To do	Notes
Design	X
Get stims	
Write manuscript	(Julia's turn to edit)
Program	Julia/Naseem/ZhaoBin
Spot check program	Janna
Make OSF page and prereg	Janna & anyone who wants to help
IRB	Janna & anyone who wants to help
Protocol & set up participant spreadsheets	
Pilot	
Set up youcanbook.me and calendars	
Print IRB and protocol	
Write analysis script in R	Janna & Naseem

Do listening effort measures measure listening effort?

Anyone who has attended a noisy sporting event or cocktail party understands the challenges of listening to speech in noise. On a basic level, background noise has detrimental consequences for the intelligibility of speech. For example, there is potential for spectral and temporal overlap between the target speech and the background noise, leading to masking of finer phonetic details. In some instances, the target speech and noise can be difficult to separate, which may result in further decreased intelligibility (Brouwer, Van Engen, Calandruccio, & Bradlow, 2012). While measures of spoken word recognition accuracy can be useful in quantifying this intelligibility decrease, these measures do not thoroughly reflect the attentional and cognitive resources needed to achieve a given accuracy. In other words, listening to spoken language in noisy listening conditions may result in both decreases in the intelligibility of speech and increases in the cognitive costs associated with processing it. The latter result is known as listening effort (LE). Given the assumption that humans possess a finite pool of cognitive resources (Kahneman, 1973), as listeners shift more resources toward listening in difficult settings, fewer remain to complete additional tasks (Pichora-Fuller et al., 2016).

Paradigms that measure LE tend to indicate that LE increases with the addition of louder background noise (Brown & Strand, 2018; Johnson, Xu, Cox, & Pendergraft, 2015; Ng, Rudner, Lunner, Pedersen, & Rönnberg, 2013; Strand, Brown, Merchant, Brown, & Smith, 2018). However, even without listening, noise may add general stress and fatigue, increasing cognitive load. There is evidence that noise, a ubiquitous source of environmental stress, would impact human performance. In a meta-analytic review, Szalma and Hancock (2011) discuss this relationship. Maximal adaptability theory, the most convincing explanation according to the authors, proposes that while humans are considerably adaptable to environmental stressors, there is a point at which humans are unable to adapt, which leads to a decrease in performance (Szalma & Hancock, 2011). Szalma and Hancock found that cognitive tasks were considerably more subject to decrements in performance than visual perceptual tasks or motor tasks. The type of noise was also relevant (speech noise was more disruptive than non-speech noise) as well as the noise schedule (intermittent noise is slightly more disruptive than continuous noise).

Current speech research has yet to fully dissociate the cognitive load associated with listening to speech in noise (LE) and the cognitive load associated with noise overall. In other words, is the widely-observed increase in LE directly due to *listening* to speech, or does the background noise produce cognitive fatigue on its own? A recent study attempted to address this question (Brown & Strand, 2018). One common class of LE paradigms involves listening to and repeating words in varying levels of background noise while simultaneously completing an additional task. As background noise increases, performance on the additional task (measured by reaction time) tends to decrease. These LE paradigms are known as dual-task measures and are cited as the most robust and reliable LE paradigm (Gagné, Besser, & Lemke, 2017). Brown and Strand used a dual-task measure in the main experiment, but prior to the main experiment, they collected pilot data in order to address whether the secondary task in this paradigm is itself sensitive to changes in noise. Participants completed the secondary task (a visual task involving spatial and numerical reasoning) in two levels of background noise while listening to and repeating words. There was no observed change in reaction time with increased background noise, so any changes in LE observed in the main experiment were attributed to listening effort rather than noise-related fatigue.

Despite these data, we cannot immediately conclude that all LE paradigms can distinguish listening effort from noise-related cognitive effort. Listening effort is operationalized in many ways, and a previous study by our lab found that these paradigms are inconsistently and weakly intercorrelated (Strand et al., 2018). In a later study, we assessed whether the relationship between audiovisual integration and listening effort was, in part, dependent on whether participants completed the dual task or the recall paradigm (LEAVIS). Recall paradigms are frequently used in LE research and require listening to and recalling series of words in varying levels of noise. In general, fewer words are recalled in more difficult listening conditions because fewer resources remain to encode and recall spoken words (Rabbitt, 1968), and the assumption is that those resources constitute listening effort. In the experiment, participants were randomly assigned to complete the experiment in the dual task or recall paradigm, and we saw two very different patterns of data. Previous research has also found weak intercorrelation between recall measures and subjective effort ratings (Johnson et al., 2015), as well as between cognitive spare capacity tests and recall measures (Mishra, Lunner, Stenfelt, Rönnberg, & Rudner, 2013). Together, these studies provide converging evidence that two measures of LE may not reflect the same underlying construct.

The detrimental effect of noise on human performance would suggest that the secondary tasks used in dual task measures may be impacted by noise. A few of the tasks require some higher-level cognitive functions (such as the noun-classification task), which tend to be most impacted by noise (Szalma & Hancock, 2011). While the results of Brown and Strand's pilot study contradict this hypothesis, we are hesitant to extend these results to all secondary tasks in dual-task paradigms. One limitation of this study, for example, is that participants completed the task under energetic masking noise. Speech noise tends to be more disruptive than non-speech noise (Szalma & Hancock, 2011), so it could be that the detrimental effect of noise simply was not observed under these conditions. Additionally, the dual-task tested in this pilot experiment does not involve words in any way, whereas recall-measures all require listening to and repeating words.

Recall measures show a robust effect of sensitivity to changes in background noise (Johnson et al., 2015; Ng et al., 2013; Pichora-Fuller, Schneider, & Daneman, 1995; Strand et al., 2018), but there is considerable evidence that working memory performance is adversely affected by noise, independent of listening. If this were the case, participants completing an orthographic memory task (i.e. a memory task with visually presented stimuli) would perform more poorly with the addition of background noise. The most robust evidence for this hypothesis comes from the irrelevant speech effect (a phenomenon later broadened to the "irrelevant sound effect"). The irrelevant speech effect (ISE) was first introduced in a study by Colle and Welsh, who found that recall of visually presented digits was impaired when an unfamiliar language was played in the background (Colle & Welsh, 1976). Multiple studies have replicated this effect since (Baddeley & Salamé, 1986; Beaman, Philip Beaman, & Jones, 1998; Norris, Baddeley, & Page, 2004; Salamé & Baddeley, 1982). The ISE can even have a retroactive effect (Norris et al., 2004). In a series of experiments, the authors demonstrated that irrelevant speech presented after the to-be-recalled list of digits had the same effect as irrelevant speech presented throughout that list. Rabbitt, whose 1968 study became a cornerstone for recall-based measures, also found a retroactive effect but for auditorily presented words (Rabbitt, 1968).

Existing theoretical frameworks about the mechanism of the ISE may provide insight into the conditions under which noise impairs working memory. Baddeley's 1986 model of working memory suggests that irrelevant speech automatically enters a working memory component called the phonological loop, which interferes with encoding of the relevant items (CITE BADDELEY). However, there was later evidence to suggest that speech itself is not the root cause of the irrelevant speech effect. Jones and Macken posited that what instead causes disruption of working memory is the physical change in the sound. In other words, irrelevant background speech adversely affects working memory capacity due to the continuous modulation of the sound. For two experiments, Jones and colleagues played a series of changing tones and found that it was as disruptive to working memory capacity as speech. Next, they found that a single syllable changing only in pitch was as disruptive to working memory performance as the changing tones were. Thus, Jones et al. posed what is now called the "changing-state hypothesis" (Jones & Macken, 1993). Further evidence for the changing-state hypothesis was provided by Tremblay, Macken, and Jones, who did a similar experiment but with bursts of white noise that changed in frequency. Serial recall of visually presented memory items were sensitive to the changes in white noise frequency for both visual-spatial items (e.g. the location of a point on a grid) and visual-verbal items (e.g. sequences of consonants). The authors concluded, again, that speech alone does not disrupt working memory capacity; instead, any continuously modulating sound has this capability (Tremblay, MacKen, & Jones, 2001). Additionally, two empirical studies compared other sounds (speech, vocal and instrumental music) to white noise and quiet. In both studies, while speech and music in the background impaired serial recall, white noise (or pink noise) did not differ significantly from quiet (Jones, Miles, & Page, 1990; Salamé & Baddeley, 1989). Using these findings, Cowan proposed that changing background noise competes with attentional resources, so fewer remain to complete the memory task (Cowan, 1998).

Together with the evidence that a modulating sounds can disrupt working memory capacity, we have reason to believe that recall-measures may not be a reliable way to quantify listening effort. However, the changing-state hypothesis suggests that the type of background noise may affect this relationship.

In the present study, we aim to address background noise as a potential confound in listening effort tasks and assess the task conditions under which this confound is present. Participants will complete a visual-verbal working memory task analogous to those used in recall tasks as well as two tasks often used as secondary tasks in dual-task LE paradigms. Of the latter two tasks, one will be verbal (e.g. making a judgement about a word) while the other will be non-verbal (e.g. making a judgement about a vibrotactile stimulus). With these task manipulations, we hope to assess whether noise-induced cognitive effort varies with tasks and whether this variance can be attributed to recall vs dual-tasks or verbal versus non-verbal tasks.

Each task will be completed in quiet, two-talker babble (informational noise), energetic noise, and energetic noise that modulates in amplitude. With these noise conditions, we can assess the impacts of noise on cognitive performance as well as test the irrelevant speech effect and changing-state hypothesis as potential causes for noise-related effort. We hypothesize that performance on the vibrotactile task will not be affected by static energetic, modulating energetic, or informational masking noise. While cognitive tasks can be affected by noise, we have some data to suggest that these non-verbal, non-recall tasks are not adversely affected by noise (see Brown & Strand, 2018). Additionally, this dual-task is arguably more of a perceptual task than a cognitive task, which tend to be less disrupted by noise regardless (Szalma & Hancock, 2011). For this reason, we also hypothesize that performance the verbal, non-recall task will not be affected by any type of background noise. For the recall task, we expect that the type of background noise will moderate the relationship between noise and working memory capacity. For informational masking, we expect that performance will decrease relative to quiet. The irrelevant sound effect has been robustly shown in the literature, and the speech used in informational masking is sufficiently non-constant that, according to the changing-state hypothesis, it should cause a disruption in working memory. However, we do not expect to see any differences between quiet and static energetic noise. According to the changing-state hypothesis, steady-state noise is not sufficient to produce a disruption in working memory capacity, and there is experimental evidence in support of this claim (Jones et al., 1990; Salamé & Baddeley, 1989). However, we do expect a small decrease in performance for modulating energetic noise (relative to quiet and static energetic noise), though this decrease will likely be smaller than the decrease with speech noise. This pattern of data is consistent with the changing-state hypothesis as well as Szalma & Hancock's finding that speech noise tends to impair performance more than non-speech noise does.

Methods

Participants

Participants (N = XX) were young adults (ages 18-28) from the Carleton College community. All had self-reported normal or corrected-to-normal vision and no known hearing impairment. All procedures were approved by the Carleton College Institutional Review Board.

Stimuli

- Want words with similar orthographic qualities to ones we've already recorded (like BD3.1 words, but not those same words).
- Modulating noise changes amplitude, but average amplitude should be equivalent to the amplitude of the static energetic noise and informational noise
- Informational noise = two-talker babble (same one used for BD)
- Presented through noise-cancelling headphones in order to combat noise from the tactor / outside noise during the quiet condition

Procedure

Participants sat at a comfortable distance from a 21.5 inch iMac computer and wore noise-cancelling headphones. All completed three tasks, each of which contained four blocks. Each block consisted of XX trials and either quiet, static energetic noise, modulating energetic noise, or informational noise. Throughout the experiment, participants wore noise-cancelling headphones in order to attenuate any sounds from the tactor or from outside the testing booth. Trials in all three tasks involved reading and repeating words aloud so that task conditions are maximally similar to those in LE paradigms. We used XX lists with XX words each, and we chose words that are orthographically similar to words used in LE paradigms. Lists of words were counterbalanced so that, while participants saw each word exactly once, words were equally likely to appear in each task and in each noise condition.

Vibrotactile task

Participants wore the tactor on their dominant hand and placed it inside a box in order to attenuate the noise. We did so in order to avoid participants taking auditory (rather than sensory) cues from the tactor during quiet or easy noise blocks. The first block of the task was for familiarization, and this block was always completed before the five experimental blocks. Participants were first presented with two short pulses, two medium pulses, and two long pulses. During familiarization, participants identified 18 randomly-ordered pulses (six of each length) by pressing the appropriate button on the box. In the event of a wrong response, the correct answer was immediately displayed on the screen. In order to pass the familiarization procedure, participants needed an accuracy of at least 75% (14 out of the 18 pulses). If this threshold was not met, the entire familiarization block repeated. Once the familiarization block was passed, participants completed the remaining four blocks. In the main experiment, each trial contained a word presented on the computer screen along with the vibrotactile stimulus. Participants responded to the tactor stimulus and then repeated the word aloud. We recorded reaction time to the tactor task, and word repetition accuracy was coded offline by research assistants.

Semantic task

Experimental setup and noise conditions were identical to those used in the vibrotactile task. At the start of each trial, a word appeared on the screen. Participants then determined whether the presented word was a noun. For all noun words, they were instructed to press a button as quickly as possible. For all other words, no button response was necessary. We recorded reaction time to the noun task, and word repetition accuracy was coded offline by research assistants.

Running memory (N-back) task

Experimental setup and noise conditions were identical to those used in the vibrotactile and semantic tasks. Participants completed five blocks of the running memory task (McCoy et al., 2005; Morris & Jones, 1990; Sommers & Phelps, 2016; Sommers, Tye-Murray, Barcroft, & Spehar, 2015; Strand et al., 2018). We chose this procedure because it is nearly identical to the working memory task we use when we employ recall-based LE measures, but the stimuli are visually presented. In each block, participants were visually presented with 16 lists with two each of lengths five through twelve. Participants were instructed to repeat words aloud immediately after presentation, and, at the end of each list, recall the last four words in each list, in any order. Word lists and conditions were counterbalanced so that while participants saw each word exactly once, lists of words appeared equally often throughout the conditions. All participant files were coded offline by research assistants.

Methods

- All within participants
- Three tasks: recall, SDT, Vibrotactile
- Four noise conditions: silence, energetic (steady), energetic (modulating), informational
- Block noise type within task type (so they do all four recalls in a block), counterbalancing everything
- For all tasks, participants verbally repeat the words in addition to doing the other thing they're supposed to do

Notes from lab meetings

- Outstanding questions
    - How can we block out noise in the quiet condition?
        - Noise cancelling headphones
        - Making a box for the tactor?
- What's the big question?
    - Is it possible that recall measures of listening effort are picking up on both listening effort and memory costs that are just due to noise and not to listening effort?
- Concerns:
    - If vibrotactile is also impaired (but not as much as recall) we can't say it's because of recall it may be because of VERBAL differences not recall vs vibrotactile
    - Could remove that concern by using a semantic dual task on orthographic words

References

Baddeley, A., & Salamé, P. (1986). The unattended speech effect: perception or memory? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *12*(4), 525–529.

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *23*(2), 459.

Beaman, C. P., Philip Beaman, C., & Jones, D. M. (1998). Irrelevant Sound Disrupts Order Information in Free Recall as in Serial Recall. *The Quarterly Journal of Experimental Psychology Section A*, Vol. 51, pp. 615–636. https://doi.org/10.1080/713755774

Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. *The Journal of the Acoustical Society of America*, *131*(2), 1449–1464.

Brown, V. A., & Strand, J. F. (2018). Noise increases listening effort in normal-hearing young adults, regardless of working memory capacity. *Language, Cognition and Neuroscience*, 1–13.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 17–31.

Cowan, N. (1998). *Attention and Memory: An Integrated Framework*. Oxford University Press.

Gagné, J.-P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing*, *21*, 2331216516687287.

Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A comparison of two methods for measuring listening effort as part of an audiologic test battery. *American Journal of Audiology*, *24*(3), 419–431.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 19, pp. 369–381. https://doi.org/10.1037//0278-7393.19.2.369

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology*, Vol. 4, pp. 89–108. https://doi.org/10.1002/acp.2350040203

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

LeCompte, D. C. (1994). Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(6), 1396–1408.

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *58*(1), 22–33.

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., & Rudner, M. (2013). Visual information can hinder working memory processing of speech. *Journal of Speech, Language, and Hearing Research*, *56*, 1120–1132.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology* , *81*(2), 111–121.

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., & Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *International Journal of Audiology*, *52*(7), 433–441.

Norris, D., Baddeley, A. D., & Page, M. P. A. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1093–1105.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37 Suppl 1*, 5S – 27S.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593–608.

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology*, *20*(3), 241–248.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*(2), 150–164.

Salamé, P., & Baddeley, A. (1989). Effects of Background Music on Phonological Short-Term Memory. *The Quarterly Journal of Experimental Psychology Section A*, Vol. 41, pp. 107–122. https://doi.org/10.1080/14640748908402355

SalamÉ, P., & Baddeley, A. (1990). The effects of irrelevant speech on immediate free recall. *Bulletin of the Psychonomic Society*, *28*(6), 540–542.

Sommers, M. S., & Phelps, D. (2016). Listening effort in younger and older adults: A comparison of auditory-only and auditory-visual presentations. *Ear and Hearing*, *37 Suppl 1*, 62S – 8S.

Sommers, M. S., Tye-Murray, N., Barcroft, J., & Spehar, B. P. (2015). The effects of meaning-based auditory training on behavioral measures of perceptual effort in individuals with impaired hearing. *Seminars in Hearing*, *36*(4), 263–272.

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., & Smith, J. (2018). Measuring listening effort: Convergent validity, sensitivity, and links with cognitive and personality measures. *Journal of Speech, Language, and Hearing Research: JSLHR*, *61*, 1463–1486.

Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: a meta-analytic synthesis. *Psychological Bulletin*, *137*(4), 682–707.

Tremblay, S., MacKen, W. J., & Jones, D. M. (2001). The impact of broadband noise on serial memory: Changes in band-pass frequency increase disruption. *Memory* , *9*(4), 323–331.

Manuscraps

Task conditions other than type of background noise may moderate the relationship between noise and performance on working memory tasks. The vast majority of irrelevant sound effect studies make use of procedures in which the same consonants are presented in a random order. In a 1990 study, Salamé and Baddeley address this and opt for more naturalistic stimuli (bi-syllabic words). They also chose to do a free-recall task (where items can be recalled in any order), whereas the irrelevant sound effect papers at the time employed serial recall (where memory items are recalled in the order that they were presented). Salamé and Baddeley found no irrelevant speech effect (SalamÉ & Baddeley, 1990) and concluded that serial recall was necessary in order to produce it. There were some theoretical justifications as to why free-recall tasks are immune to irrelevant sound. Namely, the changing-state hypothesis suggests that order information is disrupted by irrelevant sound, and subsequent studies also found that tasks not requiring rote rehearsal of serial order were not susceptible to irrelevant speech (Beaman & Jones, 1997; LeCompte, 1994; SalamÉ & Baddeley, 1990). However, Beaman & Jones did a follow-up study, also using words rather than consonants, and found that, when participants were assigned to the same task but with either free recall or serial recall, the irrelevant speech effect was present. They then did the same experiment but with longer lists of items and found the same, if attenuated, effect (Beaman et al., 1998). Both experiments are in direct contrast to the results of Salamé and Baddeley. The authors concluded that, in free-recall, order information is taken into account when recalling item information, even though participants are not instructed to do so.

Studies with methodological variation can produce different patterns of data, which poses a challenge in reconciling these findings and extending them to the listening effort literature. Relatively few studies have examined the effect of background noise on free recall versus serial recall, and recall-based listening effort paradigms tend to employ free recall. However, although participants may recall the memory items in any order, there is still some implicit order information in the task. We can also allay some concerns about memory items. Relatively few irrelevant speech effect studies have used stimuli other than consonant letters or digits, whereas recall paradigms tend to use words in addition to consonant letters and digits. However, we do know that recall paradigms correlate reasonably with one another (Strand et al., 2018) and that the irrelevant sound effect has been shown for lists of words (Beaman & Jones 2018).

	Recall	Semantic Dual Task	Vibrotactile
Silence			
Energetic, steady			
Energetic, modulating			
Informational			