably improve word performance.
- However:
  - N-gram based word prediction models latent sematic analysis or grammar tagging may offer additional way of improving Speaker/Partner influenced word prediction.
  - The impact of Speaker/Partner recency on word prediction may differ across different discourse genres (e.g., casual conversation, giving news, telling a story).
  - The impact of Speaker/Partner recency may be bolstered in real-time interactions using a speech recognition technology. Interlocutors may adapt their talk to optimize the effect of this recency technique.

Contact information: Haesik Min (hsmin@buffalo.edu)
Center for Excellence in Augmented Communication
Department of Communicative Disorders and Sciences
State University of New York at Buffalo