Application of Ideal Binary Masking to Disordered Speech

NCT05244603

12/3/2019

Study Protocol

Dysarthria and hearing loss are communication disorders that can substantially reduce intelligibility of speech and the addition of background noise adds a further challenge. This proposal utilizes an established signal processing technique, currently exploited for improved understanding of speech in noise for listeners with hearing loss, to investigate its potential application to overcome speech-in-noise difficulties for listeners understanding dysarthric speech. Successful completion of this project will demonstrate proof-of-concept for the application of this signal processing technique to dysarthric speech in noise, and inform the development of an R01 proposal to perform a large-scale evaluation of the technology, and clinically meaningful implications, in a broad range of disordered speech types and severities.

Dysarthric and healthy control phrases will be processed in six different experimental conditions. Conditions 1 and 2 will consist of the unprocessed phrases representing dysarthric (condition 1) and control speech (condition 2) in quiet. Conditions 3 and 4 will consist of the phrases representing dysarthric (condition 3) and control speech (condition 4) mixed with noise at a given signal-to-noise ratio (SNR). For a masker, a cafeteria noise (Auditec of St. Louis) will be used to represent a relatively natural listening background. Conditions 5 and 6 will consist of the same phrases representing dysarthric (condition 5) and control speech (condition 6) mixed with cafeteria noise as in conditions 3 and 4, but processed by the IBM. Speech stimuli will be processed by custom MATLAB scripts and will be generated using high-quality D/A converters and a PC.

All participants will be seated in a sound treated booth and fitted with *Sennheiser* HD280 Pro circumaural headphones. Hearing-impaired participants will not wear their hearing aids during testing, but audibility will be ensured through the spectral shaping via a digital equalizer. Participants will be told that they will be presented with phrases that will be difficult to understand, either because the phrases will be produced by a person with a speech disorder and/or background noise, and that their job is to listen carefully and try to understand what is being said. The experimenter will present the phrases, one at a time, and following each phrase, the participants will be asked to repeat back what they think they heard. The experimenter will type out the participant's verbal responses. This method was selected over more traditional "hear and type" methods to mitigate possible issues with mobility and typing competencies, particularly applicable to the older listeners. Responses will be scored for words correct, using previously established scoring criteria for the semantically anomalous phrases and a validated in-house computer program.

Multiple regression analysis will be used to test for differences in PWC scores across the conditions. Each model will control for individual age, gender, and PWC scores in quiet—thereby controlling for baseline perception capabilities.

Informed Consent

Application of Ideal Binary Masking to disordered speech

Introduction

You are invited to participate in a research study conducted by Sarah Yoho Leopold an Assistant Professor in the Department of Communicative Disorders and Deaf Education at Utah State University. The purpose of this research is to understand how people understand disordered speech in background noise, and whether or not this understanding can be improved. Your participation is entirely voluntary.

Procedures

Your participation will involve listening to speech over headphones and transcribing what was said. The study will last approximately 30 minutes. We anticipate that 400 people will participate in this research study.

Risks

This is a minimal risk research study. That means that the risks of participating are no more likely or serious than those you encounter in everyday activities. The foreseeable risks or discomforts include fatigue due to the length of testing (30 minutes). In order to minimize those risks and discomforts, you make take short breaks as needed.

Benefits

There is no direct benefit to you for participating in this research study. More broadly, this study will help the researchers learn more about hearing and speech understanding and may help future individuals with communication difficulties through a deeper understanding of how speech understanding can be improved for this population.

Confidentiality

The researchers will make every effort to ensure that the information you provide as part of this study remains confidential. Your participation is anonymous. Your identity will not be revealed in any publications, presentations, or reports resulting from this research study.

We will collect your information through assessments and questionnaires. Online activities always carry a risk of a data breach, but we will use systems and processes that minimize breach opportunities. Data will be securely stored in a restricted-access folder on Box.com, an encrypted, cloud-based storage system and/or in a locked drawer in a restricted-access laboratory. No identifiable data will be collected or stored at any time.

Voluntary Participation & Withdrawal

Your participation in this research is completely voluntary. If you do not complete the study, any data we already have from you will be destroyed as soon as possible.

Compensation

For your participation in this research study, you will receive $4. You will also receive a bonus of $1 if you demonstrate engagement in the task by responding the requested tasks. If you choose to withdraw before completion of the study, you will not receive compensation.

Findings & Future Participation

Identifiers may be removed from your data. These de-identified data may be used or distributed for future research without additional consent from you. If you do not wish for us to use your data in this way, please state so below.

IRB Review

The Institutional Review Board (IRB) for the protection of human research participants at Utah State University has reviewed and approved this study. If you have questions about the research study itself, please contact the Principal Investigator at 435-797-9239 or sarah.leopold@usu.edu. If you have questions about your rights or would simply like to speak with someone *other* than the research team about questions or concerns, please contact the IRB Director at (435) 797-0567 or irb@usu.edu.

Sarah Yoho Leopold, PhD
Principal Investigator
(435) 797-9239; sarah.leopold@usu.edu

Stephanie Borrie, PhD
Co-Investigator
(435) 797-1388; stephanie.borrie@usu.edu

Informed Consent

By continuing with this study, you agree to participate in this study. You indicate that you understand the risks and benefits of participation, and that you know what you will be asked to do. You also agree that you have asked any questions you might have and are clear on how to stop your participation in the study if you choose to do so. Please be sure to retain a copy of this form for your records.

Statistical Design and Power

The outcome measure for the current study is percent words correct (PWC). To test whether there is a difference in PWC between the experimental conditions, linear mixed effects regression analysis will be used. Each mixed effects model will control for covariates including individual age, gender, MOCA score, audiometric thresholds, frequency-specific audibility of stimuli, and QuickSIN scores. We will also assess PWC scores in quiet as both a covariate and as a potential moderator—thereby controlling for baseline perception capabilities—and intra-individual variability. Of particular interest is whether the IBM-processed speech is significantly more intelligible than unprocessed speech in noise (which would indicate that IBM processing provides benefit). These models will also test whether IBM-processed speech is not significantly different from unprocessed speech in quiet (which would provide evidence that IBM processing restores intelligibility of speech in noise to that of speech in quiet). These tests will be done systematically. The first aim includes the unadjusted (bivariate) model as well as the adjusted model discussed in the Research Strategy. Aim 2 builds on this by including important moderators.

For power, we used the *simr* R package to simulate power given the sample size and expected effect sizes. This analysis relied on Monte Carlo simulations from a known population model. All code used is provided below. For Aim 1, a minimum of 30 participants will be tested per sub-experiment, as this results in sufficient power for the planned comparisons (for alpha of .05, with 6 covariates, effect size of .5, power = .93).

For Aim 2, at least 30 participants will be tested, as the power analyses with a moderate interactive effect indicate this sample size will result in sufficient power for the planned interactions (for alpha of .05, with 6 covariates, power = .91).

The R code to run the analyses are provided below.

```
library(tidyverse)
set.seed(843)

library(simr)

data_production <- function(npart = 30,          ## number of subjects
                            nplot = 6,           ## number of repeated measures
                            effect_x = .5,       ## effect of variable x
                            sigma_s = 2,         ## variance within participant
                            sigma = 1) {         ## residual variance
     x = rep(1:nplot, times = nplot)
     z = rnorm(npart*nplot)
     a = rnorm(npart*nplot)
     b = rnorm(npart*nplot)
     c = rnorm(npart*nplot)
     d = rnorm(npart*nplot)
     e = rnorm(npart*nplot)
     f = rnorm(npart*nplot)
     standeff = rep( rnorm(npart, 0, sigma_s), each = nplot)
     stand = as.character(rep(1:npart, each = nplot))
     ploteff = rnorm(npart*nplot, 0, sigma)
     resp = 1 + effect_x*x + standeff + ploteff
     dat = data.frame(x = as.character(x), z = scale(z), stand, resp = scale(resp), a, b, c, d, e, f)
     dat
}

data_production_int <- function(npart = 30,          ## number of subjects
                                nplot = 6,           ## number of repeated measures
                                effect_x = .5,       ## effect of variable x
                                sigma_s = 2,         ## variance within participant
                                sigma = 1) {         ## residual variance
     x = rep(1:nplot, times = nplot)
     z = rnorm(npart*nplot)
     a = rnorm(npart*nplot)
     b = rnorm(npart*nplot)
     c = rnorm(npart*nplot)
     d = rnorm(npart*nplot)
     e = rnorm(npart*nplot)
     f = rnorm(npart*nplot)
     standeff = rep( rnorm(npart, 0, sigma_s), each = nplot)
     stand = as.character(rep(1:npart, each = nplot))
     ploteff = rnorm(npart*nplot, 0, sigma)
     resp = 1 + effect_x*x + effect_x*z + effect_x*x*z + standeff + ploteff
     dat = data.frame(x = as.character(x), z = scale(z), stand, resp = scale(resp), a, b, c, d, e, f)
```

```
    dat
}

d <- data_production()
fit <- lmer(resp ~ x + a + b + c + d + e + f + (1|stand), data = d)
fit2 <- extend(fit, along="stand", n=30)
fixef(fit2)[2:3] <- 0.15
fixef(fit2)[4:5] <- 0.30
fixef(fit2)[6]   <- 0.45
# summary(fit2)

broom::augment(fit2) %>%
  ggplot(aes(factor(x), .fitted, fill = factor(x), color = factor(x))) +
    geom_boxplot(alpha = .7) +
    theme_classic() +
    scale_fill_viridis_d() +
    scale_color_viridis_d() +
    theme(legend.position = "none")

sim1 <- powerSim(fit2)
sim1

d2 <- data_production_int()
fit <- lmer(resp ~ x*z + a + b + c + d + e + f + (1|stand), data = d2)
fit2 <- extend(fit, along="stand", n=30)
fixef(fit2)[14]  <- 0.2
fixef(fit2)[15]  <- 0.2
fixef(fit2)[16] <- 0.3
fixef(fit2)[17] <- 0.3
fixef(fit2)[18] <- 0.4
# summary(fit2)

broom::augment(fit2) %>%
  ggplot(aes(z, .fitted, fill = factor(x), color = factor(x))) +
    geom_point(alpha = .7) +
    geom_smooth(method = "lm") +
    theme_classic() +
    scale_fill_viridis_d() +
    scale_color_viridis_d()

sim2 <- powerSim(fit2, test = fcompare(~ x + z))
sim2
```