Sound quality impacts the speed and effort of sentence perception

i.e. Rapid reduction of listening effort from predictive language processing

Matthew Winn University of Washington

UNIVERSITY of WASHINGTON Listen Lab

SPEECH AND HEARING SCIENCES

American Auditory Society 2016 #54

INTRODUCTION

Speech communication
is more than correctly identifying words.

As we perceive stimuli,
we *predict* what comes next [1,2]

In Speech Perception,
prediction can be driven by:

Acoustics
Coarticulation lets you predict upcoming sounds [3,4]

Knowledge of a talker
You have ideas about what a talker should sound like [5]

Context
Words are easier to recognize when preceded by relevant context [6,7]

Context helps us predict and understand what we are *about to hear*

"**Sweep the floor** with a *broom*"

"Nicole thought about a *broom*"

High-context sentences are more intelligible than **low-context** sentences

Context facilitates better and faster word recognition

Questions in this study:

1. Does context reduce *listening effort*?
   (i.e. do you get "effort release" from context?)
2. If so, how quickly does it occur?
3. Can people with CIs benefit from context as quickly and as effectively as people with normal hearing?
   (i.e. does spectral degradation interfere with effort release?)

METHODS

PARTICIPANTS: 21 young listeners with normal hearing (ages 19 – 32 y)
12 listeners with cochlear implants (ages 40– 67 y)

STIMULI: Revised speech-in-noise (R-SPiN) sentence lists [6]
Each list contains 25 high-context and 25 low-context sentences.

SPECTRAL RESOLUTION:
Four testing blocks alternated in **sound quality** between normal (clear) speech and degraded (8-channel vocoded) speech.

PROCEDURE:
Listen to & repeat sentences while fixating on a monitor
(3 s silence) Stimulus (2 s silence) Response
"Stir your coffee with a spoon"

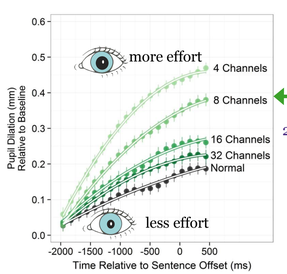
MEASUREMENT OF LISTENING EFFORT:
High-speed eye tracking was used to measure pupil dilation during each trial.
Greater pupil dilation indicates increased listening effort [8, 9]

more effort

less effort

WHY MEASURE PUPIL DILATION?

Poor spectral resolution increases overall listening effort [10; below].
This approach lets us observe the growth of effort *during* the perceptual process.



1. PREVIOUS results using noise vocoded speech showed that poor spectral resolution leads to greater overall growth of pupil dilation.
2. The CURRENT study sets out to explore whether spectral resolution has effects on the speed of processing the content of the message,

   which would show up as changes in changes in the speed and timing of pupil dilation, as a function of the message content and listening condition.

RESULTS: Semantic context reduces listening effort rapidly…

but only if signal quality is clear

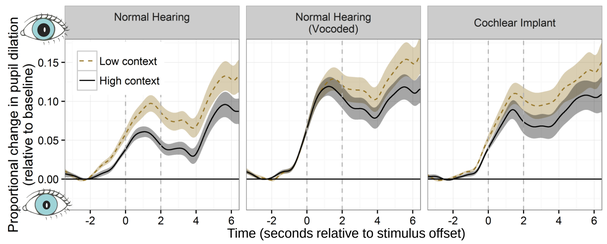


Fig.1: *Growth of pupil dilation in response to low- and high-context sentences*

Less effort is exerted for sentences with context

The magnitude and timeline of effort varies across listener groups

Growth curve analysis was implemented using a cubic polynomial fit

Slope term reflects growth of pupil dilation within each time window

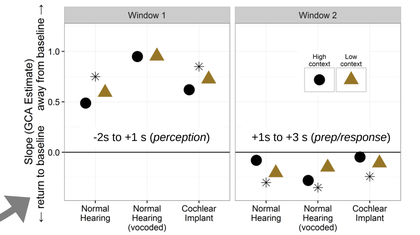


Fig. 3: *Slope of pupil size change over time for two time windows labeled in Fig 1. Asterisks indicate significant difference between slopes for low-context and high-context conditions*

How do we quantify benefit of context on listening effort?

Effort Release:
The difference between

Response for Low-context

and

Response for High-context

Asymptote, slope and latency were fit with a non-linear least-squares procedure using this three-parameter sigmoid function:

$$\text{Percent reduction} = \frac{(\text{max} * \text{group})}{1 + e^{-(\text{Time} + \text{shift} * \text{group}) * (\text{slope} * \text{group}))}}$$

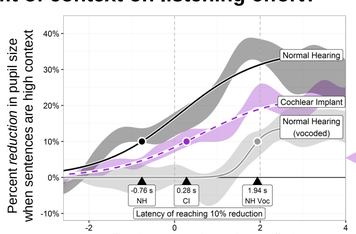


Fig. 2: *Reduction of pupil dilation when sentences were high context (difference between curves in Fig 1)*

Compared to NH listeners, we observed:
* ***Significant lower asymptote (less total effort release) for CI listeners***
* ***Significant shallower slope (slower effort release) for CI listeners***
* ***Significantly longer latency (slower onset of effort release) for CI listeners***

These effects were tested against a significance criterion of $p < 0.05$;
Stronger effects of the same direction were observed for the NH vocoder condition

Signal degradation (via cochlear implant or vocoder) delays the reduction of effort obtained from context/prediction

(see latency to obtain 10% reduction)

Benefits of context for reducing effort can be reduced and delayed even when intelligibility scores are high.

Results persist even if intelligibility is perfect

Perception can be "restored" after the sentence is heard, but the pupillary response can indicates that context was not exploited rapidly.

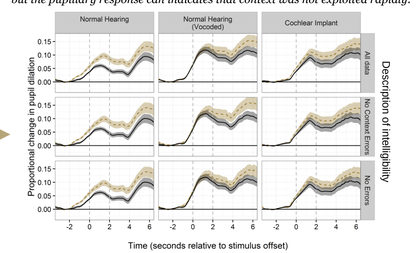


Fig. 4: *Growth of pupil dilation in response to low- and high-context sentences, separated by intelligibility performance. "Context error" is misperception of any word before the final word.*

INTELLIGIBILITY:
(scored by hand during testing)

	NH		NH vocoded		CI	
	high-context	low-context	high-context	low-context	high-context	low-context
% of sentences that contained errors on non-target ("context") words	2	2	16	22	9	15
% of sentences that contained errors on "target" words	3	3	11	38	8	32
% of target word errors preceded by "context" errors	1	1	8	9	5	6

CONCLUSIONS

- Semantic context reduces listening effort (Fig 1)
- Effort reduction from context is *rapid* for NH listeners, and *delayed (by ~ 1 second)* for CI listeners (Fig 2) & NH listeners hearing vocoded speech (Figs 2, 3)
- Delays in effort release are observed even when intelligibility is perfect (Fig 4)
- **Implications**: Benefit of context might occur only *after* a sentence has been heard, but still lead to good intelligibility in the clinic/lab,
  but in conversational speech, we don't have lengthy silent pauses after sentences for listeners to catch up and recover context;
  a brief delay in processing might cause interference between the last sentence and the next sentence
- **Methods**: Time-varying physiological measures (such as pupillometry) can capture the temporal dynamics of listening effort as it unfolds *during* the perceptual process.

Philosophy of outcome measures: Speech perception is more than just recognition of isolated units like syllables and words:
Poor signal quality can cause disruption in the *ongoing process* of prediction and restoration of words.

REFERENCES

[1] Clarke (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences.*

[2] Bar (2007) The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences.*

[3] Gow (2002). Does English coronal place assimilation create lexical ambiguity? *J Exp Psych: Human Perception & Performance*

[4] Mahr, McMillan, Saffran, Ellis-Weismer, & Edwards (2015). Anticipatory coarticulation facilitates word recognition in toddlers. *Cognition.*

[5] Winn, Rhone, Chatterjee, & Idsardi (2013) The use of auditory and visual context in speech perception by listeners with normal hearing and listeners with cochlear implants. *Frontiers in Psychology*

[6] Bilger, Neutzel, Rabinowitz, & Rzeczkowski (1984). Standardization of a test of speech perception in noise. *JSLHR.*

[7] Pichora-Fuller, Schneider, Daneman (1995). How young and old adults listen to and remember speech in noise. *JASA.*

[8] Beatty (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol Bulletin.*

[9] Zekveld, Kramer, Festen (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear & Hearing.*

[10] Winn, Edwards, Litovsky (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear & Hearing.*

[11] Mirman (2014). *Growth Curve Analysis and Visualization Using R.* New York, NY: CRC Press.

Scan to download

Email: mwinn2@uw.edu

Alan Kan provided assistance in programming the EyeLink; Brianna Vandyke assisted with data collection.
Financial Support provided by NIH-NIDCD 5R01 DC03083 (Litovsky), the UW-Madison Department of Surgery, and the NIH Loan Repayment Program