Paying Attention to Speech: The Role of Cognitive Capacity and Acquired Experience

Bar Lambez[1], Galit Agmon[1], Paz Har-Shai[1], Yuri Rassovsky[1,2,3]
and Elana Zion Golumbic*[1]

[1] The Gonda Multidisciplinary Center for Brain Research, Bar Ilan University, Ramat Gan, Israel

[2] Department of Psychology, Bar Ilan University, Ramat Gan, Israel

[3] The Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles

*corresponding author

Dr. Elana Zion Golumbic

The Gonda Multidisciplinary Center for Brain Research

Bar Ilan University, Ramat Gan, Israel

elana.zion-golumbic@biu.ac.il

+972-58-593-3840

Abstract

Managing attention in multi-speaker environments is a challenging feat that is critical for human performance. However, why some people are better than others in allocating attention appropriately, remains highly unknown. Here we investigated the contribution of two factors - Cognitive Capacity and Acquired Experience - to performance on two different types of Attention task: Selective Attention to one speaker and Distributed Attention among multiple concurrent speakers. We compared performance across three groups: Individuals with low (n=20) and high cognitive capacity (n=26), and Aircraft Pilots (n=25), who have gained extensive experience on both Selective and Distributed attention to speech through their training and profession. Results indicate that both types of Attention benefit from higher Cognitive Capacity, suggesting reliance on common capacity-limited resources. However, only Selective Attention was further improved in the Pilots, pointing to its flexible and trainable nature, whereas Distributed Attention seems to suffer from more fixed and hard-wired processing-bottlenecks.

Introduction

Natural environments are characterized by an abundance of sounds bombarding the auditory system and competing for our attention. Focusing attention appropriately on relevant speech in adverse listening settings is a feat that many individuals find extremely challenging. Moreover, different contexts may require applying different attentional strategies. Some tasks require Selective Attention, i.e., focusing on a single speech-source while ignoring all others and avoiding distraction (Beaman, Bridges, & Scott, 2007; Broadbent, 1954; Cherry, 1953; Elliott & Briganti, 2012; Zion Golumbic et al., 2013), whereas other contexts require precisely the opposite: Distributing Attention among several speakers and gleaning semantic information from all of them (Baldock, Kapadia, & van Steenbrugge, 2018; Brungart, Kordik, & Simpson, 2005; Getzmann, Golob, & Wascher, 2016; Gygi & Shafiro, 2012). Although the term "Attention" is used to describe both processes, these tasks require overcoming fundamentally different perceptual and cognitive challenges. The primary challenge of Selective Attention is separating among acoustic sources and suppressing irrelevant portions of the auditory scene, a feat that becomes progressively difficult as the acoustic and linguistic overlap between concurrent speech increases (Bronkhorst, 2015; Drullman & Bronkhorst, 2004; Kidd et al., 2016; Neely & LeCompte, 1999; Oswald, Tremblay, & Jones, 2000). In contrast, the challenge of Distributed Attention stems from the limited nature of internal processing resources, that arguably pose inherent 'bottlenecks' for processing concurrent speech, the precise nature of which is still heavily debated (Bronkhorst, 2015; Deutsch & Deutsch, 1963; Driver, 2001; Lachter, Forster, & Ruthruff, 2004; Lavie, Hirst, de Fockert, & Viding, 2004)

The current study seeks to understand whether Selective and Distributed attention to speech rely on common underlying cognitive mechanisms and to highlight

3

the similarities and distinctions between these two aspects of human attention. Specifically, we test the contribution of two factors - Cognitive Capacity and Acquired Experience - to performance on these tasks, as a mean for examining the processes underlying Selective and Distributed attention to speech.

Cognitive capacity is positively associated with performance on a range of attentionally-demanding tasks, and has been attributed to the availability of more cognitive resources as well as more effective top-down control over the allocation of these resources (Sörqvist & Rönnberg, 2014; Tsuchida, Murohashi, Katayama, & Murohashi, 2012; Wiemers & Redick, 2018). However, whether Selective and Distributed rely on shared attentional-resources, as proposed by some (Fusser et al., 2011; Kahneman, 1973; Salmela, Moisala, & Alho, 2014), or are differently affected by the availability of cognitive resources (Elliott & Briganti, 2012), remains currently unknown. Besides the contribution of cognitive capacity to performance, another fundamental question is how amenable attention is to improvement through training, and if it is possible to master attention-as-a-skill. Investigating whether performance can be further improved with experience touches upon the 'fixed' or 'flexible' nature of the processing bottlenecks and limitations they impose on performance.

In order to test the relative contributions of cognitive capacity and acquired experience to Selective and Distributed Attention to speech, here we compared performance on increasingly demanding attentional tasks, in three experimental groups. The role of cognitive capacity was assessed by comparing individuals with low- and high cognitive capacity, operationalized by an estimate of general intellectual functioning and working memory capacity. The role of acquired experience was studied by comparing performance of aircraft pilots with matched high cognitive capacity controls. Aircraft pilots afford a unique opportunity for studying the effects of acquired

4

experience on attention, as their intense professional training and daily job pose high demands for both Selective and Distributed Attention to speech (Gopher, 1982; Gygi & Shafiro, 2012). Studying performance in this highly trained population allows us probe the limits of the speech-processing and attention systems, testing the extent to which humans are able to improve their abilities as well as identifying hard-wired bottlenecks that cannot be overcome.

Methods

Participants

Participants in this study were males between the ages of 24-49 (median = 29), with normal hearing and vision, and no history of psychiatric or neurological disorders or ADHD. They were recruited from three groups: Pilots, High Cognitive Capacity group (HC), and Low Cognitive Capacity group (LC; see Table 1). The Cognitive Capacity of each participant was evaluated by testing for general intellectual functioning (g-factor Test of Nonverbal Intelligence 3rd Edition; TONI-3) and Working Memory Capacity (WMC, Wechsler Adult Intelligence Scale-Fourth Edition). Since these two measures are largely correlated (Borella, Pezzuti, De Beni, & Cornoldi, 2019), rather than distinguishing among them, here we consider them jointly as reflecting Cognitive Capacity.

The Pilots group consisted of 25 active, commercial pilots, recruited from commercial airline companies, between the ages of 25-45. All Pilots had continuous flying experience of at least 7 years, allowing for the development of expertise (Ericsson, Krampe, & Tesch-Römer, 1993). The HC group consisted of 26 participants, who were matched to the Pilots on age (p=1.0), g-factor (p=0.992), and WMC (p=1.0). The LC group consisted of 20 participants who were also matched to the Pilots on age (p=1.0), but had significantly lower g-factor and WMC than the HC and Pilot groups

5

(both p<0.001). Two participants (1 Pilot, 1 HC) were rejected from data analysis due to technical errors. Demographic data was collected using a self-report socio-demographic questionnaire.

The study was preregistered prior to commencement of data collection through the Center for Open Science (COS; https://osf.io/tk6gc), describing the specific experimental design and power-analysis for a-priori estimation of the required group-size. The original preregistration focused only on the HC and Pilot groups and recruitment of the LC group was decided upon at a later stage. The study was approved by the Ethics Committee of Bar Ilan University, and signed informed consent was obtained from each participant prior to the experiment. Participants were reimbursed for their time and travel expenses.

Table 1. Participants Demographic characteristics

	Pilots (n=24)		HC (n=25)		LC (n=20)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age (Years)	29.42	5.52	29.16	4.4	29.7	7.25
g-factor	126.54	10.99	126.5	9.66	93.4	10.42
Working Memory	13.75	2.78	13.42	2.06	9.5	2.26

Stimuli

The stimuli were comprised of a list of short Hebrew words, including mono-syllabic nouns (e.g., pitcher; "*Kad*") and digits (e.g., seven; "*Shevah*"). The words were recorded by two male and two female speakers, rendering the speakers clearly distinguishable. Audio editing of the individual words and their combination into sequences were performed using Audacity (www.audacityteam.org) and Matlab (The MathWorks, Inc). The perceived loudness of each speaker was equated offline and verified using inter-rater reliability testing. Word lengths varied between 800 ms to

6

1100 ms, and they were concatenated into 55-second long sequences, separately per speaker. The inter-stimulus-interval (ISI) between words varied across speakers between 600-750ms (but was held constant within speaker), in order to minimize common onsets and offsets. Sequences by different speakers were then combined to create diotic multi-talker scenes, according to the different experimental conditions, as explained below.

Experimental procedure

We designed two parallel word target-detection tasks, probing Selective and Distributed Attention to speech under multi-speaker conditions (Figure 1). The Selective Attention task, required participants to attend to one designated ("attended") speaker and respond to the target-word uttered only by this speaker. In the Distributed Attention condition, participants were instructed to respond to the target word spoken by <u>any</u> of the concurrently presented streams. The two tasks were matched for low-level acoustics, and the level of difficulty was parametrically varied by increasing the number of competing speakers from two to four. The same target word (tree; *"Etz"*) was used in all conditions and occurred 8 or 9 times per trial. In the Selective Attention condition additional 'catch-targets' could occur equi-probably in the unattended sequences, which participants were accordingly supposed to ignore.

Selective and Distributed trials were presented in four separate blocks (two per condition), in counter-balanced order across subjects. Each block consisted of 12 trials where the Number of Speakers was pseudo-randomized between 2 to 4, as well as two trials where only one speaker was presented (this condition was not included in the statistical analysis). The experiment was conducted in a sound-attenuated booth, and speech stimuli were presented through headphones (Sennheiser, HD 280 Pro). The experiment was programmed using Psychopy software (Peirce et al., 2019), and

responses were recorded using a response box (Cedrus RB 840). Response Accuracy and Reaction Times (RTs) were extracted for each target and used for statistical analysis.
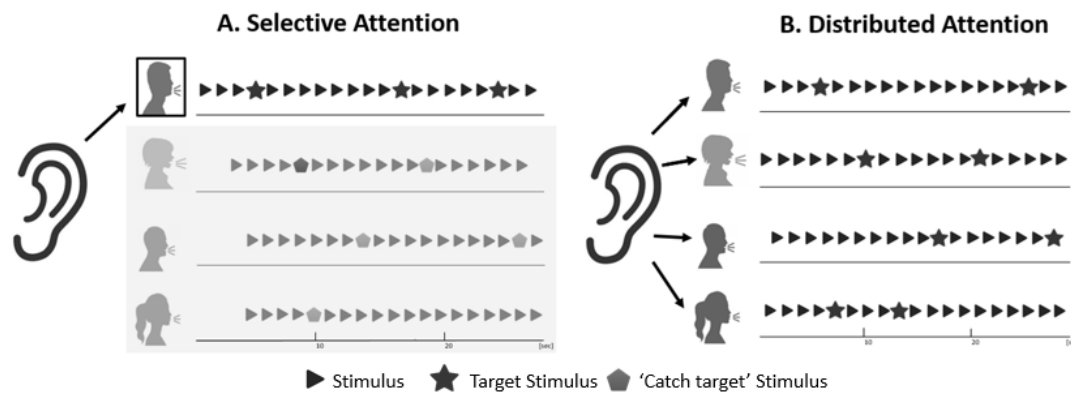

Figure 1. A. Example of the trial structure in the Selective Attention condition, in the case of four concurrent speakers. The speech stimulus of each speaker consisted of a sequence of Hebrew words (indicated by triangles). In Selective Attention blocks, participants were instructed to attend to one designated "attended" speaker and respond to a target-word (marked with star icon) uttered only by this speaker, while ignoring all the other speakers. A similar number of catch-targets (the target-word uttered by an unattended speaker, marked with a pentagon icon) were distributed among the other unattended speakers. The number of concurrent speakers varied between two and four, across trials. **B.** Example of the trial structure in the Distributed Attention condition, in the case of four concurrent speakers. The stimuli are similar to those used in A, however here participants were instructed to respond to the target-word uttered by *any* of speakers, requiring participants to distribute auditory attention among all concurrent speakers. Here too, the number of concurrent speakers varied between two and four, across trials.

Statistical Analysis

To test for differences between conditions, we fit the behavioral responses with mixed effects linear regression models, using R's lme4 package (Bates, Maechler, Bolker, & Walker, 2015): for RT data we used a linear mixed effects regression and for the accuracy data we used a generalized linear mixed effects regression. The advantage of mixed effects models is that they account for variability between subjects and correlations within the data (Baayen, Davidson, & Bates, 2008); in the case of accuracy,

they also allow accurate analyses of binomial data in repeated-measures designs (Jaeger, 2008).

The model fitted for the RTs included three factors and their interactions as fixed variables: Attention (*Selective, Distributed*), Group (*LC, HC, Pilots*) and Number of Speakers (*2, 3, 4*). The factor Attention was sum coded (i.e., the beta estimate represents deviance from the grand mean), the factor Group was treatment coded (i.e., one beta compares Pilots to HC and another beta compares LC to HC), and the factor Number of Speakers was difference coded (i.e., one beta codes for the difference between 3 and 2 speakers, and the second beta codes for the difference between 4 and 3 speakers). We included by-subject random intercepts, as well as by-subject random slopes for Attention and Number of Speakers. To better comply with the assumption of normality, we used the logarithmic transformation of RT as the dependent variable. As a pre-processing step, we removed from the data any RTs that were below 300 ms or above 2000 ms, thus removing 0.6% of the total data. P-values were obtained using the Satterthwaite approximation of degrees of freedom (Satterthwaite, 1946), which is implemented in R's lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). For accuracy, we used the same fixed variables and coding schemes. By-subject random intercepts and by-subject random slopes for attention were included. For hit trials in which RTs were either below 300 ms or above 2000 ms, accuracy was set to be 0. The reported p-values are based on asymptotic Wald tests which are included in the summary of R's glmer function.

9

Results

Performance was relatively high on all tasks, with an average of 90.5% detection rate across conditions and groups (±6.4% SD). Moreover, in the Selective Attention condition the false-alarm rate was extremely low (0.9%±1.3%; responses to the target-word spoken by unattended speakers), indicating that participants understood the difference between the Selective and Distributed tasks and followed instructions accordingly. Mixed-linear regression models were used to assess the effects of Attention Type (Selective vs. Distributed), Number of Speakers (2spk, 3spk and 4spk) and Group (LC, HC and Pilots) on accuracy and reaction times (RTs). Results across all conditions and groups are shown in Figure 2, and the full statistical results of all contrasts in the regression models are listed in Tables S1 & S2.
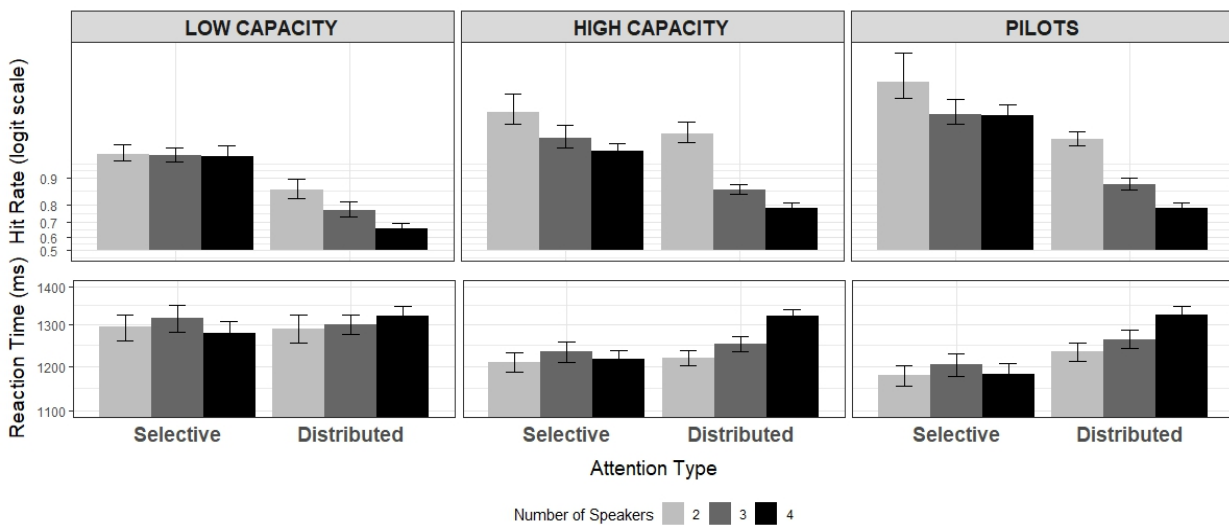

Figure 2: Mean Hit Rates (top) and RTs (bottom) on the Selective and Distributed Attention tasks, as a function of the number of concurrent speakers (2-4), for each of the three experimental groups: Low Cognitive Capacity, High Cognitive Capacity and airplane Pilots.

Comparison of performance between the two Attention Type tasks indicated that the Distributed Attention condition was more difficult than the Selective Attention condition, as manifest in reduced accuracy (β=-0.8, z=-5.6, p<0.0001) and prolonged RTs (β=0.02, t=3.5, p=0.0009). Performance also generally deteriorated as the number

of concurrent speakers increased, observed in the 3spk vs. 2spk contrast (Accuracy: β=-1.3, z=-5.9, p<0.0001; RT: β=0.02, t=3.5, p=0.0008) and the 3spk vs. 4spk contrast (Accuracy: β=-0.5, z=-3.2, p=0.001; RT: β=0.02, t=3.3, p=0.001). Some of these effects interacted with the Attention Type, indicating that increasing the number of speakers adversely affected performance on the Distributed Attention task, and less so (or not at all) in the Selective Attention task [RTs (*4spk vs. 3spk) x Attention Type*: β=0.03, *t*=6.2, p<0.0001; Accuracy (*3spk vs. 2spk) x Attention Type:* β=-0.5, *z*=-2.1, p=0.03] (Figure 3).
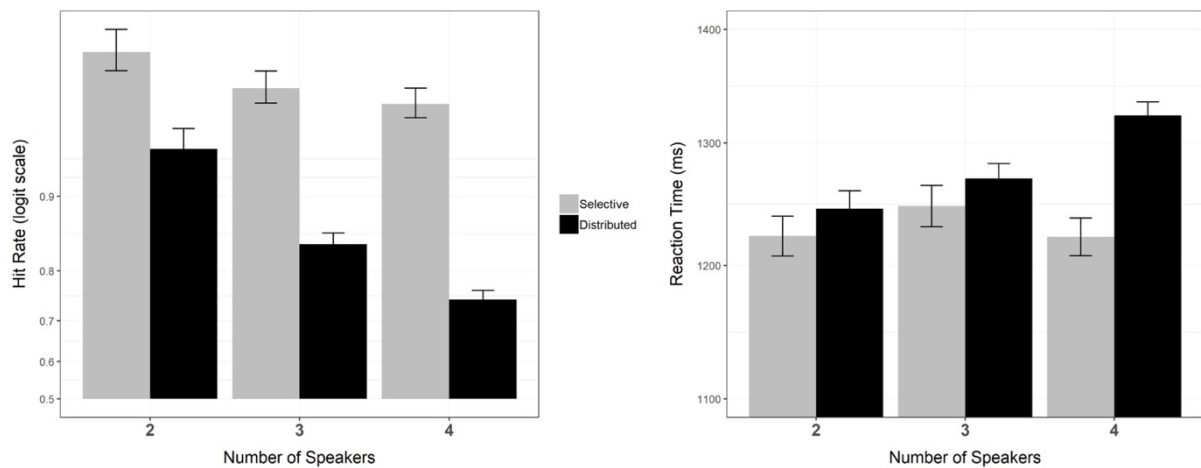

Figure 3: Mean Hit Rates (left) and RTs (right) on the Selective and Distributed Attention tasks, as a function of the number of concurrent speakers (2-4), collapsed across groups. Overall performance was worse in the Distributed Attention condition. Increasing the number of speakers had a more adverse effect on performance in the Distributed vs. Selective Attention condition. Error bars indicate SEM.

With regard to differences between the three Groups (Figure 4), the LC group demonstrated significantly lower accuracy than the HC group (β=-0.8, *z*=-3.9, p<0.0001), and a trend towards slower RTs (β=0.04, *t*=1.8, p=0.08). This contrast did not interact significantly with Attention Type, suggesting that both types of attention were similarly influenced by the difference in Cognitive Capacity between these two groups. Some additional higher-order interactions of the differences between the LC and HC groups with the Number of Speakers were also significant [(*LC vs. HC) x (2spk*

11

vs. 3spk): Accuracy β=0.9, *z*=3.5, p=0.0005; [*(LC vs. HC) x (3spk vs. 4spk):* RTs: β=-0.02, t=-2.4, p=0.02]. These interactions likely stem from floor effects observed in the LC group, who displayed overall worse performance regardless of the number of speakers.

Comparison of performance between the HC group and the Pilot group, did not reveal any main effects of Group on accuracy or RTs (RTs: p=0.7, Accuracy: p=0.07), however there was a significant interaction with Attention Type [*(HC vs. Pilots) x AttentionType,* RTs: β=0.02, *t*=2.4, p=0.02, Accuracy: β=-0.4, *z*=-2, p=0.05)]. This indicates that in the Selective Attention task the Pilot group performed better than their HC matched controls, however the two groups did not seem to differ in their performance on the Distributed Attention task.
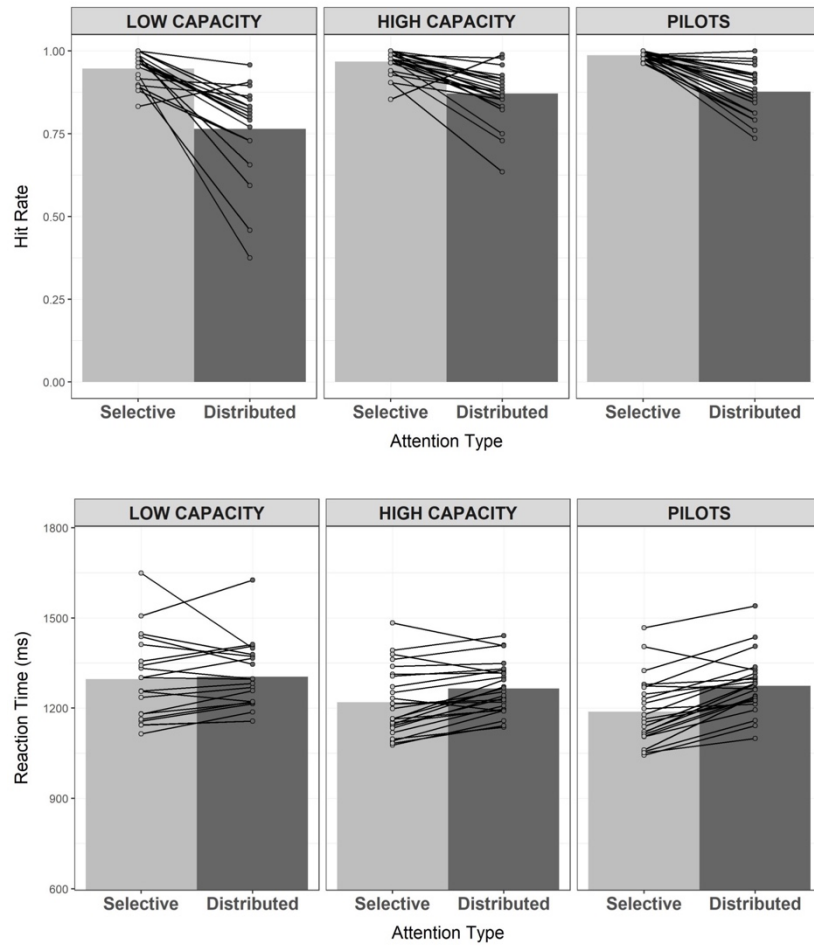
12

Figure 4: Mean accuracy (top) and RTs (bottom) in the Selective and Distributed conditions, separately for each of the three groups. Black lines connect between the scores of individual participants on the Selective and Distributed tasks.

Discussion

The current study aimed at understanding the contributions of Cognitive Capacity and Acquired Experience to performance of Selective and Distributed Attention to speech. We find that both Selective and Distributed Attention are improved in individuals with high Cognitive Capacity, suggesting some reliance on common executive resources. However, these two facets of attention seem to differ in the processing bottlenecks encountered as well as in their fixed or flexible nature. Specifically, results suggest that Selective Attention is a highly effective and malleable process that can withstand substantial adverse conditions. In contrast, Distributed Attention suffers from more severe processing bottlenecks that are arguably of a fixed nature and are less prone to improvement through training.

The comparison between Selective and Distributed Attention indicates that not only is the latter more effortful (Baldock et al., 2018), but they are differently affected by increasing the number of concurrent speakers. For Selective Attention, we find only a slight decrease in target-detection accuracy and no impact on RTs as the number of concurrent speakers increased from two to four. This ability to withstand severe acoustic and informational masking, is in line with accounts suggesting that distractor-speakers are not fully segregated among themselves but rather treated collectively as acoustic 'background' (Hausfeld, Riecke, Valente, & Formisano, 2018; Puvvada & Simon, 2017). This pattern demonstrates a high degree of resilience of the auditory and/or attention system for dealing with background noise, as long as the cognitive task is limited to processing only one speaker. In contrast, Distributed Attention does not allow designating portions of the acoustic scene as 'background', and increasing the number of speakers also increases the perceptual and cognitive-load of the task itself (Baldock et al., 2018). This resulted in a sharper decline in accuracy and prolonged RTs

14

as the range of Distributed Attention broadened from two to four speakers. This pattern indicates that attempting to process multiple concurrent speech-inputs draws upon perceptual and cognitive resources that are of a more limited nature, either at the level of the auditory system (Murphy, Spence, & Dalton, 2017) or at higher linguistic levels (Lachter et al., 2004).

The capability and limitations for parallel speech processing have been the focus of long-standing theoretical debates. However, much of this has been fueled by Selective Attention studies focusing on whether so-called unattended speech is processed (Beaman et al., 2007; Ding et al., 2018; Driver, 2001; Lachter et al., 2004; Röer, Körner, Buchner, & Bell, 2017; Wood & Cowan, 1995). The use of a Distributed Attention paradigm, as used here, provides a substantially more direct way for evaluating the ability and limitations of encoding, processing and responding to concurrent speech. Results highlight the fact that Distributed Attention suffers more substantially from processing 'bottlenecks' relative to Selective Attention, and provide insight into the nature of these bottlenecks. We now turn to discuss the role of Cognitive Capacity and Acquired Experience in mitigating these bottlenecks.

Both Selective and Distributed Attention require top-down control for allocating processing resources according to behavioral goals (Murphy et al., 2017). The Cognitive Capacity underlying top-down control varies across individuals and is strongly associated with a variety of executive processes, including selective, sustained and divided attention (Carroll, 1993; Kane & Engle, 2002). Previous studies have demonstrated that individuals with low WMC can be more susceptible to distraction (Hughes, 2014; Sörqvist & Rönnberg, 2014), incur a higher cost for attention switching (Lin & Carlile, 2015), display poorer speech-in-noise intelligibility (Keidser, Best, Freeston, & Boyce, 2015; Oberfeld & Klöckner-Nowotny, 2016), and exhibit overall

15

slower processing speed (Fry & Hale, 1996). In line with these findings, we find that the LC group displayed poorer performance on both Selective and Distributed attention tasks, relative to the HC group. Moreover, there was no apparent difference in the effect of Cognitive Capacity on the two tasks. The plethora of attention-related deficits associated with reduced Cognitive Capacity is in line with perspectives suggesting that even though the behavioral goals and required cognitive operations may differ across attentional tasks, they nonetheless rely on overlapping attentional-resources and common neural substrates (Fusser et al., 2011; Humes, Lee, & Coughlin, 2006; Kahneman, 1973; Kiyonaga & Egner, 2013; Salmela et al., 2014; Salo, Salmela, Salmi, Numminen, & Alho, 2017),

Beyond an individual's Cognitive Capacity, performance on difficult tasks can often improve substantially with training. Learning, and its accompanying neural plasticity, is one of the hallmarks of the nervous system, allowing organisms to adapt behavior to changing environmental and internal goals. Observing improvement by training is indicative of inherent flexibility of the underlying cognitive process (Buitenweg, Murre, & Ridderinkhof, 2012). Conversely, when no degree of training can improve performance beyond a certain level, this suggests a 'hard bottleneck', reflecting the upper limit of the system (Enns, Kealong, Tichon, & Visser, 2017). Comparison between the Pilot group and their matched HC controls provided a unique opportunity to assess whether the cognitive processes underlying Selective and Distributed Attention, are amenable to improvement through extensive training and experience. In the cockpit, Pilots need to meet extremely high standards of attention to multiple concurrent sources of speech, manifest both by monitoring several radio-channels (Distributed Attention), as well as focusing on one relevant radio-channel and disregarding irrelevant channels (Selective Attention). Their experience in performing

16

these tasks builds on years of intense training and constant maintenance of these skills throughout their career (Hilburn, 2004). Indeed, it is difficult to imagine a more intense training regime for the refinement of attention skills to concurrent speech in humans, and consequently the Acquired Experience of Pilots substantially exceeds that gained through any experimental training program in both length and intensity.

Here, Pilots displayed an advantage on the Selective Attention task relative to their matched HC controls, which manifest in improved accuracy and faster RTs. This suggests that besides Cognitive Capacity, experience and training can improve Selective Attention to speech even further. In line with these results, previous studies have found improved Selective Attention performance in Pilots after successful completion of a two year long flight training program (Gopher, 1982), and similar advantages have been reported in air-traffic controllers (Arbula, Capizzi, Lombardo, & Vallesi, 2016). Critically, the amenability of Selective auditory attention performance to training is not limited to these highly-selective groups. Improvement in Selective Attention performance has been demonstrated after a 4-week dichotic-listening training program in healthy adults (Soveri et al., 2013), and similar training effects have been reported in typically developing children (Murphy, Moore, & Schochat, 2015), as well as children with dyslexia (Helland et al., 2018). These findings point to the flexible nature of auditory Selective Attention, which can be improved through formal training programs and as well as frequent practice (Mishra, de Villers-Sidani, Merzenich, & Gazzaley, 2014).

In contrast, our results suggest that Distributed Attention does not enjoy similar flexibility. Despite the rigorous training and Acquired Experience of Pilots, their performance was not significantly better their matched non-pilot controls. This points to an upper-limit on the ability to distribute attention among speech, that is not

17

amenable to improvement through training. This pattern supports limited-resources perspectives of concurrent speech processing (Lachter et al., 2004) and suggests that bottlenecks may be hard-wired, possibly dictated by the a-priori availability of cognitive resources. It is important to acknowledge that the comparison between Pilots and the HC group performed here focused on individuals who can already perform the Distributed Attention task quite well, due to their high Cognitive Capacity. Future research should further look into possible effects of training on improving Distributed Attention performance in individuals with initially lower Cognitive Capacity.

To conclude, understanding why some people are better than others in allocating attention to speech in complex multi-speaker environments is extremely important, given its ecological significance to many real-life situations. This study elucidates the crucial role of cognitive capacity in the ability to process multiple concurrent speakers and distribute internal resources among them in accordance to task-goals. Moreover, it highlights the capacity and limitation of extensive training for improving different types of attentional performance. These results contribute to long-standing theoretical debates regarding the nature of 'cognitive bottlenecks' that limit the ability to deal with the influx of sensory information encountered in natural environments, and shed new light on their flexible or fixed nature.

Author Contributions Statement

BL, YR & EZG designed the study, BL & PHS collected the data, BL & GA ran statistical analyses & prepared the figures, BL & EZG wrote the main manuscript text. All authors reviewed the manuscript.

Competing Interests

The authors declare no conflict of interests.

Data Availability

The data and code for statistical analysis in R are available on the Golumbic Lab website: www.golumbiclab.org/data

References

Arbula, S., Capizzi, M., Lombardo, N., & Vallesi, A. (2016). How Life Experience Shapes Cognitive Control Strategies: The Case of Air Traffic Control Training. *PLOS ONE*, *11*(6), e0157731. https://doi.org/10.1371/journal.pone.0157731

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baldock, J., Kapadia, S., & van Steenbrugge, W. (2018). The Task-Evoked Pupil Response in Divided Auditory Attention Tasks. *Journal of the American Academy of Audiology*. https://doi.org/10.3766/jaaa.17060

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beaman, C. P., Bridges, A. M., & Scott, S. K. (2007). From Dichotic Listening to the Irrelevant Sound Effect: a Behavioural and Neuroimaging Analysis of the Processing of Unattended Speech. *Cortex*, *43*(1), 124–134. https://doi.org/10.1016/S0010-9452(08)70450-7

Borella, E., Pezzuti, L., De Beni, R., & Cornoldi, C. (2019). Intelligence and working memory: evidence from administering the WAIS-IV to Italian adults and elderly. *Psychological Research*, 1–13. https://doi.org/10.1007/s00426-019-01173-7

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, *47*(3), 191–196. https://doi.org/10.1037/h0054182

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*(5), 1465–1487. https://doi.org/10.3758/s13414-015-0882-9

Brungart, D. S., Kordik, A. J., & Simpson, B. D. (2005). Audio and Visual Cues in a Two-Talker Divided Attention Speech-Monitoring Task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(3), 562–573. https://doi.org/10.1518/001872005774860023

Buitenweg, J. I. V., Murre, J. M. J., & Ridderinkhof, K. R. (2012). Brain training in

progress: a review of trainability in healthy seniors. *Frontiers in Human Neuroscience*, *6*, 183. https://doi.org/10.3389/fnhum.2012.00183

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. *Cambridge University Press*. New York.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustic Society of America*, *25*, 975–979.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80–90. https://doi.org/10.1037/h0039515

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2018). Attention Is Required for Knowledge-Based Sequential Grouping: Insights from the Integration of Syllables into Words. *The Journal of Neuroscience*, *38*(5), 1178–1188. https://doi.org/10.1523/JNEUROSCI.2606-17.2017

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology (London, England : 1953)*, *92 Part 1*, 53–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11802865

Drullman, R., & Bronkhorst, A. W. (2004). Speech perception and talker segregation: effects of level, pitch, and tactile support with multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *116*(5), 3090–3098. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15603154

Elliott, E. M., & Briganti, A. M. (2012). Investigating the role of attentional resources in the irrelevant speech effect. *Acta Psychologica*, *140*(1), 64–74. https://doi.org/10.1016/J.ACTPSY.2012.02.009

Enns, J. T., Kealong, P., Tichon, J. G., & Visser, T. A. W. (2017). Training and the attentional blink: Raising the ceiling does not remove the limits. *Attention, Perception, & Psychophysics*, *79*(8), 2257–2274. https://doi.org/10.3758/s13414-017-1391-9

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Fry, A. F., & Hale, S. (1996). Processing Speed, Working Memory, and Fluid Intelligence: Evidence for a Developmental Cascade. *Psychological Science*,

7(4), 237–241. https://doi.org/10.1111/j.1467-9280.1996.tb00366.x

Fusser, F., Linden, D. E. J., Rahm, B., Hampel, H., Haenschel, C., & Mayer, J. S. (2011). Common capacity-limited neural mechanisms of selective attention and spatial working memory encoding. *European Journal of Neuroscience*, *34*(5), 827–838. https://doi.org/10.1111/j.1460-9568.2011.07794.x

Getzmann, S., Golob, E. J., & Wascher, E. (2016). Focused and divided attention in a simulated cocktail-party situation: ERP evidence from younger and older adults. *Neurobiology of Aging*, *41*, 138–149. https://doi.org/10.1016/J.NEUROBIOLAGING.2016.02.018

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. *Hum Factors.*, *24*(2), 173–183.

Gygi, B., & Shafiro, V. (2012). Spatial and temporal factors in a multitalker dual listening task. *Acta Acustica United with Acustica*, *98*(1), 142–157. https://doi.org/10.3813/AAA.918500

Hausfeld, L., Riecke, L., Valente, G., & Formisano, E. (2018). Cortical tracking of multiple streams outside the focus of attention in naturalistic auditory scenes. *NeuroImage*, *181*, 617–626. https://doi.org/10.1016/J.NEUROIMAGE.2018.07.052

Helland, T., Morken, F., Bless, J. J., Valderhaug, H. V, Eiken, M., Helland, W. A., & Torkildsen, J. V. K. (2018). Auditive training effects from a dichotic listening app in children with dyslexia. *Dyslexia (Chichester, England)*, *24*(4), 336–356. https://doi.org/10.1002/dys.1600

Hilburn, B. (2004). *Cognitive complexity in air traffic control: a literature review*.

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30–41. https://doi.org/10.1002/pchj.44

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *The Journal of the Acoustical Society of America*, *120*(5), 2926–2937. https://doi.org/10.1121/1.2354070

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory*

and Language, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Kahneman, D. (1973). *Attention and effort.* Prentice-Hall. Retrieved from https://books.google.co.il/books/about/Attention_and_effort.html?id=7kvuAAA AMAAJ&redir_esc=y

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Keidser, G., Best, V., Freeston, K., & Boyce, A. (2015). Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations. *Frontiers in Psychology*, *6*, 597. https://doi.org/10.3389/fpsyg.2015.00597

Kidd, G., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., Best, V., & Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *The Journal of the Acoustical Society of America*, *140*(1), 132. https://doi.org/10.1121/1.4954748

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, *20*(2), 228–242. https://doi.org/10.3758/s13423-012-0359-y

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13 (URL: http://doi.org/10.18637/jss.v082.i13)

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-Five Years After Broadbent (1958): Still No Identification Without Attention. https://doi.org/10.1037/0033-295x.111.4.880

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lin, G., & Carlile, S. (2015). Costs of switching auditory spatial attention in following conversational turn-taking. *Frontiers in Neuroscience*, *9*(MAR).

https://doi.org/10.3389/fnins.2015.00124

Mishra, J., de Villers-Sidani, E., Merzenich, M., & Gazzaley, A. (2014). Adaptive training diminishes distractibility in aging across species. *Neuron*, *84*(5), 1091–1103. https://doi.org/10.1016/j.neuron.2014.10.034

Murphy, C. F. B., Moore, D. R., & Schochat, E. (2015). Generalization of Auditory Sensory and Cognitive Learning in Typically Developing Children. *PloS One*, *10*(8), e0135422. https://doi.org/10.1371/journal.pone.0135422

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research*, *352*, 40–48. https://doi.org/10.1016/J.HEARES.2017.02.005

Neely, C., & LeCompte, D. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory and Cognition*, *27*(1), 37–44. https://doi.org/10.3758/bf03201211

Oberfeld, D., & Klöckner-Nowotny, F. (2016). Individual differences in selective attention predict speech identification at a cocktail party. *ELife*, *5*. https://doi.org/10.7554/eLife.16747

Oswald, C. J. P., Tremblay, S., & Jones, D. M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory*, *8*(5), 345–350. https://doi.org/10.1080/09658210050117762

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Puvvada, K. C., & Simon, J. Z. (2017). Cortical Representations of Speech in a Multitalker Auditory Scene. *The Journal of Neuroscience*, *37*(38), 9189–9196. https://doi.org/10.1523/JNEUROSCI.0938-17.2017

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion*, *17*(4), 740–750. https://doi.org/10.1037/emo0000274

Salmela, V. R., Moisala, M., & Alho, K. (2014). Working memory resources are shared across sensory modalities. *Attention, Perception, & Psychophysics*, *76*(7), 1962–1974. https://doi.org/10.3758/s13414-014-0714-3

Salo, E., Salmela, V., Salmi, J., Numminen, J., & Alho, K. (2017). Brain activity associated with selective attention, divided attention and distraction. *Brain Research*, *1664*, 25–36. https://doi.org/10.1016/j.brainres.2017.03.021

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110. https://doi.org/10.2307/3002019

Sörqvist, P., & Rönnberg, J. (2014). Individual differences in distractibility: An update and a model. *PsyCh Journal*, *3*(1), 42–57. https://doi.org/10.1002/pchj.47

Soveri, A., Tallus, J., Laine, M., Nyberg, L., Bäckman, L., Hugdahl, K., … Hämäläinen, H. (2013). Modulation of Auditory Attention by Training. *Experimental Psychology*, *60*(1), 44–52. https://doi.org/10.1027/1618-3169/a000172

Tsuchida, Y., Murohashi, H., Katayama, J., & Murohashi, H. (2012). Working memory capacity affects the interference control of distractors at auditory gating. *Neuroscience Letters*, *516*(1), 62–66. https://doi.org/10.1016/j.neulet.2012.03.057

Wiemers, E. A., & Redick, T. S. (2018). Working memory capacity and intra-individual variability of proactive control. *Acta Psychologica*, *182*, 21–31. https://doi.org/10.1016/J.ACTPSY.2017.11.002

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 255–260. https://doi.org/10.1037/0278-7393.21.1.255

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., … Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party." *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037