Exploring the text-level 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 speaker 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 some word prediction improvement.[5, 7, 8]

Domain-specific Word Prediction

- **Topic Priming:** the use of domain-specific predictions through the inclusion of contextual factors.
- **Contextual factors:** facilitated the topic priming for word predictors and lead to improved prediction efficiency.[2, 5, 6]
  - The knowledge of discourse genre: written texts vs. conversations, face-to-face conversation vs. phone conversation
  - Conversational topics: general and frequently shifting vs. narrow and focused
- **Utilizing web-based information**: searching and retrieving relevant language materials when a new topic is introduced (e.g., Web Crawler[4]).
- **Geographic Location**: providing information regarding context specific activities, facilitating topic priming and offering location specific vocabulary and expressions.[6]

Research Objectives

- The primary goal of the study was to investigate whether the **text-level** contribution of local linguistic context (i.e., speaker recency and partner recency) provided additional improvement over simple word prediction performance (i.e. no recency involved) .
- Secondary goals were: (a) to investigate whether discourse genre differentially affected word prediction and (b) to examine how the amount of talk per speaker influenced word prediction performance.

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

AAC-RERC SPREAD THE WORD

University at Buffalo The State University of New York

UB's Signature Center for Excellence in Augmented Communication

DynaVox Mayer-Johnson Advancing human expression and learning.

NIDRR NATIONAL INSTITUTE ON DISABILITY AND REHABILITATION RESEARCH

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

Data Conversion

Raw data

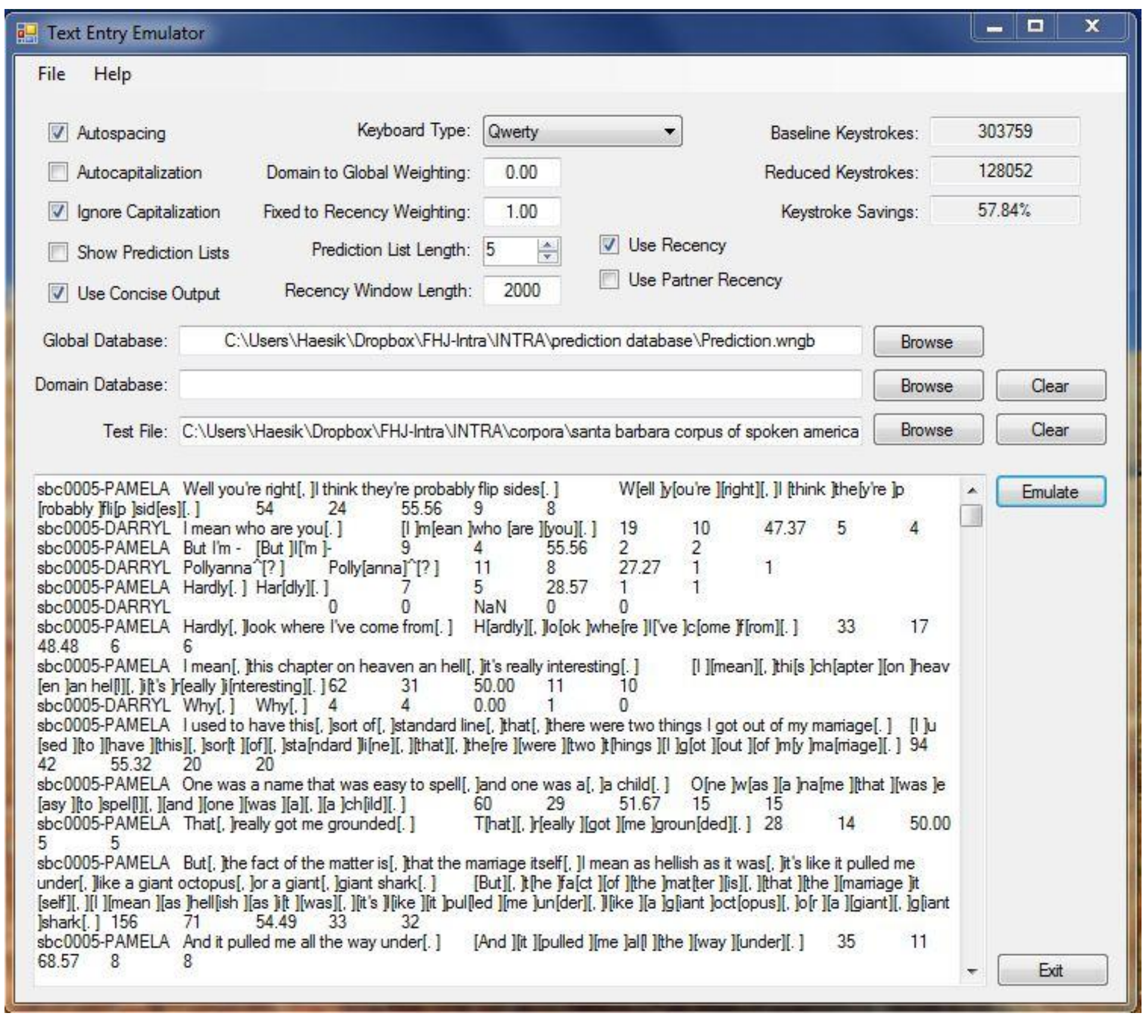
10.27	10.85	PAMELA:	Well you're right,
10.85	12.97		I think they're probably flip sides.
12.97	13.84	DARRYL:	(TSK) I mean who [are you].
13.53	13.87	PAMELA:	[But I'm] --
13.87	14.56	DARRYL:	Pollyanna?
14.56	16.95	PAMELA:	... (SWALLOW) (TSK) Ha=rdly=.
16.95	17.58	DARRYL:	@@@
17.58	18.56	PAMELA:	Hardly=,

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.

Texts 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

Corpora	Transcript ID	Speaker ID	Utterance	Baseline Keystrokes	Reduced Keystrokes	Keystroke Savings (%)	# of words
SBC	05	PAMELA	Well you're right[,]I think they're probably flip sides[.]	54	24	55.6	9
SBC	05	DARRYL	I mean who are you[.]	19	10	47.4	5
SBC	05	PAMELA	But I'm -	9	4	55.6	2
SBC	05	DARRYL	Pollyanna^[?]	11	8	27.3	1
SBC	05	PAMELA	Hardly[.]	7	5	28.6	1

The outputs of the Text Entry Emulator were processed by utterances of each speaker with keystroke savings in a given recency condition.

Compiled data

Corpora	ID	# of words	NO	S	S/P
SBC	05	435	43.4	49.5	50.3
SBC	06	847	43.0	50.0	50.3
SBC	07	412	39.2	45.1	46.1
SBC	09	436	42.5	48.4	48.8

Corpora	ID	Speaker ID	# of words	NO	S	S/P
SBC	05	DARRYL	183	41.8	46.6	47.4
SBC	05	PAMELA	252	44.6	51.5	52.4
SBC	06	ALINA	609	45.9	53.2	53.4
SBC	06	LENORE	238	31.7	37.9	38.5

Compiled by transcript ID (left) and by speaker ID (right) with average keystrokes savings ACROSS three recency (no, speaker, speaker/partner) conditions and the number of words

Research Design

- **Repeated measures analysis of covariance** was used to assess the contribution S and S/P recency within each genre (Call Home, HCRC and SBC):
  - Dependent variable: keystroke savings, Within-group factor: recency, Covariate: total number of words
- **Friedman one-way analysis of variance** was used to measure the contribution of S and S/P recency with ALS narratives:
  - Dependent variable: keystroke savings, Independent variable: recency
- **Analysis of covariance** was used to explore the genre effect among the three recency conditions:
  - Dependent variable: keystroke savings, Independent variable: genre (Call Home, HCRC and SBC), Covariate: total number of words
  - ALS narratives were excluded from the analysis due to small sample size.
- The identical analyses (repeated ANCOVA, Friedman one-way ANOVA and ANCOVA) were performed to examine how the amount of talk per speaker influenced word prediction performance.
  - Speakers were coded as either MORE or LESS based on the proportion of conversation.
- Statistical significance was set at .05 of alpha. A family-wise alpha was adopted by making the Bonferroni adjustment for follow-up contrasts.

Results

Average number of words and average keystroke savings

	Words	NO	S	S/P
ALS (N=6)	377.7	43.0	47.3	48.4
Call Home (N=120)	1,709.6	45.0	53.3	53.3
HCRC (N=124)	1126.9	42.7	53.2	54.0
SBC (N=16)	3,672.8	40.9	47.9	48.3

Overall Recency Effect (*p*<.025)

Call Home	(S/P = S) > NO
HCRC	S/P > S > NO
SBC	(S/P = S)* > NO

*It approached significance (p=.032).

Overall Genre Effect (*p*<.025)

NO	Call Home > HCRC > SBC
S	(Call Home = HCRC) > SBC
S/P	HCRC > Call Home > SBC

Average number of words and average keystroke savings – Speaker Type

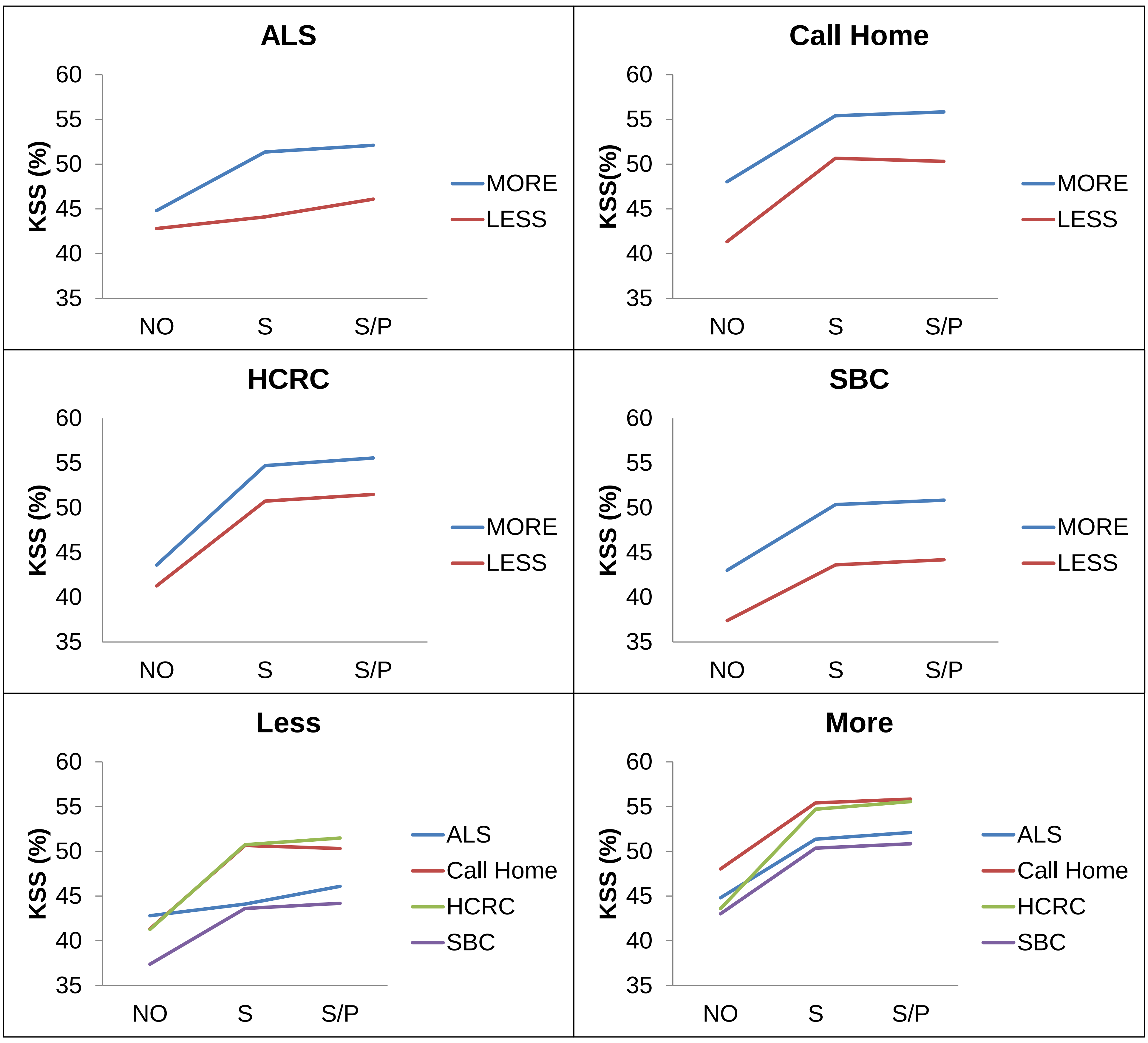
	LESS				MORE			
	Words	NO	S	S/P	Words	NO	S	S/P
ALS	82.5	42.8	44.1	46.1	295.2	44.8	51.4	52.1
Call Home	593.5	41.3	50.7	50.3	1,116.1	48.0	55.4	55.9
HCRC	386.5	41.3	50.8	51.5	776.7	43.6	54.7	55.6
SBC	1,255.9	37.4	43.6	44.2	2,416.9	43.0	50.4	50.8

Recency Effect – Speaker Type (*p*<.025)

	LESS	MORE
Call Home	(S/P = S) > NO	S/P > S > NO
HCRC	S/P > S > NO	S/P > S > NO
SBC	(S/P = S) > NO	S/P > S > NO

Genre Effect – Speaker Type (*p*<.025)

	LESS	MORE
NO	(Call Home = HCRC) > SBC	Call Home > (HCRC = SBC)
S	(Call Home = HCRC) > SBC	Call Home > HCRC > SBC
S&P	HCRC > Call Home > SBC	(Call Home = HCRC) > SBC



Summary & Future Research

- Speaker recency improved word prediction performances across all discourse genres.
- Speaker/Partner recency provided additional predictability for those with narrow and predefined topics (e.g., task-oriented conversations) and possibly face-to-face conversations.
- Conversations with narrow and predetermined topics benefited the most with the use of recency techniques.
  - Maintaining a topic on a narrow scope may potentially improve word prediction of AAC devices.
- Lack of visual access did not negatively affect word prediction performance when the scope of the topic was narrow.
  - Using word prediction while communicating over the telephone with AAC devices could be beneficial if maintaining a topic on a narrow scope can be attained.
- Speaker type in terms of amount of talk (more/less) positively influenced word prediction performance: more > less.
  - Speaker type poses a challenge for AAC users to take advantage of speaker recency.

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