4-21-2016

The Effect of Signal-Independent Information and Sentence Predictability on Speech Comprehension in Babble Noise for Typical Hearing Young Adults

Mariah N. Smart
Cedarville University, mariahsmart@cedarville.edu

The Effect of Signal-Independent Information and Sentence Predictability on Speech Comprehension in Babble Noise for Typical Hearing Young Adults

Mariah N. Smart

Cedarville University

Abstract

The influence of signal-independent information and sentence predictability on listeners' comprehension was investigated by measuring transcription accuracy of a series of English sentences spoken in multi-talker babble background noise with circumstances varying between sets of sentences. Fourteen native English speaking young adults with no self-reported hearing or language disabilities completed written transcriptions of four sets of sentences. Listeners were presented with a control set of the signal alone, followed by a set of sentences spoken in multi-talker babble noise, then a set with visual presence of the speaker included, and finally a set where content context was given in addition to the previous conditions. These sentences also alternated in predictability, as defined by the content and grammar of the sentence. Overall, results were consistent with existing literature on the topic of speech comprehension. This study shows that comprehension in noise for typical hearing native speakers is significantly influenced by a variety of factors, and so these factors should be considered during communication in noise as well as in the context of speech comprehension testing.

The Effect of Signal-Independent Information and Sentence Predictability on Speech Comprehension in Babble Noise for Typical Hearing Young Adults

Introduction

Verbal communication is a complicated process, dependent on several factors simultaneously, and typically executed successfully without conscious consideration of these factors. It is when communication is unsuccessful that the individual components must be considered. Problems originating from the speaker or the receiver of the speech signal are generally the focus of the work of health professionals, such as speech-language pathologists and audiologists. This study focuses instead on problems originating in the channel through which the speaker's message travels, specifically the effect of noise on comprehension. While it is generally understood that noise affects comprehension negatively, the goal in this study is to better understand how inclusion of some non-verbal information can combat a decrease in successful communication due to noise.

Literature Review

Understanding verbal communication, referred to as speech perception, can be considered at two levels: intelligibility and comprehension. Speech intelligibility can be defined as "the amount of speech understood from the signal alone," (Fontan, Tardieu, Gaillard, Woisard, & Ruiz, 2015, p.8) where signal refers to the verbal production of the speaker. This excludes signal-independent information such as linguistic and contextual clues which are used to interpret message content. Speech comprehension on the other hand, considers the listener's understanding of the speaker's message via signal and context. Both intelligibility and comprehension are dependent on three parts of communication: the signal produced by the speaker, the transmission channel or environment, and finally the listener. The study of perception in general is important, and it has been demonstrated that there is a significant relationship between speech perception and language development in children. Generally speaking, in assessments for fields such as speech-pathology, foreign language learning, acoustics, and audiology, intelligibility is tested more often than comprehension (Fontan et al., 2015). More recently, there has been a topic shift to consider all the cognitive abilities necessary for understanding speech. Because both are used, differentiation is necessary between testing just the reception of phonemes and basics of language, which is the focus of intelligibility, and tests which consider signal-independent information and the more general idea of speech comprehension.

Comprehension involves the use of cognitive abilities and different types of processing, and so requires more explanation than intelligibility. This use of multiple contextual clues for comprehension is called “top-down processing” and is limited by processing capacity and working memory constraints of the listener (Uslar, et al., 2013). Top-down processing is relied on more heavily during noise, when the speech signal has been degraded and intelligibility is difficult. This means that skills like working memory and sentence recognition are stronger predictors of listening comprehension in noise, while intelligibility of the speech signal alone determines comprehension in quiet (Nagaraj, 2014). Pichora-Fuller (2008) studied the use of semantic context, a top-down processing factor, in adults of different ages and found that older adults rely more on semantic content than younger adults. While it is not known whether this is due to hearing loss, a wider collection of semantic models throughout their lifetime, or another variable, the use of semantic context should be considered, especially with older populations, when testing speech comprehension. In 1968, Boothroyd developed a metric to predict the use of top-down influences based on the probability of recognizing the separate constituents of the word in question. The metric is $p_w = p^j_p$ where P_w is the probability of recognizing the word, p_p is the probability of recognizing each part or phoneme, and j is the number of independent channels of information required for recognition and is between 1 and n. While helpful, top-down processing is used not only for individual words, but on semantic and pragmatic levels which affect whole sentences and conversations. For example, semantically unpredictable and complex sentences are likely to produce different results in intelligibility testing than predictable ones, and could result in misleading information if top-down processing is not considered (Uslar, et al., 2013).

Intelligibility and comprehension are both studied almost exclusively in the context of signal degradation from background noise. The most common type of artificial degradation used is multi-talker babble (MTB) which consists of the overlapping of several unintelligible speakers to create a masking noise. This is used primarily due to its reflection of reality. MTB is often cited as the most frequently encountered environmental noise in real world situations, and so studies using MTB have more ecological validity than those which use steady-state masking such as white noise (Fontan et al., 2015). One significant discovery made during these studies is that the mental process of speech comprehension actually changes as environmental noise is introduced. The physical part of the brain in which speech perception and identification occurs has traditionally been understood to be the superior temporal cortices. Du, Buchsbaum, Grady, and Alain (2014) studied the role of the speech motor system (SMS) in speech perception tasks and found evidence to support current theories that, while primarily the area involved in speech production, it is also involved in perception. The SMS's role in perception is understood as generating internal models that predict sensory consequences of articulatory gestures and then matching these predictions with acoustic representations in sensorimotor interface areas of the brain (Du et al., 2014). Noise has been shown to speed up information processing in working memory and attention switching tasks, and can be understood as enhancing arousal and concentration for cognitive tasks (Nagaraj, 2014). It is possible that this processing speed improvement is associated with the use of multiple parts of the brain, including the SMS working in an adaptive and task-specific manner, while speech processing in optimal listening conditions only occurs within the auditory system. The use of the SMS in comprehension occurs in a bell curve along an axis of increasing background noise. In quiet conditions, there is no need for the SMS to assist in comprehension, while in extremely noisy conditions comprehension is not possible even with the use of the SMS. So the highest functioning of the SMS occurs in somewhat noisy conditions where enough stimuli are received to result in comprehension with the assistance of this alternate method of processing (Du et al., 2014). While the exact functioning of the brain during speech processing in noise or in quiet is still being studied, the inclusion of noise like MTB in testing and research has been widely accepted as one of the most informative methods of assessing intelligibility and comprehension in typically developed individuals, as well as in those with language and hearing impairments.

The presence of noise with a speech signal increases the load in the perceptual stream, and is useful in drawing out subtle deficits in a person's perception which would not be evident in a quiet context (Vance & Martindale, 2012). Because of the frequent use of signal-in-noise testing for diagnosing speech and hearing disabilities, much of the research on intelligibility and comprehension compares typically developed individuals to those with impairments. For example, dyslexic children demonstrate difficulty recognizing speech in noise "because the ability to make strong predictions about what kinds of structures to expect makes perception resistant to noise making" (Caldwell & Nittrouer, 2013, p.14). A difficulty linked to conditions with variable intonation and correctly identifying the words of focus has also been recorded in some dyslexic individuals (Hazan, Messaoud-Galusi, & Rosen, 2013). These perception deficits seem to be linked mainly to top-down processing, which is why both intelligibility and comprehension should be tested when predicting possible challenges for those with a variety of impairments. Children categorized more generally as language impaired have shown deficits in auditory processing, such as sensitivity to sound duration and frequency discrimination of tones. These same children showed difficulties in speech perception tasks including categorical perception, consonant identification, and consonant vowel syllable discrimination. While this is not true of every language impaired child, the correlation should be noted as possibly a cause-effect or other type of relationship between language impairment and processing (Vance & Martindale, 2012). Finally, it is a common belief that users of cochlear implants (CIs) are especially affected by background and environmental noise, but this has not been proven. In a comparison between dyslexic children and CI users, there was a clear disproportionate effect of noise on speech perception abilities for children with dyslexia. However, CI users declined in perception with noise at a very similar rate as typically developed children; the use of CIs just lowered perception at all levels of noise and quiet fairly equally (Caldwell & Nittrouer, 2013).

For bilinguals and situations involving nonnative speakers, varying languages and social expectations can significantly influence the perception of the listener. The use of native or nonnative language becomes relevant when considering the acceptable noise level (ANL). The ANL was developed by Nábelek, Tucker, and Letowski in 1991 as “a subjective measure of listeners’ willingness of accepting background noise” (as cited in Lu-Feng, Azcona, & Buten, 2015, p.497). This idea of putting up with noise is subjective, but the average of ANL is significantly lower when the background noise consists of babble in the listener’s native language than if the babble is in another language or made of unintelligible sounds. Understanding the masking noise makes it more distracting or easier to confuse with the speech signal (Lu-Feng, et al., 2015). Another consideration with native and nonnative language is the effect of noise on listeners of their second language. Second language learners lack some familiarity with the phonology and syntax of the language and are more affected by difficult listening conditions than native speakers, regardless of the type of babble or noise masking the signal (Caldwell & Nittrouer, 2013). Mi, et al. (2013) found that native advantage in sentence recognition is the highest at medium levels of noise-masking conditions. This is the same level of background noise which causes listeners to rely on top-down processing the most heavily; at lower levels the speech signal is clear enough to be processed alone and at higher levels the noise-masking is too severe for top-down processing techniques to be helpful. This further confirms the importance of pragmatic knowledge, for example, for accurate comprehension of speech.

Apart from the difficulties of comprehension for a nonnative speaker, culture almost always plays a role in speech perception. Studies show that social expectations can influence the listener's processing of the speech signal and will do so differently in response to variables such as age, gender, sexual orientation, race, and others. McGowan (2015) discovered that transcription of the same speech signal was more accurate when presented with a face matching the race of the actual speaker than one that did not match. In this specific study, a Chinese and Caucasian face and voice were presented to participants in combinations, and regardless of past experience with Chinese accented English, participants showed more accurate perception when given more information about the speaker. This finding is inconsistent with early research which claimed that listeners hallucinated accents in perception when they believed the speaker to be foreign. It was thought, especially in an academic setting, that social expectations would negatively influence perception and concentration of speech from foreign instructors. These more recent sociophonetic studies show that, while perception does have an influence in speech perception, it is not necessarily a negative one, and congruent social primes and voices allow for listeners to interpret the signal with improved accuracy (McGowan, 2015). This is consistent with the idea that the SMS generates internal models based on external information in order to sharpen the perceptual accuracy of the sensory system via a top-down allocation mechanism (Du et al., 2014) In fact, electroencephalography (the recording of electrical activity in different parts of the brain) suggests that this kind of information that is independent of the speech signal is not only inherent to the speech comprehension process, but is processed simultaneously with the linguistic message (Fontan et al., 2015) This heavy influence of social expectations and signal-independent information highlights the need for comprehension testing, and not just intelligibility testing, when diagnosing and studying language or hearing impairments of any kind.

Most of the existing literature surrounding speech perception in noise concerns either children, individuals with hearing or language impairments, or both. These are the individuals most severely affected by the process of speech intelligibility and comprehension; however, language processing is something every person experiences daily. This leads to the conclusion that more study regarding typically developed adults would be of value. While the traditional intelligibility test is administered through a write-down paradigm, “in which a listener transcribes a list of speech items like words or sentences” (Fontan et al., 2015, p.977), recent studies are shifting focus to the importance of comprehension for communication and all factors involved in processing.

Because of the relative newness of the focus on comprehension in the field, there are several types of signal-independent factors that have yet to be studied. In this study, I propose to observe the effect of visual information, contextual information, and sentence predictability on comprehension in background multi-talker babble noise. The demographic chosen for this study consists of typical hearing young adults. The goal is to better understand the process of comprehension for the general public in circumstances which reflect those frequently found in daily life, specifically communicating in noisy or crowded environments.

Methods

In this study, the following hypotheses were tested: (1) Signal-independent information generally improves speech comprehension for typical hearing young adults, (2) The predictability of sentences significantly affects speech comprehension for typical hearing young adults, and (3) The predictability of sentences affects speech comprehension for typical hearing young adults more significantly when fewer sources of signal-independent information are available.

In order to answer these questions, a study was designed in which participants would listen to a primary signal in varying circumstances, and transcribe what they heard. Both the primary signal and babble noise were created and recorded in a way that could be regulated, consistent, and provide a stimulus useful for this area of study. These recordings were then played, along with some verbal instruction and explanation, for a population of participants who transcribed their understanding of the primary signal. Measures were taken during each stage to ensure the validity of data collected.

Participants

Fourteen native English speakers between ages eighteen and twenty-two, currently full time students at the undergraduate level, completed this study. These individuals responded to an invitation, which was sent out to a larger number of students, asking for participants who were willing to complete a short transcription exercise. All of the participants were female, which was not an intentionally chosen subgroup but occurred as a result of the population that was available at the time. Compensation was provided in the form of a small gift. This study focused on typical hearing young adults, and each participant self-reported no history of hearing loss or other language/hearing related diagnosis with the exception of two individuals. The first received two years of speech therapy as a child and had a lisp that has since been resolved, and the other reported a constant ringing noise in the left ear which has never been officially diagnosed or treated by a professional. These conditions were noted to be considered alongside results, but the type of interference reported was not severe enough to omit the data collected from these individuals. Each participant completed an identical transcription exercise and listened to the same audio, so there were no subgroups within the population studied.

Design

A recording of twelve spoken sentences was created for this study, as well as 6-speaker multi-talker babble (MTB) noise. Both were recorded in the same room using a uni-directional audio condenser microphone with an anti-wind foam cap and a frequency response of 20Hz-20kHz. The speech signal of twelve sentences was recorded as an MP4 file using a Samsung laptop built-in 1.3 MP camera to capture video with the audio. The MTB was recorded and layered using Audacity, and edited so each speaker’s average dB level was similar (~ 68 dB). The babble noise consisted of English sentences with a similar intonation pattern and length as the sentences used in the tested speech signal, and consisted of both male and female voices. The sentences were looped in order to create two minutes of MTB, and staggered so that pauses at the beginning and end of sentences did not occur at the same time, making it more consistent and continuous. The intention in the creation of this MTB was to replicate conditions that might be common in crowds or public areas, while maintaining a consistent dB level and content to minimize unwanted variables between participants.

The sentences used for the speech signal to be transcribed were selected from the existing AzBio Minimum Speech Test Battery sentence list, a list of 165 sentences distributed by Auditory Potential, LLC used to evaluate the speech perception abilities of hearing impaired listeners and cochlear implant users (Spahr, et al., 2012; “Minimum Speech Test,” 2014). The sentences used for this study were selected from throughout the list, based on the number of words in the sentence and predictability. Four sets of four sentences were to be transcribed, and each set consisted of two sentences that were five words in length and two sentences that were seven words in length. A total of sixteen sentences were used, with no repetition or overlap of content, except in the final set where content was similar. The sentences also alternated between predictable and unpredictable. A sentence was considered predictable if the subject and verb indicated an object or other element in the remaining part of the sentence that could be logically inferred. For example, “I can type forty words per minute,” was a predictable sentence used, and “They need a lock on the refrigerator,” was an unpredictable sentence used in this study. All sentences were spoken by the same individual, a forty-eight year old male with a Midwest accent, representative of generally accepted Standard English. He was instructed to read the selected AzBio sentences in a manner that resembled natural, conversational speech. This gave the speech signal natural intonation and varied pace, as well as specific types of emphasis that conveyed or supported some sentence meanings, as would be heard in typical conversational verbal communication. A worksheet was created with a brief explanation of the speech conditions and numbered lines divided into sets for the sentences to be transcribed.

Procedure

Participants completed transcription one at a time, in a room with minimal outside noise sources or other distractions. For the convenience of different participants, two separate rooms were used. The rooms were both private and similar in size, but one had an average of 60 dB environmental noise and the other an average of 55 dB. Transcription sheets were not marked to indicate in which room they were completed, but in both rooms the environmental noise stayed consistent during the entire study. Each participant was given the same verbal and written instructions regarding what was expected of them and what they would be hearing. Information about the use of babble noise, number of sentences, and the inclusion or omission of video or context was disclosed, but participants were not informed about sentence length, predictability, or the hypotheses being tested. Participants were also instructed to write down anything they heard or thought that they heard, even if they were unsure of accuracy. In the first set, audio of the speech signal was played alone, as a control set and to familiarize participants with the process. In the second, the next set of sentences from the speech signal was played along with MTB noise at a 2:1 signal-to-noise ratio (SNR). For the third set, the signal and MTB was played at the same SNR, and video of the speaker was shown to the participant. In the final set, the signal and MTB was played at the same SNR with video shown, and the participants were informed of the context of the sentences, which in this set were all focused on the same topic. All sentences were played with sufficient pause following to allow time for transcription, in order to focus the test on comprehension and not memory or speed of writing.

Results

Data from all fourteen participants were included in the analysis, as there were no outliers and none of the testing procedures were compromised. Quantitative analysis of data was based on the averages to the second decimal place of words transcribed correctly in each set. Averages provide the clearest representation of all participants as a whole, because individual results varied between individuals based on their hearing and personal abilities, but change between sets followed a general trend. Averages were also used to compare predictable and unpredictable sentences. A word was considered correct when the entire word was written, and no partial words were counted. However extra affixes or words were not counted against a participant's score, so there was no penalty for writing incorrect words or extra words as long as the original word from the speech signal was also included. Each set of four sentences consisted of 24 words, so all correctly transcribed words were calculated as a fraction over 24.

As seen in Figure 1, in the control set an average of 23.14 (96.4%) of the spoken words were transcribed correctly. The second set, audio with babble noise, presented the lowest score with only 12.07 (50.3%) of words correctly transcribed on average. The addition of video in the third set resulted in an average of 17.21 (71.7%) words, and the addition of context in the final set resulted in an average 21.50 (89.6%) words correctly transcribed. In addition to the results of each set, percent change between sets was considered. Audio in babble noise was transcribed 46.1% less accurately than the control set. Addition of video in babble noise increased accuracy by 21.4% from audio in babble noise, a 24.7% decrease from the control. Finally, the inclusion of video and context along with audio in babble noise increased accuracy by 17.9% from the video and audio set, increased by 39.3% from audio only in babble noise, and decreased by 6.8% from the control.

Figure 2 shows the effects of predictability on transcription accuracy, with predictable and unpredictable sentences alternating within all four sets, resulting in a total of eight predictable and eight unpredictable sentences in the entire transcription exercise. Accuracy of transcription varied based on the signal independent information that changed between sets, but there was also a distinction in accuracy based on predictability regardless of set conditions. Of the forty-eight words in predictable sentences, an average of 41.64 (86.8%) words were correctly transcribed. Of the forty-eight words in unpredictable sentences, an average of 32.29 (67.3%) words were correctly transcribed. This shows a 19.5% difference in transcription accuracy between predictable and unpredictable sentences, throughout the various sets of signal independent information changes.

In Figure 3, the average transcription accuracy for predictable and unpredictable sentences is compared within each set of differing signal-independent conditions. In the control set, unpredictable sentences were 5.1% less accurately transcribed than predictable sentences. In the second set, unpredictable sentences were transcribed with 33.2% less accuracy. In the third set, accuracy dropped by 28.1% in unpredictable sentences compared with predictable sentences. In the final set, the decrease in accuracy for unpredictable sentences measured 13.7%. Control set aside, this shows a negative correlation between the number of signal-independent information sources present and the percent difference in transcription accuracy between predictable and unpredictable sentences.

Apart from accuracy of words transcribed correctly, there were other trends in the results that should be noted. In the final unpredictable sentence of Set 3, “They need a lock on the refrigerator,” seven of the fourteen participants added the word “door” as the final word of the sentence. Those who did not add “door” only transcribed zero to two words in the sentence correctly, but there were no participants who correctly transcribed a majority of the sentence who did not also add the word "door." None of the participants had any contact with each other or saw any of the previously written transcriptions, and the speech signal did not include any stuttering or other flaws which would directly cause hearing that particular word.

A final trend noted in the data was the tendency to correctly transcribe words in chunks. That is to say, scores of 5/5 or 7/7 (completely correct) or scores of 0/5 or 0/7 (completely incorrect or not transcribed) were more common than scores somewhere in the middle. On average, 9.93/16 questions were transcribed completely accurately or not at all, while 6.07/16 were transcribed part-way. It is possible this difference is influenced by the number of completely accurate transcriptions in the control, so this observation should not be considered completely dependable, as the nature of this study was not conducive to measuring this trend accurately. However, the results did show signs of commonly transcribing accurately or inaccurately in chunks of information, and this could be a valid topic for more in-depth study in future research.

Discussion

All of the hypotheses proposed at the beginning of this study were either supported by the data, or there was not enough information to conclude either a confirmation or rejection. First, it was proposed that signal-independent information generally improves speech comprehension for typical hearing young adults. The data shows a significant increase in transcription accuracy with the addition of visual information and contextual information. The most significant increase in accuracy compared to the previous set occurred with the addition of visual information. Visual information includes cues such as non-verbal pragmatic communication in expression, and watching lip movement for improved understanding of individual phonemes. The ability to see when the speaker began and finished speaking also seemed significant to improving comprehension, as some participants missed speech in the audio and babble noise set because they were unsure of exactly when the speech signal was occurring. The addition of signal-independent information improved comprehension in babble noise so much that, with both visual information and contextual information, transcription was only 6.8% less accurate than the control set. This also aligns with the research of Du et al. (2014), and the proposal that speech comprehension in noise occurs with the help of the SMS instead of the superior temporal cortices. The use of the SMS permits non-linguistic information to be processed and interpreted for speech comprehension, and the data in this study shows that addition of this type of information does in fact improve comprehension.

The second hypothesis tested in this study was that the predictability of sentences affects speech comprehension for typical hearing young adults. The study showed a 19.5% increase in transcription accuracy for all sets when transcribing predictable sentences. Predictability affects top-down processing, which was shown by Nagaraj (2014) to be relied on more heavily in noise than in silence. While this can be seen in predictability throughout all sets, the way predictability affects comprehension is seen more clearly when looking at each set individually. This consideration of both sets and predictability relates to the third hypothesis, which states that the predictability of sentences affects speech comprehension for typical hearing young adults more significantly when fewer sources of signal-independent information are available. The distance between predictable and unpredictable sentence transcription accuracy decreases with the addition of more signal-independent information sources. This then shows evidence that being able to infer sentence content from a few words of the sentence is more crucial to understanding when less information is available through other channels. With the addition of information like lip movement and context, comprehension based on predictability is not relied on as heavily. Therefore, because predictability always influenced comprehension, even in the control set, it is a factor that must seriously be taken into consideration when dealing with signal-in-noise situations.

The data concerning word addition and information chunking is less exact, but continues to support the idea of top-down processing in noise. The addition of the word "door" by half of the participants was not due to any phonetic production on the part of the speaker. This means that every one of the participants concluded that "door" was a logical word to include in the sentence, and because the signal was obstructed by the babble noise, assumed its inclusion. This addition may not have been made by younger listeners, who have a smaller corpus of sentences in their memory to use as reference for understanding, but this study does not allow for that assumption to be made. The data also seemed to suggest a trend of transcribing sentences completely accurately or not at all, or at least chunking the information into large groups of meaning when transcribing. This would also support a top-down cognitive approach to comprehension but, as previously mentioned, this was not shown clearly enough in the data to make any kind of conclusions on the topic from this study alone. At the minimum, this observation does not contradict or call into question any of the hypotheses that the rest of the study supports.

Limitations

While measures were taken in order to produce valid results, there were still several limitations to this study. Concerning the demographic studied, there were enough participants to see some trends in the data, but repeating it with more participants would allow other patterns to emerge. This could include patterns like information chunking, which seemed to occur in the data, but would require additional evidence to make any conclusions regarding this trend. Similarly, with a larger data pool it is possible that sound-specific errors might be seen. Also limiting the study was the inclusion of only female participants. The inclusion of male participants may not have any significant effect on the results, but in order to generalize conclusions to all typical hearing young adults these hypotheses should be tested on both men and women. Data was also possibly skewed due to the environment and quality of equipment used. There were no significant audio obstructions in the recordings or from environmental sources, but more exact results would be obtained using professional recording equipment in a lab or studio made specifically for controlling environmental noise and echoes.

Implications

The results of this study support existing literature on the topic and are applicable for typical hearing young adults, considering the limits of generalization previously mentioned. The significant changes in comprehension accuracy with the addition of visual and contextual signal-independent information support the consideration of these factors in a variety of contexts. When testing for comprehension, for example, accuracy should be compared with typical hearing scores in exactly the same signal-independent conditions to avoid a misdiagnosis of a hearing or language impairment. These results may also be considered when planning mass communication in crowded or noisy environments. The intentional addition of visual information or contextual information, along with the use of predictable sentences when possible, will greatly improve understanding in environments where comprehension is negatively affected by environmental noise.

Conclusions

A positive relationship has been shown to exist between accuracy of speech transcription in babble noise and the addition of signal-independent information for normal hearing young adults. Predictability of the sentence content also significantly influenced transcription accuracy, and had a greater effect when less signal-independent information was available. This supports existing literature on speech comprehension, and the use of top-down processing to understand speech in noise. This study was limited by the type and number of participants as well as the narrow focus of the research, but the data collected is reliable and informative concerning the small demographic and focus to which it applies. It would be beneficial to further study this topic on a larger scale, incorporating other variables such as age of listener, or type of speech signal.

References

Auditory Potential, LLC. (2014). *Minimum Speech Test Battery*. Retrieved from http://www.auditorypotential.com/MSTB_Nav.html

Boothroyd, A. (1968). Statistical Theory of the Speech Discrimination Score. *The Journal of the Acoustical Society of America, 43,* 362-367.

Caldwell, A. & Nittrouer, S., (2013). Speech Perception in Noise by Children With Cochlear Implants. *Journal Of Speech, Language & Hearing Research*, 56(1), 13-30.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *111*(19), 7126-7131.

Fontan, L., Tardieu, J., Gaillard, P., Woisard, V., & Ruiz, R. (2015). Relationship Between Speech Intelligibility and Speech Comprehension in Babble Noise. *Journal Of Speech, Language & Hearing Research*, *58*(3), 977-986.

Hazan, V., Messaoud-Galusi, S., Rosen, S. (2013). The Effect of Talker and Intonation Variability on Speech Perception in Noise in Children With Dyslexia. *Journal of Speech, Language & Hearing Research*, 56(1), 44-62.

Lu-Feng, S., Azcona, G., & Buten, L. (2015). Acceptance Noise Level: Effects of the Speech Signal, Babble, and Listener Language. *Journal of Speech, Language & Hearing Research*, *58*(2), 497-508.

McGowan, K. B. (2015). Social Expectation Improves Speech Perception in Noise. *Language & Speech*, *58*(4), 502-521.

Mi, L., Tao, S., Wang, W., Dong, Q., Jin, S., & Liu, C. (2013). English vowel identification in long-term speech-shaped noise and multi-talker babble for English and Chinese listeners. *Journal Of The Acoustical Society Of America*, *133*(5), EL391-EL397.

Nagaraj, N. (2014). *Explaining Listening Comprehension in Noise Using Auditory Working Memory, Attention, and Speech Tests.* (Electronic Thesis or Dissertation). Retrieved from https://etd.ohiolink.edu/

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal Of Audiology*, *47*72-82

Spahr, A., Dorman, M., Litvak, L., Van Wie, S., Gifford, R., Loizou, P., Loiselle L., Oakes T., Cook, S., (2012). Development and validation of the AzBio sentence lists. *Ear Hear, 33(1)*, 112-17.

Uslar, V. N., Carroll, R., Hanke, M., Hamann, C., Ruigendijk, E., Brand, T., & Kollmeier, B. (2013). Development and evaluation of a linguistically and audiologically controlled sentence intelligibility test. *Journal Of The Acoustical Society Of America*, *134*(4), 3039-3056.

Vance, M., & Martindale, N. (2012). Assessing speech perception in children with language difficulties: Effects of background noise and phonetic contrast. *International Journal Of Speech-Language Pathology*, *14*(1), 48-58.

Yoon, Y., Allen, J. B., & Gooler, D. M. (2012). Relationship Between Consonant Recognition in Noise and Hearing Threshold. *Journal Of Speech, Language & Hearing Research*, *55*(2), 460-473.

Figures

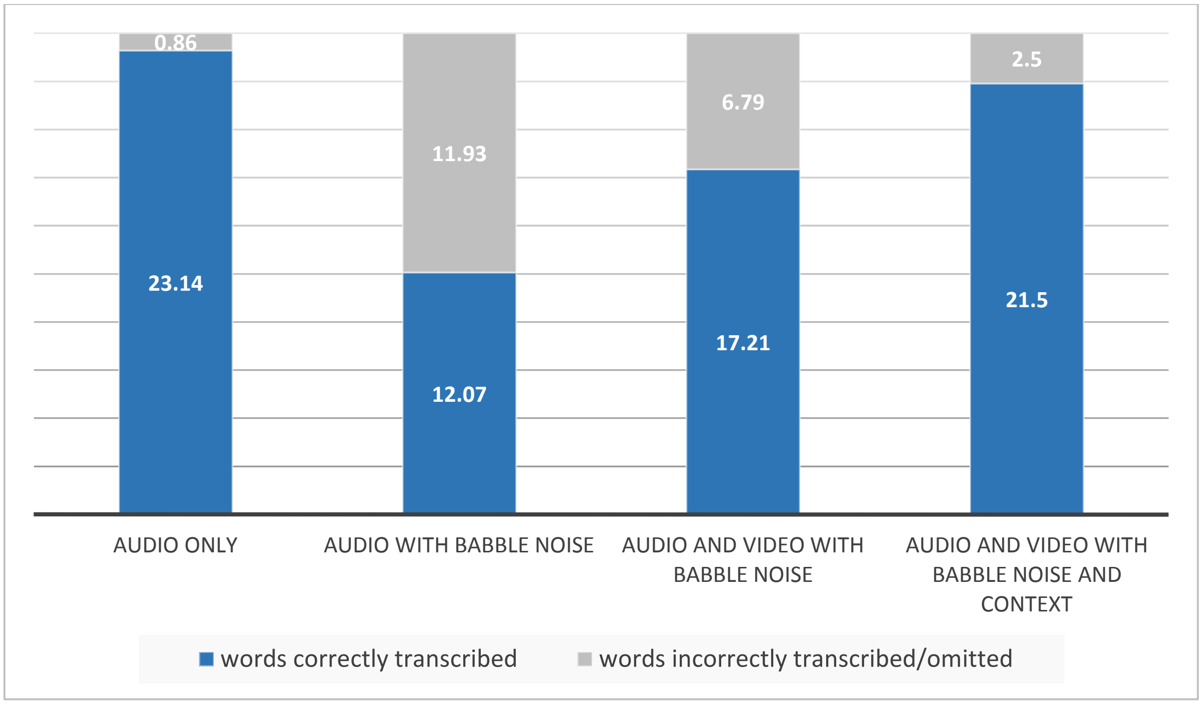

Figure 1. Average transcription accuracy by sets

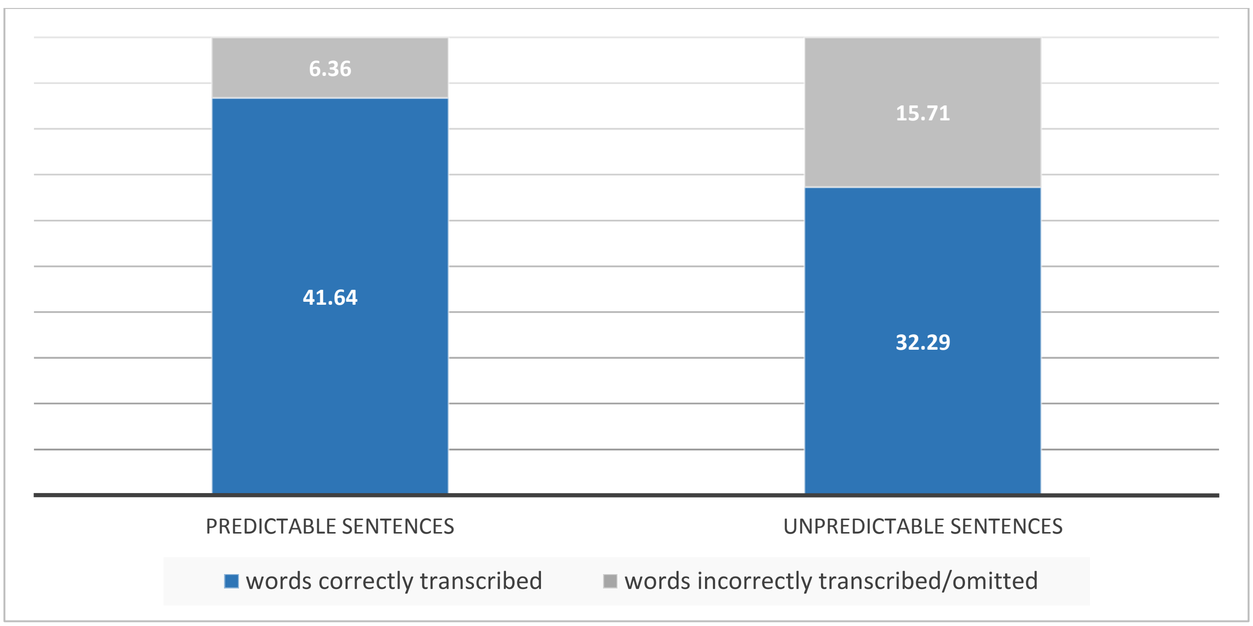

Figure 2. Average transcription accuracy by predictability

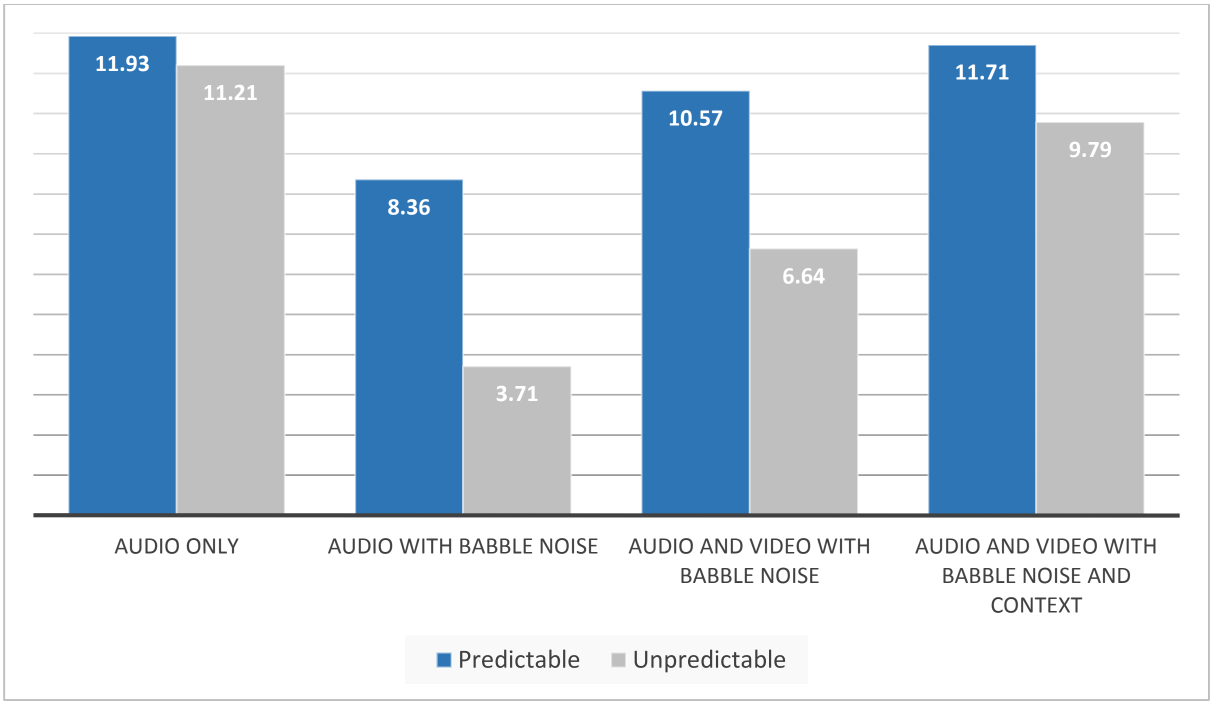

Figure 3. Average transcription accuracy by predictability within sets