Pre-Presentation Notes

Slides and presentation materials are available online at:

karlwiegand.com/defense

Disambiguation of Imprecise User Input Through Intelligent Assistive Communication

Karl Wiegand
Northeastern University
Boston, MA USA
December 2014

Thesis Statement

"Intelligent interfaces can mitigate the need for linguistically and motorically precise user input to enhance the ease and efficiency of assistive communication."

Theoretical Contributions

"...mitigate the need for linguistically and motorically precise user input..."

1. An unordered language model that bridges syntax and semantics. [Wiegand and Patel, 2012A]

2. An empirical comparison of contextual language predictors. [Wiegand and Patel, 2015B (R1)]

3. A motor movement study with current and potential AAC users. [Wiegand and Patel, 2015A]

Applied Contributions

"...to enhance the ease and efficiency of assistive communication."

1. A semantic approach to icon-based, switch AAC. [Wiegand and Patel, 2014B]

2. A continuous motion overlay module for icon-based AAC. [Wiegand and Patel, 2012B]

3. Mobile, letter-based AAC that supports conversational speeds. [Wiegand and Patel, 2014A]

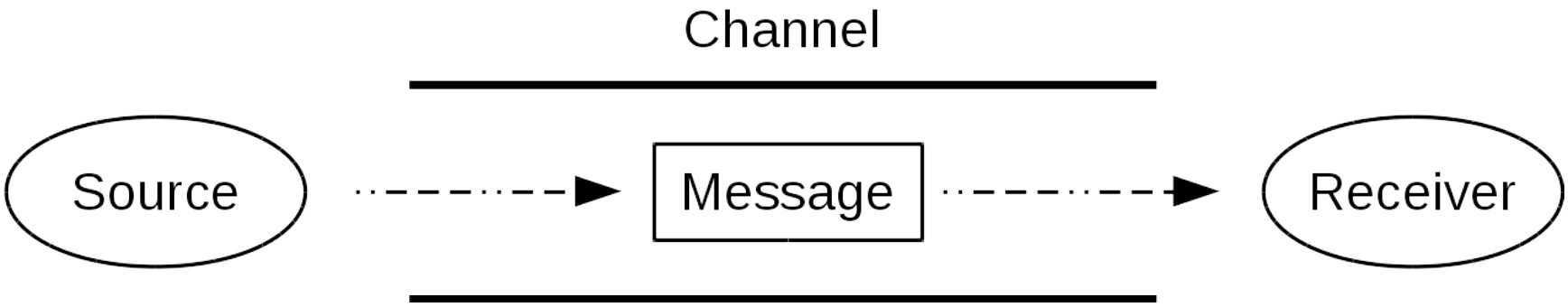
Outline

1. Assistive Communication
2. Theoretical Contributions
3. Applied Contributions
4. Summary and Conclusion

Part 1:

Assistive Communication

On Communication



- SMCR and derivatives [Shannon and Weaver, 1949]
- Affected by distortion to any component
- What if there is distortion from the Source?

Who Uses AAC?

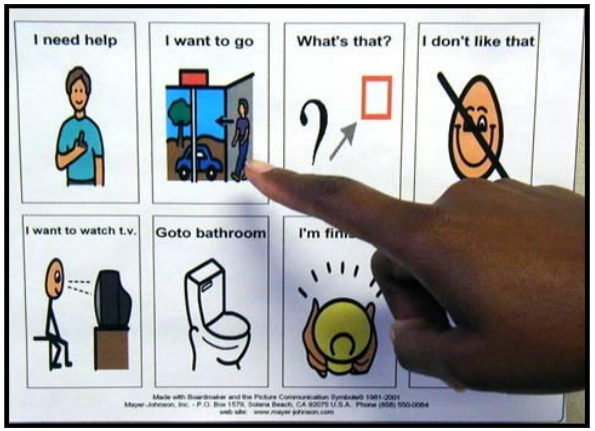
- People of all ages; ~2 million in US [NIH, 2000]
- Developmental disorders:
  - Autism, cerebral palsy...
  - 53% of people with CP use AAC [Jinks and Sinteff, 1994]
- Neurological and neuromotor disorders:
  - ALS, MD, MSA, stroke, paralysis...
  - 75% of people with ALS use AAC [Ball, 2004]

Functional Definitions

1. Target users are primarily non-speaking and may have upper limb motor impairments

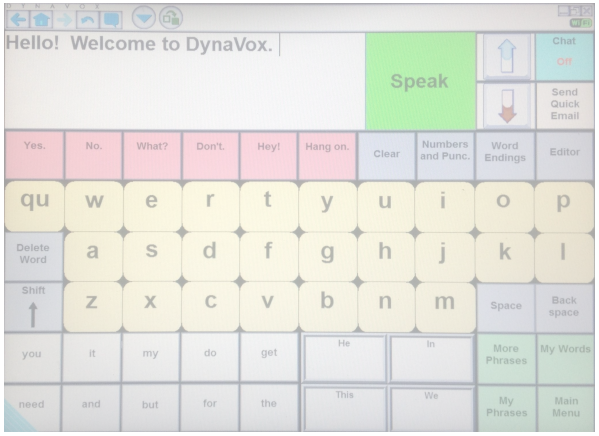
2. Target users may also have developing literacy or language impairments

Types of AAC

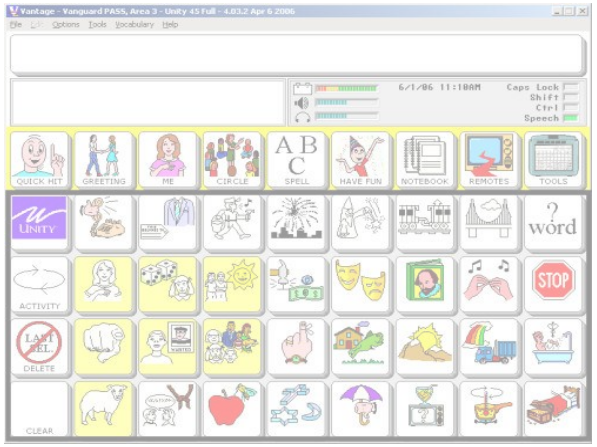


Physical Boards

Electronic Systems

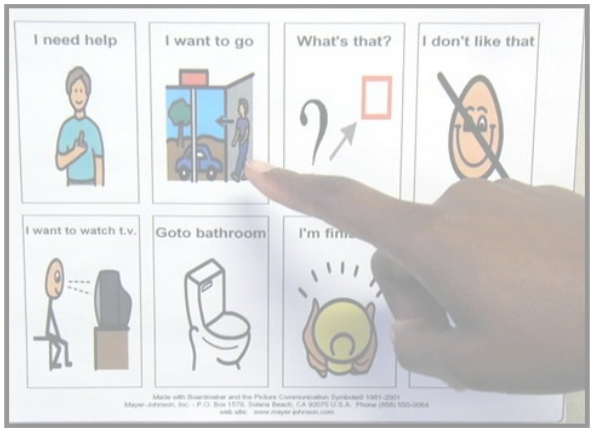


Letter-Based



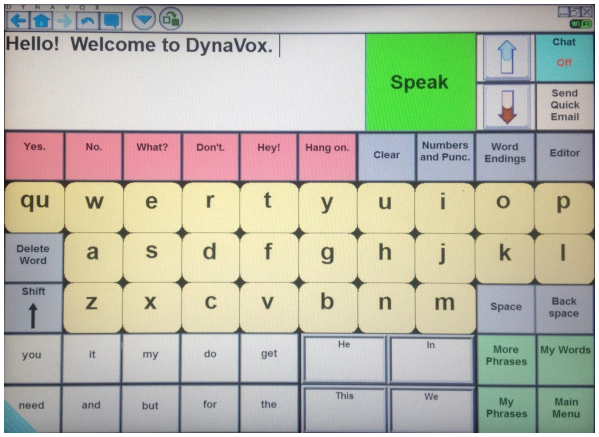
Icon-Based

Types of AAC



Physical Boards

Electronic Systems



Letter-Based

Icon-Based

On Speed of Communication

Speech is often **150 - 200** words per minute

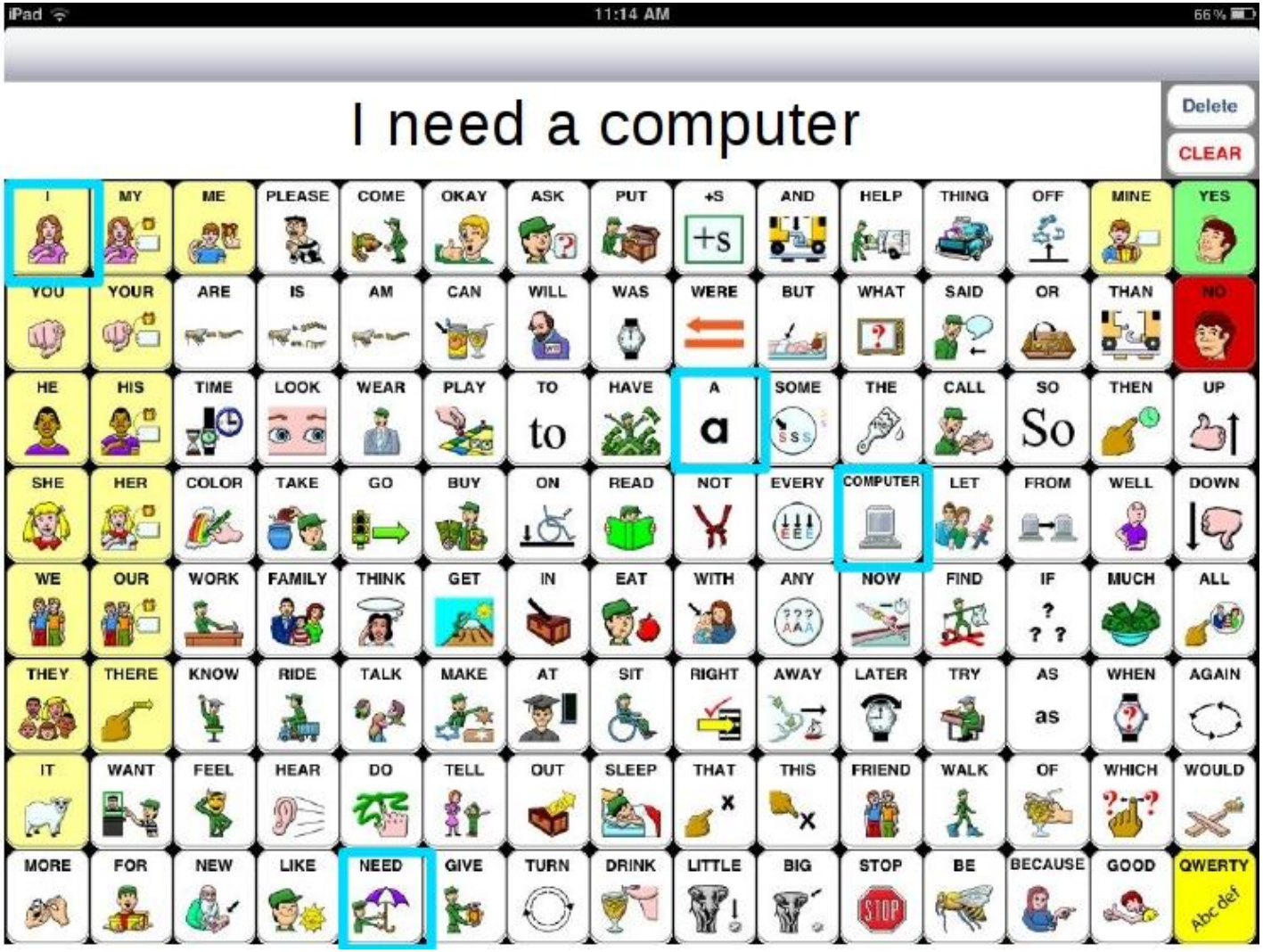
[Beasley and Maki, 1976]

vs.

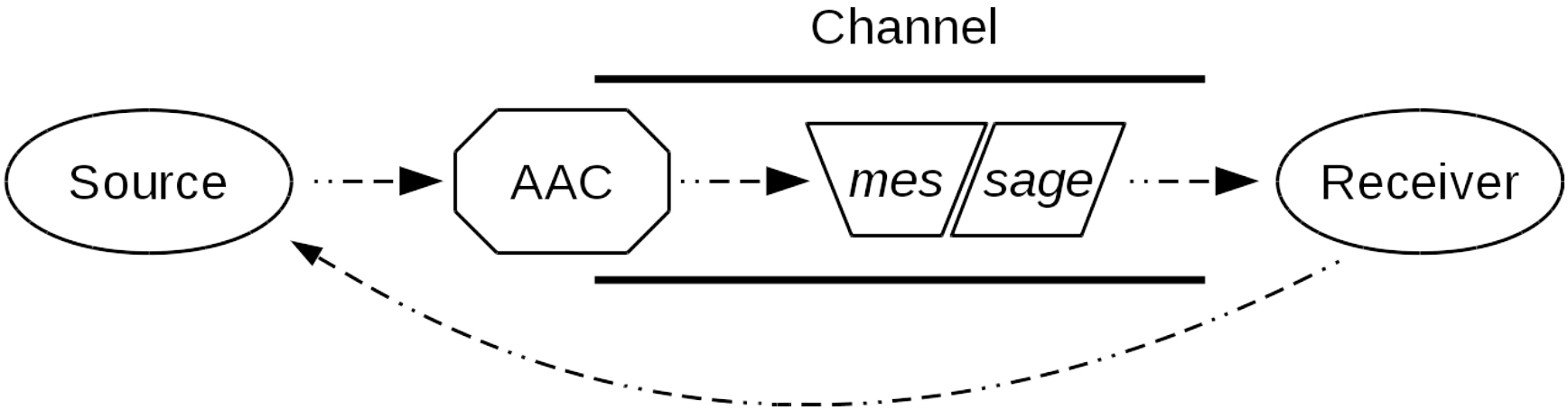
Typical AAC is **< 20** words per minute

[Higginbotham et al, 2007]

Modern AAC Application



The Problem



What is the Goal?

- Make AAC more intelligent
- "Intelligent" meaning:
  - User-specific
  - Adaptive
  - Context-sensitive

How?

By addressing some common assumptions:

1. Prescribed Order
2. Intended Set
3. Discrete Entry

Assumption 1: Prescribed Order

★ Users will select items in a specific order, such as the syntactically "correct" one.

- Users do not always select items in expected order [Van Balkom and Donker-Gimbrere, 1996]
- Using AAC devices is slow [Beukelman et al, 1989; Todman, 2000; Higginbotham et al, 2007]
- Assumptions of diminished capacity

Assumption 2: Intended Set

★ Users will select exactly the items that are desired -- no fewer or more.

- Motor and cognitive impairments may result in missing or additional selections [Ball, 2004]

- Letter-based text entry systems detect accidental and missing selections

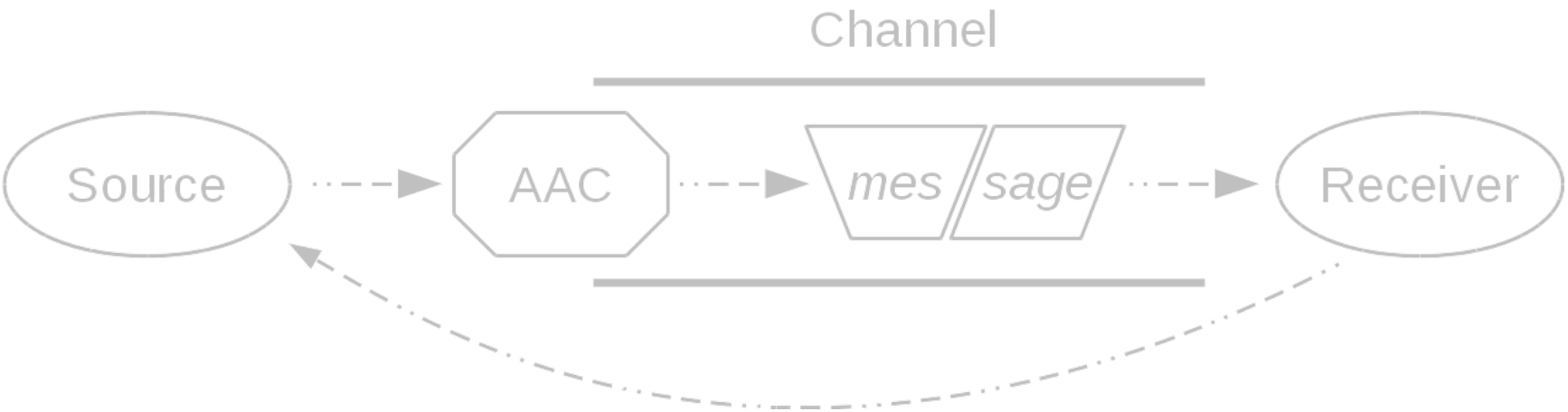
Assumption 3: Discrete Entry

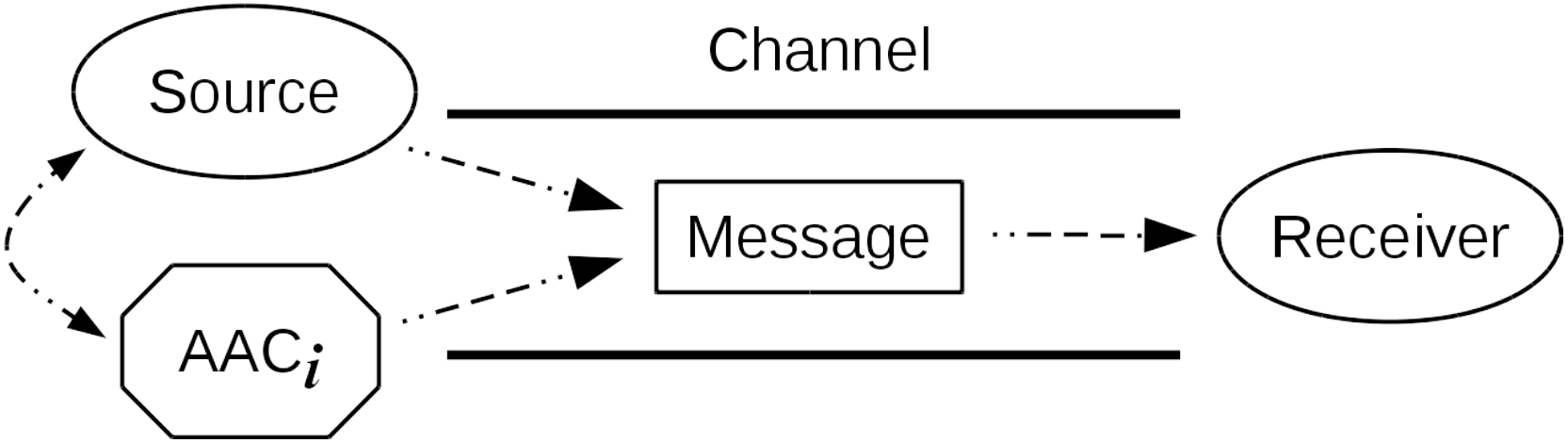
★ Users will make discrete movements or selections, either physically or with a cursor.

- Some letter-based systems have started to remove this assumption [Goldberg, 1997; Kristensson and Zhai, 2004; Kushler and Marsden, 2008; Rashid and Smith, 2008]

- Many input signals are naturally continuous

The Goal


Channel
Source
AAC
mes
sage
Receiver


Source
Channel
Message
Receiver
AAC$_i$

Part 2:

Theoretical Contributions

Theoretical Contributions

Prescribed Order	⟹	*Semantic Frames*, Semantic Grams
Intended Set	⟹	Semantic Grams, Contextual Prediction
Discrete Entry	⟹	Personalized Interaction

Theoretical Contributions

Prescribed Order ⟹ *Semantic Frames*, Semantic Grams

Intended Set ⟹ Semantic Grams, Contextual Prediction

Discrete Entry ⟹ Personalized Interaction

Addressing Prescribed Order

- Statistical MT [Soricut and Marcu, 2006]
- Semantic frames, CxG, and PAS [Fillmore, 1976]

  `Give ( Agent, Object, Beneficiary )`

- WordNet, FrameNet, "Read the Web" (NELL), Groningen Meaning Bank
- Computationally intense to obtain statistics

Motivating Questions

★ Can we create a simple and fast language model for use with semantic frames?

- Current completion and prediction strategies rely on syntactic order and word distance
  - N-grams, s-grams, skip-grams, CVSMs, etc.
  - Compansion [McCoy et al, 1998]
  - Memory-based LMs [Van Den Bosch and Berck, 2009]

★ Can utterances be predicted/completed without assuming order and distance?

Motivating Examples

Prior Input: play, video games, i, brother

Output: "My brother and I play video games."

Prior Input: play, chess, i, dad

Output: "I play chess with my dad."

Input: i, brother, ...

Output: ?

Possible Approach

- Sentences are one of the smallest units of language that are:
  - Semantically coherent
  - Semantically cohesive
  - Syntactically demarcated
- How can they be leveraged for prediction?

Semantic Grams

- A multiset of words that appear together in the same sentence.

"I like to play chess with my brother."

brother, chess (1)	brother, i (1)
brother, like (1)	brother, play (1)
chess, i (1)	chess, like (1)
chess, play (1)	i, like (1)
i, play (1)	like, play (1)

More on Sem-Grams

- Sentence Boundary Detection (SBD) is fast and relatively accurate (> 98.5%)
- Sentences provide dynamic context windows
- Sentence-level co-occurrence with uniform weight applied to all relationships in a sentence

Sem-Grams Study

- Blog Authorship Corpus
  - 140 million words from 19,320 bloggers
  - Age range of 13 - 48; balanced genders
- Split by authors: 80% training, 20% testing
- 2 n-gram and 2 sem-gram algorithms
  - Naive Bayes: N1 and S1
  - N2 (weighted adjacency) and S2 (full independence)

Method

For every test sentence:

1. Process (split, stop, stem, and check)
2. Shuffle stems
3. Remove one (target)
4. Query each algorithm for missing stem (ranked list)

Evaluation: random 2000 sentences

Score: position of target (lower score is better)

Results: Example 1

Original: “This semester Im taking six classes.”

Target Stem: class

Input Stems: take, semest, six

N1 Candidate List: next, month, class, hour, last, second, week, year, first, five, flag, ...

S1 Candidate List: class, month, year, last, time, one, go, day, get, school, will, first, ...

Results: Example 2

Original: “Hey, they’re in first, by a game and a half over the Yankees.”

Target Stem: game

Input Stems: yanke, hey, first, half

N1 Candidate List: game, stadium, like, hour, time, year, day, guy, hey, fan, say, one, two, ...

S1 Candidate List: game, got, like, red, time, play, team, sox, hour, go, fan, one, get, day, ...

Results: Example 2

Original: “Hey, they’re in first, by a game and a half over the Yankees.”

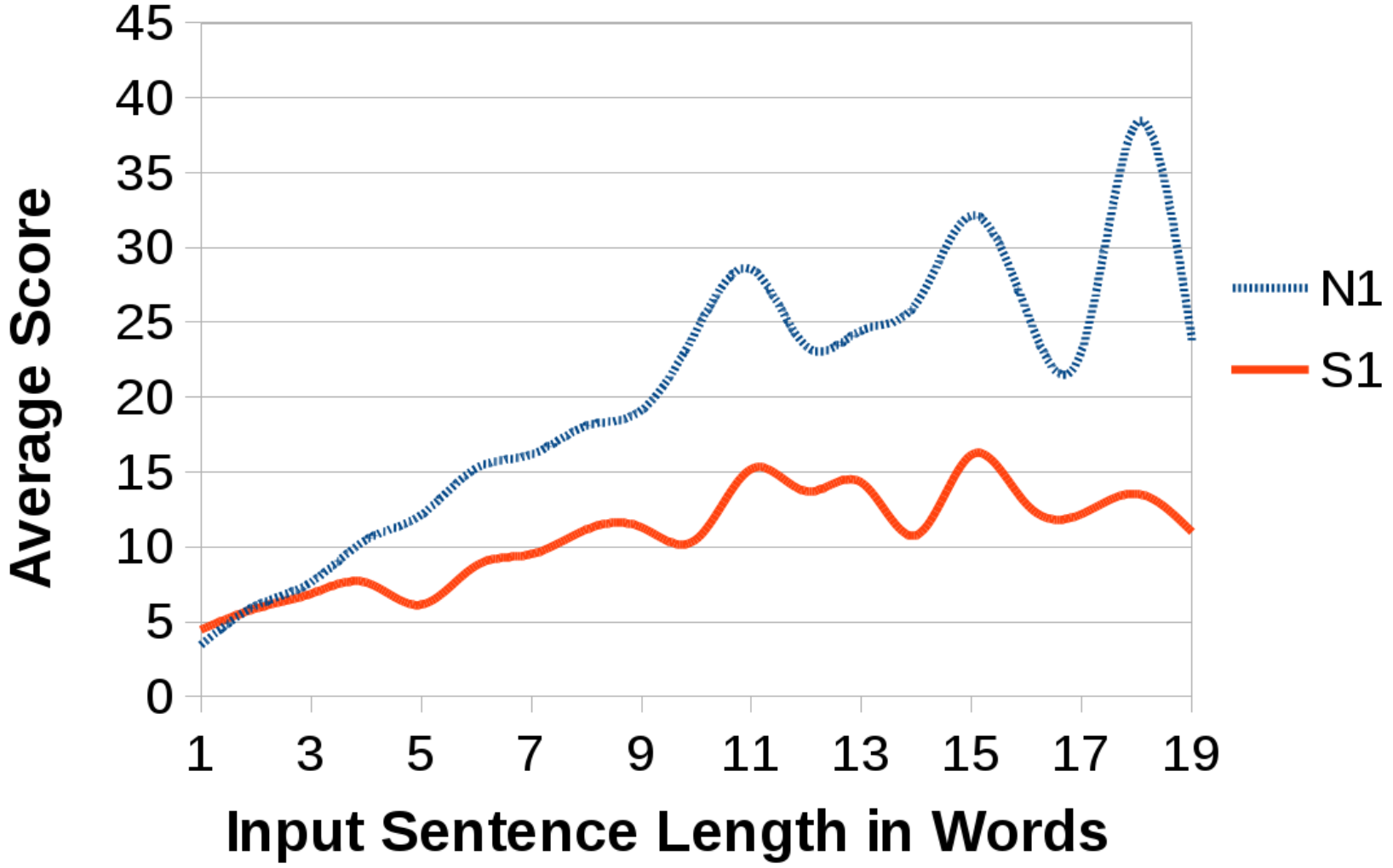
Target Stem: game

Input Stems: yanke, hey, first, half

N1 Candidate List: game, stadium, like, hour, time, year, day, guy, hey, fan, say, one, two, ...

S1 Candidate List: game, got, like, red, time, play, team, sox, hour, go, fan, one, get, day, ...

Results: Performance of Sem-Grams



Summary of Sem-Grams

- Simple, "fast" (SBD), and distance-agnostic
- More accurate than similar n-gram-based algorithms
- Alternative to more complex methods
- Natural fit for use with semantic frames

Theoretical Contributions

Prescribed Order ⟹ *Semantic Frames*, Semantic Grams

Intended Set ⟹ Semantic Grams, Contextual Prediction

Discrete Entry ⟹ Personalized Interaction

Improving Unordered Prediction

- Dropping assumption of order results in information loss
- How can we compensate?
- Devices often ask for user demographics
- Mobile AAC devices have sensors:
  - Date
  - Time
  - Location

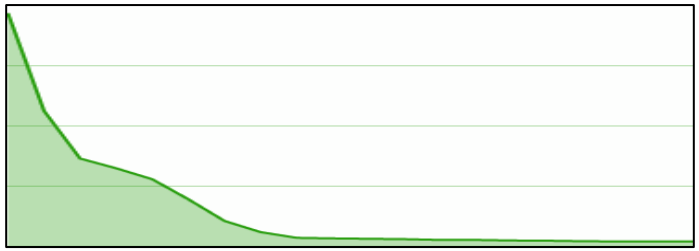
Motivating Questions

- Almost all statistical LMs require background probabilities (priors)

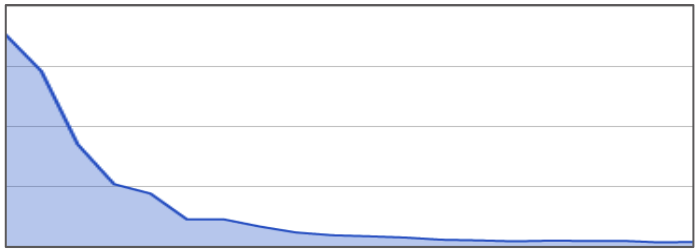
- Most systems use Google's N-Gram Corpus, Wall Street Journal, or New York Times

★ How much closer to a real user's priors can we get by leveraging context?

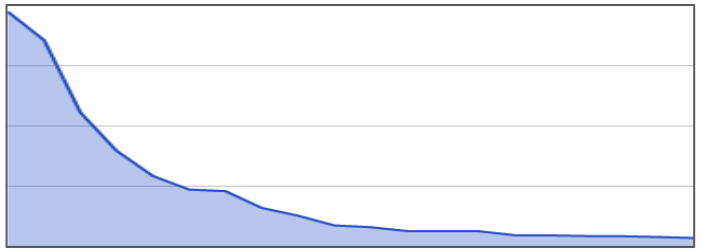
Contextual Prediction



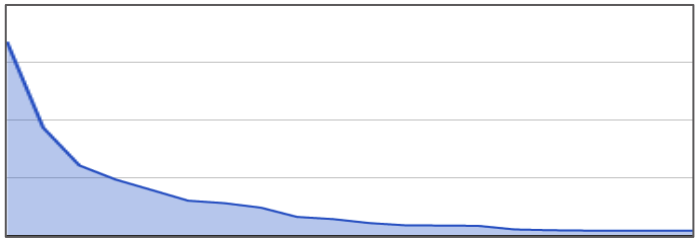
23-year-old female in Seattle



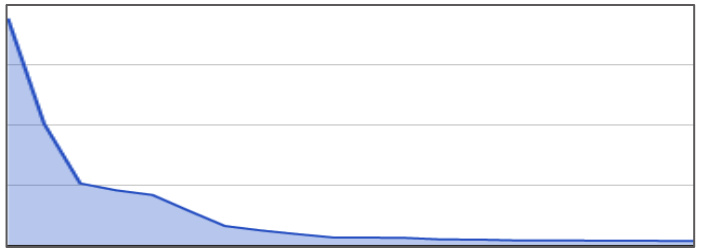
23-year-olds



Global



Seattle



23-year-old females

Contextual Prediction Study

- Blog Authorship and Yelp Academic Dataset
- Contexts: age, gender, day of the week, day of the month, month, city, and state
- Map unigrams to contexts for all authors; minimal stops and no stemming

Attribute	Blog Authorship	Yelp
Authors	19,320	130,850
Features	525,253	134,199

Method

Split by authors: 90% training, 10% testing

For every test author's unique context:

1. Obtain the true distribution (target)
2. Compare to distribution from each predictor combo based on non-target 9 folds

Metrics: Kullback-Leibler Divergence, Cosine Similarity, and Precision@20

Method Example

Target Distribution	Predictor Combos
Age: 23	Age
Gender: Female	Gender
DOW: Monday	DOW
DOM: 25 - 31	Age + Gender
Month: July	Month + City
City: Seattle	Age + Gender + City
State: Washington	...
	(48 in total)

Results: Predictors by Metric

KL Divergence	Rank	CosSim & Prec@20
DOW+DOM+Month+City	1	Gender+DOM+Month
Age+Gender+DOW+DOM+Month	2	Gender+Month
Age+DOW+DOM+Month	3	Age+Month
DOW+DOM+Month+State	4	Gender+DOW+Month
DOW+Month+City	5	Age+Gender
Age+Gender+DOW+Month	6	Age
DOM+Month+City	7	Age+DOM

. . .

(No Context)	47	31, 27	(No Context)

Summary of Context

- Contextual distributions can be more accurate than global statistics
- Location better by KL; demographics better by CosSim and Prec@20
- Some combinations consistently better:
  - Gender + DOM + Month
  - Age + Gender + DOW + Month
  - Age + Gender + DOM
  - Age + Month

Theoretical Contributions

Prescribed Order ⟶ *Semantic Frames*, Semantic Grams

Intended Set ⟶ Semantic Grams, Contextual Prediction

Discrete Entry ⟶ Personalized Interaction

Addressing Discrete Entry

- Physical path or signal characteristics
  - Rotated unistroke recognition [Goldberg, 1997]
  - Letter-based paths [Kristensson and Zhai, 2004; Kushler, 2008]
  - Relative positioning [Rashid, 2008]

- Well-received by non-disabled users

Motivating Questions

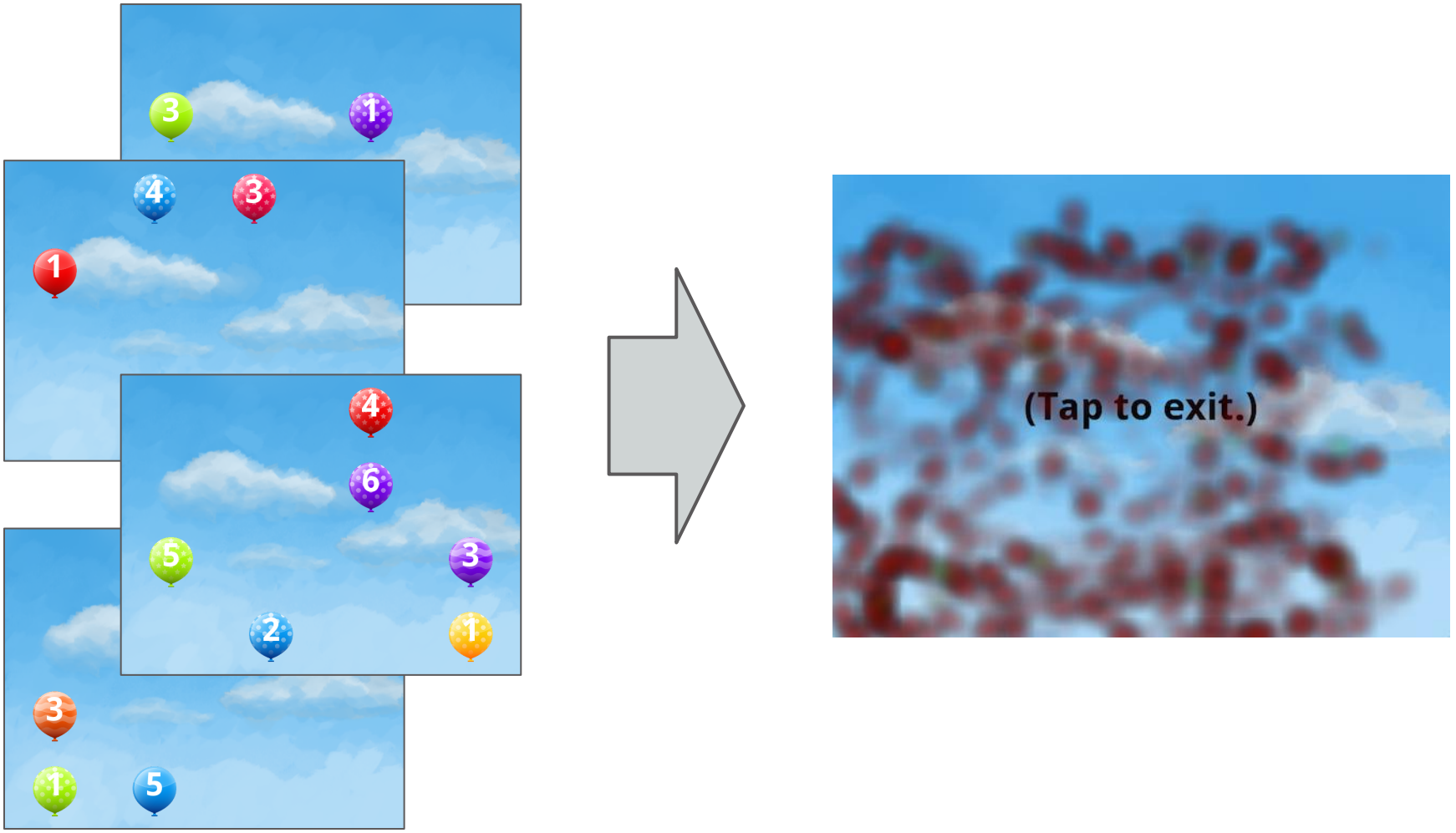
- Modern AAC now deployed on touchscreens
- Increasing research on accessibility
  - Fitts and Steering Laws [Fitts, 1954; Accot and Zhai, 1996]
  - Swabbing/sliding is easier [Wacharamanotham et al, 2011]
  - Buttons need to be bigger [Chen et al, 2013]

★ What about functional compensation?

★ Can we learn realistic, layout-agnostic interaction patterns for an individual user?

Motor Optimization GUI (MoGUI)

MoGUI Example



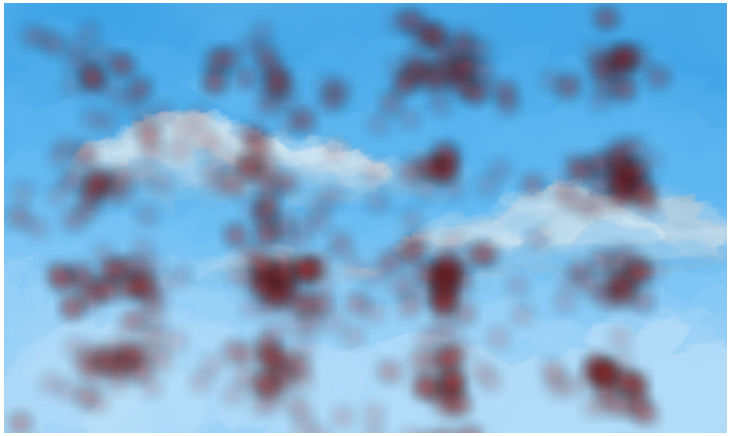
MoGUI Study

- Residents at the Boston Home
  - Current and potential AAC users
  - 10 females and 5 males
  - Ages 35 - 71 (mean of 56)
- 8 right-handed; 7 left-handed (3 due to MS)
- 2 cross-balanced sessions: taps vs. slides
- 4x4 grid = 16 locations
  - Pseudo-random shuffling (a la Latin Squares)

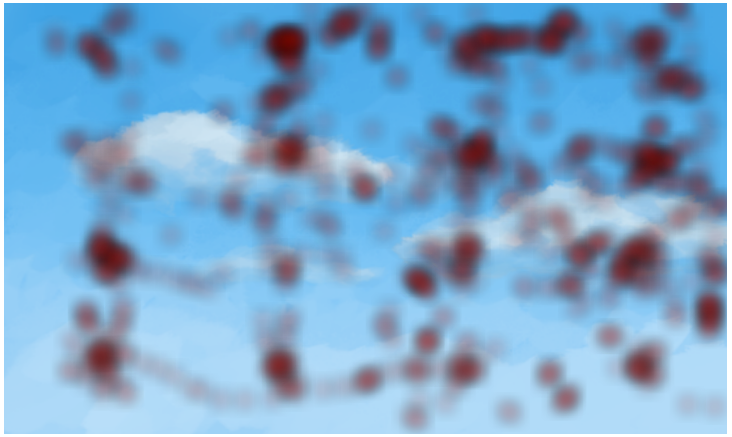
Method

- 10.1" Android tablet in comfortable, landscape position; fully reachable
- Choice of finger or stylus
- 10 levels of 3 rounds each
- 1, 2, 3, ...10 balloons per round = 165 total
- Track all hits, misses, and timing

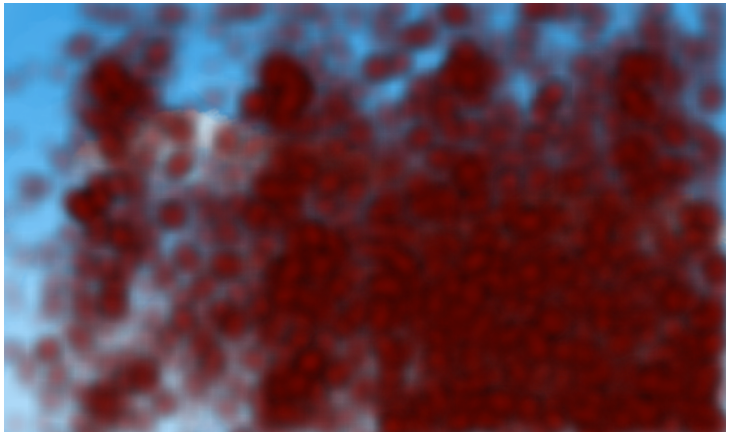
Results: Variability of Tap Misses



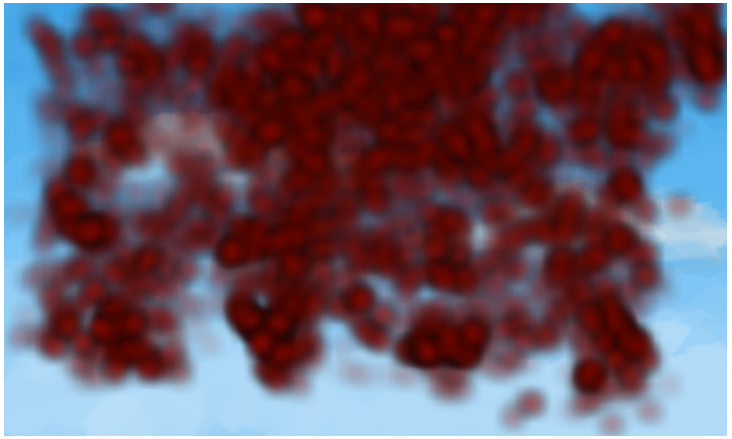
Multiple Taps



Fingers Dragging

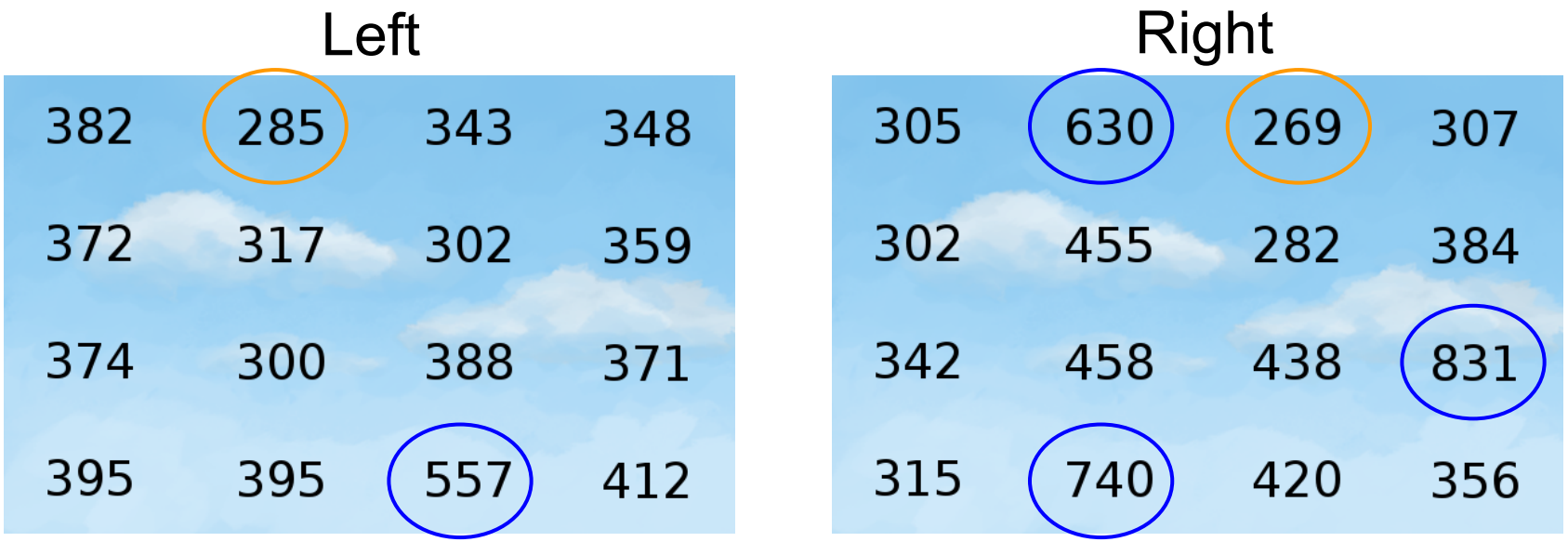


Hand Resting



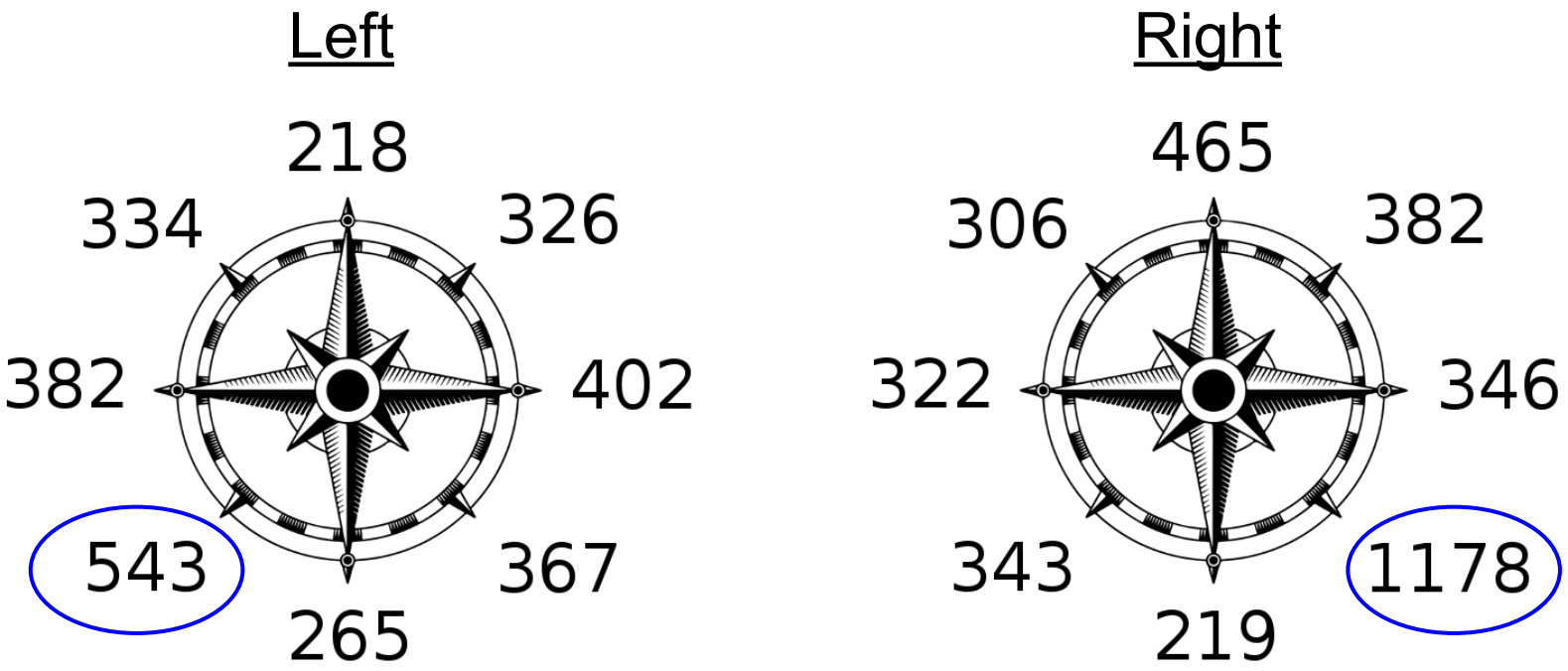
Thumb Usage

Results: Locations by Handedness



Mean speed-to-target in pixels/second

Results: Directions by Handedness



Mean speed-to-target in pixels/second

Summary of Personalization

- Sliding not significantly faster than tapping for arbitrary targets; no motor learning
  - 16% accidental slides; 43% accidental taps
- High variance in individual motor patterns; weak correlations by handedness
  - Gamified calibration
- Static improvements through personas:
  - Handedness → margins, button locations
  - Tap/slide preferences → input sensitivity

Part 3:

Applied Contributions

Applied Contributions

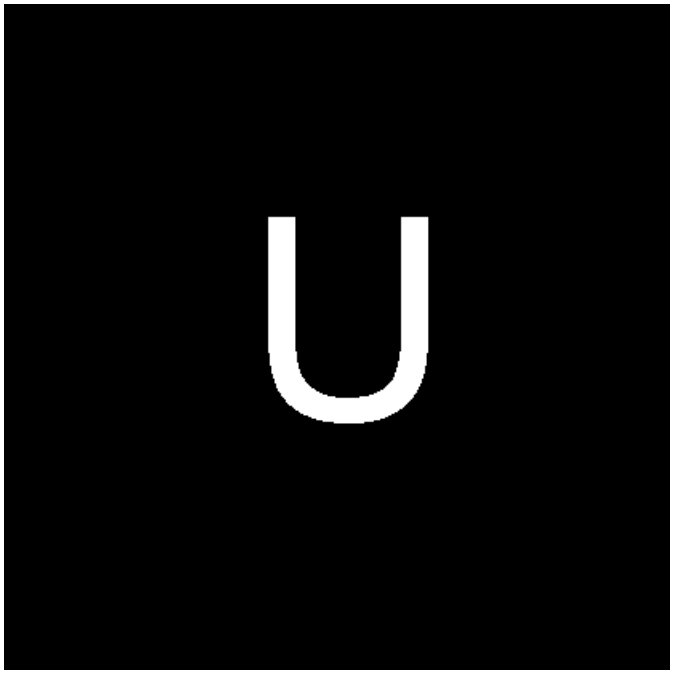
RSVP-iconCHAT	⟹	Free Order, Discrete Icons
SymbolPath	⟹	Free Order, Continuous Icons
DigitCHAT	⟹	Mobile, Mixed-Input Letters

A Collaborative Effort

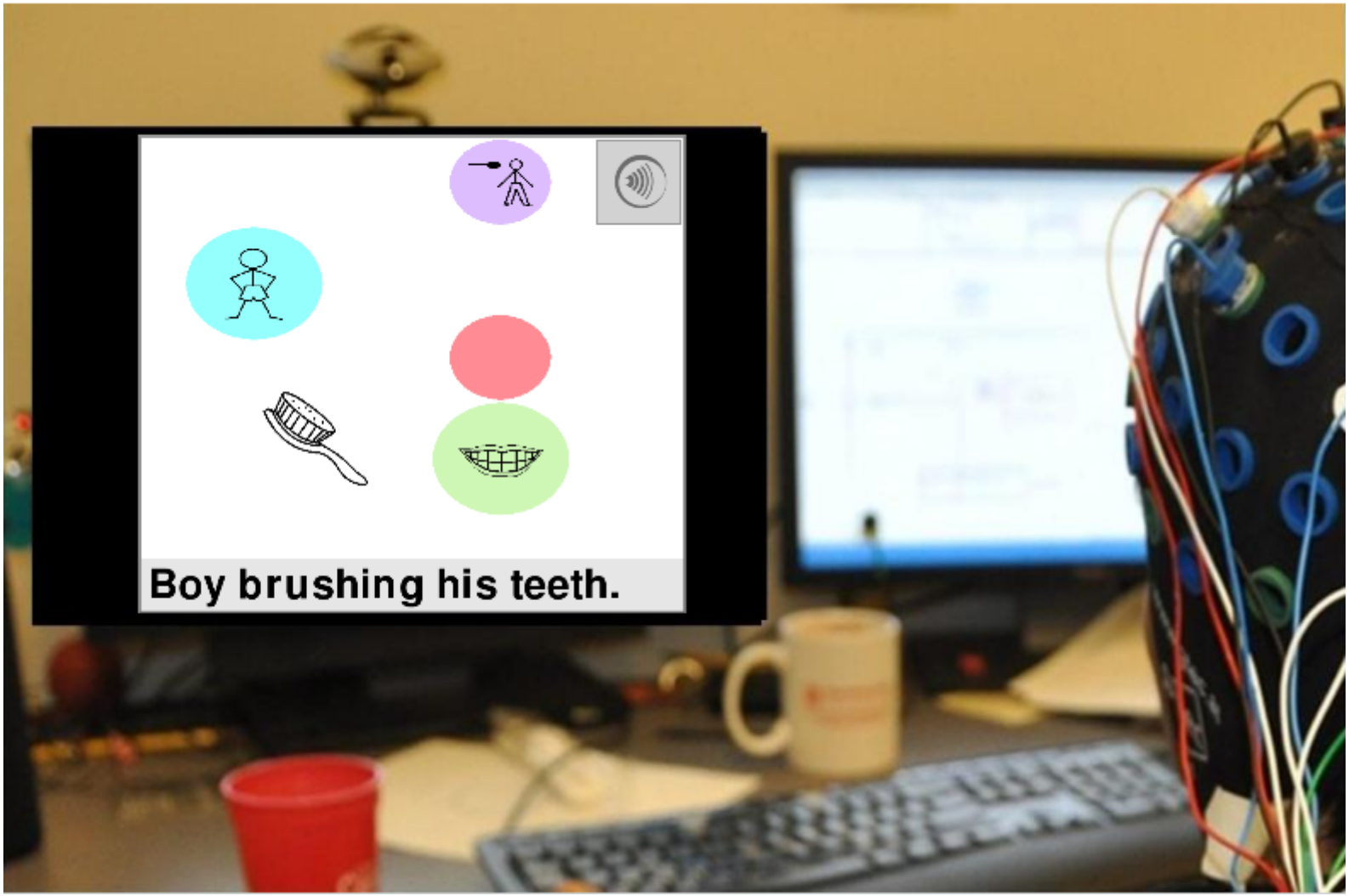
- Locked-In Syndrome (LIS)
  - Spinal injuries, ALS, tumors, strokes...
  - 1% of ischemic strokes [Smith and Delargy, 2005]
- Icon-based, switch AAC for people with LIS
  - Dr. Deniz Erdogmus and Dr. Rupal Patel
- Minimal switch/signal requirements (1+)
  - Goal of a brain-computer interface (BCI)
- Verb-first message construction [Patel et al, 2004]

Rapid Serial Visual Presentation

- Used in psychology, speed-reading, lie detection, and letter-based BCI [Orhan et al, 2012]

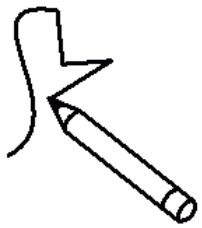


RSVP-iconCHAT

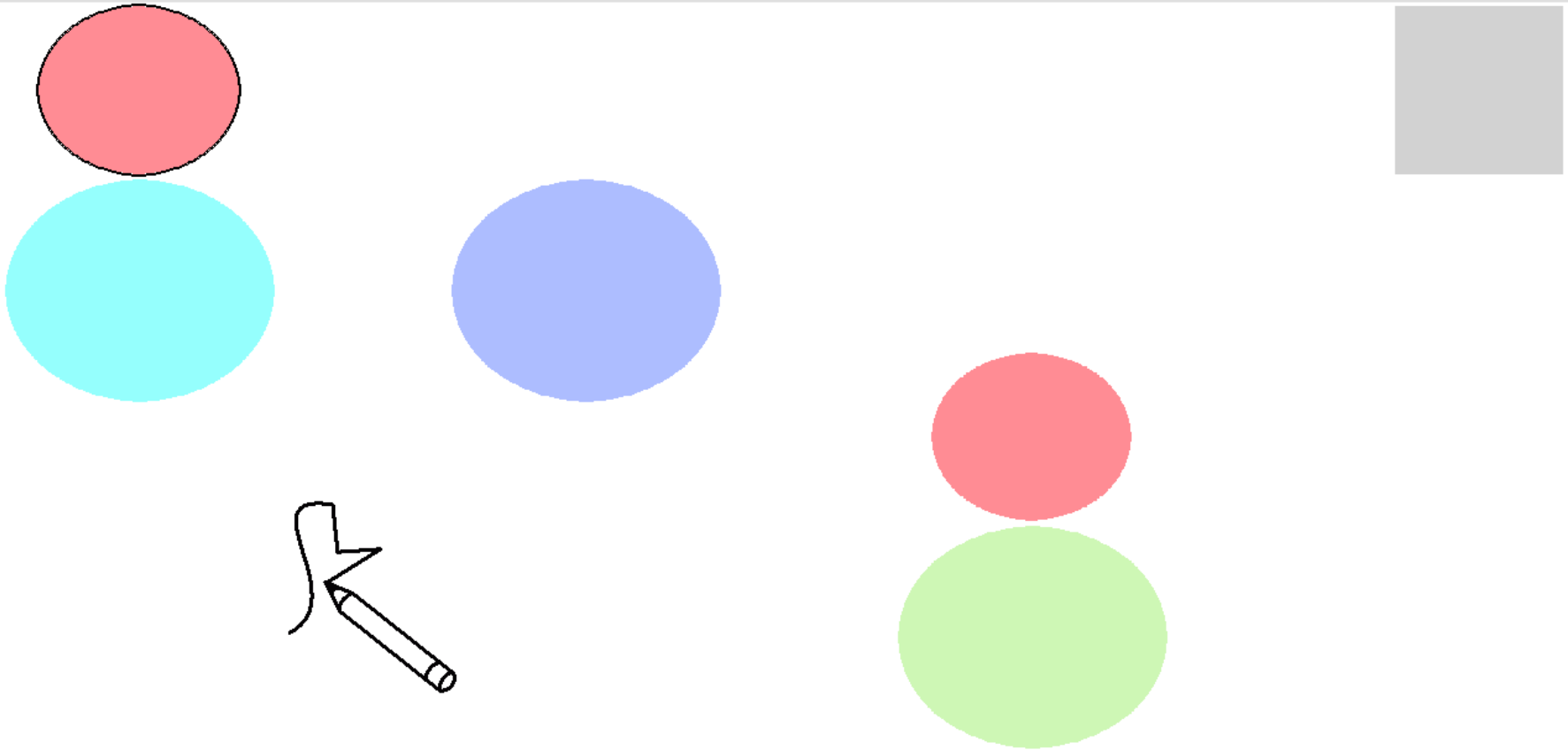


Cook.

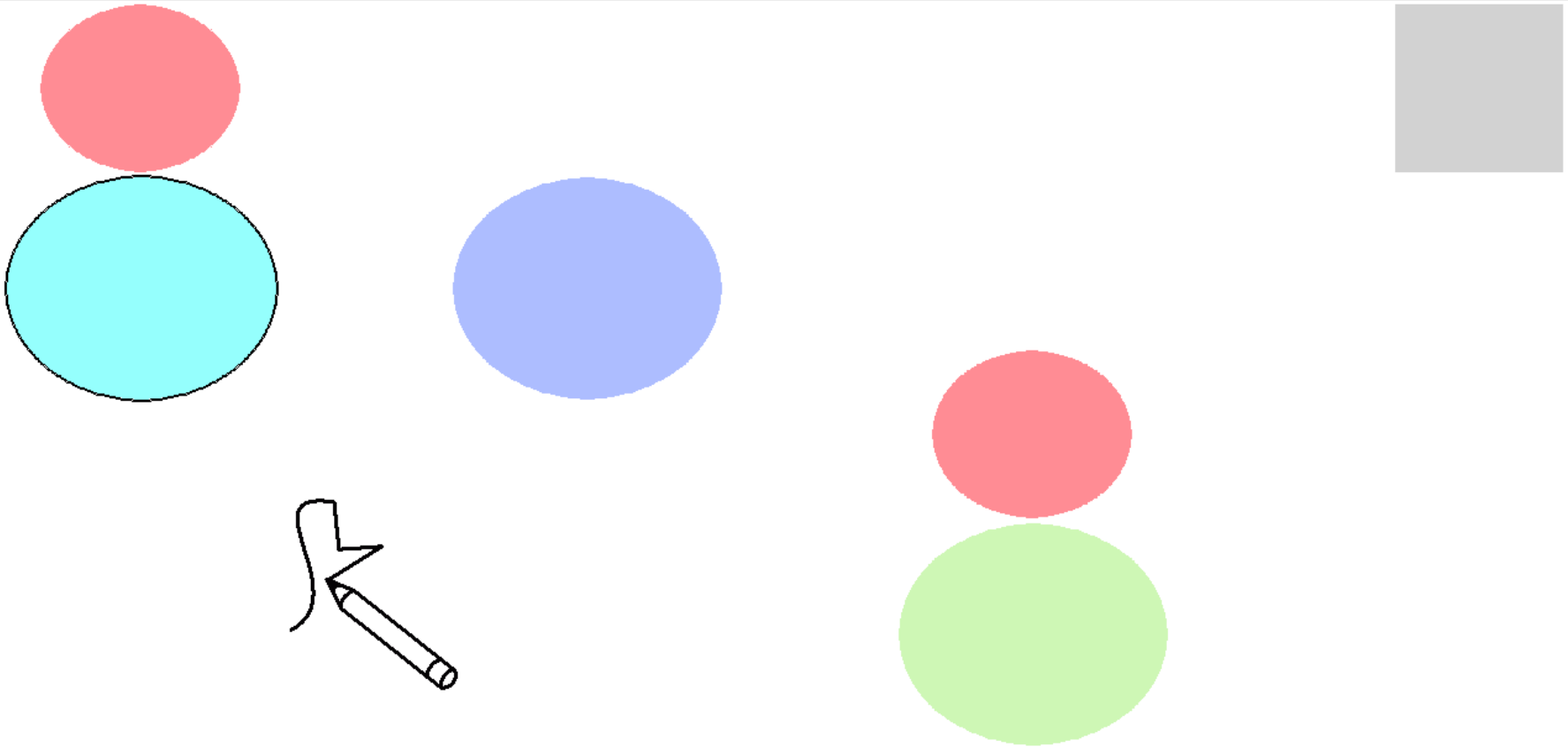
Wear.



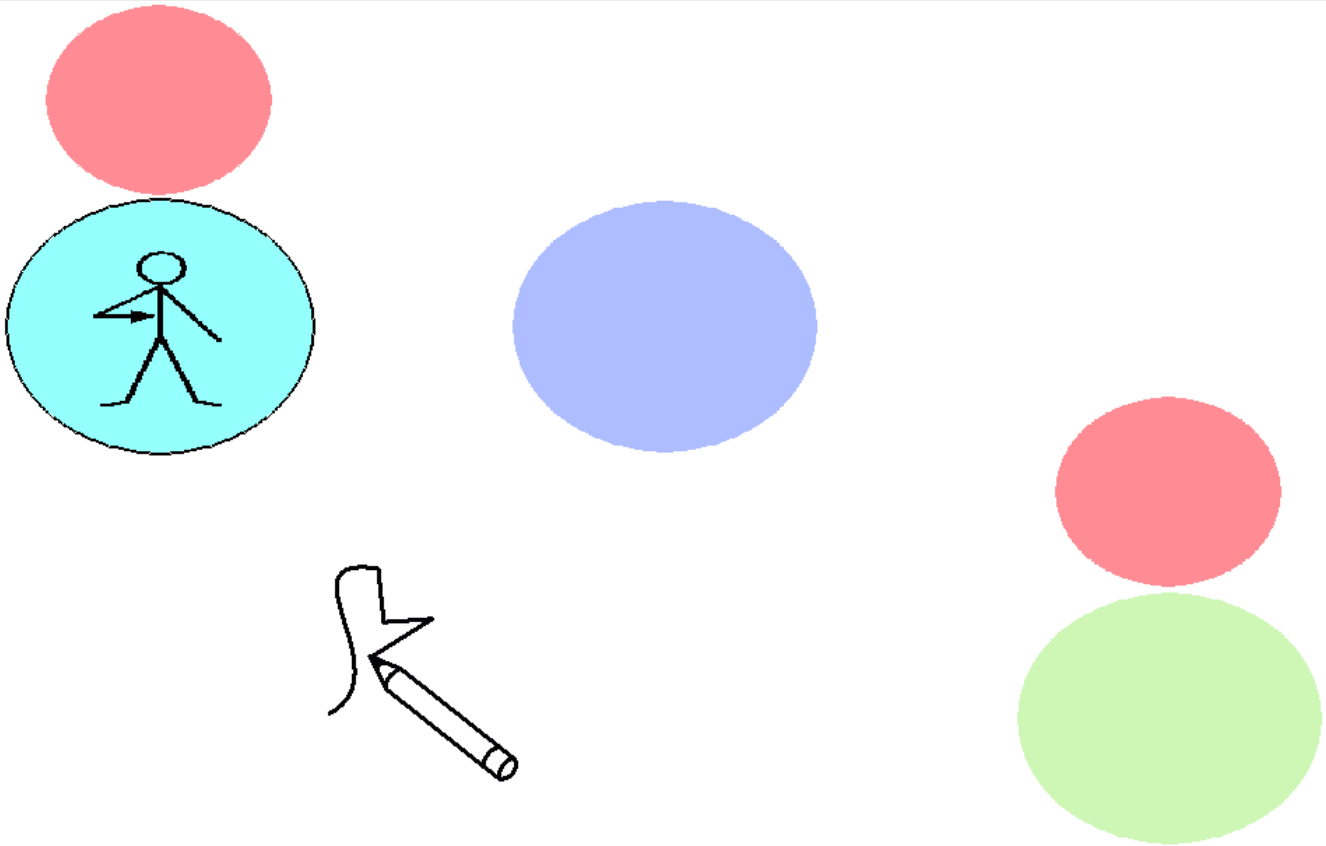
Draw.



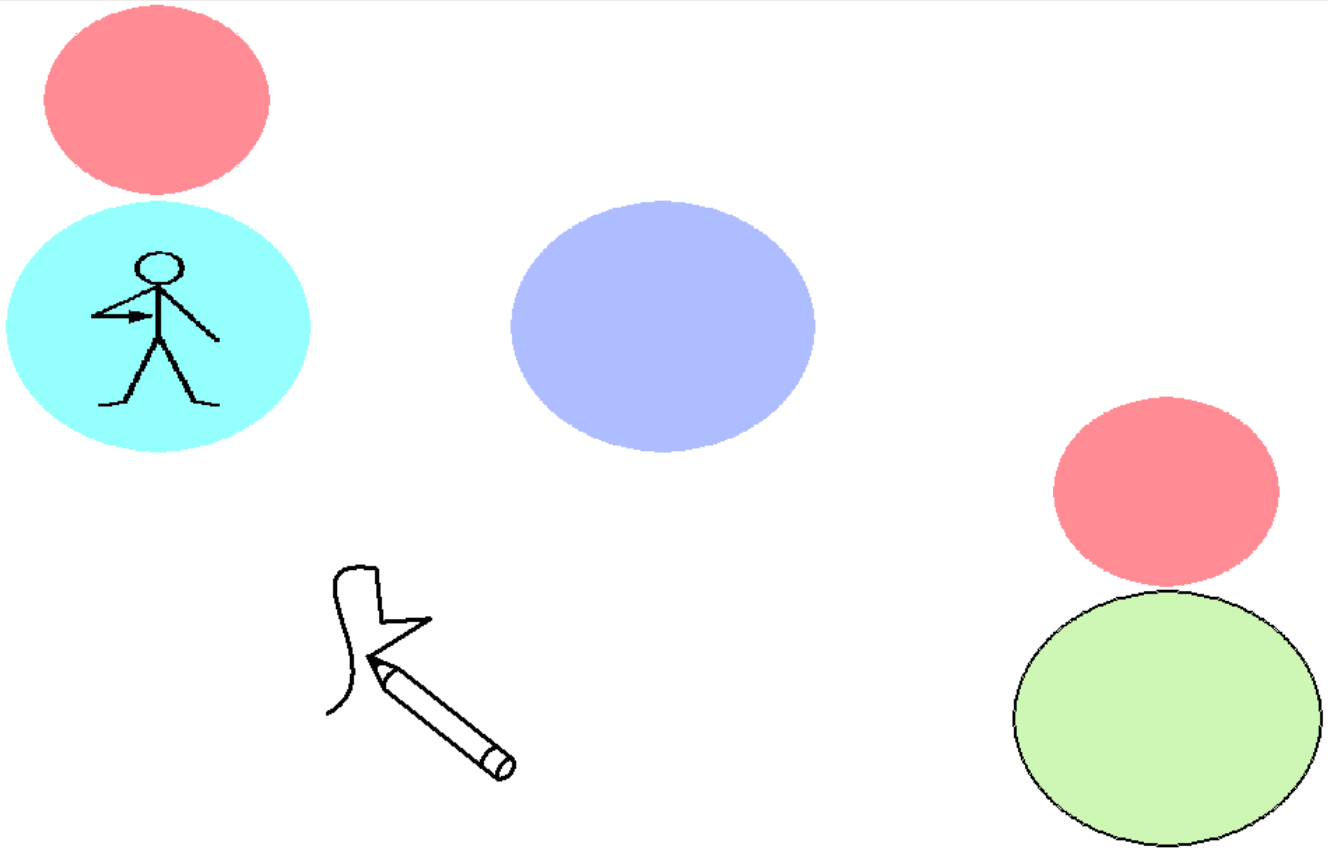
Draw.



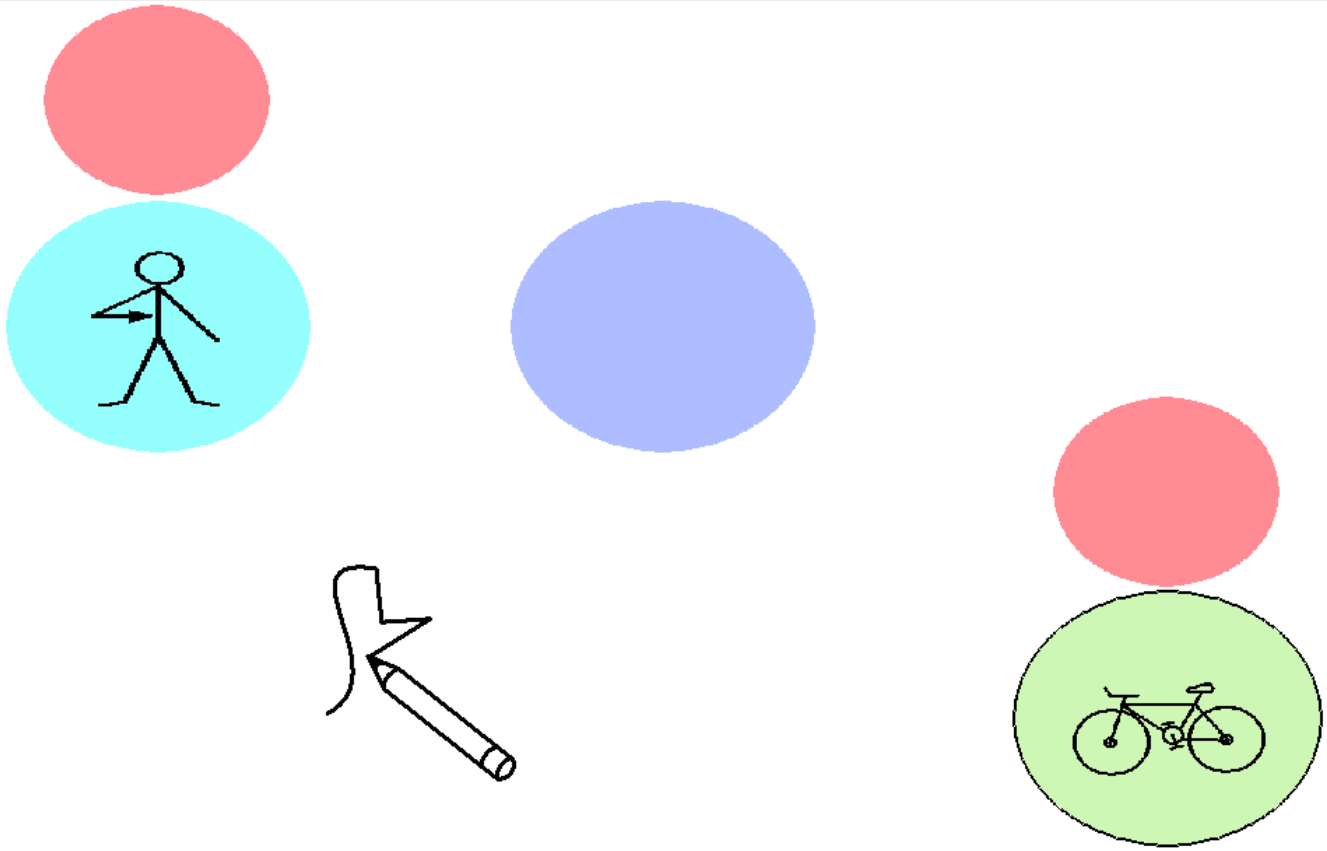
Draw.



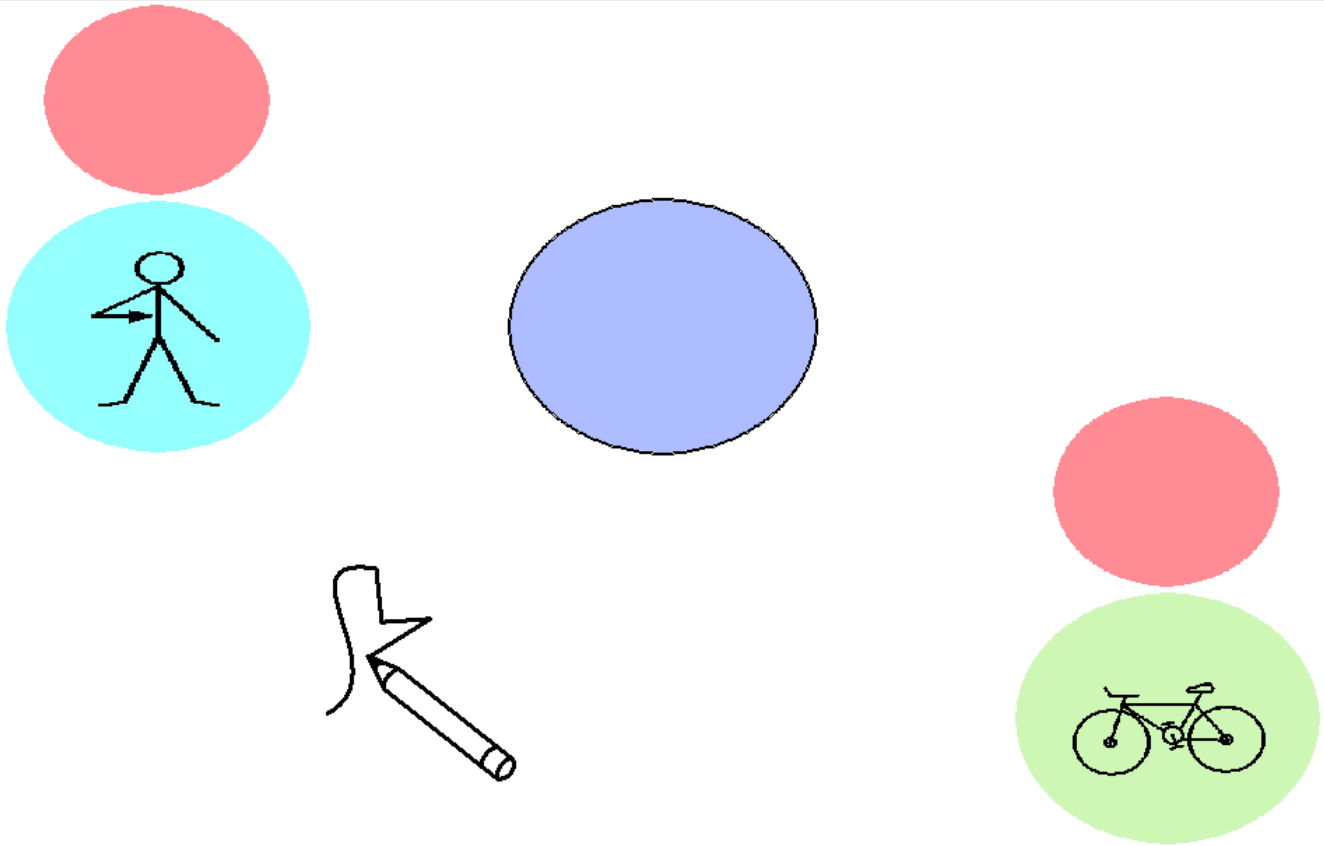
I draw.



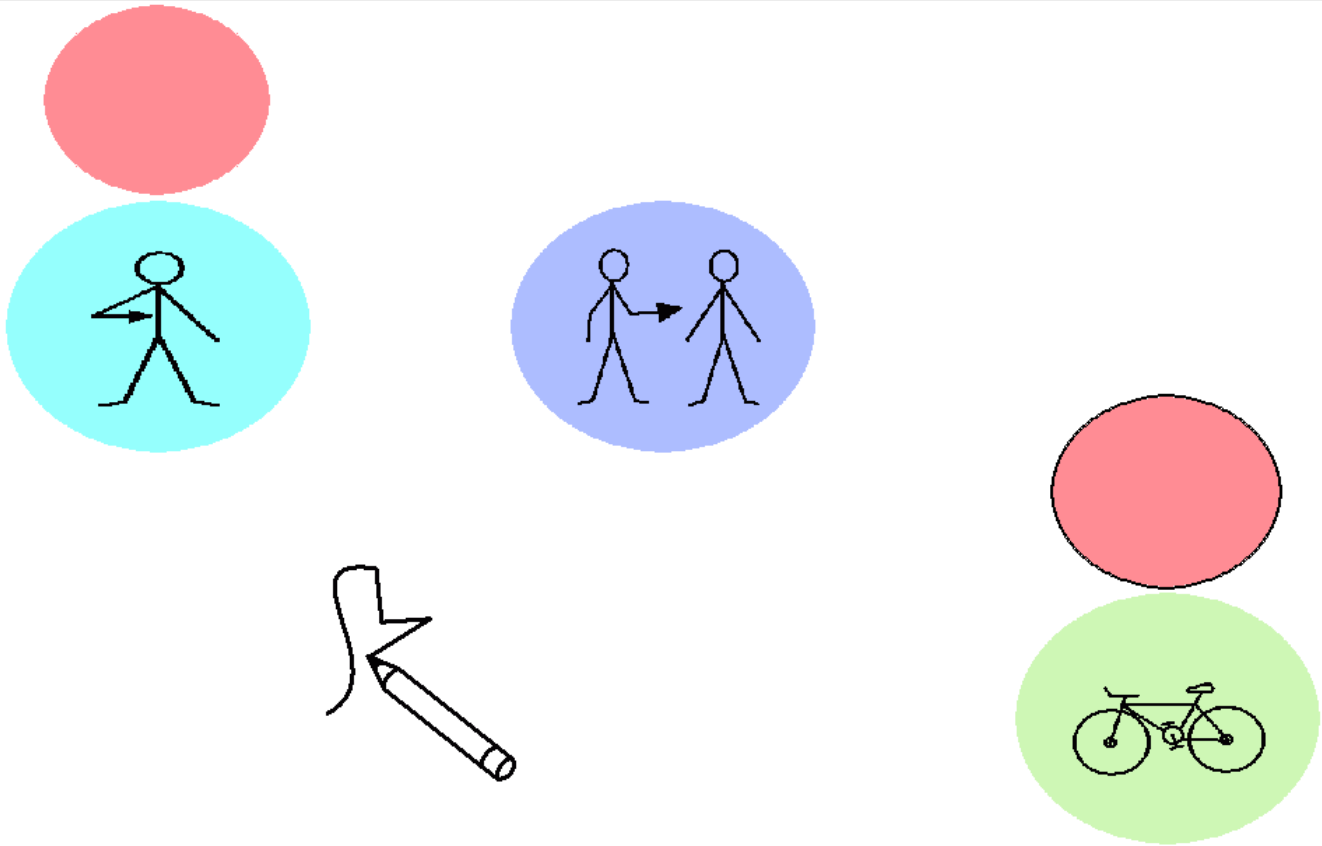
I draw.



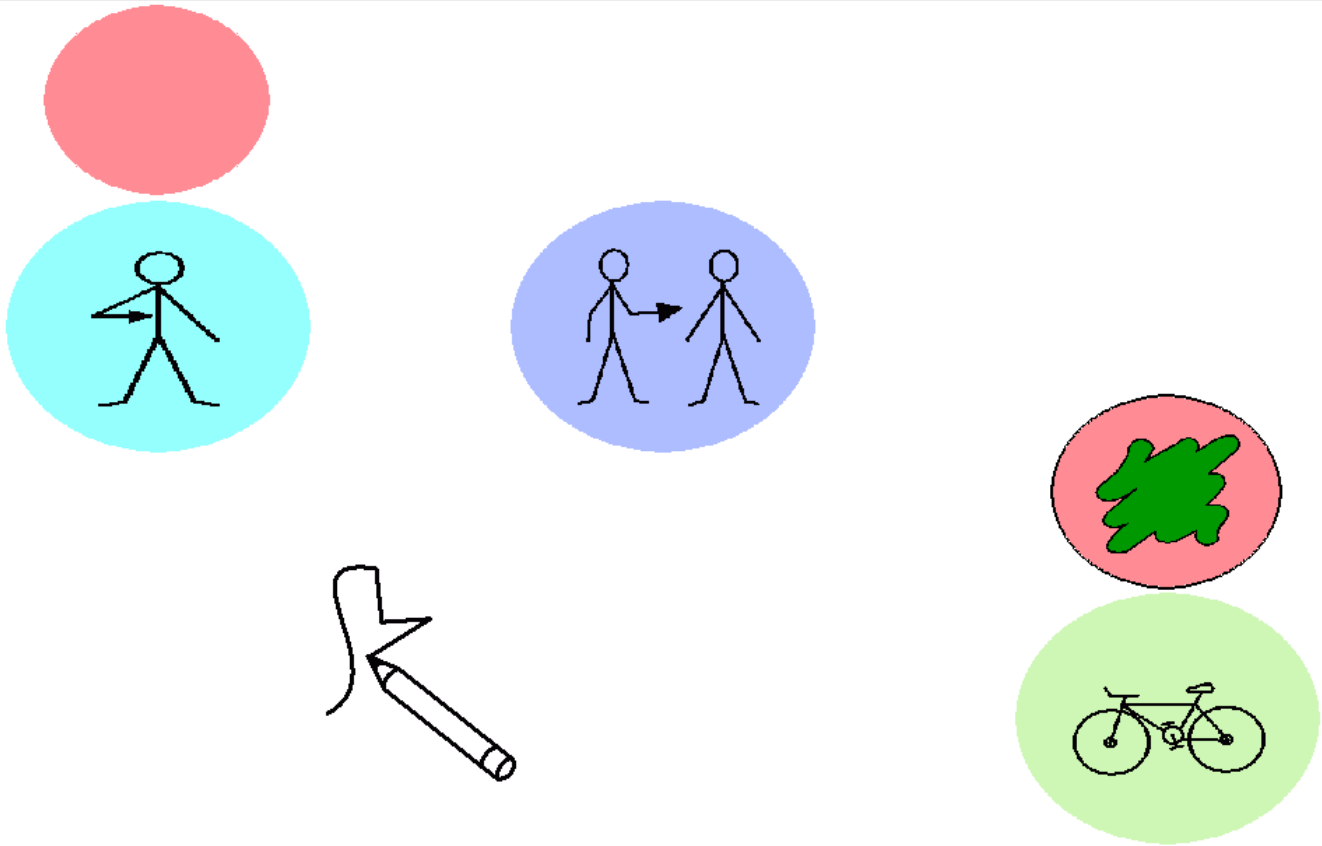
I draw a bicycle.



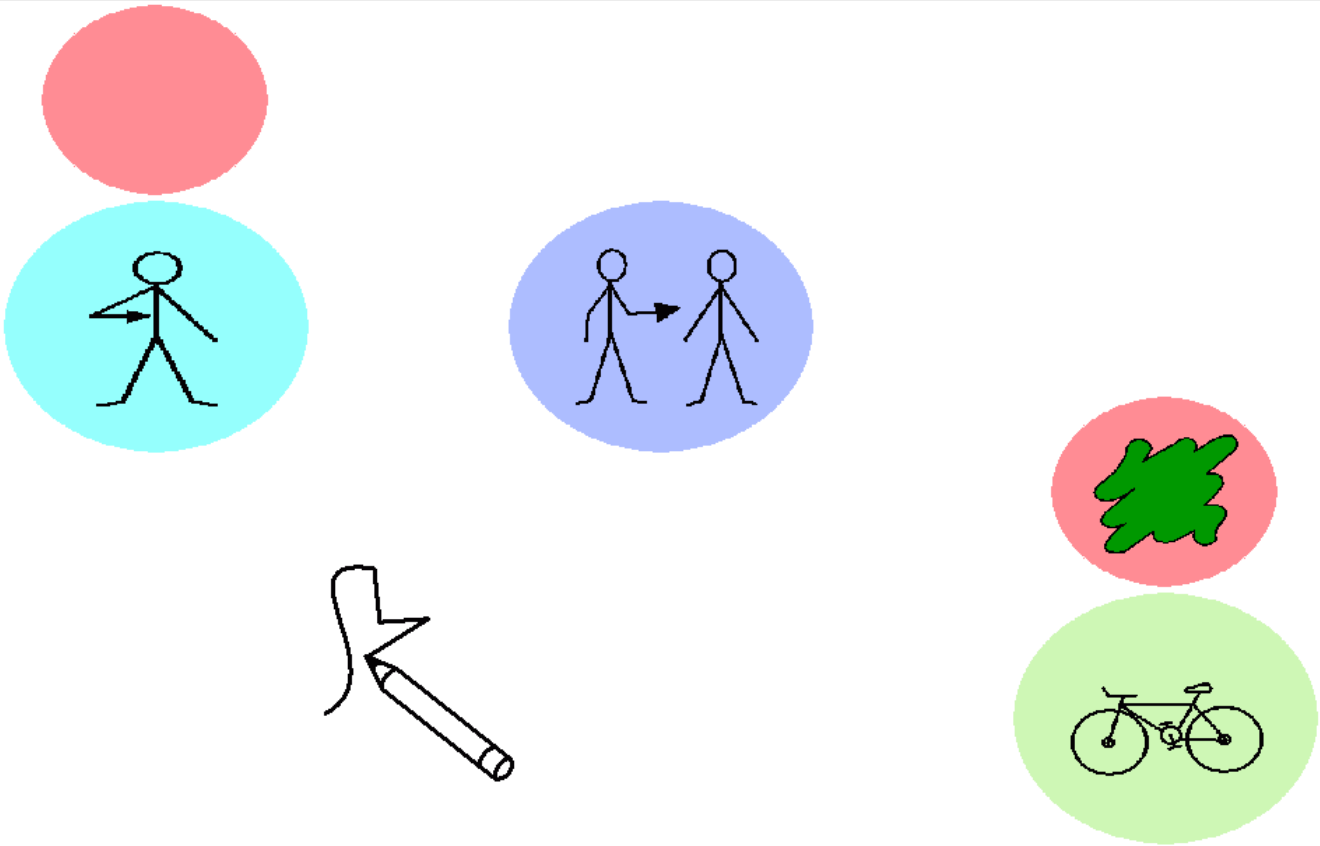
I draw a bicycle.



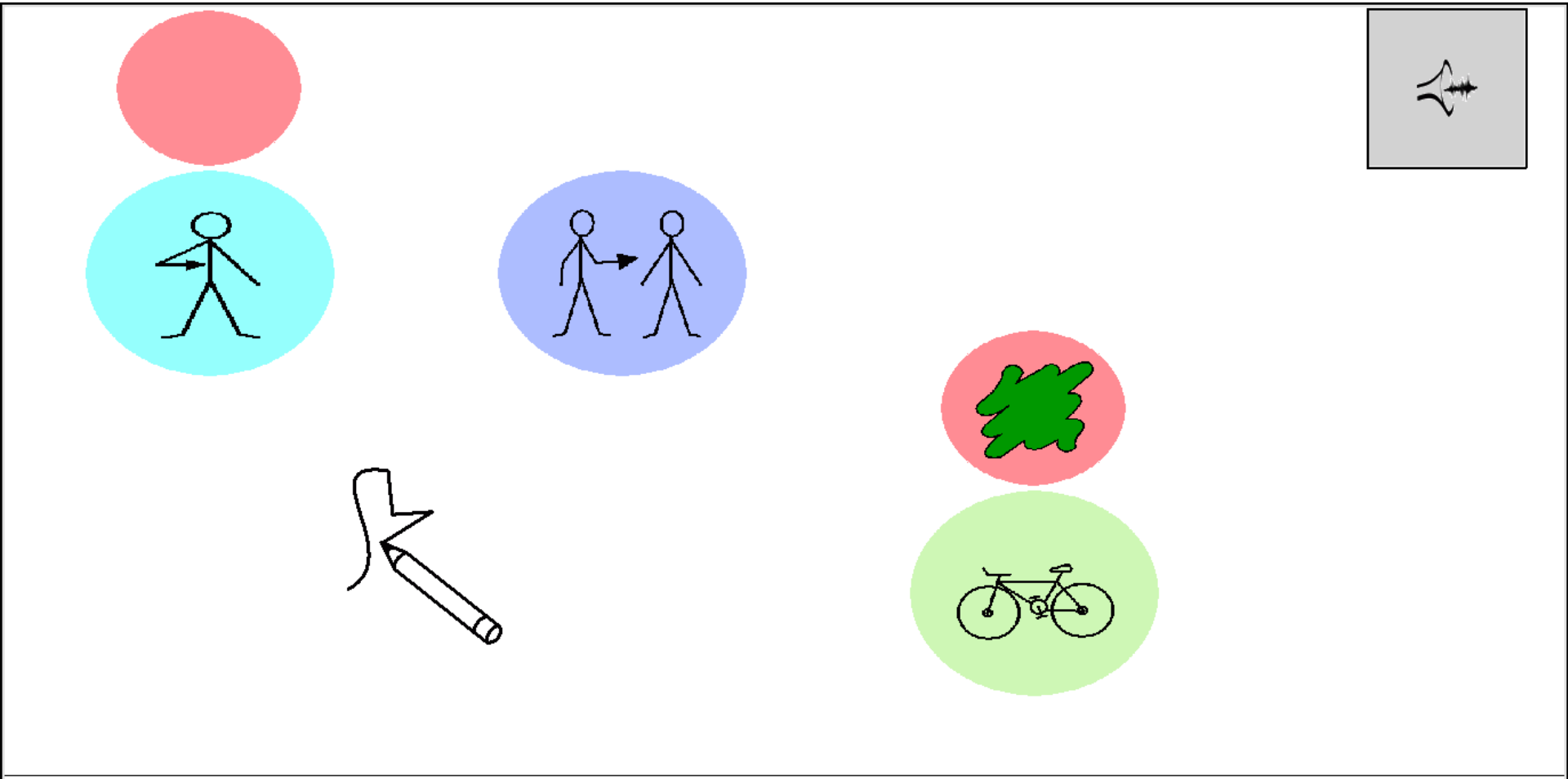
I draw a bicycle with you.



I draw a green bicycle with you.



I draw a green bicycle with you.



I draw a green bicycle with you.

Observations

- Prediction/ordering controls speed of message construction
- Natural fit for prediction via semantic grams
- Required screen space is now tied to message complexity

RSVP-iconCHAT Study

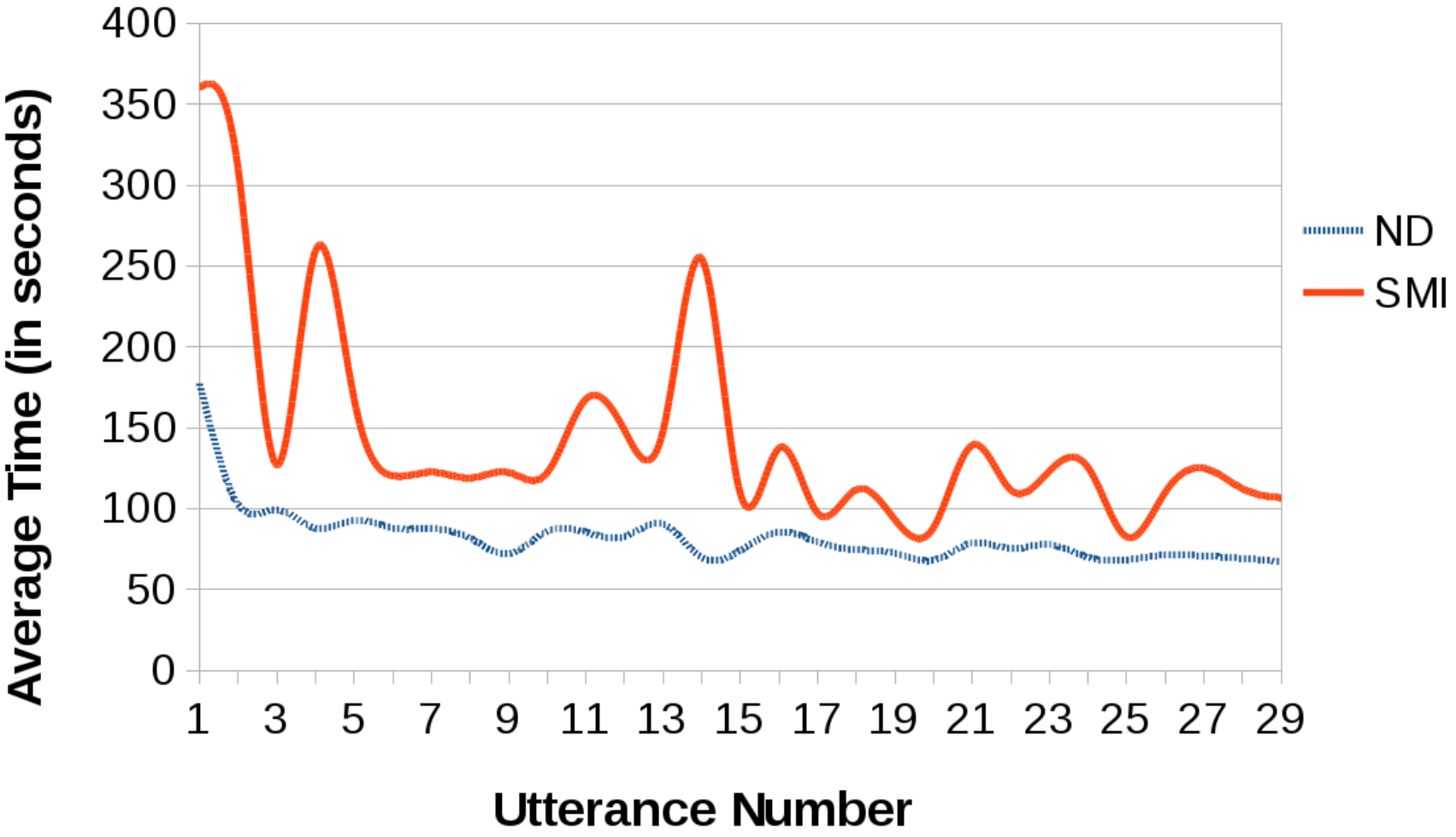
- 24 non-disabled participants (ND)
  - 14 females and 10 males
  - Ages 19 - 43 (mean of 24)
- 4 participants with speech and motor impairments (SMI)
  - 2 females and 2 males
  - Ages 33 - 56 (mean of 41)
- Space bar as switch mechanism
  - Up to 106 words in alphabetic order

Method

For every participant:

1. Introduction and 3 training cards
2. Shuffle 30 picture cards
3. Use the system to describe each card
4. RSVP starting at 700ms; adjustable at any time

Results: Construction Time



Overview of Results

- Average speed of last 5 utterances:
  - 70s (ND) vs. 107s (SMI)
- No nonsensical utterances
  - Average of 5 selections (verb + 4)
- RSVP speeds w/ positive motor response:
  - 700ms (ND) vs. 1200ms (SMI)

Summary of RSVP-iconCHAT

- Immediately applicable to mobile systems
  - Message complexity can be scaled (personalized)
- Exandable to multi-modal or analog input:
  - Push the switch harder to go faster
  - Directional switches
  - "Oops" functionality
- Involuntary responses (BCI) could leverage predictive reordering via sem-grams

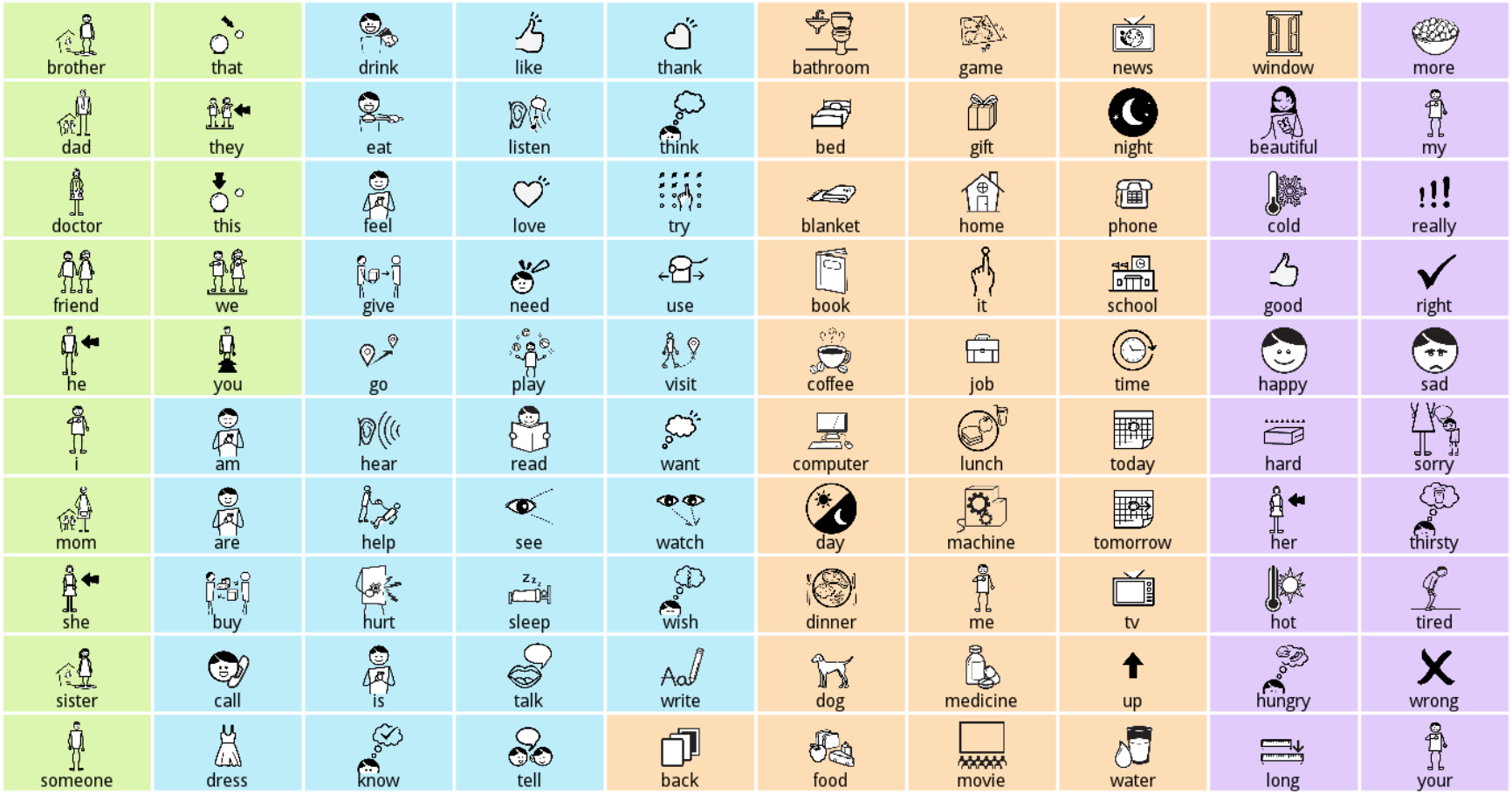
Applied Contributions

RSVP-iconCHAT ⟹ Free Order, Discrete Icons

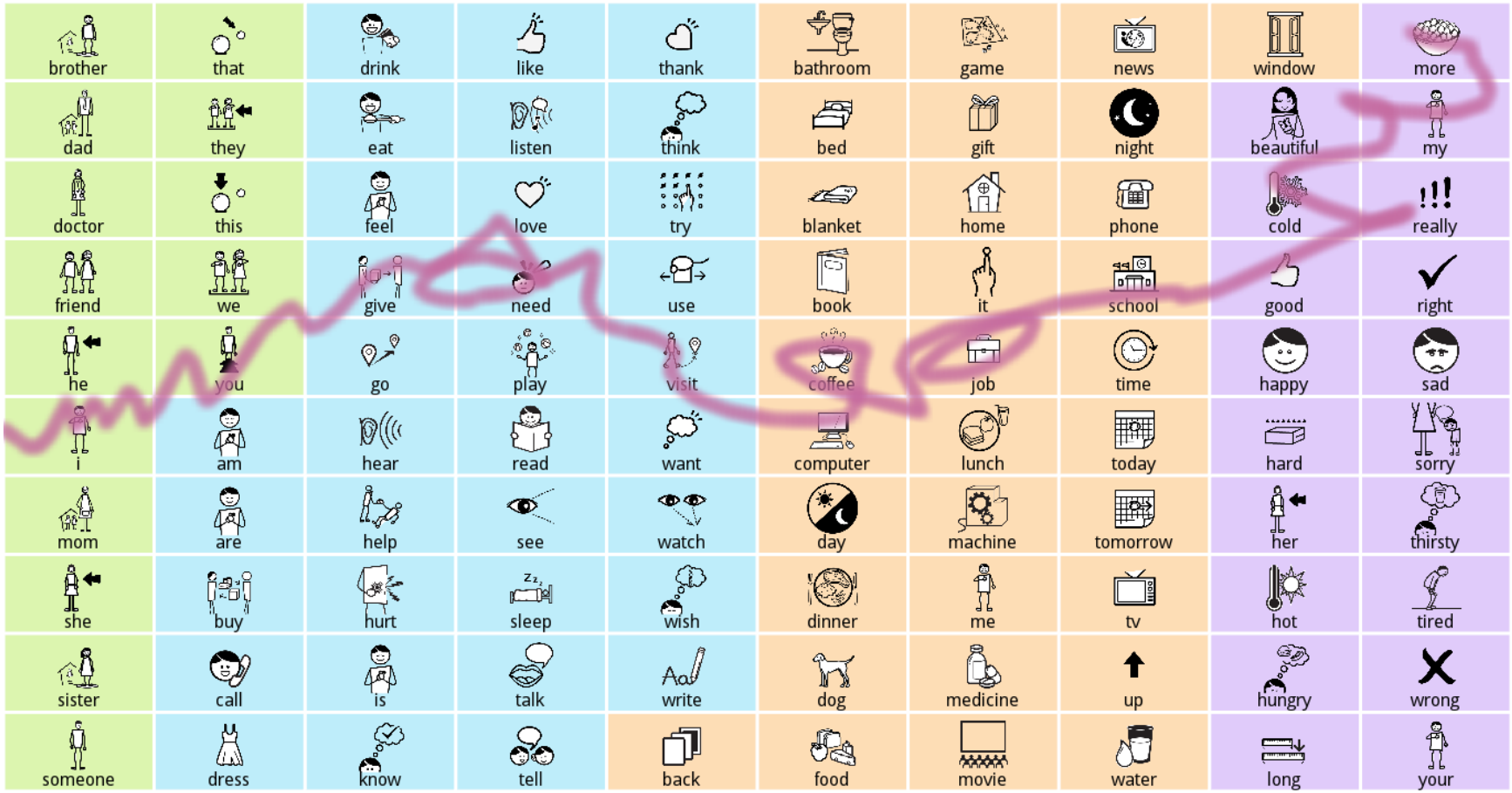
SymbolPath ⟹ Free Order, Continuous Icons

DigitCHAT ⟹ Mobile, Mixed-Input Letters

SymbolPath Motivation



SymbolPath



"I need more coffee"

Summary of SymbolPath

- Designed for people with upper limb motor impairments or developing literacy
- Semantic grams reweighted by path contour
- 75+ active users on Android

- Regular email feedback: "It's fun!"
  - Drawing and syntactic completion/generation encourages fuller utterances

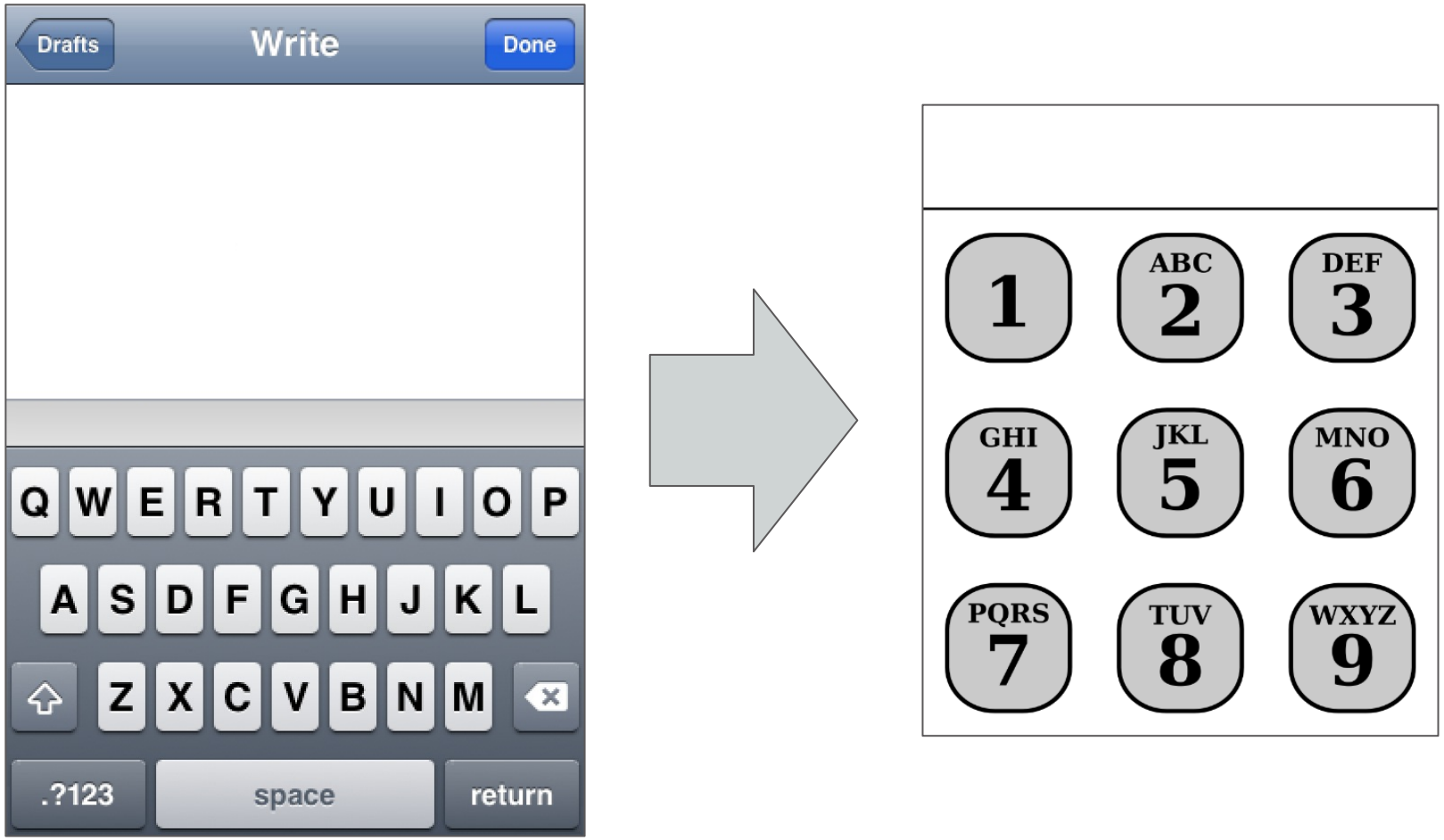
Applied Contributions

RSVP-iconCHAT ⟹ Free Order, Discrete Icons

SymbolPath ⟹ Free Order, Continuous Icons

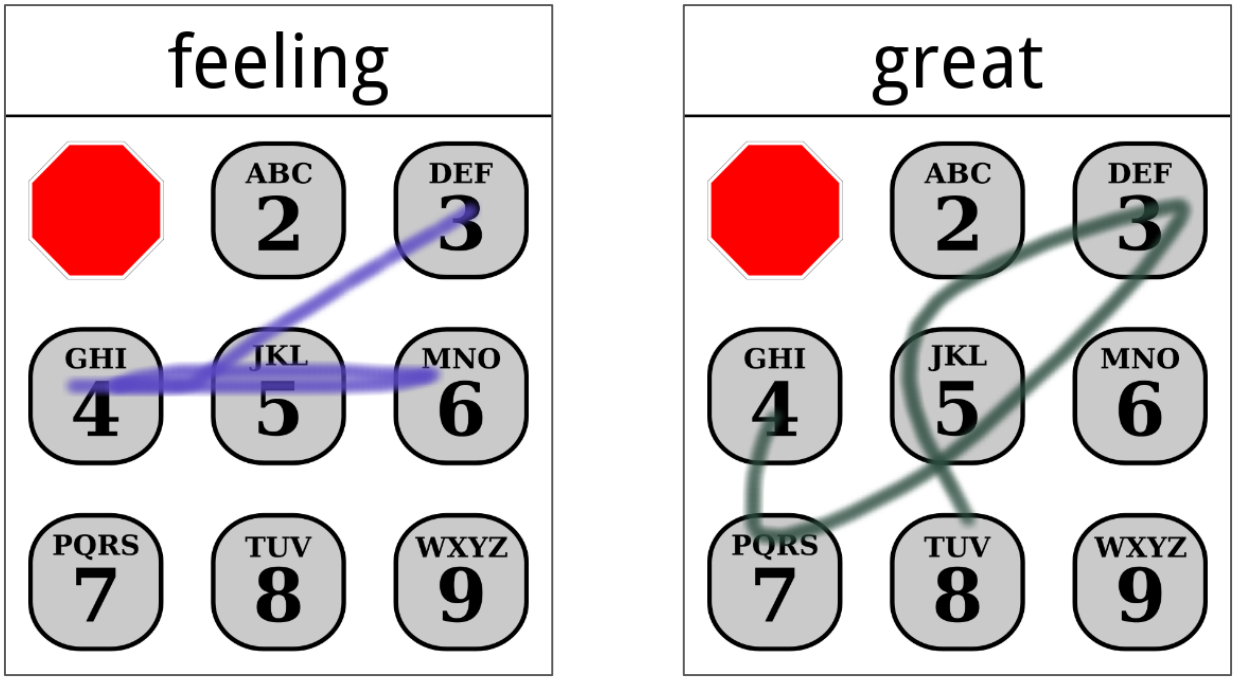
DigitCHAT ⟹ Mobile, Mixed-Input Letters

DigitCHAT Motivation



DigitCHAT

- Word-by-word, real-time construction
- Mixed-mode input and active learning

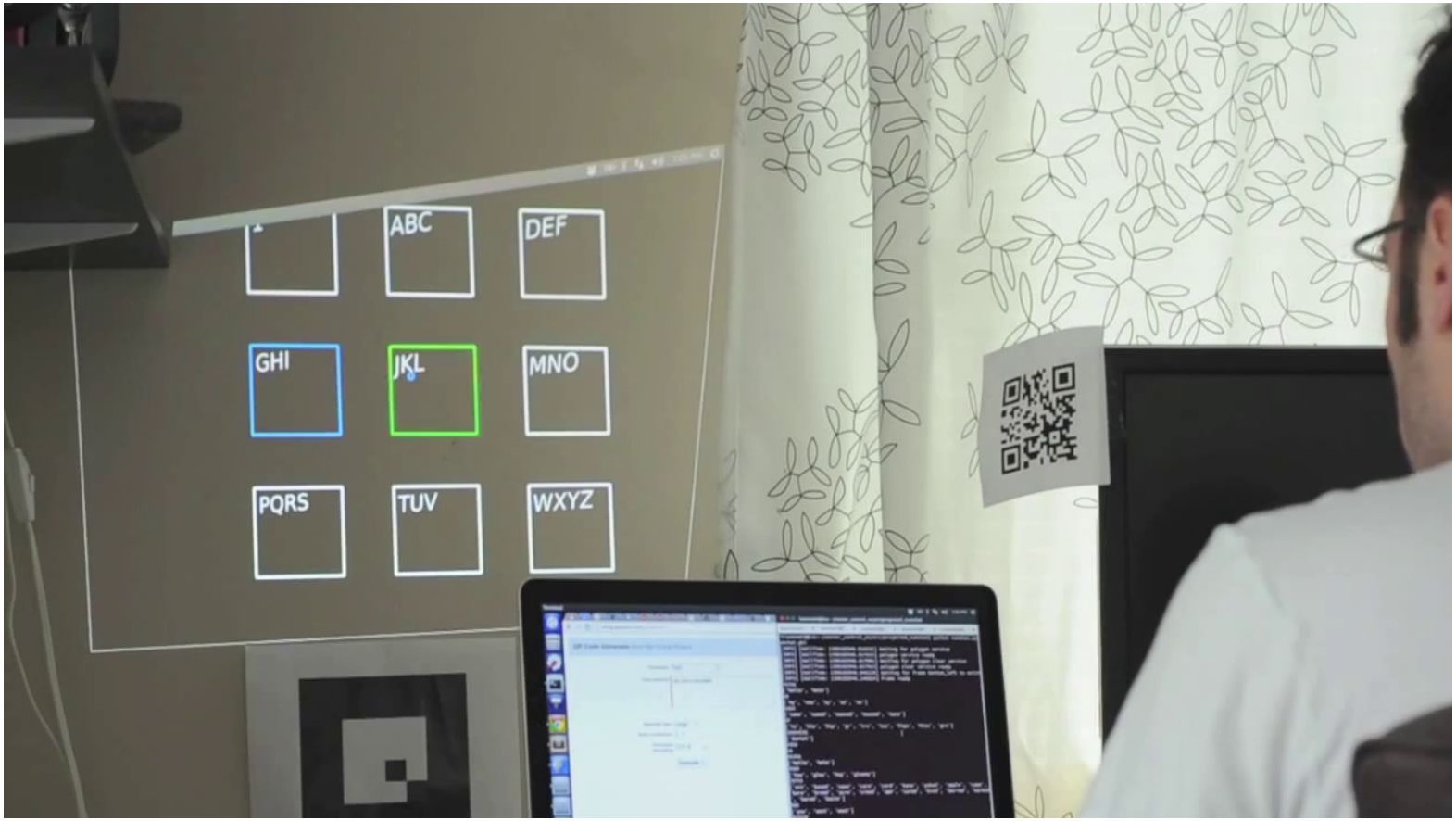


Summary of DigitCHAT

- Scalable and fast (> 45 WPM) [Silfverberg et al, 2000]
  - Compare to < 20 WPM for most AAC systems
- 15+ active users on Android

- Winner of the ACM ASSETS 2014 Text Entry Challenge

Projected DigitCHAT



Head-tracking prototype by Dan Lazewatsky and Bill Smart
(Oregon State University)

Part 4:

Summary and Conclusion

Thesis (Redux)

"Intelligent interfaces can mitigate the need for linguistically and motorically precise user input to enhance the ease and efficiency of assistive communication."

Theoretical Contributions

"...mitigate the need for linguistically and motorically precise user input..."

1. An unordered language model that bridges syntax and semantics. [Wiegand and Patel, 2012A]

2. An empirical comparison of contextual language predictors. [Wiegand and Patel, 2015B (R1)]

3. A motor movement study with current and potential AAC users. [Wiegand and Patel, 2015A]

Applied Contributions

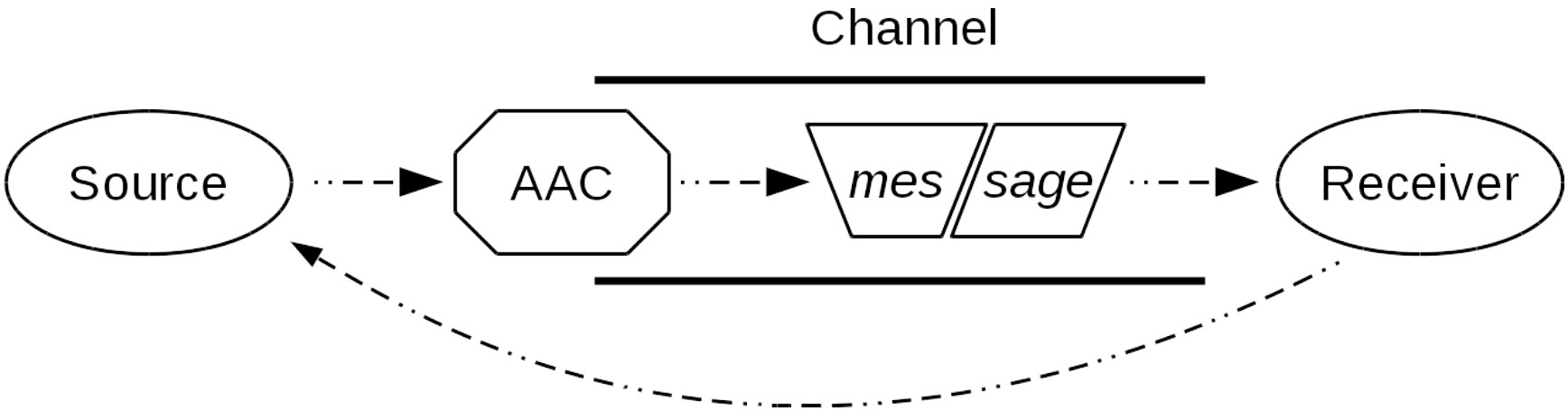
"...to enhance the ease and efficiency of assistive communication."

1. A semantic approach to icon-based, switch AAC. [Wiegand and Patel, 2014B]

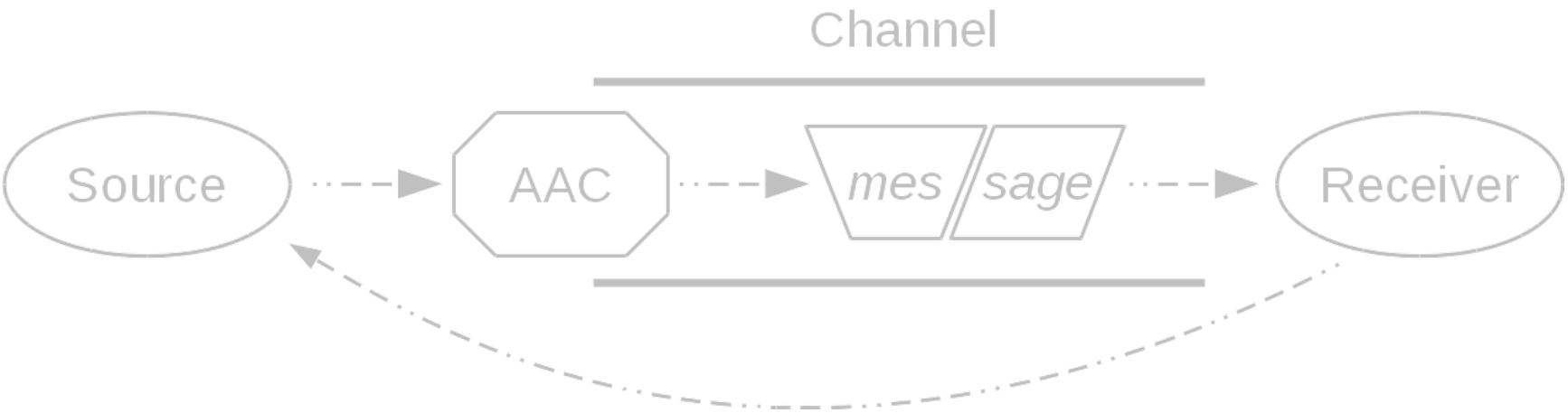
2. A continuous motion overlay module for icon-based AAC. [Wiegand and Patel, 2012B]

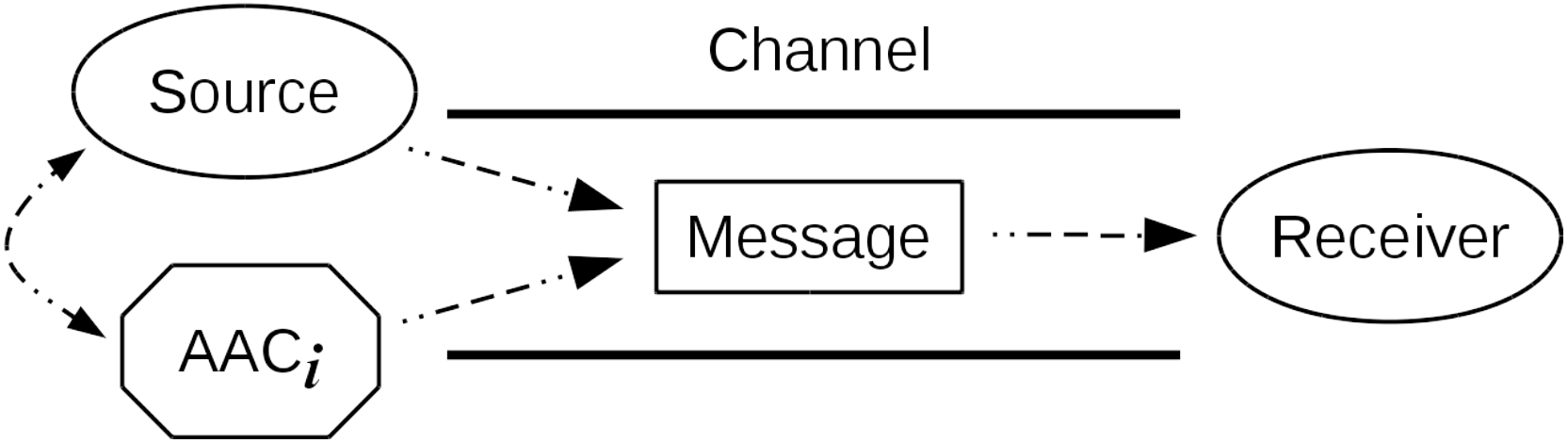
3. Mobile, letter-based AAC that supports conversational speeds. [Wiegand and Patel, 2014A]

Revisiting the Goal



Revisiting the Goal





Thank you for listening!

karlwiegand.com/defense

Special thanks to the Continuous Path Foundation and the National Science Foundation (Grants #HCC-0914808 and #SBE-0354378).

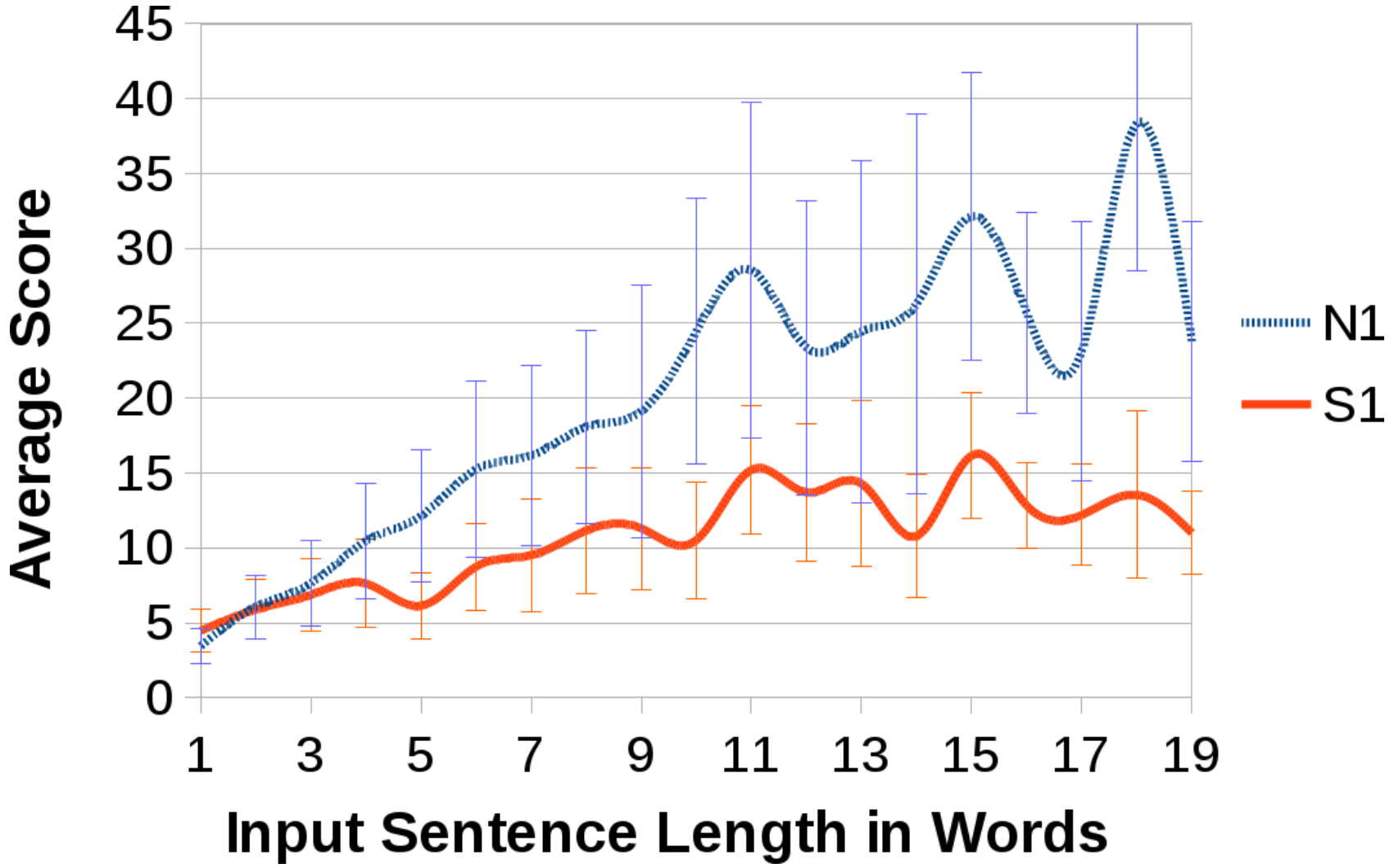
Sem-Grams: Method Details

- Test sentences truncated to 20 words
- All algorithms seeded with top 10 type-specific grams for each input word
- Maximum of 190 candidate words to rank
- Absence of target word in list was considered a "failure to predict"

Sem-Grams: Overview of Results

	N1	N2	S1	S2
# of Sentences	2000	2000	2000	2000
# Predicted	647	649	435	435
Average Score	16.26	19.70	9.04	12.67

Sem-Grams: Performance



Context: Method Details

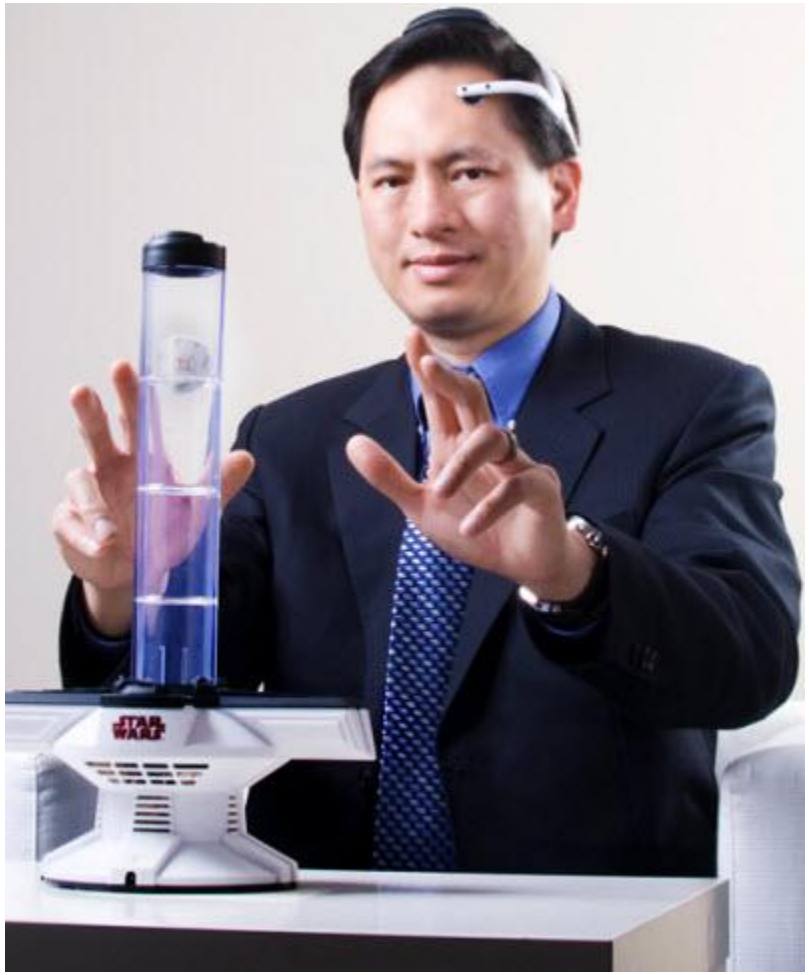
Predictor	Blog Authorship	Yelp
Age	26	-
Gender	2	-
Day of the Week (DOW)	7	7
Day of the Month (DOM)	31 (4)	31 (4)
Month	12	12
City	-	119
State	-	16

Average of 18 unique contexts per author in Blog Authorship and 4 in Yelp Dataset

MoGUI: Observations

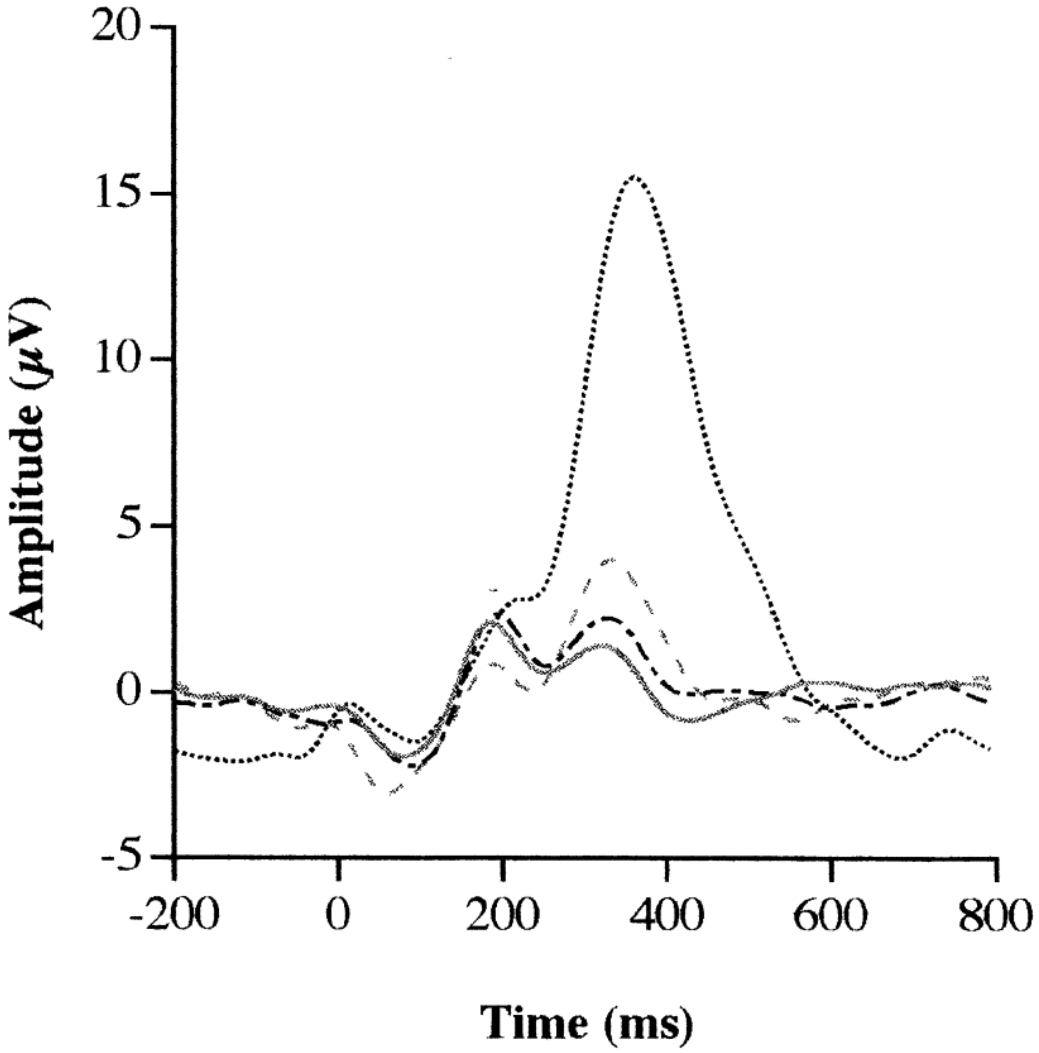
- Varied tablet and hand/arm positions
  - Tablet being held, flat/tilted on lap, on desk, tilted on table, held in wheelchair mount
  - Use of fingers, thumb, stylus, and knuckles
- Ghost tapping, spastic tapping, stylus friction, and finger humidity
- Repeated margin activation and triggering of Google Now functionality

Brain-Computer Interfaces (BCI)

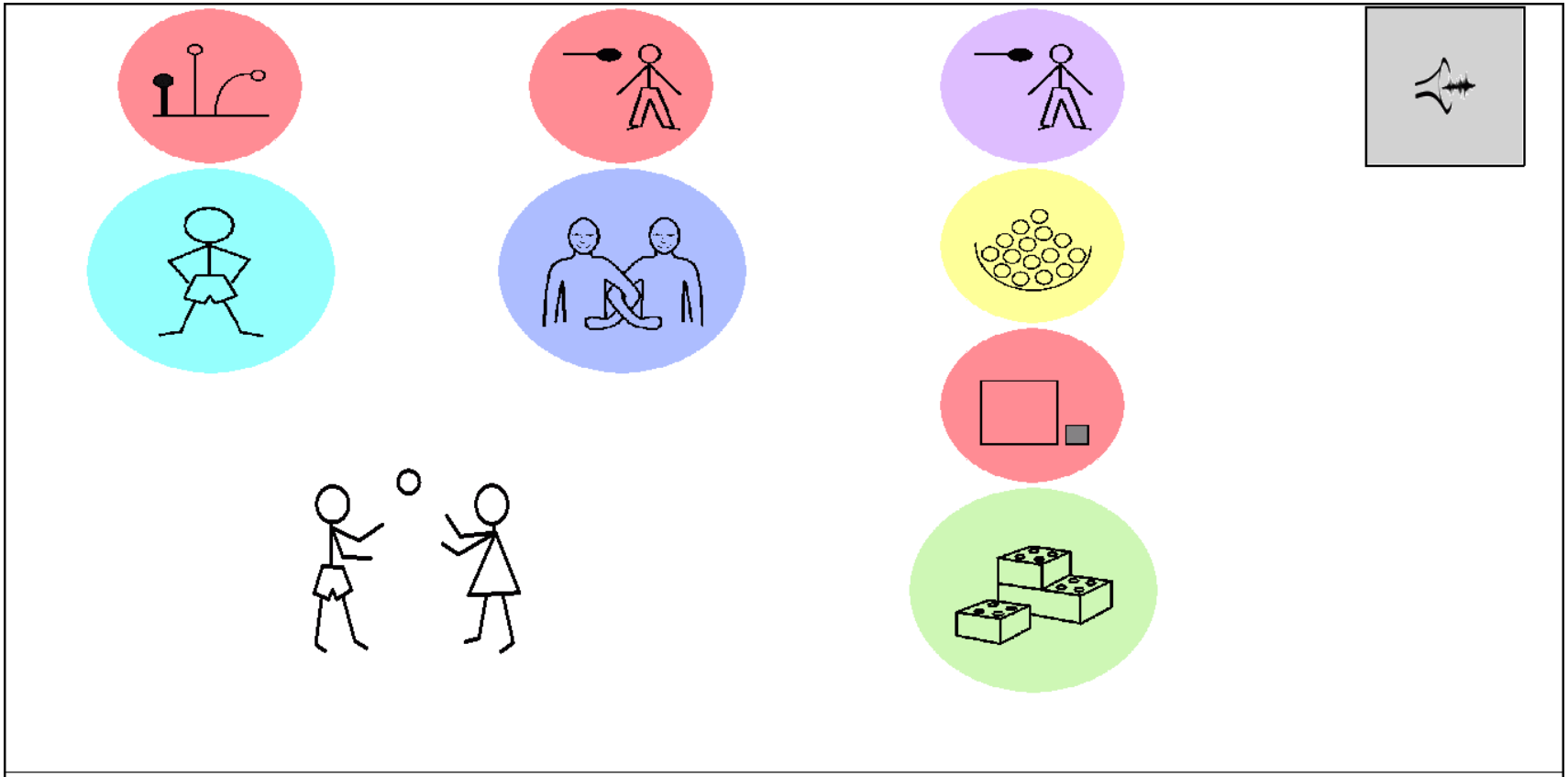


http://www.emotiv.com/
http://www.neurosky.com/

The P300 Wave

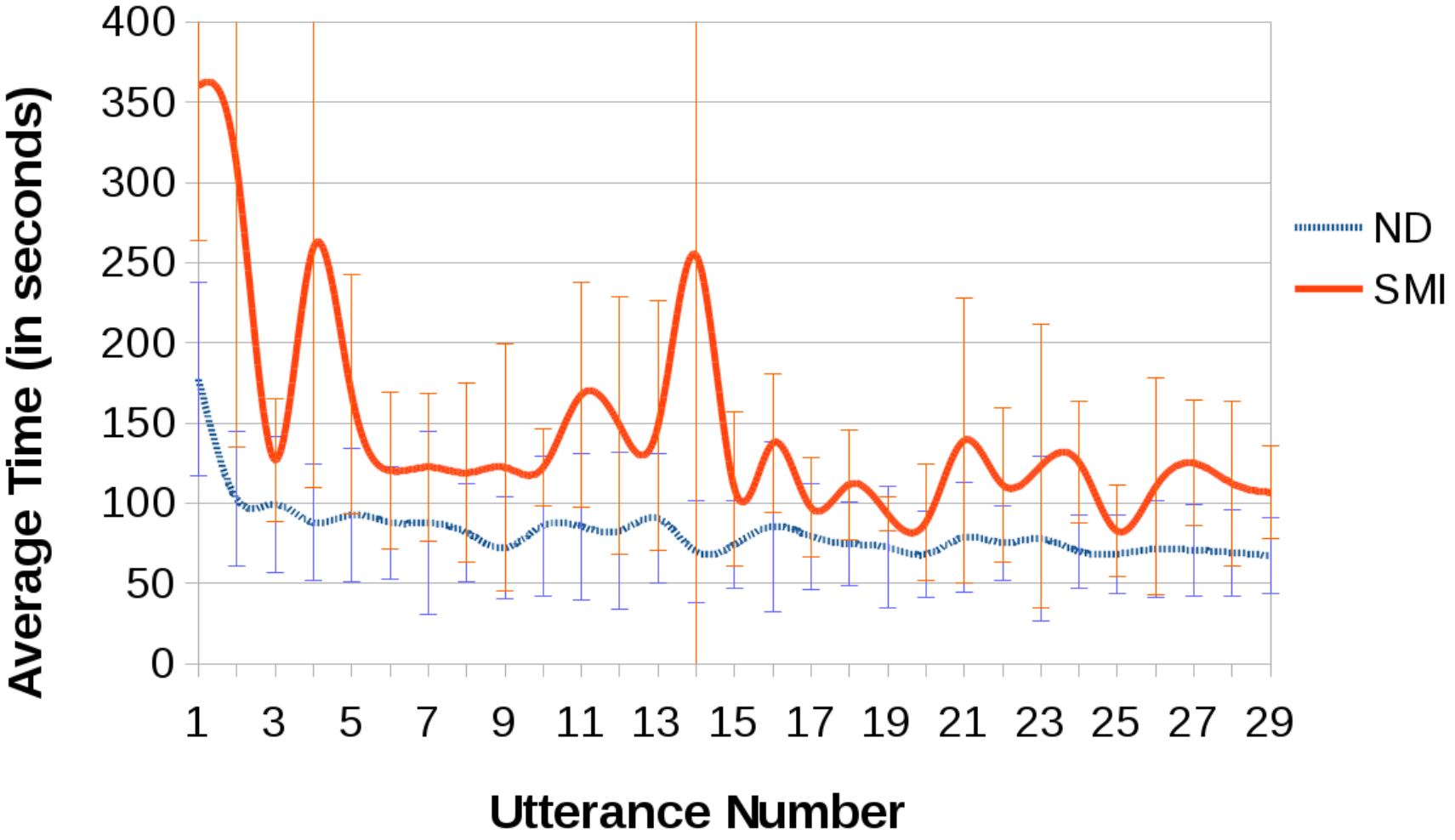


Complexity vs. Real Estate

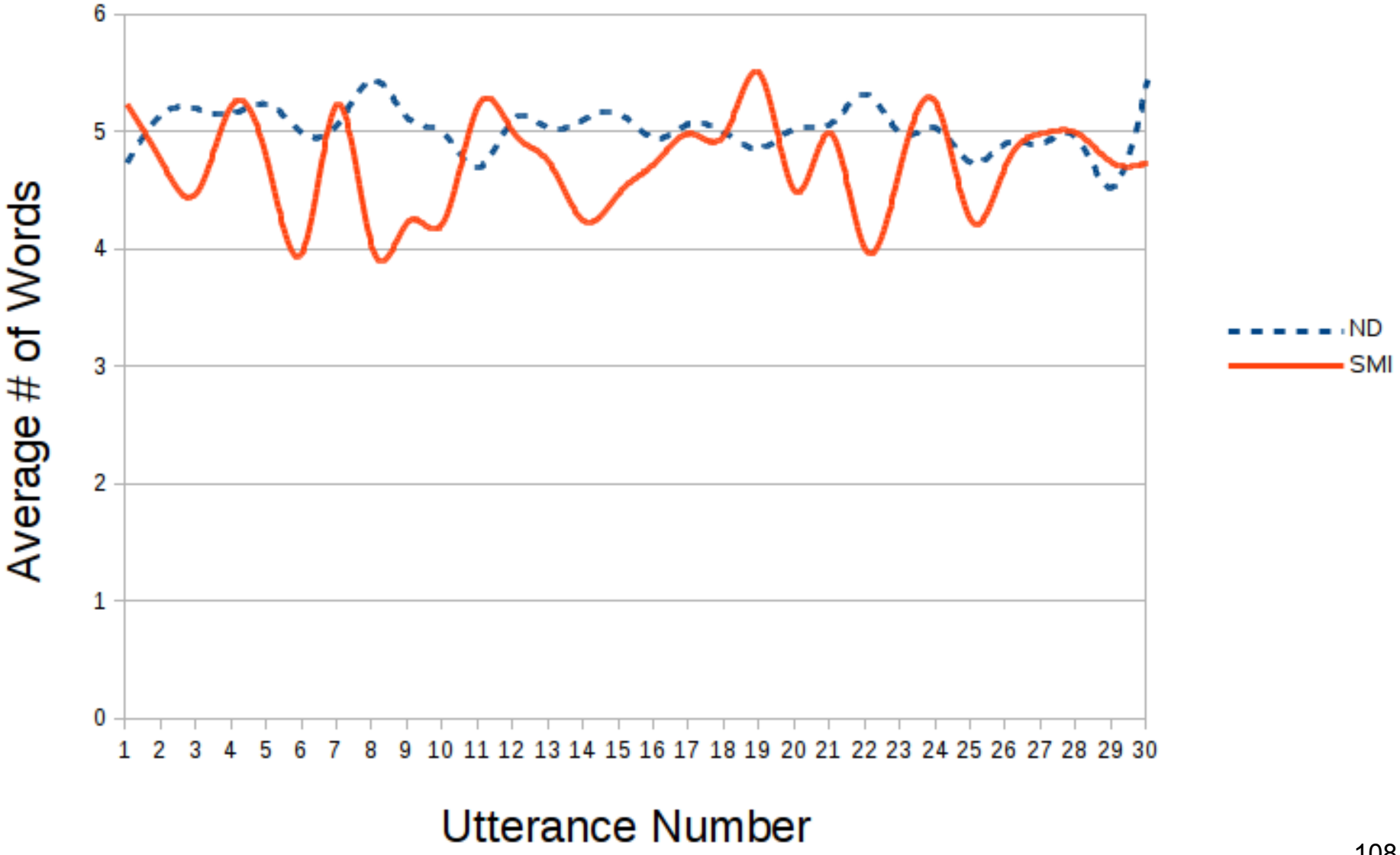


A young boy plays with his many small blocks with his friend.

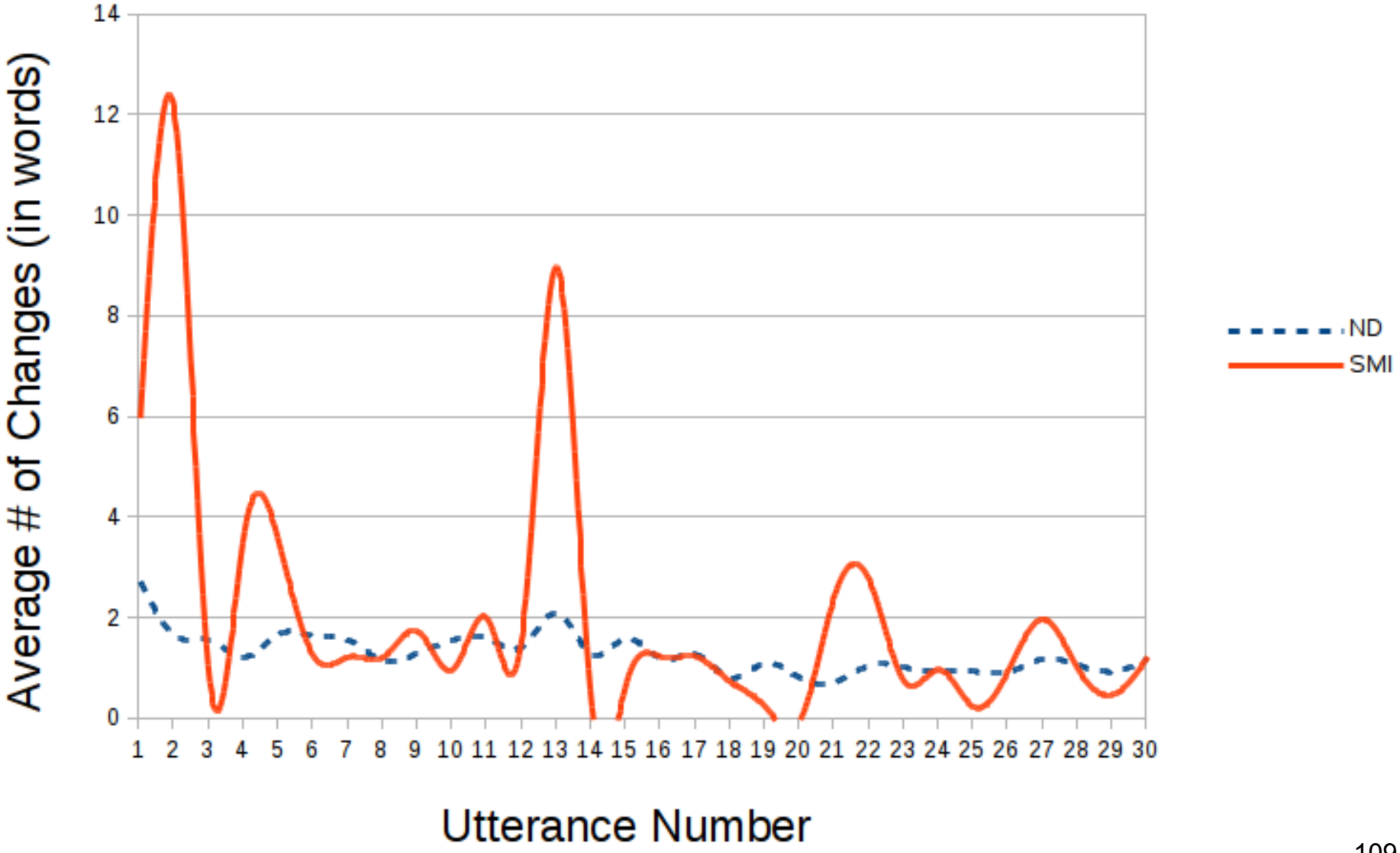
RSVP-iconCHAT: Construction Time



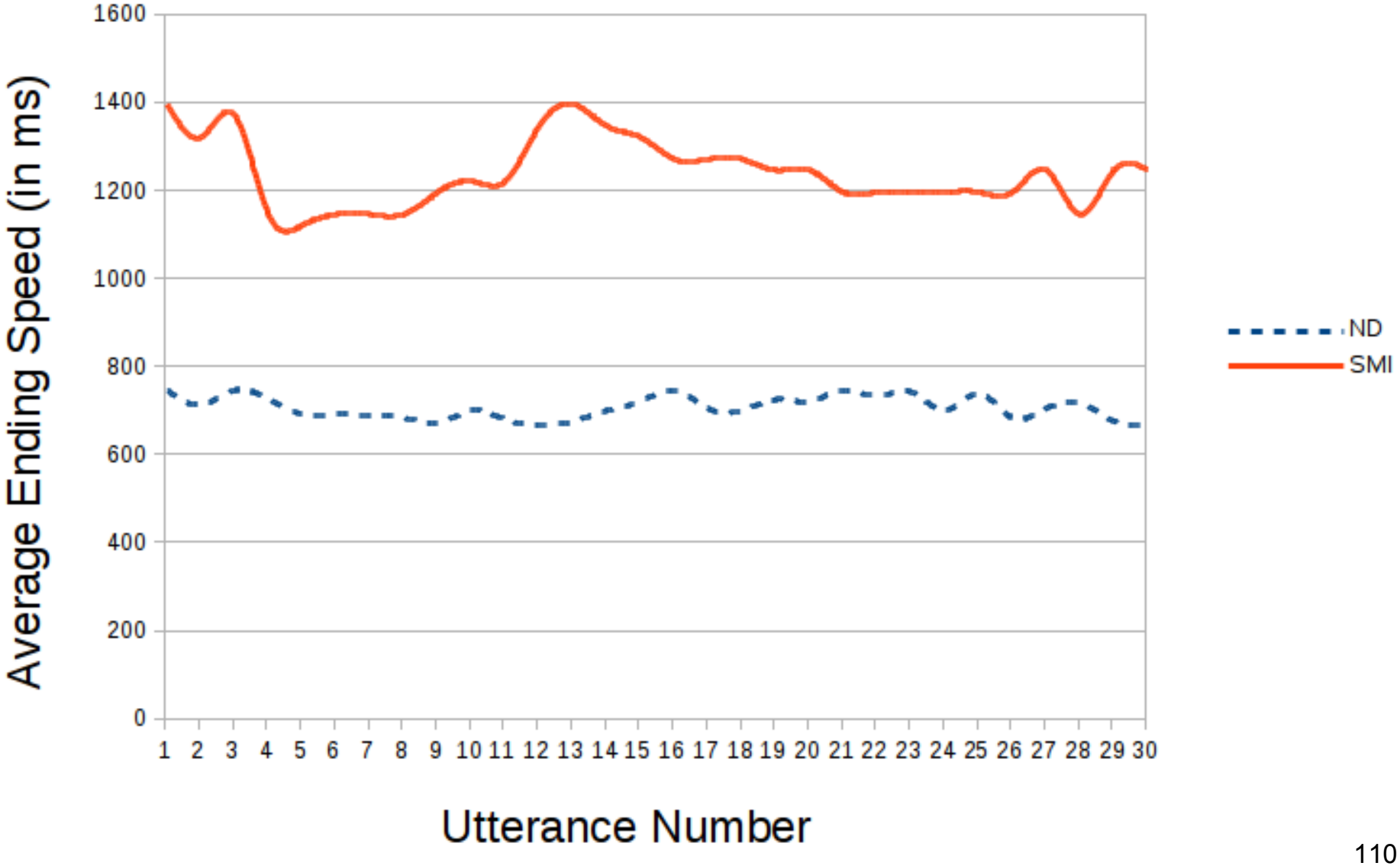
Utterance Complexity



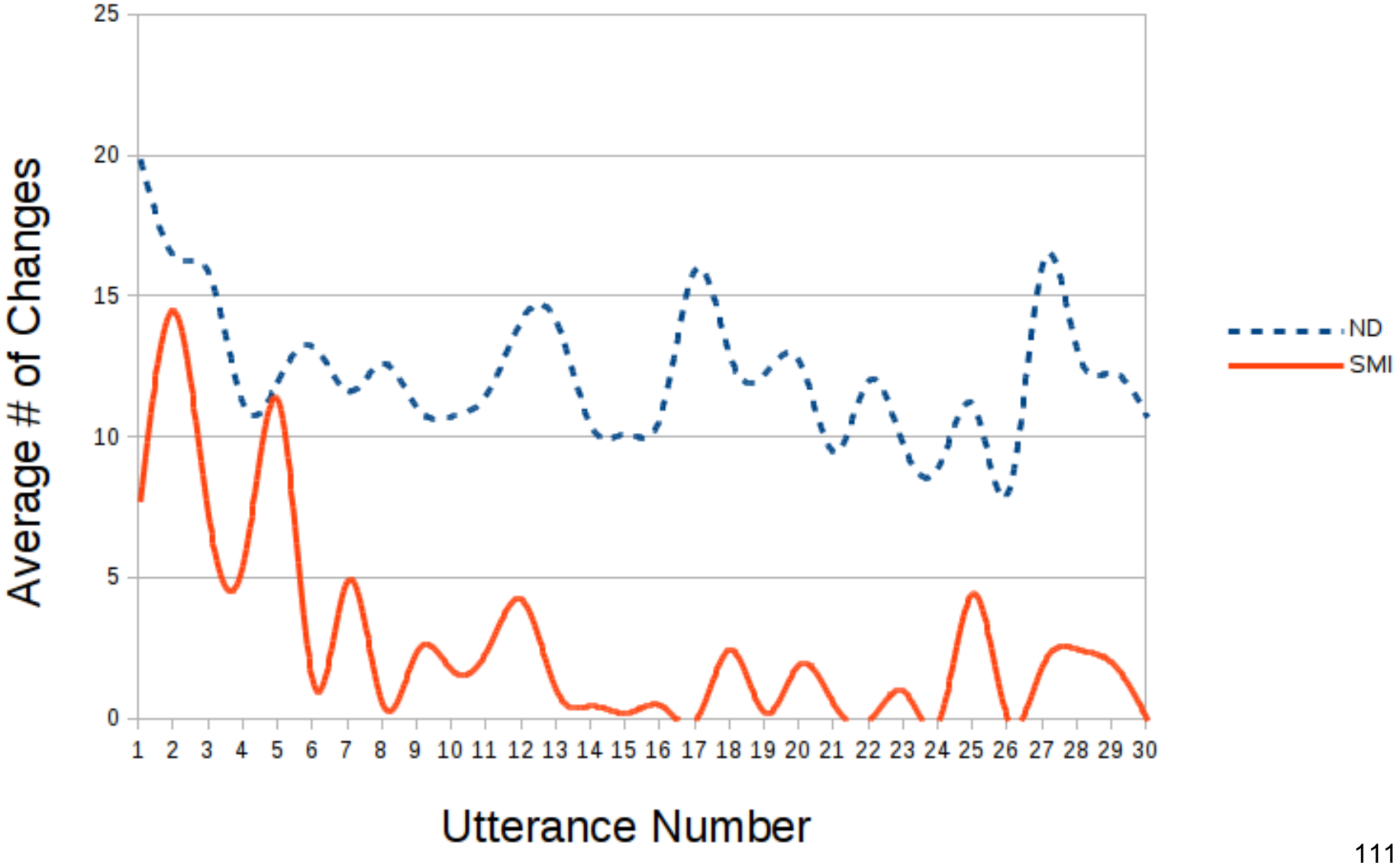
Errors and Modifications



RSVP Speed


Average Ending Speed (in ms)
1600
1400
1200
1000
800
600
400
200
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Utterance Number
ND
SMI

RSVP Speed Changes



RSVP-iconCHAT: Feedback

- All users get restless w/ alphabetic ordering
- Even alphabetic ordering can be surprising
- All users with SMI asked about other switches and multi-modal methods
- All users favorably mentioned the automatic syntax generation/modification

Information Operations Division
Code 561

"Leveraging the Information Environment"

- EW / ES (SIGINT)
  - Tactical Cryptologic Systems (TCS)
- Cyber
  - Bringing SIGINT to the Fight
  - CNO / CNE
- Military Deception / Military Information Support Operations
- Service Based Environment
  - Ozone Widget Framework (OWF)
  - APPS for the Warfighter
- Science & Technology (S&T)
  - Cognitive IO
  - Network exploitation
- Information Technology
  - System Engineering Support
  - Analysis
  - Data Fusion
  - Modeling and Simulation
- Dept of Homeland Security (DHS)