2aSC2. The role of speechreadability in the intelligibility of visual speech signals produced by Cued Speech transliterators

Jean C. Krause (jeankrause@usf.edu) and Abby N. Bennett
Department of Communication Sciences and Disorders, University of South Florida, Tampa, FL

-Background-

Cued Speech (Cornett, 1967)

- Visual-only speech communication system used by some deaf individuals
- Hand "cues" are produced in synchrony with mouth movements of speech
- Cues disambiguate phonemes confusable in speechreading (i.e. visemes)
  - Eight handshapes used to differentiate consonants
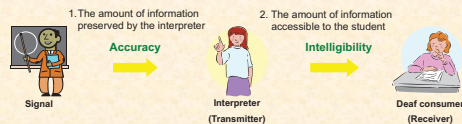 - Six placements used to differentiate vowels

Consonants

1. /d, p, ʒ/ deep azure
2. /ð, k, v, z/ the caves
3. /r, h, s/ rehearse
4. /b, m, n/ by when
5. /m, f, t/ miffed
6. /w, l, ʃ/ Welsh
7. /ŋ, g, dʒ/ thug Joe
8. /j, ŋ, tʃ/ young church

Vowels

Monophthongs: Mouth, Side — /i, ɚ, ɚ/ fir tree; /ʌ, ɔ, ɑ/ Aloha; Chin, Throat — /u, ɔ, ɪ/ too tall Ted; /u, æ, ɪ/ look at it

Diphthongs: Side-throat — /aɪ, aʊ/ time out; Chin-throat — /ɔɪ, eɪ/ Oy vay!

- When used correctly, Cued Speech allows for near-perfect reception of everyday connected speech (Uchanski et al., 1994)
→ Deaf people who use Cued Speech often access spoken information through a **Cued Speech transliterator**

-Accuracy vs. Intelligibility-

- Cued Speech transliterator = an "interpreter" who uses Cued Speech
- Little research to date measures transliterator (or interpreter) intelligibility
- **Cued Speech** transliterators are attractive candidates for initial study
- One-to-one correspondence between spoken and cued phonemes means that two aspects of the visual speech signal can be easily quantified:
  - **Accuracy:** proportion of signal correctly *transmitted* by a transliterator
  - **Intelligibility:** proportion of signal correctly *received* by a deaf consumer

1. The amount of information preserved by the interpreter — **Accuracy**
2. The amount of information accessible to the student — **Intelligibility**

Signal | Interpreter (Transmitter) | Deaf consumer (Receiver)

-Previous work-

### Intelligibility (Pelley & Krause, 2008)

Method

- **Twelve (12) Cued Speech transliterators (CSTs)** cued an 8th grade "lecture" (presented via audio recording)
- ~2700 phrases (~225 per CST) excised from transliterator videos
- Four stimulus blocks drawn from videos, such that
  1. Each transliterator was equally represented
  2. The entire lecture (~225 excised videos) could be presented in order, one phrase at a time
  3. The range of accuracy scores was as well-distributed between 0% and 100% as possible
- **Eight (8) "expert" Cued Speech receivers** viewed the stimulus blocks (expert = profoundly deaf with > 10 years of Cued Speech experience)
  - Stimulus items presented **one phrase at a time** (no repetitions)
  - Receiver typed response **verbatim**

Results

1. **Average intelligibility: 75%** (across all receivers and CSTs)
   - 82% when restricted to key words presented by CST
   - Large variation across transliterators (38 points)

Transliterator	Average Intelligibility (All words, original signal)
Lowest: CST10	52%
Highest: CST11	90%
Overall	**75%**

2. **Accuracy accounted for 26% of the variance in intelligibility** of individual stimulus items (p < 0.001)
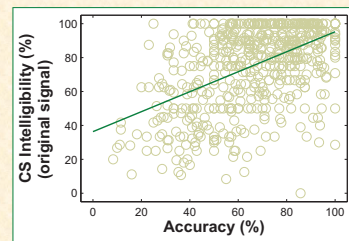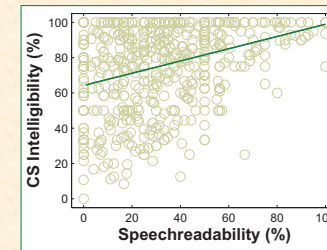

→ *Large portion of the variance remains unexplained*

-Role of speechreadability-
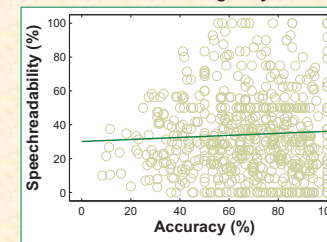
Motivation

- One factor that is likely to affect intelligibility is speechreadability
  - Mouth movements are necessary in Cued Speech
  - Talkers vary in speechreadability (Kricos & Lesner, 1982)
→ In messages produced by Cued Speech transliterators, **how does intelligibility** (for Deaf CS users) **vary with speechreadability** (for persons unfamiliar with Cued Speech)?

Participants

- Eight (8) individuals with no prior Cued Speech experience
  - Familiar with speechreading (recruited from interpreter training program)
  - Exposed to English before age 10
  - Dominant language was English for at least 10 years
  - Passed speechreading screening (>60% on Sargent Lipreading Test))

Materials

- Stimuli identical to earlier intelligibility study (Pelley & Krause, 2008)
- 12 CSTs asked to transliterate a high-school level educational film (*Life Cycle of Plants* by Films Media Group, 1989)
- Videos at **slow-conversational rate** (~88 wpm) segmented at phrase boundaries
- Four stimulus blocks drawn from videos, such that
  1. Each transliterator was equally represented
  2. The entire lecture (~225 excised videos) could be presented in order, one phrase at a time
  3. The range of accuracy scores was as well-distributed between 0% and 100% as possible

Presentation sessions

- Participants tested individually in two 2-hour presentation sessions
- Stimulus items presented **one phrase at a time**
  - Each item presented **twice consecutively** on a computer monitor
  - Stimuli periodically interspersed with excerpts from the original film, presented for context
- Participant typed response **verbatim**

Overall Results

- **Average speechreadability: 34%** (across all participants and CSTs)
  - Similar speechreadability (36%) for transliterator's key words
  - Individual participant scores ranged from 24% to 39%
  - Relatively large variation across transliterators (22 points)

Transliterator	Average Speechreadability (All words, original signal)
Lowest: CST10, CST12	27%
Highest: CST11	49%
Overall	**34%**

Relationship to intelligibility

1. **Speechreadability accounted for 13% of the variance in intelligibility** of individual stimulus items (p < 0.001)

2. **Speechreadability was uncorrelated with accuracy (p = 0.75), indicating that these two factors predict largely independent portions of the variance in intelligibility**

→ *Together, speechreadability and accuracy account for 37% of the variance in intelligibility* of individual stimulus items

Stepwise multiple regression:

Independent Variable	R^2	Delta R^2	p
Accuracy	0.26	0.26	< 0.001
Speechreadability	0.37	0.11	< 0.001

-Conclusions-

- Accounting for 26% of the variance, **accuracy has a greater contribution to intelligibility than speechreadability**
- However, **speechreadability also plays a sizeable role**
  - Accounts for 13% of the variance in this experiment
  - May account for more if measurements can be refined
    - Many reception errors are likely to be partially correct (e.g. *light* for *life*)
    - Other errors may reflect correct visemes (e.g. *trauma* for *drama*)
    - No partial credit was awarded
- Moreover, the **relative contributions of speechreadability and accuracy are largely independent**
  - CSTs with higher speechreadability are more intelligible than would be predicted from accuracy alone (and vice versa)
  - CSTs with lower speechreadability are less intelligible than would be predicted from accuracy alone
- As a result, **speechreadability and accuracy together account for 37% of the variance in intelligibility**
- **More than half of variance still unexplained**, which suggests still other sources of variance
- Sources of transliterator variability are of primary interest (could lead to improved training and evaluation methods)
- **Transliterator factors that are likely to affect intelligibility include...**
  - "Speaking" rate
  - Visual prosody
  - Facial expressions and non-manual behaviors
  - Cueing style: clear vs. conversational and highly co-articulated

-Future Work-

- Assess speechreadability data at phoneme-level (and viseme-level)
  - Example: *light* for *life*   Word-level: 0% vs. Phoneme-level: 67%
- Isolate and analyze other transliterator factors, such as rate, that may also affect intelligibility
- Extend experiments to other visual communication options used by deaf individuals: Signing Exact English, American Sign Language, etc.
- Compare results across communication options in order to
  - Increase understanding of intelligibility of visual signals
  - Gain insight into modality-independent aspects of perception

-Acknowledgments-