A Whispered Christmas: Phonetic Expectations and Type of Masking-Noise Influence Auditory Verbal Hallucinations

Mark Scott[*1]
[1]Memorial University of Newfoundland – Department of Psychology, Qatar University

1 Introduction

Auditory Verbal Hallucinations (AVH) are most commonly associated with schizophrenia, being one of the key diagnostic symptoms [1]. However, c. 9.6% of people with no known psychological issues also report this experience [2]. A common technique for inducing hallucinations in healthy volunteers is to present them with masking noise and suggest the presence of sound under the noise. The most famous demonstration of this is the 'White Christmas' effect in which participants are played stochastic 'white' noise and told to indicate when they hear the song 'White Christmas' underneath the noise. Participants often report hearing the song, despite there being nothing in the audio signal but noise [3]. Many replications have demonstrated this effect with speech [4–6].

The 'White Christmas' effect induces hallucinations through sensory expectation and masking. However, there has been little research conducted on how the phonetic details of the sensory expectation or how the type of masking noise modulate the effect. The current experiment addresses this gap by examining two types of phonetic expectation (expecting normal modal speech vs expecting whispered speech) and two types of masking noise (multitalker babble vs spectrally-matched speech-shaped stochastic noise). Masking relations likely play a key role as the 'White Christmas' effect is similar to *phoneme restoration*, in which expectation causes people to hear speech-sounds that are absent, but only if there is masking that *would* have obscured the missing sound [7]. It is predicted that there will be more AVH for multitalker babble than for speech-shaped noise masking, as multitalker babble is a more effective masker. Similarly, there should be more AVH for expecting whispered than for expecting spoken voice, as whisper is quieter and easier to mask. An interaction is also predicted by which AVH should be disproportionately more common when expecting a normal voice under babble noise, as the periodicity of normal voice should be particularly masked by babble.

2 Method

The experiment was conducted at Qatar University. Participants sat in front of a computer screen wearing sound-isolating headphones. On each trial, participants were presented with a target word in Arabic and were told to listen for this word during the noise that followed. Whenever they heard the full word (not just part of it), they were to hit the spacebar. They were told that the word could appear multiple times during a trial, so they could hit the spacebar more than once per trial. They were also told that the word may not occur at all, so it was fine not to hit the spacebar. When ready to start a trial, the participant hit the 'enter' key at which point either multitalker babble or spectrally-matched speech-shaped noise played for thirty-five seconds. The noise factor (**Babble** vs **Speech-shaped** noise) was run *within-subjects*. There were 14 trials in the experiment (7 each of **Babble** and **Speech-shaped** noise). The voice conditions (**Spoken** vs **Whisper**) were run *between-subjects*. Participants in the **Spoken** condition were told they would hear a spoken voice under the noise, participants in the **Whisper** condition were told they would hear a whispered voice. The experiment, including instructions and practice, lasted about twenty minutes.

Two lists of seven Arabic words were created as the 'to-be-listened-for' sounds, one for each noise condition. The assignment of word-list to noise condition was reversed for half of the participants, so the assignment was counterbalanced over the full run of the experiment. The words were all chosen to be concrete, easily imageable nouns. The words in these lists were selected to use phonetically similar sounds (roughly equal numbers, between the lists, of e.g., nasals, sibilants, stops). This phonetic balancing was an extra precaution to minimize possible variability (within each run of the experiment) in how easy these words might be to hallucinate. The lists of words are provided in table 1 :

Table 1: Word lists for the two conditions.

List One		List Two	
Arabic	**English gloss**	**Arabic**	**English gloss**
tuffah	'apple'	*katef*	'shoulder'
maktab	'desk'	*minshaar*	'saw'
shooka	'fork'	*furshaah*	'brush'
korsi	'chair'	*kitab*	'book"
mismaar	'nail'	*sikkin*	'knife'
shafra	'razor'	*sariir*	'bed'
bedla	'suit'	*dulab*	'cupboard'

The multitalker babble was created using audio from the Qatari Arabic Corpus. Twenty-one files were selected without music or other non-speech sounds. They were normalized to the same intensity and transitions were cut at crossing points in the waveform to avoid pops at the joins. These sounds were then overlapped to create a multitalker babble of c. 30 voices. While individual words could not be distinguished in this babble, as an extra precaution the transcripts of the recordings were searched for the fourteen target words and none of them occurred in the sections used.

The speech-shaped noise was created by Praat script [8]. This measured the long-term average spectrum and intensity profile of the multitalker babble and applied these to stochastic noise of equal duration. The two noise sounds thus had

*. mark.a.j.scott@gmail.com

equal intensity profiles and average spectra.

There were 28 Arabic-speaking female participants (14 each for **Spoken** and **Whisper** conditions). Participants were paid or given course credit for their participation.

3 Results

A mixed ANOVA found a main effect of **Noise Type** (F[1,26] = 10.9, p = 0.003) and a marginal main effect of **Voice Type** (F[1,26] = 3.42, p = 0.075]). There was no significant interaction of **Noise Type** by **Voice Type**. These results are shown in figure 1.

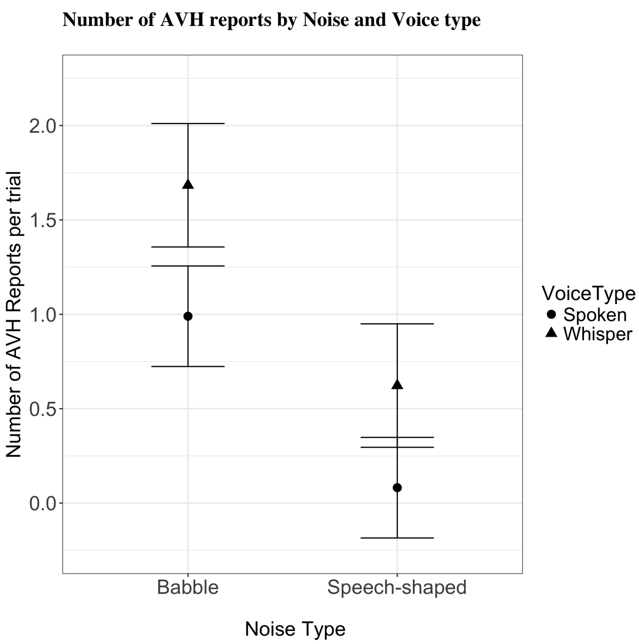

Figure 1: Results.

4 Discussion

These results show that speech hallucinations are influenced by both phonetic expectation and type of masking noise. With multitalker babble masking-noise, the rates of AVH are, as predicted, higher. Similarly, when the participant is told to expect whisper, hallucination rates are (marginally) higher. These results fit with findings on phoneme restoration, in which people can be induced to hear a missing speech-sound based on the surrounding context as long as there is sufficient masking that the absent sound *would* have been obscured had it been present. Multitalker babble is a more effective masker than stochastic noise, and whisper is more easily masked than modal voice. Contrary to prediction, there was no interaction between phonetic expectation (voice-type) and masking (noise-type). Attention may play a role in this, given that attentional demands for detecting a sound in noise are greater when the sound is similar to the noise – as occurs for detecting normal speech in babble. However, perhaps the data set is simply too small to reveal an interaction.

These findings are consistent with a proposed mechanism of hallucination in which a perceptual *schema* of a sound is compared against auditory input looking for a match. When there is sufficient similarity, the person reports hearing the sound [9]. Multitalker babble contains bits of speech with phonemes that will occasionally match part of the word being listened for, and so this should increase the likelihood of a listener reporting a 'match' with the expected target word.

Given the links between hallucination and imagery, these results are likely related to research showing that ambiguous speech sounds can be shaped by expectations provided by speech imagery [10, 11].

5 Conclusions

This experiment shows that both phonetic expectation and type of masking noise influence speech hallucinations in the 'White Christmas' effect. In line with predictions based on the phoneme-restoration effect, greater masking (multitalker babble) and a more easily masked expected-sound (whisper) induce more AVH. No interaction was found between these factors. These are preliminary data – a larger and more detailed data set has been collected and is being analyzed.

Acknowledgments

Participants were run by research assistants Maryam Moustafa Aref and Rofida Hamid Ibrahim. This research was funded by Qatar University grant QUUG-CAS-ELL-17_18-7

References

[1] Chris D. Frith. *The cognitive neuropsychology of Schizophrenia*. Lawrence Erlbaum Associates, Hove, 1992.

[2] Kim van Slobbe-Maijer. *Auditory hallucinations in youth : occurrence, clinical significance and intervention strategies*. PhD thesis, University of Groningen, 2019.

[3] Theodore Xenophon Barber and David Smith Calverley. Empirical evidence for a theory of "hypnotic" behavior : Effects of pretest instructions on response to primary suggestions. *The Psychological Record*, 14(4) :457–467, 1964.

[4] Steven R. Feelgood and Andrew J. Rantzen. Auditory and visual hallucinations in university students. *Personality and Individual Differences*, 17(2) :293–296, August 1994.

[5] Charles Fernyhough, Kirsten Bland, Elizabeth Meins, and Max Coltheart. Imaginary companions and young children's responses to ambiguous auditory stimuli : implications for typical and atypical development. *Journal of Child Psychology and Psychiatry*, 48(11) :1094–1101, 2007.

[6] Samantha Hartley, Sandra Bucci, and Anthony P. Morrison. Rumination and psychosis : an experimental, analogue study of the role of perseverative thought processes in voice-hearing. *Psychosis*, 9(2) :184–186, 2017.

[7] Makio Kashino. Phonemic restoration : The brain creates missing speech sounds. *Acoustical Science and Technology*, 27(6) :318–321, 2006.

[8] Paul Boersma and David Weenink. Praat, a system for doing phonetics by computer. *Glot International*, 5(9/10) :341–345, 2001.

[9] Ulric Neisser. *Cognition and reality : Principles and implications of cognitive psychology*. Henry Holt & Co., New York, 1976.

[10] Mark Scott, H. Henny Yeung, Bryan Gick, and Janet F. Werker. Inner speech captures the perception of external speech. *Journal of the Acoustical Society of America Express Letters*, 133(4) :286–293, 2013.

[11] Mark Scott. Speech imagery recalibrates speech-perception boundaries. *Attention, Perception & Psychophysics*, 78(5) :1496–1511, 2016.