Research Plan

Title: The Social Effect of Linguistic Alignment on Speech Production and Comprehension

Researcher: Emma Guarini

Rationale:

The cognitive processes the brain uses to process information are still widely unknown and require more research to fully understand. Previous studies have analyzed language as a way to learn more about the brain's ability to comprehend and subsequently produce speech. One aspect of language, called linguistic alignment, is a measure of one's tendency to use the same vocabulary as a partner during conversation and has been shown in previous research to improve communication effectiveness. Specifically, those who align to a conversational partner more tend to complete tasks faster and with fewer errors. It is unknown, however, if linguistic alignment is more beneficial to the speaker or listener in conversation. This research seeks to resolve this by measuring alignment levels of participants completing simple tasks. This research also seeks to determine if there is a social aspect of alignment, which has yet to be studies in previous research. Determining what prompts people to align their speech and who benefits from this alignment had the potential to increase knowledge of cognitive function, improve communication effectiveness, and understand the underlying decision-making processes of the subconscious brain.

Research questions

- Is linguistic alignment more beneficial for speech production of the speaker or speech comprehension of the listener in conversation?
- Does a positive perception of the speaker improve linguistic alignment in conversation?

Hypotheses

- It is hypothesized that participants will align to the experimenter the most when they are under load while naming the pictures because it would make the task easier for them, indicating a production benefit of alignment.

- It is hypothesized that participants who rated the experimenter more positively will align more because they are more trusting of the experimenter, supporting a social aspect of alignment.

Procedures

Experiment Set-up

- Based on the results of the pilot survey done in phase one which determine the preferred names for a series of pictures, 18 critical and 30 filler pictures will be selected to use in the experiment.
- Critical pictures will be those that had two common names produced on the pilot survey.
- Filler pictures will be those that had only one common name produced on the pilot survey.
- Six card packs will be created with three critical and five filler pictures each.

- Four extra pictures will also be put in each pack in order to ensure that card selection does not become significantly easier after many of the cards have been eliminated.
- A table will be set up so the experimenter can sit on one side and the participant can sit on the other side with a divider blocking one person from seeing the other person's cards.
- Both the participant and experimenter will have a 2x4 grid in from of them.
- The experimenter will also have the six card packs, the sheets with the card order for the participant to read, the sheet with the card order for the experimenter to read, a sheet to record the results of the digit span test and memory load throughout the experiment, and a computer to control the recordings on their side of the table.
- Before the experiment begins, both the participant and experimenter will put on their headset microphones which will be used to record the speech that will later be transcribed.

Digit Span Test

- A digit span test will be conducted to determine how many digits the participant is able to remember without any other tasks and how many digits they will receive in the memory load during the picture naming tasks.
- The experimenter will read two digits and ask the participant to repeat the digits right away.
- The experimenter will record what the participant said and if they said the exact same numbers as the experimenter had.
- The experimenter will repeat the process with a new two-digit number

- After two sets of two-digit numbers were said and recalled, the experimenter will repeat the same process with two sets of three-digit numbers, followed by two sets of four-digit numbers, etc.
- The process will be repeated until either the participant was unable to recall the correct digits for both numbers of the same length or they had successfully recalled all the digits through both 9-digit numbers.
- The experimenter will record the participant's digit span which is the number of digits in the longest number they were able to remember (nine if they were able to recall all the numbers)
- In order to determine how many digits the participant will receive as their memory load, their digit span will be divided by two and rounded up to the nearest whole number.

Picture Naming

- The participant will then complete six rounds of the experiment.
- In each round:

    1. The experimenter will place the 12 cards from pack one on the table in front of the participant.
    2. The experimenter will name eight of the cards in a pre-specified order for the participant to place on their grid.
    3. The experimenter will check that the participant has placed the cards in the correct order and subsequently collect the cards off the grid.
    4. The experimenter will hand the participant a sheet with the eight cards in a new order.

5. The participant will name the cards in that order for the experimenter to place on their grid.
6. The participant will return the sheet with the card order and the cards on the experimenter's grid will be collected.

Memory Load

- The number of digits given to the participants will depend on how many numbers they were able to remember on the digit span test as described above.
- There are four different experimental groups and each group will receive and recall the memory load at a different time.
- The participants in the "production load" group will be told their digits prior to step 5 in which the participant names the pictures aloud to the experimenter. They will be asked to remember these digits and recall them when they finish naming the pictures.
- The participants in the "production no load" group will be told their digits prior to step 5 but asked to repeat the digits right away.
- The participants in the "comprehension load" group will be told their digits prior to step 2 in which the experimenter names the pictures aloud to the participant. They will be asked to remember these digits and recall them when the experimenter finished naming the pictures.
- The participants in the "comprehension no load" group will be told their digits prior to step 2 but asked to repeat the digits right away.
- The experimenter will always record the digits the participant said and if they were the correct digits (in the correct order).

- Participants remain in the same group for the entirety of the experiment and therefore will always receive their digits at the same time
- The number of digits the participant receives will stay the same through all six rounds, however, the digits themselves will change each round.

Post Experiment Questionnaire

- A laptop displaying the post experiment questionnaire will be handed to the participant and the participant will be asked to fill out the survey to the best of their ability.
- The first section of the questionnaire will ask questions regarding demographics such as age, gender, and language history.
- The next section will ask the participant to rate the difficulty of the task on a scale of one-seven with one being extremely easy and seven being extremely difficult.
- The following section asks the participant to choose how much they agree or disagree with statements regarding the likability of the experimenter. Statements include “this person is knowledgeable” and “this person is friendly”. Participants will choose how much they agree on a scale of one-seven with one being very strongly disagree and seven being very strongly agree
- The final section of the questionnaire will ask the participant what they thought the experiment was testing and how much they thought they were using the same name as the experimenter to describe the same pictures.

(see attached for the full post experiment questionnaire)

- When the experiment has been completed, recordings will be saved into a secure folder on the computer.
- The data from the memory load sheet regarding how well the participant was able to remember their numbers during the picture naming task will be transferred into excel.
- The survey results will be transferred into excel.

Risk and Safety

- There are no potential risks or safety hazards for either the experimenter or the participant at any point in the experiment.

Data Analysis

- The speech recordings of each participant will be transcribed in excel by recording the names the participant uses when naming the pictures in each round.
- In order to assess if the participant used the same name that the experimenter had previously used, VLOOKUP will be utilized to compare the participant's names with the list of names the experimenter was instructed to use.
- Each term will be given a "1" if they used that same name as the experimenter or a "0" if they used a different name than the experimenter.
- Based on the participants total number of ones and zeros (for only the critical images), they will be given a total alignment score which will be turned into a percentage. 0% would indicate the participant never used the same name as the experimenter to name the critical images, 100% would indicate that the participant always used the same name as the experimenter to name the critical images.

- Alignment percentages will be averaged for all of the participants in each of the four experimental groups.
- These final alignment scores will be compared to determine if there was an effect of the time of load on alignment.
- Alignment values will also be looked at as a function of demographics such as age and gender which were recorded in the post experiment questionnaire.
- Each participant's perceived likability rating of the experimenter will be scored by giving each of the five choices a numerical value (1=very strongly disagree, 7=very strongly agree).
- These five numbers will be added to obtain the participant's final likability score out of 35 (higher numbers indicate the participant thought the experimenter was more likable).
- Excel will be used to graph the likability scores of each participant vs. how much that participant aligned to determine if there is a correlation between the two values.

Bibliography

Abel, J., & Babel, M. (2016). Cognitive Load Reduces Perceived Linguistic Convergence Between Dyads. *Language and Speech, 60*(3), 479-502. doi:10.1177/0023830916665652

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2). doi:10.1016/s0010-0277(99)00081-5

Branigan, H. P., Pickering, M. J., Pearson, J., Mclean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition, 121*(1), 41-57. doi:10.1016/j.cognition.2011.05.011

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482-1493. doi:10.1037//0278-7393.22.6.1482

Brennan, S. E., & Hanna, J. E. (2009). Partner-Specific Adaptation in Dialog. *Topics in Cognitive Science, 1*(2), 274-291. doi:10.1111/j.1756-8765.2009.01019.x

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & TylÃ©n, K. (2012). Coming to Terms. *Psychological Science, 23*(8), 931-939. doi:10.1177/0956797612436816

Gordon, P. C., Hendrick, R., & Levine, W. H. (2001). Memory-load interference in syntactic processing. *PsycEXTRA Dataset*. doi:10.1037/e537102012-414

Krahmer, E., Goudbeek, M., & Theune, M. (n.d.). Referring expression generation in interaction: A graph-based perspective. *Natural Language Generation in Interactive Systems,* 126-148. doi:10.1017/cbo9780511844492.006

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies Short Papers - HLT '08*. doi:10.3115/1557690.1557737

Human participants research

1. **Participants**

Participants will be within the age ranges of 13-18 and 23-28 years old in order to assess high school students and adults 10 years older. Some of these participants will be minors. Participants

will be of all genders and racial/ethnic compositions with the hope of collecting data from a diverse group of people representative of the population as a whole.

2. Recruitment

The high school aged participants (ages 13-18) will be recruited from local high school in the Westchester, NY area. They will be asked to participate via email, flyer, or personal solicitation. The adult participants (ages 23-28) will be recruited at a local research institution via email and their new hire network resource group.

3. Methods

Participants will first be asked to complete a digit span test in which their ability to remember series of numbers will be tested. The digit span test is a common test used in many experiments to assess memory span and is available for anyone who wishes to use it. The digit span test will take approximately three minutes to finish. Participants will then complete all six rounds of the main portion of the experiment. In each round, the experimenter will read a series of eight images while the participant places picture cards on a grid in front of them. The participant will then be given a sheet with the pictures in a different order and asked to read them for the experimenter to place on their grid. During each round, the participants will also be asked to remember a string digits to be recalled at a later point in the round. The number of digits range from one-five and depends on how many numbers the participants was able to recall during the digit span test. The point in each round that the participant will be asked to remember the numbers and the length of time they will remember the numbers depends of which experimental group they are in. There are four different participant groups. The participants in the first group,

called production load, will be given their numbers prior to naming the pictures aloud to the experimenter and asked to recall the numbers when they finish naming the pictures. The participants in the second group, called production no load, will also be given their numbers prior to reading the pictures aloud to the experimenter but they will be asked to recall the numbers right away. The participants in the third group, called comprehension load, will be given their numbers prior to the experimenter naming the pictures aloud to the participant and will be asked to recall the numbers when the experimenter finishes naming the pictures. The participants in the fourth group, called comprehension no load, will also be given their numbers prior to the experimenter naming the pictures aloud to them but will be asked to recall the numbers right away. In each of the six rounds, the participant will be given different pictures and different numbers. Each round will take approximately three minutes to complete. After all six rounds are completed, the participant will be given a post experiment to fill out on a laptop. The questionnaire includes questions regarding demographics – such as age, gender, and language history; the perceived task difficulty; the likability of the experimenter; and what the participant thought the experiment was testing. The questionnaire will take approximately 5 minutes to complete. In total, the experiment will take approximately 25 minutes for each participant to complete. The participants will take part in the experiment individually and will only complete the experiment once.

4. Risk Assessment

Participants will sacrifice 25 minutes of their time to take part in the experiment. They may feel bored or tired during the experiment. There are no other physical, psychological, social, or legal risks involved in participating in the experiment. The adult participants (ages 23-28) will be

given eight-dollar lunch vouchers to use at the research institution in compensation for completing the experiment. The high school participants will receive no monetary or other compensation for participating. All participants will benefit society by contributing to a study that will improve understanding of communication effectiveness.

5. **Protection of Privacy**

No identifiable information including names, phone numbers, email addresses, etc. will be collected. The age and gender of participants will be collected for demographic purposes only. All data collected will both confidential and anonymous. Data from each participant will be distinguished by subject number for analysis only. The speech recordings for each participant will be stored and transcribed in Box and can only be accessed by the student researcher and her mentor. All data will be aggregated for analysis and publication/presentation and no individual responses will be shared.

6. **Informed Consent Process**

All participants will be given a consent form that must be signed before beginning the experiment. Minors will also need to obtain parental consent before beginning the experiment. The consent form will describe the purpose of the study, scope of activity, testing conditions, time requirements, potential risks and benefits, and inform that participant that their participation is voluntary and that they have the right to stop at any time.