The Dissertation Committee for Pak Wing Lam Certifies that this is the approved version of the following dissertation:

Listening Under Pressure: The Downside of Motivation

Committee:

Bharath Chandrasekaran, Supervisor

Craig A. Champlin

Maya L. Henry

Elizabeth D. Peña

Zenzi M. Griffin

Listening Under Pressure: The Downside of Motivation

by

Pak Wing Lam

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin
December 2017

To my mom, Lai Shan To.

Acknowledgements

This dissertation would not be possible without…

Four most important mentors in my life: Dr. Li Sheng, Dr. Bharath Chandrasekaran, Dr. Griffin, and Dr. Elizabeth Pena. Thank you so much for your kind support and making my intellectual pursuit easier.

An army of professors who are full of wisdom: Dr. Craig Champlin, Dr. Barbara Davis, Dr. Thomas Marquardt, Dr. Courtney Byrd, Dr. Maya Henry, Dr. Liu Chang, Dr. Lisa Bedore, Dr. Harvey Sussman, Dr. Amanda Zappler, Dr. Angela Carey, Ms. Joan Balash, and Dr. Gloria Zhang. Thank you so much for your care and showing me what it means to be a professor.

A team of fantastic clinical supervisors: Dr. Jessica Franco, Dr. Madhu Sundarrajan, Ms. Cydney Medford, Dr. MaryAnne Nericcio, Ms. Sharon Ciccarelli, and Ms. Stephanie Karger. Thank you for helping me to be a better clinician.

A superb cohort of classmates and colleagues: Jingjing Guan, Hoyoung Yi, Ying Hao, Ying Lu, Zilong Xie, Rachel Reetzke, Han Yi, Mirza Lugo-Neris, Megann McGill, Zoi Gkalitsiou, Stephanie Grasso, Alisa Baron, Rachel Tessmer, Natalie Czimskey, Geoff Coalson, Michelle Easter, Diana Cruz, Natalya Rich, Maryann Zúrrica, Yao Du, Man Yang, and Alex Rosenfeld. Thank you for helping me through the dark hours. I should have gone to the Happy Hours more. The next round is on me.

My fantastic friends: Kwan's family and Paul Paul, Dickson Lao, Cwing Ho and Kenny, Chow's family, Bird Lei and Anna, Kam Hon Hoi, Ivy Tsoi, Mung-Ting Chung, Alex AK, Kevin and Joey, Amigo Yeung, Ryan Wang, Skylar Laai, Michael Ng, Esther Liu, Wei-Jin Zheng, Koit, Karen and Caryn, Qianping Gu, Pumpki Su, Ray Cheung, Patrick Chu, Schola Lam, Nancy Shing, Kay Wong, Ei Leen, Helen Wong, Casper Lai, and hall-mates from Hall D. Thank you for bringing me out of the laboratory and feed me with good food!

My wonderful mentees: Heather Cook, Marle Vargas, Nicole Tsao, Aislynn Fulton, Jasmin Singh, Jessica Hernandez, Lauren Kelley, Hannah Leung, Kayla Wiltz, Catherine Xu, Emily Sun, Jade Holder, and Chelbi Orrick. Thank you for showing me how to be a better mentor.

My mom: Thank you teaching me persistence, compassion, and integrity.

My dear Katie Kei. Thank you for staying around and sharing life (and food)!

Finally, I would love to thank Boji Lam (Oh yeah! That's me!).
Thank you for not giving up and keep moving on.

Let's move on!

Listening Under Pressure: The Downside of Motivation

Pak Wing Lam, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Bharath Chandrasekaran

Abstract: The desire for self-improvement is critical to human performance and learning outcomes. Paradoxically, however, being subjected to increased performance pressure can also result in "choking under pressure". No studies have experimentally examined the extent to which motivation impacts native speech processing. This dissertation manipulated performance pressure in listeners, and systematically examined its impact on three speech-processing experiments. Sixty adult native English listeners and 45 non-native listeners with poorer English proficiency completed three speech processing experiments, twice – once to establish a baseline, and again to measure changes in performance. In these experiments using native English speech, listeners detected (illusionary) sound changes, categorized phonemes under lexical interference, and recognized words in noises. After baseline testing, half of the participants in each language group were instructed to work, with a fictitious partner, towards a performance-contingent monetary reward; the other half, as controls, simply performed the tasks a second time. This study demonstrated a negative impact of performance pressure on native listeners in all experiments. Relative to the controls, the motivation group were more susceptible to illusions, failed to ignore lexical interference despite prior exposure, and recognized fewer words in cognitively-demanding listening situations. Unexpectedly, relative to native listeners, non-native listeners perceived it as less important to perform well, and those who were in the high performance-pressure group requested significantly greater amount of money for improvement. These language-group differences in task-related attitudes might be a confounding factor that moderate the effect of motivation. By illustrating a complex interaction among motivation, listener status, and performance-induced demands, this dissertation highlights the importance of motivation in speech science.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

We process a massive amount of speech signals that are extremely fast, variable, and fleeting, often existing for only tenths of seconds (Holt & Lotto, 2010). How the brain manages this significant computational challenge remains unclear. Because speech processing is generally viewed as unconscious and effortless in native listeners, the role of cognition in speech processing has been largely ignored, until recently (Arlinger, Lunner, Lyxell, & Pichora- Fuller, 2009; Pichora-Fuller et al., 2016; Rönnberg et al., 2013). However, this view of speech processing as unconscious and effortless is not entirely correct, and is partly based on the “artificial normality” of the tightly-controlled laboratory settings in which speech processing is often studied (Mattys & Liss, 2008; Mattys, Davis, Bardlow, & Scott, 2012). When one’s speech-processing system faces elevated challenges – such as being subjected to perceptual illusions, presented with ambiguous signals, or placed in adverse listening conditions – cognitive processes may play a key role in supporting speech processing. Challenging listening conditions are not a rarity; rather, they are the norm in non-laboratory situations (e.g., chatting in a noisy café, asking for directions on a busy street, etc.). However, the mechanistic role of cognitive processes in speech processing in various listening environments remains poorly understood.

In cognitive science, motivation has been extensively studied as a key mediator of cognitive processes (e.g., Beilock & Carr, 2005; Beilock & DeCaro, 2007; Gimmig, Huguet, Caverni, & Cury, 2006; Maddox & Markman, 2010; Markman, Maddox, & Worthy, 2006; for a review, see Botvinck & Braver, 2015). Because motivation is so critical to human behaviors, it has, since the beginning of the last century, been the center of investigative efforts in such diverse fields as performance pressure (e.g., Baumeister; 1984; Wankel, 1972), learning science (e.g., Dörnyei, 1994; Yerkes & Dodson, 1908), and rehabilitation science (e.g., Baekeland, & Lundwall, 1975; Miller, 1985). Though a large body of research on human performance has shown the relevance of motivation, this important psychological dimension has not received much attention in speech processing. The recent development of the *Framework for Understanding Effortful Listening* (FUEL) by speech and hearing scientists confirms motivation is key to understanding "when and how much effort we expend during listening in everyday life" (Pichora-Fuller et al., 2016). However, current knowledge about the influence of motivation on speech processing is extremely limited. This dissertation aims to fill this significant gap by experimentally manipulating levels of motivation in adult listeners to examine its impact on speech-processing performance.

1.1 MOTIVATION AND ITS PARADOXICAL EFFECTS ON PERFORMANCE

A key question relating to virtually all research on motivation concerns what factors determine the *optimal* level of mental or physiological arousal (or drive) needed to *maximize* certain performance outcomes. This key question is particularly relevant to speech processing in challenging conditions, where uncertainties, ambiguities, and interferences can be significant. Yerkes and Dodson (1908) elegantly depicted the complex relationship between performance and motivation as a U-shaped parabola, in which performance improves as levels of motivation increase, until the optimal, task-dependent "sweet" spot is reached; however, performance declines when levels of motivation are either too low (e.g., boredom) or too high (e.g., performance pressure).

This dissertation targeted one end of Yerkes and Dodson's parabola. Specifically, we examined the extent to which one's desire for good performance can create performance pressure and, ironically, cause one to underperform (Baumeister, 1984). This phenomenon, termed as "choking under pressure", is an active area of research in sports (Baumeister, 1984; Beilock & Carr., 2001), mathematics (Beilock & Carr, 2005; Beilock & DeCaro, 2007; Beilock, Kulp, Holt, & Carr, 2004), and intelligence testing (Gimmig, Huguet, Caverni, & Cury, 2006). A common thread across these studies is that the drive towards superior performance sometimes creates psychological pressure that, rather than assisting, can negatively impact performance. While the choking effect has not been studied in relation to speech processing, it should be relevant to countless real-life scenarios – for example, a college student in his/her last semester, who needs to pass a foreign-language listening assessment to satisfy a language requirement needed to graduate; an air-traffic controller who have to maintain accurate communications with pilots in challenging listening conditions; or a soldier in a warzone, for whom clearly understanding the orders given by his/her superior is literally a matter of life and death.

Studies on the "choking under pressure" effect share a general theme – that performance pressure can be detrimental to performance. Current understanding about the choking effect posit that it impairs performance via two mechanisms that are not mutually exclusive (Beilock & Carr, 2001; Yu, 2015). *Explicit monitoring theory* posits that performance pressure elevates one's awareness of one's accurate control of skill processes to perform correctly (Baumeister, 1984; Lewis & Linder, 1997). Elevated attention to execution at the step-by-step level disrupts proceduralized performance, resulting in the "choking under pressure" effect (Beilock and Carr, 2001; Kimble & Perimuter, 1970; Lewis & Linder, 1997, Masters, 1992). *Distraction theory* posits that performance pressure creates worrying thoughts and impairs cognitive processes, by competing for limited cognitive resources needed for task execution (Beilock et al., 2004; Maddox & Markman, 2010). For studies that support the distraction theory, cognitive processes and resources (e.g., working memory) play mediating roles. Working memory (WM) often predicts one's ability to control the maintenance of information and attention, which is crucial for such higher-order cognitive activities as problem solving, mental arithmetic, and language processing (Daneman & Merikle, 1996; Linck, Osthus, Koeth, & Bunting, 2014; for a review, see Engle, 2002). However, multiple studies have shown that performance pressure can weaken the association between cognitive ability and performance, highlighting the “choking under pressure” effect (Beilock & Carr, 2005; Beilock & DeCaro, 2007; Gimmig et al., 2006). Because distraction theory posits a strong association between performance pressure and cognitive processes, this framework is particularly relevant to speech processing in challenging listening conditions.

CHAPTER 2 PRESSURE MANIPULATION AND ITS IMPACT ON COGNITIVE AND LINGUISTIC PERFORMANCE

How can motivation be manipulated so that it pushes participants to one end of Yerkes and Dodson's parabola and creates performance pressure? To create performance pressure, Beilock and colleagues (2004) have adopted a highly effective script that *combines* monetary rewards and peer pressure. Specifically, the researchers notified participants their rewards were performance-contingent, and that poor performance would cause their (fictitious) partners to lose their rewards. Additionally, they informed participants that both accuracy and speed contributed equally to performance scores. Using this script, Bilocok and Carr (2005) found that performance differences between individuals with better and poorer WM were present only in low performance-pressure conditions; the performance gap between the groups narrowed because only individuals with better WM suffered from performance decrements in high performance-pressure conditions. Beilock and DeCaro (2007) replicated the distraction effect of performance pressure by showing that having better WM did not predict better performance in solving arithmetic problems in high performance-pressure conditions. Taken together, targeting one's desire to approach a reward and fear of disappointing potential partners is highly effective in introducing performance pressure and elevating the likelihood of a "choking" effect.

Within the framework of distraction theory, Beilock and colleagues argued that performance pressure prompts worrying thoughts about the importance of success and the consequences of failure, which occupy cognitive resource and interferes with cognitive processes (e.g., WM). Worrying thoughts create a dual-task environment for cognition, wherein an individual must pay attention to the primary task at hand (e.g., arithmetic computation), while simultaneously controlling for any interfering worries. Converging supports for the distraction theory and the effectiveness of Beilock and colleagues' script come from research on categorical learning and decision making (e.g., Maddox, Koslov, Yi, & Chandrasekaran, 2016; Maddox & Markman, 2010; Markman, Madox, & Worthy, 2006). These studies often instruct participants to categorize sets of stimuli that may or may not be captured by verbalized rules. Using a similar script, these studies found that worrying thoughts compete for cognitive resource (e.g., WM) and impair the learning of rule-based stimuli, which relies on explicit, hypothesis-driven, frontal processes (DeCaro, Thomas, Albert, & Beilock, 2011; Maddox et al., 2016; McCoy, Hutchinson, Hawthorne, Cosley, & Ell, 2014). Furthermore, consumption of cognitive resource prompted participants to adopt implicit, procedural-based learning, which benefits the acquisition of category-boundary stimuli that cannot be optimally captured by verbalized rules (Markman et al., 2006).

Note that Beilock and colleagues' manipulation of motivation and performance pressure, which is widely used in ensuing studies on decision making and category learning, combines monetary rewards and peer pressure. While monetary rewards represent positive goals participants would desire, peer pressure represents negative consequences participants would avoid. How motivation creates pressure and impacts performance is highly complex (Maddox et al., 2016; Maddox & Markman, 2010) that some studies begin attempt to examine in greater mechanistic details in adults (Maddox, Filoteo, Glass, & Markman, 2010; Maddox & Markman, 2010; Worthy, Markman, & Maddox, 2009) and children (Worthy, Brez, Markman, & Maddox, 2011). It has been argued that performance outcomes in high performance-pressure condition are the result of an interaction between *global* final-outcome motivation (e.g., getting bonus, receiving praise, avoiding pay cut) and *local* trial-by-trial motivation (e.g., gaining or losing point on a trial-by-trial basis) (Maddox & Markman, 2010; Worthy et al., 2010). When global motivation aligns with local motivation (e.g., gain point on each trial to get an additional bonus at the end), a regulatory match takes place and enhances performance that requires explicit, conscious, and likely effortful rule-based processing. When global motivation does not align with local motivation (e.g., gain point on each trial to avoid failing a partner), a regulatory mismatch takes place and prompts the cognitive system to shift to adopting implicit, procedural-based processing.

This dissertation does not aim to tease apart the impact of regulatory match and mismatch on speech processing. As a first step to examine the impact of motivation and performance pressure on speech processing, this study adopted Beilock and colleagues' motivation manipulation because it has been proved to be highly effective in creating performance pressure (Beilock & Carr, 2005; Beilock & DeCaro, 2007; Maddox et al., 2016). Using the same manipulation, Maddox and colleagues (2016) studied nonnative speech-category learning, which is among the first in speech science that examines the extent to which performance pressure influences speech processing. However, learning speech categories in foreign language is unlike speech processing in one's native language. While the former is argued to be greatly challenging for adults (e.g., Flege, 1987), native-speech processing is a relatively automatic process. Additionally, the two processes may differ in levels of complexity. While learning foreign speech categories relies primarily on processing of acoustic-phonetic information, native-speech processing is a complex multi-leveled ensemble, running from lower-level processing of acoustic information (e.g., voice onset time) to higher-level processing of linguistic information (e.g., parsing of syntax and comprehension).

A challenge in examining the effect of performance pressure on speech processing is that the role of cognitive processes at various levels of speech processing is largely unclear. There is some converging evidence that higher-level processing of speech may be increasingly dependent on cognitive processes. For instance, better WM was found to predict better ability to follow complex spoken directions (Engle, Carullo, & Collins, 1991), to selectively attend to target speech sounds (Conway, Cowan, and Bunting, 2001), and to make more anticipatory spoken language processing (Huettig & Janse, 2016). However, it remains unclear the extent to which performance pressure impacts speech processing that may involve cognitive ability.

2.1 THE IMPACT OF PRESSURE ON LINGUISTIC ABILITY

Because few studies have examined the effect of performance pressure on speech processing, one can only infer its impact by considering studies on different modalities (e.g., reading). Rai and colleagues (2011, 2015) adopted an experimental approach in which they manipulated psychological pressure in readers and examined its impact on speed and accuracy in processing complex inferences in non-native reading – a rare attempt in language science. Note that their studies did not manipulate performance pressure, but psychological pressure in a broad sense - the participants were instructed to read a tongue twister aloud, while the researchers video-recorded their performance. The researchers found their pressure manipulation had no effect on accuracy, but *increased* reading time. Notably, individuals with poorer WM were more strongly affected by the pressure than were individuals with better WM. In a later study, Rai and colleagues (2015) additionally sat an experimenter behind the participants in the same room, to introduce evaluative pressure. This study also recruited native-reader controls and extended their earlier finding by showing that pressure did not impact accuracy in non-native readers or native readers. Additionally, the reading time among non-native readers with poorer WM *increased* under pressure, and a similar pattern was found in native-readers. Note that findings from Rai and colleagues’ studies on reading are at odds with Beilock and colleagues’ studies on mathematics. While performance pressure impaired participants' mathematical accuracy with a selectively stronger effect in individuals with better WM, Rai and colleagues (2011, 2015) suggested WM had a "protective" effect against more general psychological pressure.

The extent to which findings from mathematics and reading generalize to speech processing is unclear. Speech signals are rapid, highly variable, and fleeting in nature. As a result, speech processing often requires the auditory system to work at its upper limit (Christiansen & Chater, 2016; Miller & Taylor, 1948). In contrast, arithmetic operations and reading comprehension are usually self-paced, which allow participants extended time for complete processing. Additionally, speech processing seldom takes place in ideal conditions, without any distractions or interference. Due to the distinctive nature of speech signals, speech processing can impose significant challenges on the brain, even for native listeners.

Notably, Rai and colleagues' (2011) target populations were late bilinguals, not native readers. In fact, most studies examining the role of motivation and associated performance pressure in speech and language processing (e.g., Dörnyei, 1994; Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1991) have focused on non-native second-language learners, and offer little information about the impact of performance pressure on native speakers. Additionally, most studies conducted to date have been correlational in nature, rather than experimental. From a practical standpoint, this focus on non-native speakers is supported by the fact that language assessments, which may introduce performance pressure, are primarily designed to evaluate the performance of non-native speakers, not native speakers. Additionally, the focus on non-native speakers may be fueled by the dominant view that language processing is more subject to language breakdowns in non-native speakers than in native speakers. There is a great amount of evidence that poorer language proficiency is associated with decreased automaticity (e.g., Segalowitz & Hulstijn, 2005) and a greater amount of cognitive effort in language processing (e.g., Meschuan & Hernandez, 2006). For non-native speakers who are yet to be native-like, language processing is expected to be more cognitively demanding, and to take an extended time to complete. Because there is a higher likelihood of language breakdowns among non-native speakers, it is reasonable to hypothesize non-native speakers may be more vulnerable to the deteriorating effect of performance pressure than native speakers.

However, it is also possible that native speakers and non-native speakers may perceive an assessment differently, depending on whether the assessment is presented in one's native language or L2. Gimmig and colleagues (2006) found participants experienced elevated levels of performance pressure when a task was introduced as a measure of academic success, which is relevant to participants' personal interests. Based on this finding, one could hypothesize that administering a task in one's native language may introduce heightened personal interest and a sense of self-relevance. Because an individual might have a greater desire to perform well on tasks presented in his/her native language than on tasks presented in non-native language, native speakers may place more importance on their good performance than non-native speakers, making native speakers more susceptible to the deteriorating effect of performance pressure.

2.2 PURPOSE AND SPECIFIC AIMS

The role of cognition in speech processing is poorly understood (Arlinger et al., 2009; Pichora-Fuller et al., 2016). Though motivation mediates cognitive processes (for a review, see Botvinck & Braver, 2015) and is critical to understanding human behaviors (Ryan & Deci, 2000), no studies have experimentally manipulated motivation in listeners to examine its impact on speech-processing performance, in a systematic manner. This dissertation aims to fill this significant gap. Specifically, we experimentally manipulated listeners' motivation levels to create performance pressure, aiming to examine changes in performance on three experimental tasks that have been widely used to examine speech-processing phenomena.

Experiment 1 studied verbal transformation, which highlights the adaptability of the speech-processing system when repeatedly fed out-of-context speech signals (Warren, 1961a). Experiments 2 studied the *Ganong* effect (Ganong, 1980), which highlights the interaction between phonetic processing and higher-order lexical knowledge. Finally, experiment 3 studied speech processing in adverse listening conditions (e.g., noise) that are ubiquitous in everyday livings, and that introduce significant computational challenges to the speech-processing system. Findings from studies on these speech-processing tasks and have been highly influential in the development of speech-processing models (e.g., Ganong, 1980; Mattys et al., 2012; Norris, McQueen, & Cutler, 2000); however, motivation is seldom studied as a variable to examine its impacts on the manifestation of these speech phenomena.

By integrating research paradigms from cognitive science (e.g., Beilock & Carr, 2005) and speech-hearing science (*Ganong* 1980; Mattys & Wiget, 2011; Warren, 1961a), this dissertation aims to examine:

1. The extent to which performance pressure influences the amount of auditory illusions (Experiment 1)
2. The extent to which performance pressure influences listeners' consistency in categorizing phonemes and ignoring lexical influence (Experiment 2)
3. The extent to which performance pressure influences listeners' ability to process speech in noises that differ in their amount of lexical interference (Experiment 3)

CHAPTER 3: PROCEDURES

In this section, we describe the background of the participants, the specific procedures and tasks that participants were administered, and the script used to manipulate performance pressure in participants. Finally, we report participants' responses to a feedback questionnaire administered at the end of the experiment. Examination of participants' responses to the questionnaire provides information about the effectiveness of the performance-pressure manipulation and task-related attitudes.

3.1 PARTICIPANTS

Sixty-seven native listeners of English and 50 non-native listeners, aged 18-35, were recruited from the University of Texas at Austin. All participants provided written informed consent and received monetary compensation for their participation. Participants reported having no history of language or hearing problems, and reported normal or corrected-to-normal vision. All participants underwent an audiological test that included both air and bone conduction, to ensure thresholds ≤ 25dB Hearing Level (HL) at 250Hz, 500Hz, 1000 Hz, 2000 Hz, and 8000 Hz, for each ear. To minimize the interaction between cultural backgrounds and MiP for good performance, the L2 learners recruited in this study encompassed a wide variety of linguistic and cultural backgrounds (Cantonese, French, Korean, Mandarin, Portuguese, Spanish, Thai). Levels of English proficiency were determined by age of acquisition (AoA), the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), and performance (accuracy and response time) on a lexical decision task (Lemhöfer, & Broersma, 2012).

A total of 62 native listeners and 47 non-native listeners completed the entire study, which was conducted in two sessions over two separate days (Day 1 and Day 2). Participants were randomly assigned to either the control or the performance-pressure group. Because performance pressure is the focus of this study, more non-native listeners (n=27) were assigned to the performance-pressure group, to approach the number of participants in the native-listener performance-pressure group (n=32). At the end of Session 2, participants assigned to the performance-pressure group were asked whether they believed in the existence of the partner. Two native listeners and two non-native listeners did not believe in the motivation manipulation (i.e., the presence of a partner), and thus were removed from further analyses. Final analyses thus included 55 participants in the motivation group (30 native listeners, 25 non-native listeners), and 50 in the control group (30 native listeners, 20 L2 non-native listeners).

3.2 DAY 1 AND DAY 2 OF THE STUDY

Table 3.1 describes the procedure of this study. There were two, two-hour sessions conducted on two separate days. Over 90% of participants completed both sessions within a week; all participants completed both sessions within three months. Because the participants were young adults without any reported neurological impairments or history of language impairments, their hearing, language, and cognitive status were assumed to remain stable over the three-month period of the study. All background measures (i.e., hearing screening, WM, lexical decision, and questionnaires) were administered on Day 1, while the three listening tasks (i.e., verbal transformation, the *Ganong* effect, speech recognition in noise) were conducted on Day 2. The study was broken into two sessions to reduce participants fatigue during the experiments. Finally, for each participant, both sessions were scheduled for the same time of day, whenever possible (e.g., 9am on Day 1 and Day 2 for participant A, 11am for participant B, etc.), to increase the likelihood participants would have comparable energy levels across both sessions.

Table 3.1: General experimental procedure on Day 1 and Day 2.

	Control group 30 native listeners of English 20 non-native listeners	**Motivation/Pressure group** 30 native listeners of English 25 non-native listeners
Day 1	a) Audiological testing; b) Working memory assessment (operation span), c) Language proficiency assessment (Lexical decision), d) Language questionnaires	
Day 2	**Baseline** • Verbal Transformation • The *Ganong* effect • Speech processing in noise ❖ Three listening tasks in a counter-balanced manner	**Baseline** • Verbal Transformation • The *Ganong* effect • Speech processing in noise ❖ Three listening tasks in a counter-balanced manner
	3-minute break	
	Participants were told that • “Let’s do some more”	Participants were told that • “Let’s do some more” • “The computer has calculated your score during baseline” • “Improve 20% on the target task to earn a bonus $10” • “Both you and your partner have to improve to get the bonus” • “Your partner has already improved” • “You will be video-taped”
	Second attempt • Verbal Transformation • The *Ganong* effect • Speech processing in noise ❖ Three listening tasks in the same order during baseline	**Second attempt** • Verbal Transformation • The *Ganong* effect • Speech processing in noise ❖ Three listening tasks in the same order during baseline
	Feedback questionnaires (Appendix A)	

3.2.1 Day 1

At the beginning of the session on Day 1, all participants gave their informed consent to participate in the study. Participants also completed measures of their cognitive ability (WM) and language proficiency (lexical decision), and filled in a language questionnaire.

3.2.1.1 Working memory (Complex operation span)

This dissertation measured WM because previous studies found that performance pressure interacted with WM capacity (e.g., Beilock & Carr, 2005; Beilock & DeCaro, 2007). Because the primary aim of this dissertation is to examine the impact of performance pressure of speech-processing ability, not the interaction between performance pressure and WM, we obtained this measure to ensure the listener groups were comparable on this possibly confounding factor.

To measure WM, this dissertation used operation span (OSpan), a complex, widely-used WM measure in cognitive science (Beilock & Carr, 2005; Beilock & DeCaro, 2007; Conway et al., 2005; Linck et al., 2014). In the computerized version of the OSpan, administered via E-prime, participants were given a series of simple arithmetic operations (e.g., 2+3). For each operation, the participants pressed designated buttons to indicate whether it were true or false, after which they were provided with a letter they had to recall later; for example, in the operation (2+3=5?, M), participants should respond "true," and memorize the letter "M." After the completion of a series, participants were prompted with a 4 X 3 matrix of letters, and asked to click on the letters they had been given to memorize, in the correct sequential order. The OSpan consists of 15 recall sequences, with each sequence length ranging from three to seven letters. The memory span (maximum span: 75) was calculated by adding up the total number of letters a participant would correctly recall in sequential order. For example, if a participant correctly recalled a sequence of five letters, five points were added to the span; however, if a participant incorrectly recalled even one letter of the sequence, zero point were added to the span. The design of the OSpan resembled the conceptualization of WM as processing (i.e., arithmetic operation), maintenance (i.e., memorizing the letter), and recall (i.e., clicking on the matrix of letter).

3.2.1.2 Language proficiency

This dissertation measured language proficiency using the Lexical Test for Advanced Learners of English (LexTALE). LexTALE predicts English lexical knowledge, and correlates it with general English proficiency (Lemhöfer, & Broersma, 2012). LexTALE includes 40 real words and 20 nonwords. Participants are shown a series of letter strings, and asked to press the corresponding button to indicate if the letter strings were existing English words or not. Lemhöfer and Broersma (2012) indicated the word frequency of some test items is so low a celling effect is highly unlikely.

3.2.1.3 Questionnaires

To ensure the performance-pressure and control groups were comparable on confounding variables that may influence levels of psychological pressure, we administered: (a) the *trait* scale of the State-Trait Anxiety Inventory, which measures trait anxiety (Spielberger, Gorsuch, & Lushene, 1970); (b) the brief resilience scale (Smith et al., 2008); and, (c) a brief version of the Fear of Negative Evaluation Scale (Leary, 1983). By ensuring the groups were comparable on these three measures, which index anxiety traits and stress reactivity, this dissertation assumed elevated psychological pressure levels in the motivation groups at the end of the experiment could be more readily attributed to motivation manipulation.

Table 3.2 shows all listener groups were comparable in terms of WM, trait anxiety, resilience to stress, and fear of negative evaluation (all ps ranged from .09 to .77). Native listeners and non-native listeners differed in all subjective and objective language proficiency measures, including the amount of daily English usage, self-rated English proficiency, lexical decision accuracy, and lexical decision response time (all ps <.001, effect size ranged from .16 to .40). The performance-pressure and control groups did not differ in these language measures within either native listeners or non-native-listeners groups (all ps>.3).

Table 3.2: Background information of participants.

	Native listeners		Non-native listeners	
	Control (N=30)	Pressure (N=30)	Control (N=20)	Pressure (N=25)
Age	21.6 (3.1)	19.5 (1.7)	21.4 (3.2)	22.0 (3.4)
Working memory (Operation span)	44.0 (16.3)	45.4 (19.7)	45.5 (15.7)	52.4 (14.2)
Age of Acquisition (English)	nil	nil	6.8 (2.5)	6.6 (3.2)
Daily English Usage (%) [a]	95 (7)	95 (10)	71 (24)	64 (20)
Self-rated English proficiency	9.6 (0.5)	9.4 (0.7)	7.9 (1.3)	7.9 (1.6)
Self-rated other (first) language proficiency [a]	2.1 (4.5)	1.3 (4.3)	8.6 (2.1)	8.9 (1.3)
Lexical decision time (ms)	834 (167)	832 (170)	1023 (359)	1071 (370)
Lexical decision accuracy (%)	92.9 (5.0)	94.1 (4.9)	81.9 (10.2)	81.7 (12.5)
Trait Anxiety [b]	42.2 (10.4)	43.9 (9.2)	45.4 (11.0)	47.6 (8.1)
Resilience [c]	27.4 (6.1)	25.6 (6.0)	25.9 (5.6)	25.0 (4.2)
Fear of negative evaluation [d]	35.5 (11.2)	37.4 (10.0)	33.0 (10.0)	38.8 (9.4)

a. the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007)
b. the trait scale of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970)
c. the brief resilience scale (Smith et al., 2008)
d. the brief version of the Fear of Negative Evaluation Scale (Leary, 1983)

3.2.2 Day 2

3.2.2.1 Hearing Screening

On Day 2, all participants completed the three listening tasks twice (i.e., baseline performance vs. the second attempt), separated by a three-minute break. The listening tasks administered during baseline performance and the second attempt did not differ in structure or instructions. Furthermore, the stimuli used in Experiments 1 and 2 were the same for baseline performance and the second attempt. For Experiment 3 (speech-in-noise processing), there were two sets of stimuli (Set A and Set B). Participants were administered either Set A or Set B first during baseline performance, in a counter-balanced manner. Note that only Experiment 3 measured percentage correct of participants' response, while Experiment 1 and Experiment 2 did not target accuracy but consistency of responses. As a result, Experiment 1 and Experiment 2 used the same set of stimuli while Experiment 3 used two different sets of stimuli. The order of the three listening tasks during baseline testing was also counter-balanced. During the second attempt, the three listening tasks were administered in the same order as they had been during baseline testing (Table 3.1). Details about stimulus properties and task instructions are provided in the corresponding sections, below.

After the three-minute post-baseline testing break, both native-speaker controls and L2 controls were simply instructed to perform the three listening tasks again. For the motivation groups, participants were further informed they had the chance to earn bonus money if they improved their performance by 20% during their second attempt. To introduce performance pressure, this dissertation used a script that replicated Beilock and Carr's (2005) script as closely as possible to create a scenario that introduced both peer pressure and social evaluation. Notably, this scenario has repeatedly been shown to elevate performance pressure (Beilock & Carr, 2004, 2005; Beilock & DeCaro, 2007). The script adapted in this study was as follows.

3.2.2.2 Script to introduce psychological pressure

"Before the break, you identified sounds, completed a speech in noise task, and a picture-matching task. For each task, the computer uses both your accuracy and response time to calculate your scores. Now you will do the tasks again. The structure and procedure of each task will be the same as what you saw before, but with different items. Importantly, your performance will decide how much money you can get at the end.

During the following session, the computer will randomly pick one of the three *tasks as the target. You will not know which task is the target until the end. If you improve your score on the target task by 20%, relative to your performance before the break, you will get a $10 bonus. At the end of this session, I will show you which task the computer chose, and your performance.*

However, getting the bonus money is a team effort. You will pair up with another *participant. For either of you to receive the extra $10, both of you must improve on the target task. If either you or your partner fails, neither of you will get the bonus money. You and your partner do not have to match on the target task.*

(Each participant was then given a set of eight cards, each of which had one participant ID. Researchers informed the participants that each card indexed a participant who had completed the study, and was yet to have a partner. The participants then picked their partner by picking a card at will. Researchers then checked the performance of the partner picked by the participant. In fact, the "potential partner" did not exist, although this was not known to the participant).

Our record shows that your partner has improved by 20% on the target task. This means you have a chance to earn the $10 bonus. However, if you fail to improve *by the required percentage, neither of you will receive the extra $10, and we will need to inform your partner. If you have no further questions, you will do the listening tasks now. Your performance will be video-taped so local language teachers and professors can examine your performance."*

At the end of the session on Day 2, this dissertation measured participants' performance pressure levels using the *state* scale from the State-Trait Anxiety Inventory (Spielberger et al., 1970). Participants were also administered a feedback questionnaire (Appendix A), which asked them to report, on a nine-point scale, their: (a) perceived importance of performing at a high level; and, (b) perceived pressure to improve their performance during their second attempt. The participants also reported: (c) the amount of improvement they thought they have made; and, (d) how much money they deemed sufficient to motivate them to improve their performance during their second attempt (max = \$10).

For each psychological dimension, we conducted separate analyses using linear, mixed-effect modelling (LMER) with the lme4 package 1.1-13 (Bates, Maechler, & Bolker, 2012) in RStudio 1.0.143. The analysis started with the maximum model, which included language group (i.e., native listeners vs non-native listeners), pressure (i.e., performance pressure vs control), attempt (i.e., baseline vs second attempt), and their interactions as fixed effects. The model was refined by removing the factors exhibiting the highest p-value, one at a time, while retaining the hierarchical rule of interactions. Likelihood ratio comparisons were performed to confirm that including each factor did not improve the amount of variance explained (Baayen, Davidson, & Bates, 2008).

We first report the analysis on perceived pressure to improve performance during the second attempt. There was a simple effect of Pressure [β= 1.6, SE=.4, χ^2 (1) = 15.6, p<.00001] (Table 3.3, Fig 3.1). We further show the motivation manipulation influenced perceived amount of improvement. There was a simple effect of Pressure [β= -9.0, SE=-3.1, χ^2 (1) = 8.2, p<.01], indicating the performance-pressure group reported less amount of improvement than did the controls (Fig 3.2). There was also a simple effect of Task [β= -4.2, SE=-1.8, χ^2 (1) = 5.6, p<.05], indicating that all listener groups reported less amount of improvement in sound identification (i.e., verbal transformation, the *Ganong* effect) than they did for speech recognition in noise.

Notably, relative to native listeners, non-native listeners placed less importance on performing at a high level during the second attempt, for all tasks [β= -.9, SE=-.3, χ^2 (1) =

12.1, $p<.001$] (Fig. 3.3). There was also a simple effect of Task [β= -.3, SE=--.1, χ^2 (1) = 4.5, $p<.05$], indicating that all listener groups reported less important to perform at a high level in sound identification (i.e., verbal transformation, the *Ganong* effect) than they did for speech recognition in noise.

Finally, when asked to suggest the amount of money they wanted, the raw scores indicated the non-native performance-pressure group wanted the greatest amount of money, while the native-listener performance-pressure group wanted the least, with both control groups being at mid-range (Fig. 3.4). Statistical analyses showed a significant Language Group X Pressure interaction [χ^2 (1) = 5.1, $p<.05$]. We analyzed the two-way interactions using the lsmeans package 2.27-2, which was designed to obtain the least-squares means and to test linear contrasts for linear and generalized mixed models (Lenth, 2016). The analysis indicated the non-native performance-pressure group wanted more money than the native-listener performance-pressure group [β= 3.1, SE=.7, $p <.001$], while the two control groups did not differ from each other ($p>.4$). There was also a simple effect of Task [β= .8, SE=-.2, χ^2 (1) = 12.7, $p<.001$], indicating that all listener groups reported less important to perform at a high level in sound identification (i.e., verbal transformation, the *Ganong* effect) than they did for speech recognition in noise.

Table 3.3: Participants’ responses to a final self-assessment questionnaire at the end of the whole experiment on Day 2.

		Native listeners		Non-native listeners	
		Control (N=30)	Pressure (N=30)	Control (N=20)	Pressure (N=25)
State anxiety [a]		30.2 (6.1)	39.4 (11.2)	32.4 (6.7)	38.4 (8.2)
Importance	Identify sounds [b]	6.9 (1.4)	6.7 (1.7)	5.7 (1.7)	6.1 (1.6)
	SPIN [c]	6.9 (1.4)	7.4 (1.5)	5.9 (1.5)	6.4 (1.6)
	Average	6.8 (1.2)	7.1 (1.4)	6.2 (1.2)	6.2 (1.0)
Perceived Pressure	Identify sounds [b]	5.5 (2.3)	7.1 (2.0)	4.9 (2.4)	6.6 (2.0)
	SPIN [c]	5.5 (2.2)	6.7 (1.8)	4.6 (2.3)	6.7 (1.8)
	Average	5.3 (2.0)	6.6 (1.7)	4.4 (2.1)	6.3 (1.6)
Perceived Improvement (%)	Identify sounds [b]	21 (21)	11 (11)	18 (20)	10 (21)
	SPIN [c]	24 (22)	15 (14)	24 (23)	14 (16)
	Average	20 (16)	13 (10)	22 (22)	12 (12)
Money to motivate ($)	Identify sounds [b]	6.3 (3.3)	4.8 (2.7)	6.3 (3.1)	7.9 (2.7)
	SPIN [c]	4.8 (2.9)	4.2 (2.6)	6.0 (3.3)	7.3 (2.8)
	Average	5.1 (2.6)	4.3 (2.5)	5.5 (2.2)	7.1 (2.7)

a. the trait scale of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970)
b. verbal transformation and the *Ganong* effect
c. Speech processing in noise

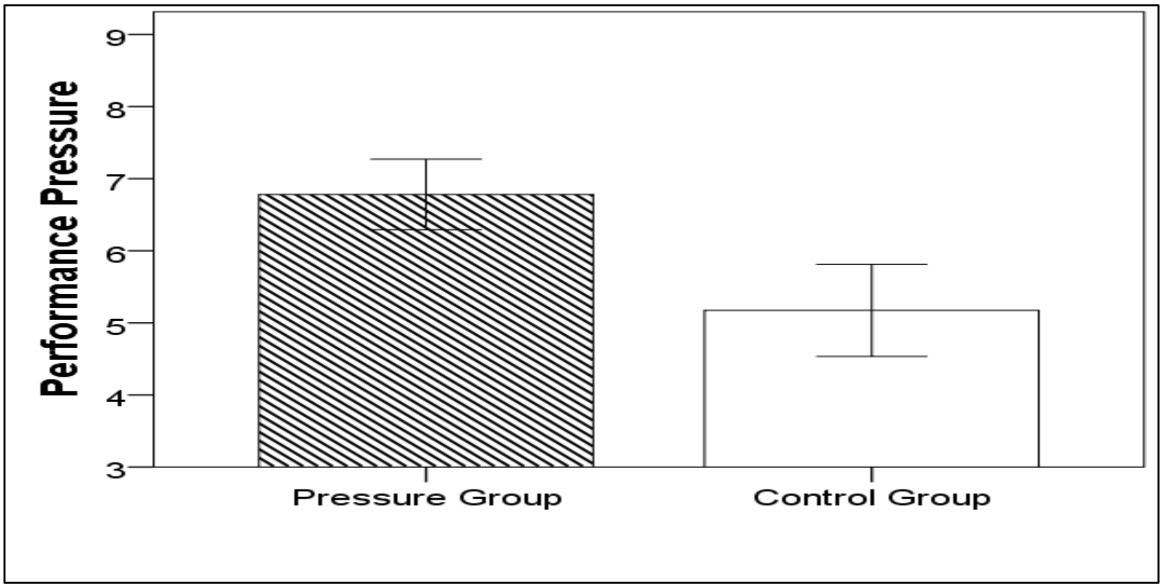

Figure 3.1: Levels of perceived pressure to improve performance during the second attempt. Error bars indicate 95% confidence interval.

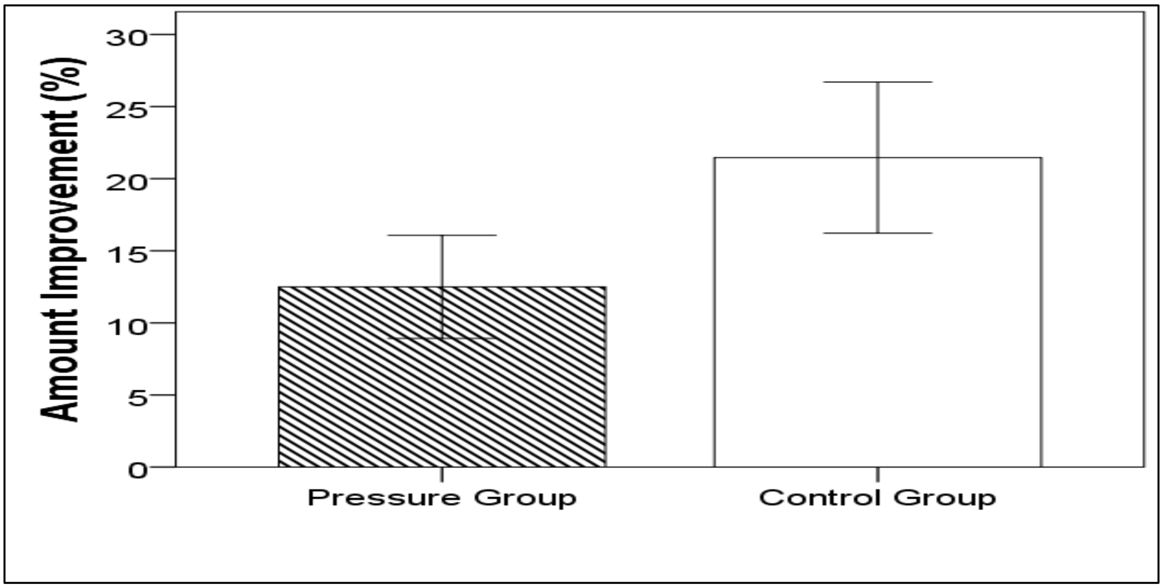

Figure 3.2: Amount of perceived improvement during the second attempt. Error bars indicate 95% confidence interval.

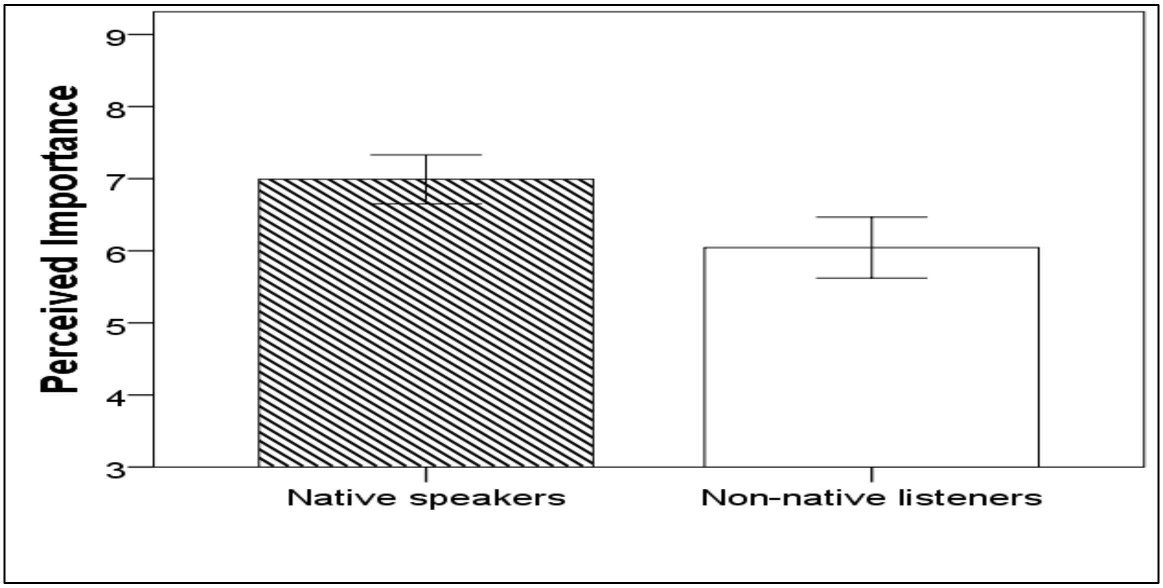

Figure 3.3: Levels of perceived importance to perform at a high level during the second attempt. Error bars indicate 95% confidence interval.

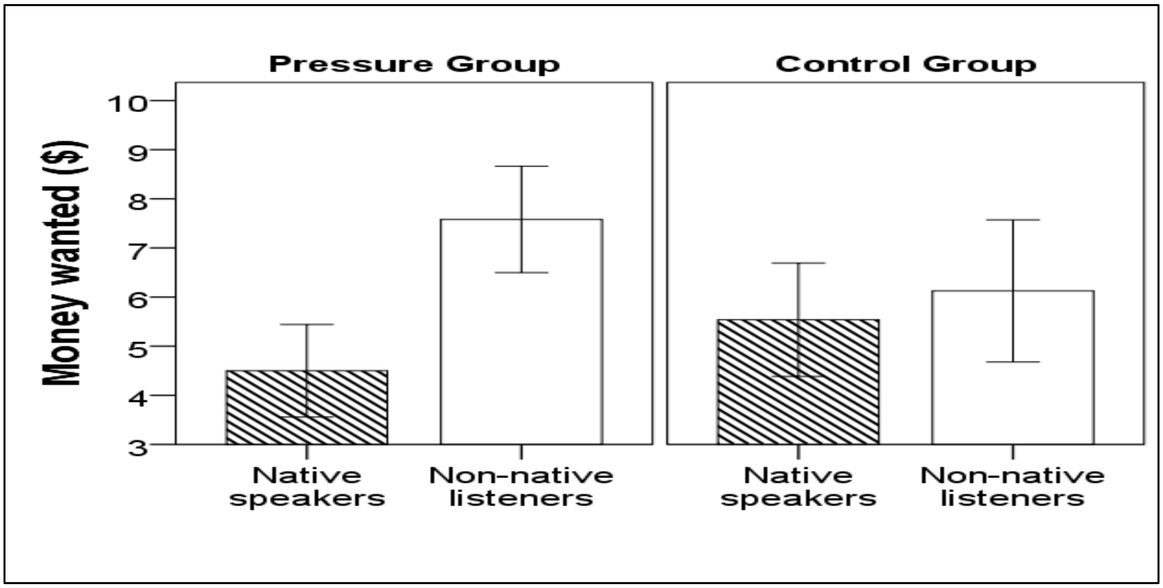

Figure 3.4: Minimum amount of money wanted to improve performance during the second attempt. Error bars indicate 95% confidence interval.

The finding of language-group differences in a) perceived importance of good performance and b) amount of money wanted between the native-listener and non-native performance-pressure groups were unexpected and might be confounding factors in examination of performance pressure. Gimmig and colleagues (2006) found that level of performance pressure is a function of perceived relevance of participants’ personal interests. Because language-group differences in perceived importance may reflect differences in perceived relevance of personal interests, it is unknown whether native listeners and non-native listeners experienced comparable levels of performance pressure. Language-group differences in amount of money wanted might also indicate differences in task commitment. Taken together, because language-group differences in task-related attitudes may be confounding factors that modulate the effect of performance pressure, the performance the non-native listeners was not analyzed further in subsequent sections.

CHAPTER 4: VERBAL TRANSFORMATION

Perceptual illusions, which indicate temporary breakdowns in veridical perception and a deterioration of perceptual accuracy, provide an effective method to examine the mechanistic details of speech processing (Pitt & Shaof, 2002; Warren, 1968; for a review, see Warren, 1996). In the auditory domain, verbal transformation is an extensively-studied illusionary phenomenon (Pitt & Shaof, 2002; Sato et al., 2004; Sato, Schwartz, Abry, Cathiard, & Loevenbruck, 2006; Warren, 1961a, 1968; Warren and Gregory, 1958; Warren & Warren, 1966). In a typical task used to elicit a verbal transformation effect, participants listen to the same spoken word or syllable repeatedly at a fast rate for an extended time and respond whenever perceptual changes occur. Though the spoken stimulus never changes, listeners begin to perceive abrupt and compelling changes in the verbal forms at some point during the stimulus presentation (Warren, 1961a, b). Their perceptions sometimes transform back and forth to the veridical percept (e.g., *tress* → *dress* → *tress*, etc.), in addition to occasional idiosyncratic transformations. The verbal transformation effect has been replicated in various populations, such as children (Warren & Warren, 1966), young and aging adults (Warren, 1961b), and individuals affected by schizophrenia who report hallucinatory experiences (Bullen, Hemsley & Dixon, 1987; Catts, Armstrong, Norcross & McConaghy, 1980; Haddock, Slade, & Bentall, 1995; Slade, 1976).

Despite years of study, the mechanisms underlying the verbal transformation effect remain poorly understood. Verbal transformation may result from two co-occurring processes (Warren 1996; Bashford, Warren, Lenz, 2006, 2008) – repetition of speech signals to satiate memory representation (i.e., habituation-induced fatigue), followed by a shift in the criteria used to categorize the speech signals. When alternate representations are deemed more plausible for matching the input than the original representation, listeners report hearing perceptual changes (for a review, see Warren, 1996). An alternate explanation roots in spreading activation. Within this framework, repetition-induced habituation lowers the activation level of the original memory representation of the input stimulus. At the same time, activation of the input stimulus spreads to its structurally-similar phonological (or semantic) neighbors. The memory representations of alternate percepts progressively increase in their activation levels and compete for recognition with the original representation as alternate percepts (Munson & Brinkman, 2004). More recent studies suggest verbal transformation is associated with perceptual coherence (e.g., Stachurski, Summers, & Roberts, 2015, 2017). For instance, Stachurski and colleagues (2015) showed that removal of formant transitions and abrupt changes in F-0 contour in the speech stream increased numbers and/or forms of verbal transformations.

Verbal transformation has been traditionally treated as a "pure" perceptual phenomenon. However, because the instructions used in the verbal transformation paradigm are suggestive of changes in the speech signals (i.e., report changes whenever you hear one), some studies have examined the extent to which changes in instructions influence the number of verbal transformations. For instance, when participants were informed they would be presented with the same word repeatedly, verbal transformations continued to occur, though at a reduced frequency (for a review, see Natsoulas, 1967). In another study, Taylor and Henning (1963) informed one group of participants the changes would be real words in English, while telling another group the changes would be a repertoire of both real and nonsense words. Taylor and Henning found their instructions had no effect on the total number of real-word transformations, while the unrestricted group reported hearing significantly more nonsense words. Notably, studies have also shown that instructional bias is stronger in atypical populations than in healthy individuals. Haddock and colleagues (1995) found instructional bias ("the word *may/will* change to other words") only influenced hallucinating patients, and not healthy controls. Taken together, these studies illustrate some biasing effects of instructions on verbal transformation; however, instructional bias cannot fully account for the robustness of this illusionary phenomenon.

Few studies have directly examined the extent to which prior task-exposure might influence listeners' tendency to report transformations. No studies have examined the influence of performance pressure on the verbal transformation effect, though distraction theory posits that worrying thoughts induced by performance pressure may impact performance in a manner similar to a dual-task design (Beilock & Carr, 2005). Lexical activation is key to driving the verbal transformation effect under both satiation-criterion-shifting framework and spreading-activation framework. In high performance-pressure condition, the need to suppress worrying thoughts may prevent listeners from attending to the speech stream, slowing down lexical activation in listeners and resulting in a weakened transformation effect. In contrast, the desire for good performance may prompt listeners to search for changes in the speech stream intentionally, disrupting the formation of perceptual coherence of speech stream and elevating the transformation effect.

4.1 METHODOLOGY

4.1.1 Stimuli

Multiple productions of the word *tress* were recorded by a female native speaker of American English in a sound-attenuated booth, with a sample rate of 44 kHz and sample resolution of 16 bits. The clearest token was selected and combined repeatedly to create a three-minute audio clip with approximately two tokens per second, making a total of 330 identical tokens.

4.1.2 Procedure

The participants listened to the audio clip, presented at 60dB, through Sennheiser HD-280 Pro headphones in a sound-attenuated room. Following Warren (1961a, b), participants were instructed to tell the researchers what the voice said as soon as the clip started, and then keep listening to the clip. Participants were instructed to press a button and call out what they heard each time the words seemed to change, without worrying about being right or wrong. Participants' verbal responses were audio-taped.

4.2 RESULTS

Regarding baseline performance, the number of verbal transformations ranged from zero to 34 times for the controls, and from one to 42 times for the performance-pressure group.

4.2.1 Changes in number of verbal transformations

We first report the analysis on the changes in number of verbal transformations reported during baseline performance and the second attempt. The following analyses were used linear mixed-effect modelling (LMER) with the lme4 package 1.1-13 (Bates, Maechler, & Bolker, 2012) in RStudio 1.0.143. The analyses started with the maximum model, which included pressure (i.e., performance pressure vs control), attempt (i.e., baseline vs 2nd attempt), and their interactions as fixed effects. By-participant intercepts were included in the model as random effects. The model was refined by removing, one at a time, factors that exhibited the highest p-value, while retaining the hierarchical rule of interactions. Likelihood ratio comparisons were performed to confirm that including a given factor did not improve the amount of variance explained (Baayen, Davidson, & Bates, 2008).

Model comparison indicated a critical two-way Pressure X Attempt interaction [χ^2 (1) = 4.6, p<.05]. We analyzed the two-way interactions using the lsmeans package 2.27-2, which was designed to obtain the least-squares means and to test linear contrasts for linear and generalized mixed models (Lenth, 2016). The analysis indicated that both controls and motivation group reported greater number of verbal transformation during the second attempt than they did during baseline testing [performance-pressure group: β= 9.2, SE=1.8, p<.0001; controls: β= 3.8, SE=1.8, p <.05]. Notably, the motivation group also reported more transformations than did the controls during the second attempt [β= 9.5, SE=3.6, p <.05] while both groups performed comparably during baseline (p>.25) (Fig. 4.1).

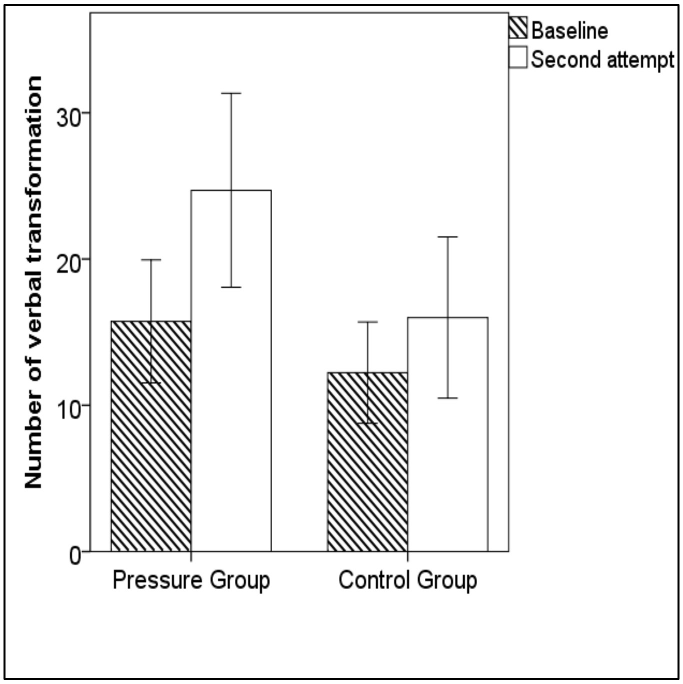

Figure 4.1: Total number of verbal transformations during baseline and the second attempt. Error bars indicate 95% confidence interval.

4.2.2 Onset latency of the first transformation

We then reported the analysis on onset latency of the first transformation, with the same analytic approach described in Section 4.2.2. The dependent variable was the number of repetitions of the stimulus “tress” before the first transformation.

Model comparisons indicated the presence of simple effect of Attempt [β= -21.1, SE= -5.1, χ^2 (1) = 14.3, p<.001], indicating both performance-pressure and control groups reported the first transformation earlier during the second during than they did during baseline testing.

4.3 DISCUSSION

Verbal transformation is a compelling, yet poorly-understood illusionary phenomenon in the speech-processing domain (Warren, 1961a; Pitt & Shaof, 2002; Stachurski et al., 2015, 2017). By manipulating performance pressure in native listeners, this study yielded two findings. First, the performance-pressure group reported more verbal transformations than baseline performance and the control group in high performance-pressure conditions, suggesting a robust effect of performance pressure. Second, the control group also reported more verbal transformations during their second attempt, though the amount of changes was statistically smaller than the motivation group. Taken together, Experiment 1 showed that native listeners generally exhibited an elevated susceptibility to auditory illusions when asked to perform the task again.

Verbal transformation has been viewed as a continuous process of organizing and evaluating speech signals repeatedly presented out of context (Warren, 1968; Warren & Meyers, 1987). When repetitive presentation of input stimulus satiates memory representation (i.e., fatigue), the criteria used to categorize speech signals shift to locate a closer match with the input stimulus (Warren 1996; Bashford et al., 2006, 2008). The critical finding of this study is that motivation exerted a robust effect on the number of verbal transformations reported in native listeners. Note that all transformations reported by listeners were essentially mismatches between perception and reality, indicating perceptual inaccuracy. Though the participants were told not to worry about right or wrong responses in a typical task used to elicit the verbal transformation effect, they were never guaranteed changes existed – which they, in fact, did not. While the instructions were suggestive of changes, our data showed a small number of listeners (only one of 60 participants) never reported any verbal transformations during either baseline performance or their second attempt. Additionally, multiple studies have shown auditory illusions cannot be solely explained by instructional bias. For instance, Natsoulas (1965) showed verbal transformations persisted, even after listeners were told the stimulus on the tape was always the same. Haddock and colleagues (1995) showed changes in instruction exerted a negligible effect on healthy adults. In their study, they compared the effect of two instructions that should influence the *perceived* likelihood of sound changes among listeners (i.e., "the word *may* change to other words" vs. "the word *will* change to other words"). While Haddock and colleagues found changes in instruction influenced only hallucinating patients, our data showed performance pressure exerted a robust effect on healthy individuals.

Why did listeners report a more illusionary perception in high performance-pressure conditions? It is likely that performance pressure, in combination with the suggestive nature of the instructions, invited listeners to search for changes within the speech stream. Recent studies suggest verbal transformation can be influenced by perceptual incoherence (e.g., Stachurski et al. 2015), perceptual segregation (e.g., Stachurski et al., 2017), whether the signal is perceived as speech or non-speech (Pitt & Shoaf, 2002), or by top-down lexical knowledge (Billig et al., 2013). While these studies have focused on stimulus properties and lexical knowledge, our findings suggest performance pressure also plays a role in verbal transformation. Our findings of an elevated susceptibility to illusionary perception does not align with the predictions from distraction theory. When verbal transformation is primarily driven by activation of input stimulus that triggers either satiation or spreading activation of structurally similar neighbors, the presence of distracting, worrying thoughts should interfere with activation of input stimulus.

Instead, an elevated susceptibility to illusions may align with the explicit monitoring theory of performance pressure, which holds that motivation for good performance directs one's attention to automatic processes and disrupts implicit performance. Relative to low performance-pressure condition where poor performance would not inflict any negative consequence, listeners might consciously scrutinize their own performance and analyze each activated lexical items with greater cautions. Elevated attention to activated lexical items might, paradoxically, induce faster satiation of the mental representation and cause earlier criterion-shift. An outstanding question is the extent to which one's tendency to search for changes in a speech stream impacts the formation of perceptual coherence of speech signals. When listeners intentionally search for changes, these changes in intention may disrupt the formation of speech coherence, increasing the tendency of perceptual regrouping of segments, and thus yielding a greater transformation effect.

Interestingly, Experiment 1 showed that, besides the influence of performance-contingent rewards and peer pressure, the controls also reported more verbal transformations during the second attempt than they did during baseline testing. It is important to emphasize that the motivation group did report more transformation than the controls did during the second attempt. However, the finding of an elevated susceptibility to illusions in controls is noteworthy because this group was not prompted to improve their performance by either monetary rewards or pressure to work harder for a partner. Though our motivation manipulation, through extrinsic motivators (i.e., money, partners), exerted an unambiguous effect on listeners' perceived pressure to improve their performance, responses from the controls suggest it could not fully account for changes in performance. In fact, more than half of the controls (63%) reported more verbal transformations during their second attempt.

A more likely reason is that native listeners might experience an inherent need for good performance when they were presented with a task and stimuli again. Table 3.3 and Figure 3.3 highlighted that, though the controls were not given any extrinsic incentives, they did not differ from the performance-pressure group in perceived importance of good performance. In other words, the controls' performance during their second attempt might have been driven by inherent need. The native-listener performance-pressure group's performance was then driven by a combination of intrinsic need and performance pressure (i.e., monetary rewards). The current study cannot tease apart the influence of intrinsic and extrinsic motivation on illusionary listening; however, the Pressure X Attempt interaction in this study suggests the performance-pressure group might have been more strongly influenced by motivation than the controls. In other words, the effect of intrinsic and extrinsic motivation on illusionary listening might be at least partly additive.

Previous studies on verbal transformation have viewed it as a perceptual phenomenon, and paid little attention to the role of motivation. This dissertation has shown that motivation-induced performance pressure exerted an unambiguous and robust effect on this widely-studied phenomenon. This dissertation's findings highlight the need to address motivation in verbal transformations, and possibly in other speech-processing phenomena. Based on the finding that auditory streaming and verbal transformation share common brain networks (e.g., Kashino and Kondo, 2012), examining the influence of motivation on these networks’ activation patterns may shed new light on the interaction between hearing and cognition.

CHAPTER 5: THE *GANONG* EFFECT

The so-called *Ganong* effect, which indexes degree of lexical activation and its suppression during phonetic processing (e.g., Ganong, 1980; Mattys & Wiget, 2011; Pitt & Samuel, 2006), has been central to a long-lasting debate about the role of top-down lexical influence in speech processing (McClelland et al., 2006; Norris et al., 2000). During a typical task used to elicit the *Ganong* effect, listeners are instructed to identify the initial phoneme presented in a syllabic context. The phonemic stimuli vary in small increments on an acoustical continuum (e.g., voice onset time), with the ends of the continuum forming either an existing word, or a nonword with the carrier syllable. In the original experiment, listeners categorized target phonemes embedded at the initial position of two word contexts, /kɪs/ (“kiss”) and /gɪft/ (“gift”). Though listeners were explicitly told to ignore the word contexts and focus on the word-initial phoneme, listeners often responded to an identical token differently, and based their responses on the word context (Ganong, 1980). For phonemes embedded in a giss-kiss continuum (nonword vs. real word), listeners reported a bias for /k/ responses. For phonemes embedded in a gift-kift continuum (real word vs. nonword), listeners reported a bias for /g/ responses.

An unresolved controversy is whether top-down lexical influence originates from an interactive speech processing system (e.g., TRACE; McClelland & Elman, 1986) or an autonomous speech processing system (e.g., Merge; Norris, McQueen, & Cutler, 2000). The interactive view argues for a lexical influence on pre-lexical representation of the speech sounds, assuming automatic bi-directional feedback from activation within multiple representational layers (i.e., acoustic/articulatory features, phonemes, words). In contrast, the autonomous view argues for a supervision-free bottom-up process and that lexical influence arises only at later decisional stages. While the interactive view assumes a facilitative or compensatory effect of top-down influence on perceptual performance, the autonomous view questions the extent to which top-down influence could benefit speech processing (for a review, see McClelland, Mirman, & Holt, 2006 and Norris et al., 2000). Early studies of the *Ganong* effect adopted a linguistic approach, focusing on the extent to which the properties of lexical contexts (e.g., word length, acoustic structure, etc.) influence phoneme categorization (Burton, Baum, & Blumstein, 1989; Connine, Clifton, & Cutler, 1987; McQueen, 1991; Pitt & Samuel, 2006). Recent studies have found that the *Ganong* effect may be influenced by cognitive factors, such as attention and top-down control (e.g., Mirman, McClelland, & Holt, 2006; Mirman, McClelland, Holt, & Magnuson, 2008).

In a series of studies, Mattys and colleagues have shown that the *Ganong* effect is associated with divided attention in dual-task conditions (Mattys & Wiget, 2011), cognitive decline associated with aging (Mattys, & Scharenborg, 2014), and weakened executive control due to induced anxiety (Mattys, Seymour, Attwood, & Munafò, 2013). For instance, Mattys and Wiget (2011) asked participants to decide whether a red square was present on a visual display and, at the same time, categorize the initial phoneme embedded in the –iss and –ift continuum. Relative to conditions where listeners were administered only the phonemic categorization task, divided attention in the dual-task conditions significantly increased the amount of lexical bias in listeners. In a later study, Mattys and colleagues (2013) showed that anxiety, which affects prefrontal brain regions (Bishop, 2009) associated with attentional control (Buschman & Miller, 2007), elevated lexical bias at a magnitude comparable to Matys and Wiget's (2011) dual-task design. Lam and colleagues (2017) further showed that adult native listeners who exhibited a greater *Ganong* effect also exhibited selectively poorer word recognition performance in adverse listening conditions, such as when competing talkers are present and introduce significant lexical interferences to the listeners. Taken together, the *Ganong* effect might reflect individual differences in susceptibility to lexical interferences and difficulties ignoring lexical influence during phonetic processing.

Few studies have directly examined the extent to which prior task-exposure might reduce lexical influence on phonetic processing. Also, no studies have examined the influence of performance pressure on the *Ganong* effect, though distraction theory posits that worrying thoughts induced by performance pressure may impact performance in a manner similar to a dual-task design (Beilock & Carr, 2005). According to Mattys and colleagues' studies, both divided attention and induced anxiety distract listeners from the task at hand (i.e., phonemic categorization), and/or compete for limited cognitive resources, thus elevating lexical influence. According to distraction theory, the need to suppress worrying thoughts in high performance-pressure conditions may reduce the amount of resource reserved for suppressing lexical influence. As a result, listeners might fail to focus on the target phonemes flexibly, resulting in an elevated lexical bias as the "choking under pressure" effect.

5.1 METHODOLOGY

5.1.1 Stimuli

Preparation of the stimuli followed the procedures described in Mattys and Wiget (2011), and contained three, eight-step continua: *gift* - *kift* (real word vs. nonword), *giss* - *kiss* (nonword vs. real word), and *gi* - *ki* (nonword vs. nonword). The *gift* - *kift* and *giss* - *kiss* continua were included to estimate the magnitude of lexical influence, while *gi* - *ki* was included to assess lexical-influence-free phoneme categorization. Phoneme categorization performance in the *gi* - *ki* condition provided information about whether differences in the amount of lexical influence on phoneme categorization were due to differences in strength of lexical influence *per se*, or to differences in acoustic-phonetic processing of /*g*/ and /*k*/ (e.g., Mattys & Wiget, 2011, Mattys et al., 2013; Mattys & Scharenborg, 2014). Multiple productions of the words *gift* and *kiss* were recorded, in a sound-attenuated booth by a female native speaker of American English, with a sample rate of 44 kHz and sample resolution of 16 bits. The clearest tokens were selected, and then split into initial consonant */g/* - */k/*, vowel */ɪ/*, and coda */ft/* and */s/*. We chose a vowel that originated from one of the *kiss* tokens, because it exhibited relatively neutral coarticulation. We created an eight-step continuum of the initial consonants */g/* - */k/* out of */k/* by editing out the aspiration, using Praat (Boersma & Weenink, 2010). The final continuum had the following voice onset time (VOT) values: 15 ms, 23 ms, 28 ms, 33 ms, 38 ms, 43 ms, 48 ms, and 56 ms. The difference in VOT increments was a strategic decision; the end points are further away to ensure clear end points. The continuum in the middle is denser to capture the VOT area of uncertainty. The eight-step */g/* - */k/* continuum was then recombined with the vowel */ɪ/*, and coda of either */ft/* or */s/*, making a total of 24 syllables. The average duration of the three continua of *gift* - *kift*, *kiss* - *giss*, and *gi* - *ki* were 521 ms, 468 ms, and 193 ms, respectively. The procedures and properties of the stimuli were consistent with those used in previous studies (Mattys and Wiget, 2011; Mattys et al., 2013).

5.1.2 Procedure

During testing, the stimuli were binaurally presented to participants over Sennheiser HD-280 Pro headphones, at a comfortable, fixed listening level (~70dB SPL). Each of the 24 syllables that formed a VOT continuum (in the contexts of *–i, –ift,* and *–iss*) was presented randomly five times, resulting in a total of 120 trials. Before the experiment, participants were explicitly instructed to focus on the initial consonant, and to ignore the meaning of the syllables. On each trial, after stimulus presentation, participants were asked to decide whether the first sound of the syllable was /*g*/ or /*k*/. Following the procedures used in Mattys and Wiget (2011) and Mattys and colleagues (2013), the magnitude of susceptibility to lexical influence was calculated as the average percentage difference in identification in the *giss-kiss* continuum from the counterparts in the *gift-kift* continuum, across the eight steps. The method to derive the amount of lexical influence is identical to the method published in Lam and colleagues (2017). Figure 5.1 illustrates the *Ganong* effect by showing data from two participants published in Lam and colleagues (2017), one with high lexical influence (Fig 5.1, left panel) and one with low lexical influence (Fig. 5.1, right panel).

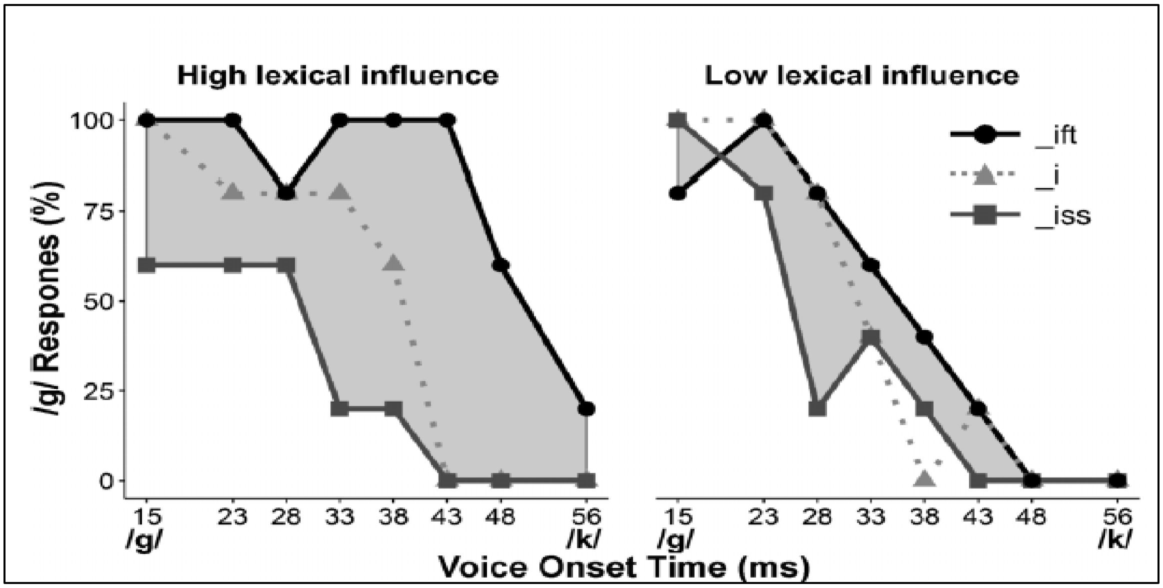

Figure 5.1: Percentage of /g/ responses for the gift-kift (solid black line with circles), gi-ki (light gray dotted line with triangles), and giss-kiss (dark gray solid line with squares) continua, from an exemplar participant with high susceptibility to lexical influence (left panel) and an exemplar participant with low susceptibility to lexical influence (right panel). The magnitude of lexical influence is calculated as the average proportion differences of /g/ response between the gift-kift continuum and giss-kiss continuum across the eight steps of voice onset time (VOT).

5.2 RESULTS

The average percentage differences in identification in the giss - kiss continuum from their counterparts in the gift - kift continuum across the eight steps was used to measure the magnitude of susceptibility to lexical influences (Lam et al., 2017; Mattys and Wiget, 2011; Mattys et al., 2013). Percentage differences in /g/ responses suggest inconsistencies in listeners' responses to the same token. Regarding baseline performance, the amount of susceptibility to lexical influences ranged from 9.4% to 100% for the controls, and from 9.4% to 96.9% for the performance-pressure group.

5.2.1 Changes in susceptibility to lexical influences

We first report the analysis on the changes in susceptibility to lexical influences during baseline performance and the second attempt. The following analyses used linear, mixed-effect modelling (LMER) with the lme4 package 1.1-13 (Bates, Maechler, & Bolker, 2012) in RStudio 1.0.143. The analysis started with the maximum model, which included performance pressure (i.e., performance pressure vs control), attempt (i.e., baseline vs second attempt), and their interactions as fixed effects. By-participant intercepts were included in the model as random effects. The model was refined by removing the factors exhibiting the highest p-value, one at a time, while retaining the hierarchical rule of interactions. Likelihood ratio comparisons were performed to confirm that including each factor did not improve the amount of variance explained (Baayen, Davidson, & Bates, 2008).

The model comparison indicated the presence of a two-way Pressure X Attempt interaction [χ^2 (1) = 4.8, p<.05]. No other simple effects or interactions were significant (all ps >.1). The two-way Pressure X Attempt interaction indicated the controls were less susceptible to lexical influences during their second attempt, than during baseline performance [β= -15.7, SE=.3.4, p<.0001], while the performance-pressure group members were comparably susceptible to lexical influences at both times (p=.13) (Fig. 5.2). Finally, there were no group differences in susceptibility to lexical influences during either baseline performance or the second attempts (both ps>.25).

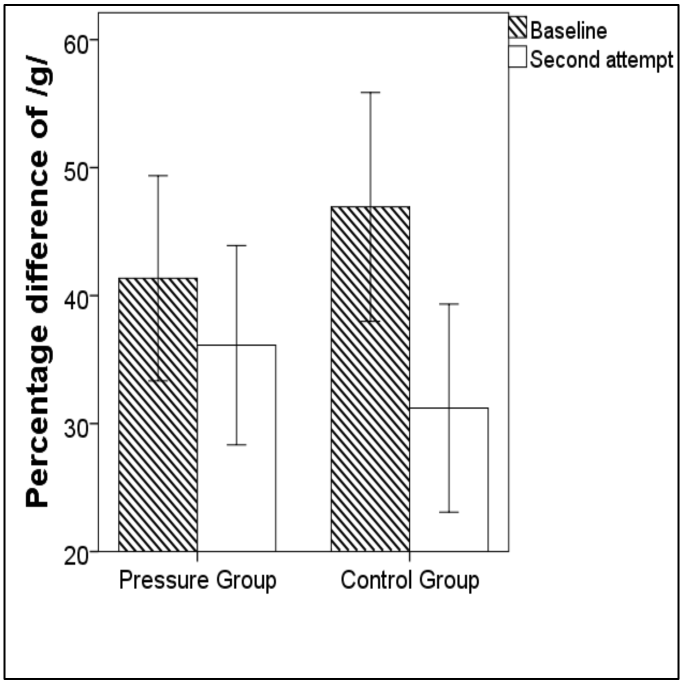

Figure 5.2: Susceptibility to lexical influences during baseline performance and the second attempt. Susceptibility was calculated by the percentage difference of /g/ responses in -iss and -ift context. Error bars indicate 95% confidence interval.

The analysis also examined participants' phonemic categorization when the target phoneme was embedded in the -i context, where lexical influence was minimized. The dependent variable was the average percentage of /g/ response across the eight steps of voice onset time. The analysis showed that the native listeners categorized the target phonemes with excellent consistency during baseline performance and the second attempt, at the group level (Fig. 5.3). None of the variables of interest (i.e., performance pressure, attempt) predicted the percentage of /g/ response when the phoneme was embedded in the -i context (all ps >.5).

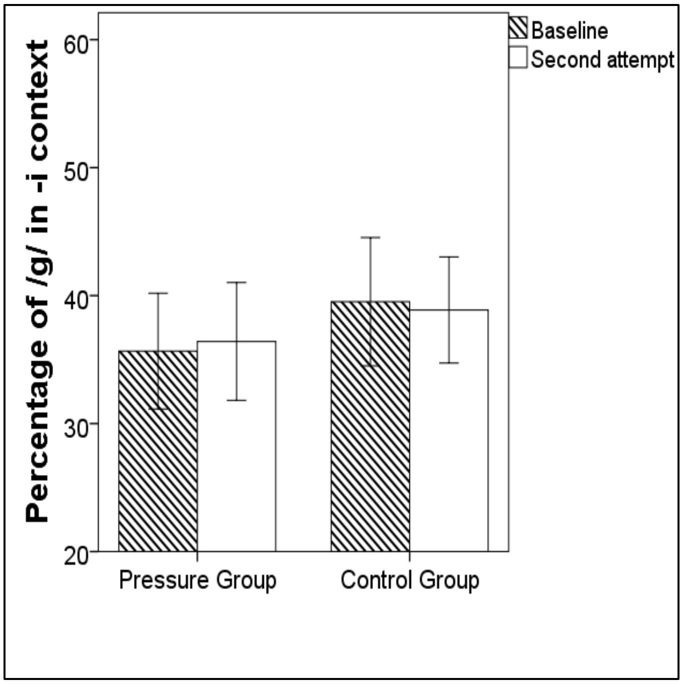

Figure 5.3: Percentage of /g/ responses in -i context during baseline performance and the second attempt. Error bars indicate 95% confidence interval.

5.2.2 Changes in response time to categorize initial phonemes

We then report the analysis on response time of participants to categorize initial phonemes during baseline performance and the second attempt. The analysis adopted the same analytic approach described in Section 5.2.1. The analysis indicated that attempt was the only significant predictor of response time. Both performance-pressure and control groups responded to the task faster during their second attempt than during baseline performance [β = -104.6, SE=26.8, χ^2 (1) = 13.8, p<.001].

5.3 DISCUSSION

Lexical influences on speech processing are pervasive, yet poorly understood (for a review, see Samuel, 2011). The *Ganong* effect has been viewed as a phenomenon illustrating the interplay between lexical activation and phonetic processing (Mattys & Wiget, 2011). On the one hand, the interactive view holds that phonetic processing and lexical activation exhibits an early, bi-directional interaction (McClelland & Elman, 1986; McClelland, Mirman, & Holt, 2006). On the other hand, the autonomous view maintains that lexical activation exerts no influence on pre-lexical processing until later decisional stages (Norris et al., 2000). Experiment 2 manipulated performance pressure and examined the choking effect on the *Ganong* effect in native listeners, yielding two primary findings. First, the controls exhibited a practice effect and improved in the ability to suppress lexical influence. Second, listeners failed to take advantage of prior exposure when they were placed in high performance-pressure conditions.

When acoustic-phonetic segments are ambiguous, perception is often driven by lexical factors (Ganong, 1980; for a review, see Samuel, 2001). However, whether phonetic details can be fully integrated at a later stage of processing may also depend on the extent to which lexical influence can be suppressed (Gow, Segawa, Ahlfors, & Lin, 2008; Mattys & Wiget, 2011), which results in the *Ganong* effect. Previous studies have shown changes in stimuli properties can manipulate listeners' attention to either the target phonemes or entire words, and modulate the amount of lexical bias (McClelland, Mirman, & Holt, 2006; Mirman, McClelland, Holt, & Magnuson, 2008; Norris et al., 2000). More recent studies have put greater effort into understanding the role of cognition in individual differences in lexical bias (Lam et al., 2017; Mattys et al., 2013; Mattys & Scharenborg, 2014; Mattys & Wiget, 2011). A general theme of such studies is that lexical bias is linked to difficulties in suppressing lexical influence. For instance, a greater amount of lexical bias has been linked to conditions where attentional control may be compromised, such as cognitive decline associated with aging (Mattys & Scharenborg, 2014), or divided attention when people need to perform multiple tasks simultaneously (Mattys & Wiget, 2011).

Experiment 2 showed native listeners could reduce lexical bias and improve their ability to ignore lexical interferences upon prior exposure. This finding highlights that the dynamic interplay between lexical activation and phonetic processing could change and decrease lexical influence over time, with practice. With prior exposure, listeners might guide their attention to the target phonemes more effectively, and improve the consistency with which they categorized phonemes. Reduction in lexical bias showed native listeners could, with practice, could shift their attention flexibly according to specific listening conditions and instructions (i.e., focus on the phonemes and ignore the meaning of the syllables).

The critical finding of Experiment 2 is that improvement was much reduced in high performance-pressure conditions. When native listeners were made aware of the importance of good performance, the practice effect was absent at the group level, though listeners categorized phonemes faster. According to distraction theory, performance pressure might introduce worrying thoughts and create a condition resembling a dual-task environment (Beilock & Carr, 2005). The need to suppress interfering thoughts might have prevented listeners from shifting their attention flexibly to the target phonemes, resulting in a much-reduced practice effect. One may argue reduced practice effect may result from a strategic decision; because most stimuli are perceived as ambiguous in the *Ganong* paradigm, performance pressure might prompt listeners to respond with the phoneme that best fits the syllabic context, as a perceivably more reliable approach. In other words, the absence of a practice effect may be the result of changes in response strategies, instead of a weakened ability to suppress lexical interference or ignore irrelevant information. However, the strategic approach should instead lead to a much greater increase in lexical bias that goes beyond the baseline performance, which is not what Experiment 2 found.

Note that our findings are at odds with those of Mattys and colleagues. While Experiment 2 showed a reduced practice effect, Mattys and Wiget demonstrated an elevated lexical bias. As a follow-up study, lexical bias was also elevated when listeners were subjected to induced anxiety (Mattys et al., 2013). Both Experiment 2 and Mattys and colleagues' study point to a deteriorating effect of distractions (pressure-induced worrying thoughts, divided attention, anxiety) during speech processing; however, the magnitude of the deteriorating effect may depend on the sources of distraction. While distraction theory posits that worrying thoughts create a dual-task environment for cognition, Mattys and Wiget (2011) adopted an actual dual-task design by asking listeners to categorize phonemes, while performing a visual search at the same time.

In Mattys and Wiget (2011), the dual-task condition required participants to decide whether a red square were present in a 2 X 2 or 6 X 6 matrix on the screen, in *every* trial. Participants were pressed for time to respond to the visual search task, because each visual stimulus lasted a mere half-second. In their later study, where anxiety was induced, the manipulation was delivered by increasing atmospheric carbon dioxide via an oro-nasal face mask (Mattys et al., 2013). The dual-task design and induced anxiety might have created distractions at a consistently pervasive level, throughout the entire task. In the study on induced anxiety, changes in the composition of inhaled air may have induced temporary physiological changes. In contrast, the amount of worrying thoughts induced by performance pressure may fluctuate throughout the study, due to within-task adaptation. Because worrying thoughts may not exert an equal effect across time, the deteriorating effect of performance pressure might reduce the practice effect but yet to be strong enough to elevate lexical bias.

Taken together, Experiments 2 showed the dynamic nature of the *Ganong* effect that native listeners can improve their performance and ignore lexical interference, with practice. However, the amount of benefits could be reduced when there may be concurrent, worrying thoughts in high performance-pressure conditions. Future studies on the *Ganong* effect may examine the extent to which instruction and task design interact with performance pressure. By providing listeners feedback about whether their responses on the current trial are consistent with their pervious responses, listeners might experience performance pressure more consistently throughout the experiment, and thus increase their lexical bias. How speech processing relates to one's ability to remain goal-directed and suppress irrelevant information remains relatively unclear in both interactive and autonomous models (Connine & LoCasto, 2000). Our findings suggest the need to consider cognitive factors, such as ability to ignore irrelevant information, in understanding the intricate interaction between bottom-up and top-down processes.

CHAPTER 6: SPEECH RECOGNITION IN NOISE

Speech recognition in noise presents a significant challenge to the perceptual, linguistic, and cognitive systems (Mesgarani & Chang, 2012; Xie, Maddox, Knopik, McGeary, & Chandrasekaran, 2015). Though motivation is increasingly viewed as a useful dimension for studying listening in challenging, effortful conditions (Pichora-Fuller et al., 2016), the extent to which motivation-induced performance pressure impacts speech recognition in noise is extremely limited. The challenge presented by background noise originates from at least two sources – energetic masking, and informational masking (Brungart, 2001). Maskers with significant energetic masking (EM) overlap the target speech spectrally and temporally. Informational masking (IM), often characterized as interference from noise once its EM effect is accounted for (Cooke, Lecumberri, & Barker, 2008), is particularly an issue in environments where competing talkers are present, such as cocktail parties (Cherry, 1953; for a more recent review, see Bronkhorst, 2000). The differentiation between EM and IM is important, because both exert distinct effects on perceptual and cognitive systems (e.g., Zekveld, Rudner, Johnsrude, & Rönnberg, 2013). Current understandings about EM and IM suggest EM renders target speech partially inaudible and degrades the neural representation thereof at the level of the auditory periphery. In contrast, IM interferes with speech processing at higher auditory, linguistic, and cognitive levels (Arbogast, Mason, & Kidd, 2002; Freyman, Balakrishnan, & Helfer, 2004; Xie et al., 2015). These central interferences include increased lexical interferences, greater levels of cognitive/memory load, and attention distraction from the target (Cooke et al., 2008; Mattys et al., 2012).

Noises with competing talkers introduce significant higher-level lexical interferences and IM. Prior exposure to the masker speech (Van Engen & Bradlow, 2007;

Van Engen, 2012) and/or higher degrees of similarities between the masker and the target speech (Brouwer, Van Engen, Calandruccio, & Bradlow, 2012) predict greater decrements in speech recognition accuracy, due to elevated interference. When the linguistic information in the maskers is perceivable, the goal of the auditory system is to attend to the target speech selectively to improve word recognition, an ability termed auditory object selection (Shinn-Cunningham, 2008; for a recent review, see Shinn-Cunningham, Best, & Lee, 2017). Emerging evidence suggests cognitive factors (e.g., WM, attention) contribute to individual differences in speech recognition, particularly in conditions involving competing talkers that require listeners to ignore irrelevant background linguistic information (e.g., Chandrasekaran, Van Engen, Xie, Beevers, & Maddox, 2015; Xie et al., 2015; Zekveld et al., 2013).

Limited information is available about the extent to which prior noise-exposure improves speech recognition performance in noise as a function of noise types. Also, no studies have examined the influence of performance pressure on speech recognition in noise. Distraction theory posits that worrying thoughts induced by performance pressure create distraction and/or compete for cognitive resource (Beilock & Carr, 2005). In high performance-pressure conditions, the need to suppress worrying thoughts may exert additional challenge for the speech processing system to suppress lexical interferences in competing-talker conditions and selectively focus on the target speech. In contrast, for noises that are predominantly EM and do not introduce lexical interferences (e.g., pink noise), successful word recognition might be less dependent on cognitive processes and ability to suppress irrelevant information. Unlike word recognition in competing-talker conditions, word recognition in EM-predominant maskers may involve fewer vacillations between lexical alternatives. Thus, worrying thoughts from performance pressure may have relatively less effect on word recognition in EM than in IM.

To examine the effect of performance pressure on speech-recognition ability in challenging listening conditions, this study employed four types of noises that differ in their amount of lexical interferences and associated IM: one-talker babble (1T); two-talker babble (2T); two-talker babble-like speech (reversed speech; 2T-Reversed); and pink noise. Notably, multiple studies have shown that these four noise types exert distinct effect on speech-recognition ability (e.g., Chandrasekaran, et al., 2015; Freyman et al., 2001; Hygge et al., 1992; Rosen et al., 2013; Xie et al., 2015; Van Engen, 2012; Zekveld et al., 2013). However, no studies have examined the extent to which performance pressure modulate the effect of these maskers on speech-recognition ability. Among these four types of noise maskers, 1T and 2T contained the greatest amount of interfering linguistic information and significant IM, with 2T being the more challenging masker (Miller, 1947; Freyman, Balakrishnan, & Helfer, 2001; Rosen, Souza, Ekelund, & Majeed, 2013). Two-talker reversed speech degrades the 2T masker drastically, via a time reversal procedure (Saberi and Perrott, 1999); specifically, 2T-Reversed preserves partial phonetic information of 2T, yet renders all words embedded in the original masker unintelligible. In other words, 2T-Rerversed is spectro-temporally comparable to 2T, but absent intelligible words. Note that reversed speech maskers are not necessarily less distracting than forward speech maskers, though all intelligible words are removed (Hygge, Rönnberg, Larsby, & Arlinger, 1992). Finally, pink noise is one of the most widely used maskers representing EM (Howard-Jones & Rosen, 1993). Among the four maskers used in this study, pink noise was the “purest” type of EM. Additionally, pink noise has fewer momentary masker energy fluctuations than 2T-Rerversed, which preserves the fluctuations of the 2T masker. Fluctuations in masker energy allow “dip listening” in listeners, because of “glimpsing” acoustic information, which reduces the amount of EM (Howard-Jones and Rosen, 1993; Miller and Licklider, 1950; Rosen, et a., 2013). Finally, this experiment also manipulated signal-to-noise ratio to examine if the negative impact of performance pressure on performance in competing-talker conditions, if found, would be exacerbated in more challenging noise levels.

To summarize, Experiment 3 examined the extent to which performance pressure impacts speech recognition ability in native listeners. To that end, it manipulated levels of performance pressure in listeners, and measured their ability to process speech in noises that varied in amounts of IM and EM.

6.1 METHODOLOGY

6.1.1 Stimuli

6.1.1.1 Target sentences

The target stimuli were taken from the Revised Bamford–Kowal–Bench (BKB) Standard Sentence Test (Bamford & Wilson, 1979). For the current study, a total of sixty-four BKB sentences (identical to the sentence stimuli selected by Chandrasekaran and colleagues (2015)) were selected. The sentence stimuli were divided into two sets of stimulus sentences (Set A and Set B), each with a total of 128 keywords for scoring (32 sentences X four keywords per sentence). Set A and Set B were comparable in their distribution of syntactic structures, the total number of words in each stimulus sentence ($p>.5$), and the word frequency of each keyword ($p>.25$ for log-transformed word frequency for each keyword). The stimuli sentences were recorded by a female native speaker of American English, in a sound-attenuated booth at Northwestern University (Van Engen, 2012). For each target sentence, the RMS amplitude was equalized to 60dB SPL, using Praat (Boersma & Weenink, 2010). The presentation level of the stimulus sentences was confirmed, using an audiometer.

6.1.1.2 Maskers

N-talker babble tracks were created using the following procedures. Eight native speakers of American English, all females, were recruited to produce 30 simple English sentences. Recordings were conducted in a sound-attenuated booth at Northwestern University (Van Engen et al., 2008). Each talker's sentences were first equalized for RMS amplitude, then concatenated to create 30-sentence strings. There was no silence between sentences. One string was used as the one-talker masker track; to generate the two-talker masker track, the string from a second talker was mixed with the first. To generate time-reversed two-talker maskers, the two-talker masker track was reversed along the temporal dimension. To generate steady-state pink noise, this experiment used the Noise Generator option in Audacity (Audacity Developer Team, 2008). All masker tracks were truncated to 50 seconds, and their RMS amplitude equalized to 64dB and 68 SPL, using Praat (Boersma & Weenink, 2010). The presentation levels of the masker tracks were confirmed, using an audiometer.

6.1.1.3 Mixing targets and maskers

For each stimulus set, the 32 target stimuli and maskers were mixed to generate signal-to-noise ratios of −4dB and -8dB (i.e., the noise was 4dB or 8dB higher than the target stimuli). Each target stimulus audio clip was mixed with two corresponding one-talker masker tracks, two-talker masker tracks, two-talker time-reversed masker tracks, and pink noise at the two SPLs (i.e., 64 and 68dB SPL). Each final stimulus was composed as follows: 500 ms of noise before the onset of the target sentence; the target sentence and noise together; and 500 ms of noise after the offset of the target sentence. To summarize, the final stimuli in each set (i.e., Set A and Set B) consisted of 256 stimuli mixed with: a.) one-talker babble (32 sentences X 2 SNRs); b.) two-talker babble (32 sentences X 2 SNRs);

c.) two-talker time-reversed (32 sentences X 2 SNRs); and, d.) pink noise (32 sentences X 2 SNRs).

6.1.2 Procedure

In a counter-balanced manner, participants were administered either stimulus Set A or Set B during baseline testing, and then another set during the second attempt. Each stimulus set was presented in eight experimental conditions: 4 (Type of noise: one-talker babble, two-talker babble, two-talker time-reversed, or pink noise) × 2 (SNR: -4dB, -8dB). In each listening condition, the participants listened to four unique target sentences, binaurally presented through Sennheiser HD-280 Pro headphones. These target sentences were randomly chosen by the computer from the full set of 32 target sentences, none of which were ever repeated across experimental conditions within participants. Hence, the target sentences in each condition were randomized across participants. The target sentences across the eight listening conditions were mixed, and presented to the participants in random order. As in previous studies (Lam et al., 2017; Xie et al., 2014), participants were asked to type out the target sentence after each stimulus presentation. If they were unable to understand the entire sentence, they were asked to report any intelligible words and make their best guess. If they were not able to make out any words, they were instructed to type "X." The task was self-paced, and participants had unlimited time to respond. Responses were scored as accurately-typed keywords (four per sentence). Keywords with added or omitted morphemes were scored as incorrect.

6.2 RESULTS

Regarding baseline performance, the overall percentage of correct responses, collapsing both noise and SNR, ranged from 25.0 to 70.3% for the native-speaker controls, and from 23.4 to 76.6% for the native-speaker motivation group.

6.2.1 Overall speech recognition performance in noise

This analysis employed a generalized, linear, mixed-effect model to analyze word recognition ability in noise, using the lme4 package 1.1-13 (Bates, Maechler, & Bolker, 2012) in RStudio 1.0.143. Keyword identification (correct or incorrect) was a dichotomous dependent variable. Motivation (motivation vs controls), Noise (1T, 2T, 2T-Reversed, and pink), SNR (-4dB VS -8dB), attempt (baseline performance vs the second attempt), and their interactions were included as fixed effects. By-participant intercepts were included in the model as random effects. The model was refined by removing, one at a time, the factors exhibiting the highest p-value, while still retaining the hierarchical rule of interactions. Likelihood ratio comparisons were performed to confirm that including each factor did not improve the amount of variance explained (Baayen, Davidson, & Bates, 2008).

Model comparison indicated a four-way Language X Motivation X Noise X SNR X Attempt interaction [χ^2 (3) = 21.1, p<.001]. To tease apart the four-way interaction, the following analysis used the lsmeans package in R, which was designed to obtain the least-squares means, and to test linear contrasts for linear and generalized mixed models (Lenth, 2016).

6.2.2 Between-group comparisons

In this section, we report between-group comparisons and focus on performance difference between the performance-pressure and control groups.

For 1T maskers, the performance-pressure group outperformed the native-speaker control group at both -4dB and -8dB levels during baseline performance [-4dB: β= .5, SE=0.2, p<.05; -8dB: β= .4, SE=0.2, p<.05]. In contrast, there were no group differences during the second attempt (both ps>.1) (Fig. 6.1).

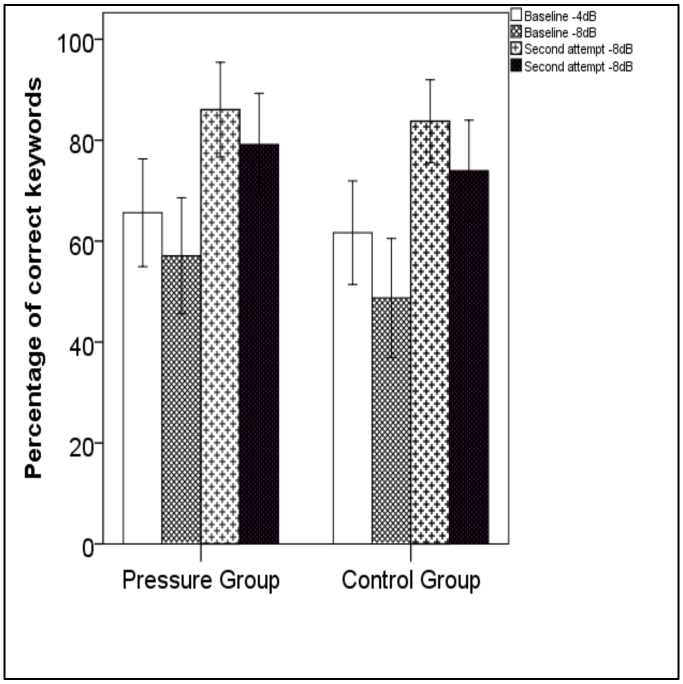

Figure 6.1: Word recognition performance in one-talker babble during baseline performance and the second attempt. Error bars indicate 95% confidence interval.

For 2T maskers, the performance-pressure group and the controls performed comparably during baseline performance (both ps>.3). During the second attempt, the performance-pressure group outperformed the control group at the easier -4dB level [β= .5, SE=0.2, p<.05] but underperformed the controls at the more challenging -8dB level [β= -.6, SE=0.2, p<.01] (Fig. 6.2).

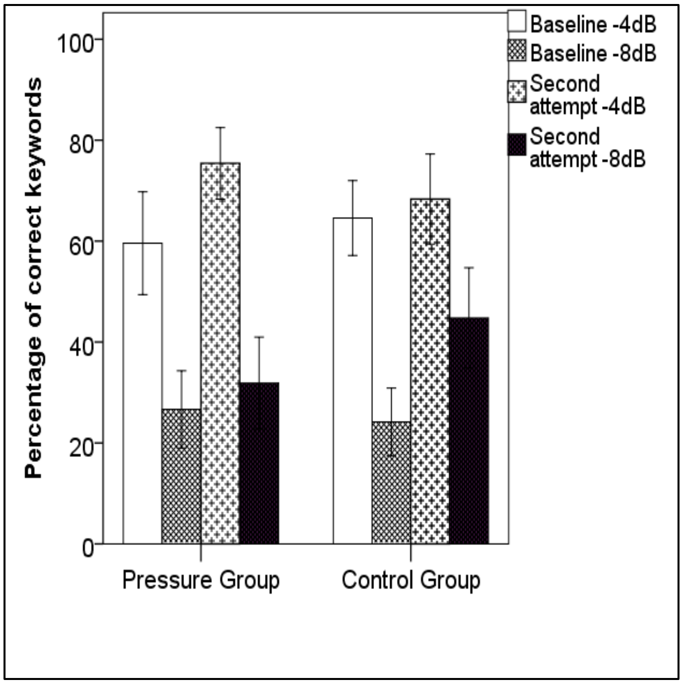

Figure 6.2: Word recognition performance in two-talker babble during baseline performance and the second attempt. Error bars indicate 95% confidence interval.

For 2T-reversed maskers, the performance-pressure group and the controls performed comparably during baseline performance (both ps>.1). During the second attempt, the performance-pressure group underperformed the controls at -4dB [β= -.6, SE=0.2, p<.05] (Fig. 6.3).

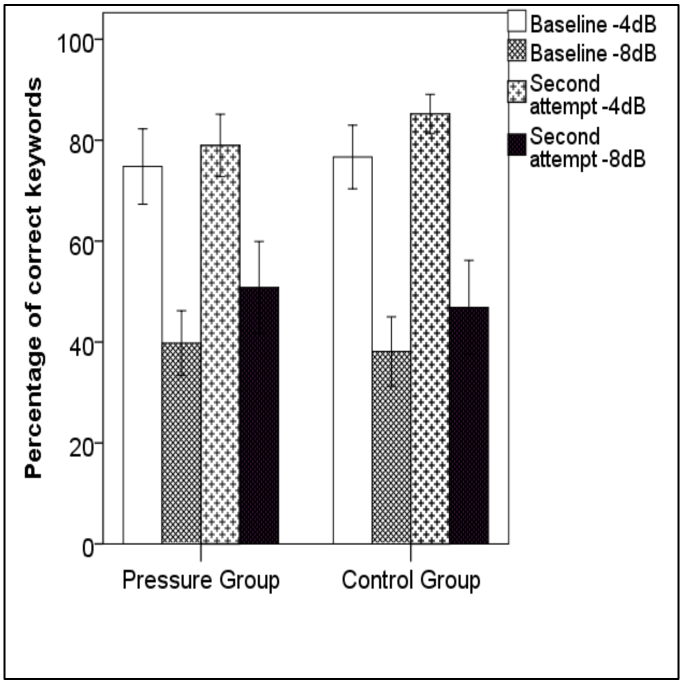

Figure 6.3: Word recognition performance in two-talker reversed masker during baseline performance and the second attempt. Error bars indicate 95% confidence interval.

For pink noise, no group differences were found between the performance-pressure group and the controls at either SNRs during baseline performance or the second attempt (all ps>.5) (Fig. 6.4).

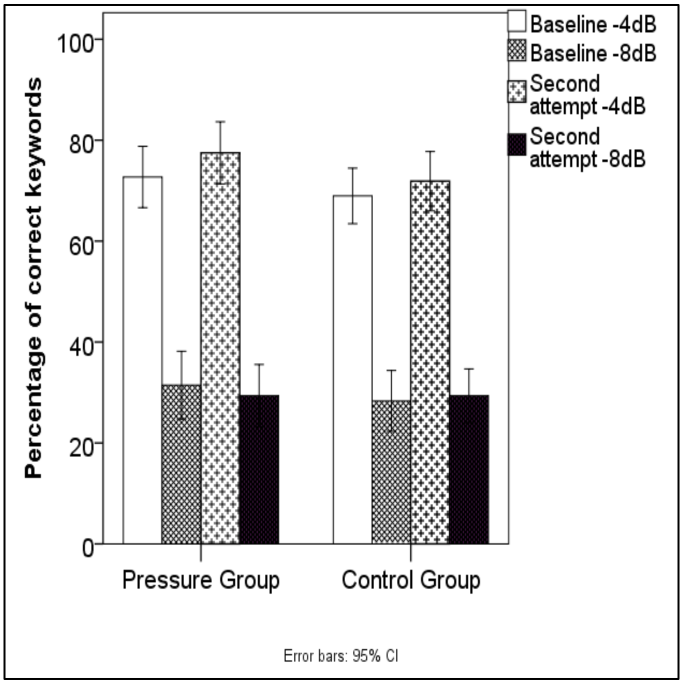

Figure 6.4: Word recognition performance in pink noise during baseline performance and the second attempt. Error bars indicate 95% confidence interval.

To summarize, during the second attempt, the performance-pressure group's superior one-talker babble performance during baseline performance no longer existed. While the performance-pressure group outperformed the control group in 2T masker at -4dB, the performance-pressure group underperformed the controls in the same masker at the more challenging -8dB level, and in two-talker time-reversed conditions at -4dB.

6.2.3 Within-group comparisons

Here we report within-group comparisons, with a focus on changes in performance between baseline performance and the second attempt. The following section reports each noise type systematically.

For 1T maskers, the controls improved at -4dB and -8dB [-4dB, β= 1.5, SE=0.2, p<.0001; -8dB: β= 1.3, SE=0.2, p<.0001]. Similarly, the performance-pressure group improved at both SNRs [-4dB, β= 1.1, SE=0.2, p<.0001; -8dB: β= 1.2, SE=0.2, p<.0001] (Fig. 6.2).

For 2T maskers, the controls improved controls improved at the more challenging -8dB level [β= 1.0, SE=0.2, p<.0001], but not at -4dB (p=.17). In contrast, the performance-pressure group improved at both -4dB and -8dB [-4dB, β= .9, SE=0.2, p<.0001; -8dB: β= .4, SE=0.2, p<.05] (Fig. 6.2).

For 2T-reversed maskers, the controls improved at both -4dB and -8dB [-4dB, β= .6, SE=0.2, p<.001; -8dB: β= .4, SE=0.1, p<.01]. In contrast, the performance-pressure group improved at the more challenging -8dB level [β= .6, SE=0.1, p<.0001], while their improvement at the easier -4dB level approached statistical significance (p=.053) (Fig. 6.3).

For pink noise, the controls did not improve at any SNRs (both ps >.5). The performance-pressure group improved at -4dB [β= .3, SE=0.2, p<.05] but not at the more challenging -8dB (p>.5) (Figure 6.4).

6.3 DISCUSSION

Speech processing seldom occurs in ideal listening conditions (Mattys et al., 2012). This study examined speech recognition in noises that introduce either significant energetic masking (EM) (i.e., two-talker reversed “2T-Reversed”, pink) or significant linguistic interference and information masking (IM) (i.e., one-talker babble “1T”, two-talker babble “2T”) at two signal-to-noise ratios (SNR: -4dB, -8dB). This study also manipulated motivation-induced performance pressure to examine its impact on spoken-word recognition ability in adverse listening conditions. There were two primary findings. First, improvement in word recognition ability was a function of noise types. Among all noises, listeners improved in 1T maskers consistently at both SNRs. Second, there was a highly-complex interaction between performance pressure and speech-processing ability in noise. Note that there were no performance decrements in high performance-pressure conditions. However, performance pressure influenced performance differences between the pressure group and controls, and that influence was dependent on noise types and SNRs. Depending on listening conditions, native listeners under performance pressure sometimes performed worse than the controls (e.g., two-talker babble at -8dB, two-talker reversed at -4dB).

Only for 1T masker, native listeners demonstrated significant improvement at *both* the -4dB and -8dB levels upon prior exposure. Multiple studies have shown listeners can apply top-down control to focus on cues that facilitate sound detection and target selection, such as talker characteristics (e.g., Culling, Hodder, & Toh., 2003), timing (e.g., Wright and Fitzgerald 2004; Varghese, Ozmeral, Best, & Shinn-Cunningham. 2012), and spatial location (e.g., Arbogast and Kidd 2000; Kidd, Mason, Brughera, & Hartmann, 2005). When given a second chance to process speech in 1T maskers, listeners might use various cues to improve object selection, especially the onset of the target speech (i.e., 0.5 sec after the masker speech). Regarding IM-predominant maskers, 1T is known to be less challenging than 2T masker (e.g., Freyman et al., 2001, 2004; Rosen et al., 2013), though both 1T and 2T maskers introduce significant lexical interferences to listeners. Comparing 1T to 2T masker, greater fluctuations in masker energy in the former masker might allow more “glimpsing” opportunities for acoustic information of the target speech, facilitating listeners’ usage of cues to perform target selection.

Regarding distinctions between EM and IM maskers, findings from behavioral studies have shown performance in significant IM maskers is not associated with speech intelligibility in significant EM maskers (Van Engen, 2012). Scott and colleagues’ (2004) neuroimaging study further shows the processing of EM and IM maskers are neurally dissociable. Additionally, decreases in SNR also exert distinct effect on speech-recognition ability in EM- and IM-predominant maskers. For instance, for primarily-EM maskers, the intelligibility of target speech declines very rapidly for SNRs below -3dB (e.g., Brungart, 2001). Taken together, one-talker maskers differ from other noise types used in this study (2T, 2T-reversed, pink) in terms of loci and amount of interference (i.e., auditory periphery vs linguistic interference), opportunities of "glimpsing," and associated levels of perceptual and cognitive demands. These noise-type differences might modulate the amount of prior-exposure benefits listeners receive at different SNRs.

A significant finding in this study is that the relationship between the "choking" effect and performance on noise is highly complex. Notably, our data show an important driving force for that relationship centers around group differences in 2T maskers and their variants. In contrast, performance pressure had a much simpler effect on performance in other maskers, with the performance-pressure and control groups performing comparably in 1T maskers and pink noise during the second attempt. Two-talker maskers not only introduce significant lexical interference, it is among the most challenging of forward speech maskers (Freyman et al., 2001; 2004; Rosen et al., 2013). Speech masker effectiveness is determined by the number of talkers involved (Rosen et al., 2013). When fewer talkers are present, IM is more predominant than EM (Boulenger, Hoen, Ferrange, Pellegrino, & Meunier, 2010), as there is more audible linguistic information one must ignore (Van Engen & Bradlow, 2007).

However, as mentioned earlier, studies have shown 2T is especially challenging to people's hearing, relative to 1T. Performance declines significantly from 1T to 2T (Freyman et al., 2001), where word recognition performance reaches its breakpoint (Rosen et al., 2013), and improves when more talkers are added to the 2T masker (Freyman et al., 2004). When the effectiveness of a masking condition is treated as the net effect of IM and EM (Agus, Akeroyd, Gatehouse, & Warden, 2009; Rosen et al., 2013), one's performance improves with the addition of talkers to 2T, because the gain of IM release outweighs the loss of reduced glimpsing (Rosen et al., 2013). In other words, 2T can be treated as the most effective of all forward speech maskers, given other variables (e.g., talker sex and identity, intensity, etc.) remain constant. On the one hand, relative to 1T, 2T allows fewer opportunities for “dip listening”; on the other hand, relative to pink noise, 2T introduces significant amounts of linguistic interference. When combined with performance pressure, 2T maskers may produce a robustly deteriorating effect to native listeners, when the challenge is significant at lower SNR (i.e. -8dB).

Note that the performance-pressure group outperformed the controls in 2T masker at the -4dB. In other words, performance pressure might not exert a unanimously-negative effect on speech recognition ability in noise. Interestingly, the performance-pressure group underperformed the controls in 2T-reversed masker at the same SNR. If the performance-pressure group’s poorer performance were caused by a combination of the deteriorating effect of performance pressure and masker distractibility, our findings of a reversed pattern in group differences in 2T and 2T-reversed conditions at -4dB would suggest the distractibility of speech masker may not simply depend on the presence of intelligible words. This interpretation would align with Hygge and colleagues (1992) that meaningless speech maskers via time-reversal were no less distracting than forward meaningful speech maskers, for both normal-hearing and hearing-impaired individuals, in audio-only conditions, despite the removal of all intelligible words. The presence of an unintelligible, but apparently speech-like masker in the background may invite listeners to make sense out of it, especially in high performance-pressure condition where listeners’ ability to selectively focus on the target speech may be compromised. Because 2T-reversed maskers sound more novel than 2T maskers, the former may become more distracting than the latter, in high performance-pressure conditions.

Our findings might shed light on current discussions about IM and EM in adverse listening conditions (Brungart, 2001; Mattys et al., 2012). Relative to EM, IM is poorly defined, and is primarily a “classification by exclusion” (Carlile & Corkhill, 2015). By showing a significant improvement at both -4 dB and -8dB for *only* 1T masker but not in EM-predominant masker (i.e., pink noise), this study provides further information about the influence of IM-EM distinction on the benefits of prior exposure to noise. Our findings also shows performance pressure exerted relatively little effect on performance in pink noise – one of the most widely-used maskers for representing EM (Howard-Jones & Rosen, 1993). In contrast, performance pressure exerted a potentially stronger influence on performance on maskers that introduce significant IM and elevate difficulties to selectively attend to target speech. The selective effect of performance pressure may result from participants’ weakened ability to suppress lexical interferences, due to the presence of distracting, worrying thoughts. When distraction theory posits a strong association between performance pressure and cognitive ability, the findings of a selective choking effect on 2T maskers and its variate resonate with emerging evidence that cognitive factors contribute to individual differences in speech recognition in competing-talker conditions (e.g., Chandrasekaran, et al., 2015; Xie et al., 2015; Zekveld et al., 2013).

Speech-in-speech-processing abilities differ substantially across listeners (e.g., Gilbert, Tamati, & Pisoni, 2013; Wilson, McArdle, & Smith, 2007; Xie et al., 2015; Chandrasekaran et al., 2015). Our findings highlight the importance of studying motivation-induced performance pressure in adverse listening conditions, which may provide greater mechanistic details about the role of cognition in hearing. In current listening-effort models, motivation plays a critical, yet insufficiently-defined role (Pichora-Fuller et al., 2016). Our study highlight the importance of studying performance pressure to clarify the role of motivation in challenging listening conditions.

CHAPTER 7: GENERAL DISCUSSION

Of all auditory signals, speech is the most complex for the human brain to process (Saberi & Perrott, 1999). The need for effective communication and the costs of miscommunication would *motivate* listeners to process speech signals quickly and accurately. Though motivation is crucial to understanding speech processing (Pichora-Fuller et al., 2016) and critical to all human behaviors (Ryan & Deci, 2000), current speech-processing models are mostly built upon data collected from conditions in which participants' poor performance had no negative consequences. There is a non-negligible gap between laboratory and countless real-life scenarios (e.g., high-stake assessments, cafés and restaurants, public transportation, warzones, etc.) – while motivation is seldom examined as a variable in the former, motivation is the key to success in the latter. To advance speech and hearing science, the critical gap between laboratory and reality needs to be filled. To this end, this dissertation manipulated levels of motivation in listeners to create performance pressure and examine its impact on three widely-studied speech-processing experiments, including verbal transformation (Experiment 1), the *Ganong* effect (Experiment 2), and speech recognition in noise (Experiment 3).

This dissertation examined a specific type of motivational effect – the "choking effect" (Baumeister, 1984; Beilock & Carr., 2001). The choking effect highlights a common, yet paradoxical effect of motivation – i.e., that one's desire for good performance creates performance pressure and causes underperformance (Baumeister, 1984; Beilock & Carr, 2005; Gimmig et al., 2006; Maddox et al., 2016; Maddox & Markman, 2010; Markman et al., 2006). Beilock and colleagues have shown the introduction of performance-contingent rewards and peer evaluation is highly effective in creating performance pressure and the choking effect (e.g., Beilock & Carr, 2005; Beilock &

DeCaro, 2007; Beilock et al., 2004). Using the same motivation manipulation adopted by Beilock and colleagues, this dissertation showed, for the first time, the paradoxical effect of motivation-induced performance pressure on multiple speech-processing experiments in native listeners.

Table 7.1 Summary of findings of three speech-processing experiments

Changes in performance upon prior exposure		
	Pressure Group	Control Group
Verbal transformation	❖ Reported greater number of verbal transformations ❖ Magnitude of increase was greater than that of the controls	❖ Reported greater number of verbal transformations
Ganong effect	❖ No change in performance	❖ Reduced lexical bias
Speech recognition in noise	❖ In one-talker masker, improved accuracy at both -4dB and -8dB ❖ In two-talker masker, improved accuracy at -4dB but not -8dB ❖ In two-talker time-reversed masker, improved accuracy at -8dB. Improvement at -4dB approached significant ❖ In pink noise, improved accuracy at -4dB but not -8dB	❖ In one-talker masker, improved accuracy at both -4dB and -8dB ❖ In two-talker masker, improved accuracy at -8dB but not -4dB ❖ In two-talker time-reversed masker, improved accuracy at -4dB and -8dB ❖ In pink noise, no improvement at any SNRs
Between-group comparisons during the second attempt		
Verbal transformation	❖ The performance-pressure group reported greater number of verbal transformations than the control group	
Ganong effect	❖ No group differences	
Speech recognition in noise	❖ In one-talker masker, no group differences during the second attempt. Note that the performance-pressure group outperformed the controls at both SNRs during baseline ❖ In two-talker masker, the performance-pressure group outperformed the controls in two-talker masker at -4dB, but underperformed the controls at the more challenging -8dB ❖ In two-talker time-reversed masker, the performance-pressure group underperformed the controls at -4dB ❖ In pink noise, no group differences	

Table 7.1 summarizes the findings of the three experiments. Specifically, Experiment 1 showed both performance-pressure and control groups became more prone to perceptual illusions, with the performance-pressure group showing a more elevated susceptibility. Experiment 2 showed phonetic processing in the performance-pressure group remained susceptible to lexical interference in high performance-pressure condition, while the controls benefited from prior exposure. Finally, experiment 3 showed the performance-pressure group recognized fewer words accurately than did controls in cognitively demanding listening conditions and at challenging noise levels (two-talker babble at -8dB, two-talker time-reversed speech at -4dB) when performance pressure was high. To summarize, this dissertation has demonstrated a choking effect in native listeners in all experiments.

7.1 THE "CHOKING UNDER PRESSURE" EFFECT ON SPEECH PROCESSING IS DIFFERENT FROM THE PRESSURE EFFECT ON READING

Young, healthy native listeners were expected to process speech signals in their native language effortlessly, so why did they choke under pressure? It should first be pointed out that the choking effect does not necessarily lead to a unanimous underperformance in between-group comparisons. For instance, regarding the *Ganong* effect, the performance-pressure and control groups did not differ in the amount of lexical bias during the second attempt. Instead, the choking effect on the performance-pressure group was presented as an absence of practice effect that was present in the controls. Regarding speech recognition in noise, the performance-pressure group outperformed the controls in two-talker masker at -4dB though the superior performance was reversed at the more challenging SNR.

Note that the choking effect demonstrated by this dissertation is at odds with previous research on reading (Rai et al., 2011, 2015). Per Rai and colleagues (2011, 2015), psychological pressure increases reading time, but has no effect on accuracy. In contrast, this dissertation on speech processing has shown performance pressure influenced both accuracy and consistency in responses besides response time. Note that this dissertation and Rai and colleagues (2011, 2015) differ not only in modalities (listening vs. reading), but also in how psychological pressure was manipulated. While Rai and colleagues targeted pressure in a general sense (e.g., tongue twisters, video-taping), this dissertation specifically targeted motivation-induced performance pressure by introducing performance-contingent rewards and peer evaluation. Besides differences in pressure manipulation, differences in the pressure-effect on speech processing and reading may also be rooted in the distinct nature of speech signals (fast, variable, fleeting).

Regarding reading, readers can modify their strategies and associated cognitive processes that might affect cognitive resource availability (Linderholm & van den Broek, 2002). For instance, Rai and colleagues (2015) showed reading performance was associated with speed-accuracy trade-off strategy; it is unclear to what extent trade-off strategy would facilitate speech processing. Because speech signals persist for only mere tenths of seconds (Holt & Lotto, 2010), the strategies available to reading might be less applicable to speech processing. Our data showed that, performance pressure might exert a more robust and deteriorating effect on spoken speech processing accuracy than previous research has shown it has on reading (Rai et al., 2011, 2015), which might be due to the temporal constraints imposed by speech signals, and the lack of opportunities of strategy usage,

Though this dissertation has shown motivation can create performance pressure and exerts a widespread choking effect on speech processing – as it does on other cognitive processes such as arithmetic operations and intelligence testing (Beilock & Carr, 2005; Gimmig et al., 2006), this dissertation acknowledges the significant challenges in examining the mechanistic details of the impact of motivation-induced performance pressure on speech processing. A significant challenge to examine the performance-pressure effect on speech processing is that the role of cognitive processes in speech processing is poorly understood (Arlinger et al., 2009). While reading research allows use of techniques such as the “think-aloud” procedure to examine cognitive processes via participant-provided conscious reports (Kucan & Beck, 1997; Magliano & Millis, 2003), the validity of using verbal reports to study speech processing has yet to be established. Even for studies adopting highly-effective paradigms for manipulating cognitive processes (e.g., dual tasks, induced anxiety), the role of cognitive processes, and their impacts on speech processing remain ambiguous (Mattys et al., 2013; Mattys & Wiget, 2011).

7.2 THE NEGATIVE IMPACT OF THE "CHOKING UNDER PRESSURE" EFFECT IN NATIVE LISTENERS

This dissertation showed that the choking effect on speech processing might sometimes differ from what predicted by distraction theory. Specifically, findings from the verbal transformation effect (Experiment 1) might align more closely to the predictions of explicit monitoring theory. When performance pressure was high, listeners’ might process the speech stream more carefully and analyze the activated lexical percepts more closely. Elevated attention to the speech stream may increase rates of lexical activation of the input stimulus (i.e., “tress”) and frequency of satiation. Furthermore, listeners might intentionally look for changes in the speech stream by adopting an analytic approach, which might disrupt the formation of perceptual coherence, which is associated with verbal transformation (Stachurski et al., 2015, 2017). Explicit monitoring theory and distraction theory differ primarily in their loci of interference, with distraction theory highlighting the additional consumption of limited cognitive resources. It is unclear why reduced cognitive resources might elevate perceptual illusions in motivation groups.

One might argue worrying thoughts (or inner voice) about poor performance might interfere with listeners' ability to attend to repetitive speech signals (i.e., "tress"), as listeners' occasional tendency to "wander" between worrying thoughts and the task might make them lose track of the speech stream. Due to the suggestive nature of the instructions used in the verbal transformation paradigm, listeners might report hearing changes more frequently, as a compensatory strategy. This is unlikely. Warren (1961a) systematically examined the effect of listening conditions (e.g., clear vs faint speech) and found faint or masked speech reduced, instead of increased, verbal transformations. Though degraded speech signals increases the challenge of keeping track of speech signals, listeners do not report more illusions, despite the suggestive nature of the instructions. Additionally, while aging is associated with poorer speech processing, due to sensory and/or cognitive decline (e.g., Dubno, Dirks, & Morgan, 1984; Helfer & Freyman, 2014; Humes, 1996), older adults report fewer verbal transformations (Warren, 1961b). Taken together, elevated auditory illusions found in high performance-pressure conditions cannot be readily explained by cognitive interferences, nor by difficulties in keeping track of the speech stream. Future studies could examine verbal transformation in dual-task conditions (e.g., Mattys & Wiget, 2011), and compare their findings with our data, to examine the role of divided attention in verbal transformation.

Our data showed the choking effect on the *Ganong* effect and speech recognition in noise may align more closely with the predictions of distraction theory, which targets cognitive processes. Emerging evidence suggests performance on the *Ganong* effect and speech recognition in noise is associated with individual differences in cognitive ability. For the *Ganong* effect, Mattys and colleagues showed susceptibility to lexical influences is mediated by factors related to attention and cognitive control (Mattys et al., 2013; Wiget

& Scharenborg, 2014; Mattys & Wiget, 2011). Our data further show native listeners failed to improve their ability to suppress irrelevant information, despite prior exposure.

Regarding speech recognition in noise, emerging evidence shows energetic masking (EM) and information masking (IM) exert distinct effects on perceptual and cognitive systems (e.g., Zekveld et al, 2013), with performance in IM maskers being more strongly associated with the ability to suppress irrelevant information (Chandrasekaran, et al., 2015; Lam et al, 2017; Xie et al., 2015; Zekveld et al., 2013). Our data show performance pressure had little effect on speech-processing performance for pink noise masker, which is among the most widely-used maskers for representing EM (Howard-Jones & Rosen, 1993). In contrast, the complex interaction between MiP and noise was primarily driven by performance differences between the performance-pressure and control groups in two-talker masker and its variant, which introduce significant IM (e.g., Van Engen & Bradlow, 2007) and challenges to listeners (Freyman et al., 2001, 2004; Rosen et al., 2013).

Taken together, our data show cognition can infiltrate and influence the manifestation of multiple widely-studied speech-processing experiments. This dissertation's data highlight that cognitive factors, such as motivation, exert wide-spread influence on speech processing, even for tasks predominantly viewed as a perceptual phenomenon (e.g., verbal transformation).

7.3 LANGUAGE-GROUP DIFFERENCES IN PERCEIVED IMPORTANCE OF GOOD PERFORMANCE

An unexpected finding of this study is language status as a confounding factor, which might moderate the performance-pressure effect. Note that our motivation manipulation unambiguously elevated levels of performance pressure in both native and non-native listeners, who became more conservative than the controls in reporting the amount of improvement they thought they had made during their second attempt. Nevertheless, our motivation manipulation had a negligible effect on manipulating the perceived importance of good performance. For both native and non-native listeners, the performance-pressure and control groups did not differ in the importance they attached to performing at a higher level during the second attempt. In contrast, listeners' language status was more powerful than motivation manipulation in influencing perceived importance of good performance, with native listeners feeling it was more important to perform at a high level during their second attempt, compared to non-native listeners. Critically, this perception also applied to native-listener controls, who were not prompted to improve their performance through monetary incentives, nor pressured to work harder for a partner.

Why did the native listeners and non-native listeners differ in perceived importance of high-level performance during the second attempt? Gimmig and colleagues (2006) showed tasks relevant to participants' personal interests (e.g., assessments of academic success) might elevate participants' desire to perform well and, perceived importance of good performance. Because all speech-processing tasks in this dissertation were presented in non-native listeners' weaker second-language, non-native and native listeners might differ in perceived likelihood of success and failure. While non-native listeners might be more prepared to fail, native listeners might have greater expectations of success. Another possibility is that the size of the monetary reward might not have been equally motivating to native and non-native learners.

The two possibilities are not mutually exclusive; however, participants' responses to the final feedback questionnaire shed more light on the second. By asking participants the minimum amount of bonus money needed to motivate them to improve their performance during the second attempt, this dissertation could estimate the perceived value of the monetary incentives actually provided. It was assumed the value of the monetary incentives given (i.e., $10) was inversely related to the amount of money suggested by the listeners at the end of the experiment. Instead of simply asking participants to rate the values of the given monetary incentives, money was adopted as a more direct index for its being a concrete, universal, and highly quantifiable concept, shared among listeners of different cultural backgrounds.

The non-native performance-pressure group wanted more money than the native-listener performance-pressure group did, on all speech-processing tasks. One might argue the amount of money wanted might simply index task-associated effort – i.e., due to differences in language proficiency, non-native listeners found the tasks more difficult than native speakers, and thus wanted more money in compensation. This interpretation is not entirely correct. Note that the native-listener and non-native-listener controls did not differ in the amount of money wanted, though the non-native-listener controls should expend greater task-associated effort than the native-listener controls due to poorer language proficiency. In other words, the amount of money wanted cannot be merely an index of task-associated effort, but may also index differences in the perceived values of the motivators (e.g., money) provided.

Our motivation manipulation "pushed" the two performance-pressure groups in opposite directions, while the two control groups stayed at the mid-range in terms of the amount of money wanted. As task difficulty interacts with motivation (Broadhurst, 1959; Weiner, 1979), over-challenging tasks and lower self-perceived competence might reduce task commitment (e.g., Arnold, 1976; Hughes, Sullivan, & Lou Mosley, 1985). Heightened task difficulty and lower self-perceived competence might have reduced non-native listeners' task commitment to improve their performance during the second attempt, relative to native listeners. Furthermore, increased levels of risk might have joined with lower self-perceived competence to prompt non-native listeners to refrain from committing to a high-stake situation (i.e., poor performance causing their partner to lose bonus money). As a result, our motivation manipulation failed to elevate perceived importance in the L2-motivation group.

Taken together, the native and non-native listeners might differ in both the value they place on external motivators (i.e., monetary rewards, peer pressure) and their internal need for good performance. On the one hand, native listeners had greater internal need to improve their performance. As a result, the native-listener control and performance pressure groups reported comparable level of perceived importance of good performance, even though native-listener controls had no prospect of monetary reward nor fears of negative consequences. When native listener were placed in high-stakes situations where poorer performance would inflict negative consequences on them and their partners, the performance-pressure group experience additional drive for good performance, creating a "choking under pressure" effect. On the other hand, non-native listeners experienced lower self-relevance in their weaker second-language, and thus perceived it as less important to perform well. Additionally, they might have found the incentive insufficient to make them commit to high-stakes situations. As a result, the non-native performance pressure group did not experience elevated perceived importance of good performance, even though failures would result in negative consequences for a partner who is unknown to them.

7.4 LIMITATIONS

This study has two limitations. Like other studies on performance pressure, this study inferred changes in performance-pressure levels via self-report (e.g., Beilock & Carr, 2005). Though this study found motivation manipulation elevated levels of perceived pressure, constrained perceived improvement, and a consistently negative impact of performance pressure on performance in all experiments in native listeners, this dissertation did not administer any physiological measures of pressure (e.g., cortisol). Changes in physiological measures may provide a sensitive measure in examining the association between changes in pressure level and changes in performance. Additionally, this dissertation did not administer measures of cognitive effort (e.g., pupillometry), which might provide additional methods to examine whether motivation manipulation modulates effort. With an objective measure of effort, researchers might examine whether performance declines despite (or as a result of) more conscious and effortful processing.

The second limitation is that this study did not aim to tease apart the interaction between global and local motivation on speech processing. With a primary aim to create a "choking under pressure" effect, this study has adopted a design that is equivalent to Beilock and colleagues' (2004, 2005, 2007), which introduces a repertoire of stressors. Future studies should begin to tease apart global and local motivation to examine its impact on speech processing in greater mechanistic details.

7.5 THEORETICAL IMPLICATIONS

In challenging listening conditions (e.g., repetitive, ambiguous, noisy), motivation is an important source of individual differences in speech-processing performance (Pichora-Fuller et al., 2016). Motivation determines the amount of mental effort a listener is willing to invest, and, in its turn, increases the likelihood of success. This is the upside of motivation. Following the elegant U-shaped parabolic discovered by Yerkes and Dodson (1908), and advances in studies on motivational pressure (Baumeister, 1984; Beilock and Carr, 2001, 2005; Beilock & DeCaro, 2007; Gimmig et al., 2006; Kimble & Perimuter, 1970; Lewis & Linder, 1997, Maddox & Markman, 2010; Markman et al., 2006; Masters, 1992), this dissertation has shown motivation can also cause listeners to underperform.

Cognitive energy and effort are currently priority areas in hearing science research, with a focus on understanding the interaction between motivation and effort (for a review, see Pichora-Fuller et al., 2016). Data from this dissertation highlight the downside of motivation. On the one hand, greater motivation might predict greater willingness to expend effort. On the other, increases in effort might not necessarily predict success in speech processing, which is the goal of the human auditory system. A key area for future research is to determine the *optimal* level of cognitive effort required for effective ("good enough") speech processing, which would be relevant to perceived well-being and life quality.

Regarding studies on motivational pressure, distraction theory have provided useful frameworks for understanding the detrimental effect of motivation on activities that have a well-defined and identifiable cognitive component (e.g., reading, mathematics). Data from this dissertation highlight the distinct natures of speech signals, which imposes significant challenge for distraction theory to fully capture the choking effect in speech processing.

7.6 EDUCATIONAL IMPLICATIONS

Current studies on motivation and linguistic processing have focused on second-language learners and applications of motivational principles in second-language classrooms. Data from this dissertation have shown that native speakers and non-native speakers might differ in perceived importance of good performance, depending on the situation (e.g., preventing a partner from losing money). Language-group differences might moderate the effect of motivation. However, information about the role of motivation in linguistic processing in native language is, ironically, extremely limited. The gap in current knowledge is partly driven by the predominant view that processing in non-native language would be more effortful and less automatic than in native language (e.g., Meschuan & Hernandez, 2006; Segalowitz & Hulstijn, 2005), rendering processing in non-native language more vulnerable to language breakdowns. This resonates with the wide-spread language anxiety found in second-language classrooms, a key area of research in second-language education (Bekleyen, 2009; Elkhafaifi, 2005; Horwitz, 2001; 2010; Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1994; Zhang, 2013). Language anxiety highlights the importance of studying perceived competence and fears of evaluation of one's communicative performance under uncertain conditions.

However, language anxiety is not restricted to second-language learners. Rai and colleagues (2015) observed individual differences in language anxiety also exist in native speakers, and predict self-rated L1reading proficiency. Our data show native speakers and non-native speakers learners do not differ in proficiency alone, but also in their perceptions of the importance of good performance, which may moderate the effect of motivational pressure on speech and language processing. Future studies on motivation and speech and language processing should expand their scope to include native speakers.

7.7 CLINICAL IMPLICATIONS

For vulnerable populations who have difficulties reading or retrieving words, hearing in noise, or speaking in public, intervention implies intensive exposures to activities that are cognitively effortful, emotionally draining, and perhaps occasionally painful. Motivation is thus clinically significant. However, the influence of motivation on speech and language processing and production remains poorly understood in rehabilitative science. This dissertation raises a critical question for clinical practitioners – i.e., what kinds of motivators are beneficial to clients, and why? Answering this question is crucial to a wide variety of clinical populations, as motivation relates to treatment adherence, which is a central challenge in healthcare (Vermeire et al., 2001). Our data show perceived importance and competence may moderate the influence of motivation on perceived importance of good performance, which might influence performance outcomes. Thus, perceived competence may be a crucial psychological dimension in rehabilitation, as it is associated with task commitment in typical populations (e.g., Arnold, 1976; Hughes et al., 1985). A key area of research is to examine the relationship between severity of deficits and perceived competence and how this relationship mediates the influence of motivation, in greater mechanistic detail. However, the complex relationship between perceived competence and motivation may manifest differently in various clinical populations.

CONCLUSION

The influence of motivation is largely anecdotal in speech and language science. This study represents one of the most systematic investigations of the extent to which motivation impacts multiple widely-studied speech-processing phenomena in adult native and non-native speakers of English. Notably, this study has shown motivation not only might not improve listening performance, it may, ironically, impair speech processing in native listeners. By illustrating the complex interaction amongst motivation, task demands, and individual differences among listeners, this study highlights that motivation is a critical dimension for inquiry in the emerging science of cognitive hearing, and speech and language science at large.

Appendix A

1. **On a scale of 1 to 9, how important was it for you to perform at a high level in the second round?** (1 = Not at all important to me; 9 = Extremely important to me)
   a. Sound identification task: ___________________
   b. Task with noises: ___________________

2. **On a scale of 1 to 9, how much pressure did you feel to improve your performance in the second round?** (1 = Very little pressure; 9 = Extreme pressure)
   a. Sound identification task: ___________________
   b. Task with noises: ___________________

3. **Answering in percentage, how much do you think your performance changed in the second round, relative to the first round**?
   a. Sound identification task: ______________%
   b. Task with noises: ______________%

4. **For each task, what is the minimum amount of bonus money that would motivate you to improve your performance in the second round? Your answer should be within $1 to $10.**
   a. Sound identification task: ___________________
   b. Task with noises: ___________________

References

Agus, T. R., Akeroyd, M. A., Gatehouse, S., & Warden, D. (2009). Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise. *The Journal of the Acoustical Society of America*, *126*(4), 1926-1940.

Arbogast, T. L., & Kidd Jr, G. (2000). Evidence for spatial tuning in informational masking using the probe-signal method. *The Journal of the Acoustical Society of America*, *108*(4), 1803-1810.

Arbogast, T. L., Mason, C. R., & Kidd Jr, G. (2002). The effect of spatial separation on informational and energetic masking of speech. *The Journal of the Acoustical Society of America*, *112*(5), 2086-2098.

Arlinger, S., Lunner, T., Lyxell, B., & Kathleen Pichora- Fuller, M. (2009). The emergence of cognitive hearing science. *Scandinavian journal of psychology*, *50*(5), 371-384.

Arnold, H. J. (1976). Effects of performance feedback and extrinsic reward upon high intrinsic motivation. *Organizational Behavior and Human Performance*, *17*(2), 275-288.

Audacity Developer Team. (2008). Audacity (Version 1.2.6) [Computer software]. *Available: audacity.sourceforge.net/download.*

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*(4), 390-412.

Baekeland, F., & Lundwall, L. (1975). Dropping out of treatment: a critical review. *Psychological bulletin*, *82*(5), 738.

Bamford, J., & Wilson, I. (1979). Methodological considerations and practical aspects of the BKB sentence lists. In J. Bench & J. Bamford (Eds.), *Speech-hearing tests and the spoken language of hearing-impaired children* (pp. 148–187). London: Academic Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68(3)*, 255-278.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339-2349.

Bashford Jr, J. A., Warren, R. M., & Lenz, P. W. (2006). Polling the effective neighborhoods of spoken words with the verbal transformation effect. *The Journal of the Acoustical Society of America*, *119*(4), EL55-EL59.

Bashford Jr, J. A., Warren, R. M., & Lenz, P. W. (2008). Evoking biphone neighborhoods with verbal transformations: Illusory changes demonstrate both lexical competition and inhibition. *The Journal of the Acoustical Society of America*, *123*(3), EL32-EL38.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: linear mixed-effects models using S4 classes.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, *46*(3), 610.

Bekleyen, N. (2009). Helping teachers become better English students: Causes, effects, and coping strategies for foreign language listening anxiety. *System,* 37(4), 664-675.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure?. *Journal of experimental psychology: General*, *130*(4), 701.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and “choking under pressure” in math. *Psychological science*, *16*(2), 101-105.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 983.

Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General*, *133*(4), 584.

Billig, A. J., Davis, M. H., Deeks, J. M., Monstrey, J., & Carlyon, R. P. (2013). Lexical influences on auditory streaming. *Current Biology*, *23*(16), 1585-1589.

Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature neuroscience*, *12*(1), 92-98.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.1. 25) [Computer program]. Retrieved January 20, 2010.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual Review of Psychology, 66*.

Boulenger, V., Hoen, M., Ferragne, E., Pellegrino, F., & Meunier, F. (2010). Real-time lexical competitions during speech-in-speech comprehension. *Speech Communication*, *52*(3), 246-253.

Broadhurst, P. L. (1959). The interaction of task difficulty and motivation: The Yerkes-Dodson law revived. *Acta Psychologica*, *16*, 321-338.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica*, *86*(1), 117-128.

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.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101-1109.

Bullen, J. G., Hemsley, D. R., & Dixon, N. F. (1987). Inhibition, unusual perceptual experiences and psychoticism. *Personality and individual differences*, *8*(5), 687-691.

Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 567-575.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science*, *315*(5820), 1860-1862.

Catts, S. V., Armstrong, M. S., Norcross, K., & McConaghy, N. (1980). Auditory hallucinations and the verbal transformation effect. *Psychological medicine*, *10*(1), 139-144.

Chandrasekaran, B., Van Engen, K., Xie, Z., Beevers, C. G., & Maddox, W. T. (2015). Influence of depressive symptoms on speech perception in adverse listening conditions. *Cognition and Emotion*, *29*(5), 900-909.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, *25*(5), 975-979.

Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*.

Connine, C. M., Clifton Jr, C., & Cutler, A. (1987). Effects of lexical stress on phonetic categorization. *Phonetica*, *44*(3), 133-146.

Connine, C. M., & LoCasto, P. A. (2000). Inhibition. *Behavioral and Brain Sciences, 23(03),* 328.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, *8*(2), 331-335.

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America, 123*(1), 414-427.

Culling, J. F., Hodder, K. I., & Toh, C. Y. (2003). Effects of reverberation on perceptual segregation of competing voices. *The Journal of the Acoustical Society of America*, *114*(5), 2871-2876.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review, 3(4), 422-433

DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S. L. (2011). Choking under pressure: multiple routes to skill failure. *Journal of Experimental Psychology: General*, *140*(3), 390.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation.

Ditzinger, T., Tuller, B., & Kelso, J. S. (1997). Temporal patterning in an auditory illusion: the verbal transformation effect. *Biological cybernetics, 77*(1), 23-30.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The modern language journal*, *78*(3), 273-284.

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America*, *76*(1), 87-96.

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *Modern Language Journal*, 206-220.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *The Journal of Educational Research*, *84*(5), 253-262.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages?. *Applied linguistics*, *8*(2), 162-177.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *109*(5), 2112-2122.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *115*(5), 2246-2256.

Gimmig, D., Huguet, P., Caverni, J. P., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence. *Psychonomic Bulletin & Review*, *13*(6), 1005-1010.

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test. *Journal of the American Academy of Audiology, 24*(1), 26-36.

Gow, D. W., Segawa, J. A., Ahlfors, S. P., & Lin, F. H. (2008). Lexical influences on speech perception: A Granger causality analysis of MEG and EEG source estimates. *Neuroimage, 43*(3), 614-623.

Haddock, G., Slade, P. D., & Bentall, R. P. (1995). Auditory hallucinations and the verbal transformation effect: the role of suggestions. *Personality and Individual Differences, 19*(3), 301-306.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and aging*, *7*(1), 56.

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging, 6*(4), 587 - 594.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of experimental psychology: Learning, memory, and cognition*, *17*(1), 163.

Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, *136*(2), 748-759.

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, & Psychophysics, 72*(5), 1218-1227.

Horwitz, E.K. (2001). Language anxiety and achievement. *Annual review of applied linguistics*, *21*, 112-126.

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, *43*(2), 154-167.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern language journal*, *70*(2), 125-132.

Howard- Jones, P. A., & Rosen, S. (1993). Uncomodulated glimpsing in ‘‘checkerboard’’ noise. *The Journal of the Acoustical Society of America*, *93*(5), 2915-2922.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80-93.

Hughes, B., Sullivan, H. J., & Lou Mosley, M. (1985). External evaluation, task difficulty, and continuing motivation. *The Journal of Educational Research*, *78*(4), 210-215.

Humes, L. E. (1996). Speech understanding in the elderly. *Journal-American Academy of Audiology*, *7*, 161-167.

Hygge, S., Rönnberg, J., Larsby, B., & Arlinger, S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *Journal of Speech, Language, and Hearing Research*, *35*(1), 208-215.

Kidd, G., Mason, C. R., Brughera, A., & Hartmann, W. M. (2005). The role of reverberation in release from masking due to spatial separation of sources for speech identification. *Acta acustica united with acustica*, *91*(3), 526-536.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological review*, *77*(5), 361.

Kucan, L., & Beck, I. L. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of educational research*, *67*(3), 271-299.

Kondo, H. M., & Kashino, M. (2007). Neural mechanisms of auditory awareness underlying verbal transformations. *Neuroimage*, *36*(1), 123-130.

Lam, B. P., Xie, Z., Tessmer, R., & Chandrasekaran, B. (2017). The Downside of Greater Lexical Influences: Selectively Poorer Speech Perception in Noise. *Journal of Speech, Language, and Hearing Research*, 1-12.

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371-375.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325-343.

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J Stat Softw*, *69*(1), 1-33.

Lewis, B. P., & Linder, D. E. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin*, *23*(9), 937-944.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. Psychonomic bulletin & review, 21(4), 861-883.

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of educational psychology*, *94*(4), 778-784.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1-36.

MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language learning*, *41*(4), 513-534.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language learning*, *44*(2), 283-305. Markman, A. B., Maddox, W. T., & Worthy, D. A. (2006). Choking and excelling under pressure. *Psychological Science*, *17*(11), 944-948.

Maddox, W. T., Filoteo, J. V., Glass, B. D., & Markman, A. B. (2010). Regulatory match effects on a modified Wisconsin Card Sort Task. *Journal of the International Neuropsychological Society*, *16*(2), 352-359.

Maddox, W. T., Koslov, S., Yi, H. G., & Chandrasekaran, B. (2016). Performance pressure enhances speech learning. *Applied psycholinguistics*, *37*(6), 1369-1396.

Maddox, W., & Markman, A. B. (2010). The motivation–cognition interface in learning and decision making. *Current Directions in Psychological Science*, *19*(2), 106-110.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *21*(3), 251-283.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical representations in children with SLI: Evidence from a frequency-manipulated gating task. *Journal of Speech, Language, and Hearing Research*, *51*(2), 381-393.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940-967.

Markman, A. B., Maddox, W. T., & Worthy, D. A. (2006). Choking and excelling under pressure. *Psychological Science*, *17*(11), 944-948.

Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word recognition. *Cognition, 25*, 71–102.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

Mattys, S. L., & Liss, J. M. (2008). On building models of spoken-word recognition: When there is as much to learn from natural "oddities" as artificial normality. *Attention, Perception, & Psychophysics*, *70*(7), 1235-1242.

Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: a comparative analysis of perceptual, lexical, and attentional factors. *Psychology and aging*, *29*(1), 150.

Mattys, S. L., Seymour, F., Attwood, A. S., & Munafò, M. R. (2013). Effects of acute anxiety induction on speech perception: are anxious listeners distracted listeners?. *Psychological science*, *24*(8), 1606-1608.

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*(2), 145-160.

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception?. *Trends in Cognitive Sciences*, *10*(8), 363-369.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86.

McCoy, S. K., Hutchinson, S., Hawthorne, L., Cosley, B. J., & Ell, S. W. (2014). Is pressure stressful? The impact of pressure on the stress response and category learning. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 769-781.

McQueen, J. M. (1991). The influence of the lexicon on phonetic categorization: stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 433.

Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, *485*(7397), 233-236.

Miller, G. A. (1947). The masking of speech. *Psychological bulletin*, *44*, 105-129.

Miller, G. A., & Licklider, J. C. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*(2), 167-173.

Miller, G. A., & Taylor, W. G. (1948). The perception of repeated bursts of noise. *The Journal of the Acoustical Society of America*, *20*(2), 171-182.

Miller, W. R. (1985). Motivation for treatment: A review with special emphasis on alcoholism. *Psychological bulletin*, *98*(1), 84.

Mirman, D., McClelland, J. L., Holt, L. L., & Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science*, *32*(2), 398-417.

Munson, B., & Brinkman, K. N. (2004). The influence of multiple presentations on judgments of children's phonetic accuracy. *American Journal of Speech-Language Pathology*, *13*(4), 341-354.

Natsoulas, T. (1965). A study of the verbal-transformation effect. *The American journal of psychology*, *78*(2), 257-263.

Natsoulas, T. (1967). What are perceptual reports about?. *Psychological bulletin*, *67*(4), 249.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299-325.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... & Naylor, G. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.

Pitt, M. A., & Samuel, A. G. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1120.

Pitt, M. A., & Shoaf, L. (2002). Linking verbal transformations to their causes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 150.

Rai, M. K., Kalia, R. K., Singh, R., Gangola, M. P., & Dhawan, A. K. (2011). Developing stress tolerant plants through in vitro selection—an overview of the recent progress. *Environmental and Experimental Botany*, *71*(1), 89-98.

Rai, M. K., Loschky, L. C., & Harris, R. J. (2015). The effects of stress on reading: A comparison of first-language versus intermediate second-language reading comprehension. *Journal of Educational Psychology*, *107*(2), 348.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, *7*.

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*, *12*(49), 263.

Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013). Listening to speech in a background of other talkers: Effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America*, *133*(4), 2431-2443.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, *398*(6730), 760-760.

Samuel, A. G. (2011). Speech Perception. *Annual Review of Psychology*, *62*, 49-72.

Sato, M., Baciu, M., Lœvenbruck, H., Schwartz, J. L., Cathiard, M. A., Segebarth, C., & Abry, C. (2004). Multistable representation of speech forms: a functional MRI study of verbal transformations. *NeuroImage*, *23*(3), 1143-1151.

Sato, M., Schwartz, J. L., Abry, C., Cathiard, M. A., & Loevenbruck, H. (2006). Multistable syllables as enacted percepts: a source of an asymmetric bias in the verbal transformation effect. *Attention, Perception, & Psychophysics*, *68*(3), 458-474.

Schwartz, R. G., Scheffler, F. L., & Lopez, K. (2013). Speech perception and lexical effects in specific language impairment. *Clinical linguistics & phonetics*, *27*(5), 339-354.

Scott, S. K., Rosen, S., Wickham, L., & Wise, R. J. (2004). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *The Journal of the Acoustical Society of America*, *115*(2), 813-821.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. *Handbook of bilingualism: Psycholinguistic approaches*, 371-388.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in cognitive sciences*, *12*(5), 182-186.

Shinn-Cunningham, B., Best, V., & Lee, A. K. (2017). Auditory Object Formation and Selection. In *The Auditory System at the Cocktail Party* (pp. 7-40). Springer International Publishing.

Slade, P. D. (1976). An investigation of psychological factors involved in the predisposition to auditory hallucinations. *Psychological Medicine*, *6*(1), 123-132.

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. *International journal of behavioral medicine*, *15*(3), 194-200.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Manual for the state-trait anxiety inventory.

Stachurski, M., Summers, R. J., & Roberts, B. (2015). The verbal transformation effect and the perceptual organization of speech: influence of formant transitions and F0-contour continuity. *Hearing research*, *323*, 22-31.

Stachurski, M., Summers, R. J., & Roberts, B. (2017). Stream segregation of concurrent speech and the verbal transformation effect: Influence of fundamental frequency and lateralization cues. *Hearing Research*, *354*, 16-27.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of America*, *26*(2), 212-215.

Taylor, M. M., & Henning, G. B. (1963). Verbal transformations and an effect of instructional bias on perception. *Canadian Journal of Experimental Psychology*, *17*, 210.

Van Engen, K. J. (2010). Similarity and familiarity: Second language sentence recognition in first-and second-language multi-talker babble. *Speech Communication*, *52*(11), 943-953.

Van Engen, K. J. (2012). Speech-in-speech recognition: A training study. *Language and Cognitive Processes*, *27*(7-8), 1089-1107.

Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native-and foreign-language multi-talker background noise a. *The Journal of the Acoustical Society of America*, *121*(1), 519-526.

Varghese, L. A., Ozmeral, E. J., Best, V., & Shinn-Cunningham, B. G. (2012). How visual cues for when to listen aid selective auditory attention. *Journal of the Association for Research in Otolaryngology*, *13*(3), 359-368.

Warren, R. M. (1961a). Illusory changes of distinct speech upon repetition—the verbal transformation effect. *British Journal of Psychology*, *52*(3), 249-258.

Warren, R. M. (1961b). Illusory changes in repeated words: Differences between young adults and the aged. *The American journal of psychology*, *74*(4), 506-516.

Warren, R. M., & Warren, R. P. (1966). A comparison of speech perception in childhood, maturity, and old age by means of the verbal transformation effect. *Journal of Verbal Learning and Verbal Behavior*, *5*(2), 142-146.

Warren, R. M. (1968). Verbal transformation effect and auditory perceptual mechanisms. *Psychological bulletin*, *70*(4), 261.

Warren, R. M. (1996). Auditory illusions and perceptual processing of speech. *Principles of experimental phonetics*, 435-466.

Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *The American journal of psychology*.

Warren, R. M., & Meyers, M. D. (1987). Effects of listening to repeated syllables: Category boundary shifts versus verbal transformations. *Journal of Phonetics*. Volume 78, Issue S1

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of educational psychology*, *71*(1), 3.

Wilson, R. H., McArdle, R. A., & Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*, *50*(4), 844-856.

Worthy, D. A., Brez, C. C., Markman, A. B., & Maddox, W. T. (2011). Motivational influences on cognitive performance in children: Focus over fit. *Journal of Cognition and Development*, *12*(1), 103-119.

Worthy, D. A., Markman, A. B., & Maddox, W. T. (2009). What is pressure? Evidence for social pressure as a type of regulatory focus. *Psychonomic bulletin & review*, *16*(2), 344-349.

Wright, B. A., & Fitzgerald, M. B. (2004). The time course of attention in a simple auditory detection task. *Attention, Perception, & Psychophysics*, *66*(3), 508-516.

Xie, Z., Maddox, W. T., Knopik, V. S., McGeary, J. E., & Chandrasekaran, B. (2015). Dopamine receptor D4 (DRD4) gene modulates the influence of informational masking on speech recognition. *Neuropsychologia*, *67*, 121-131.

Xie, Z., Yi, H. G., & Chandrasekaran, B. (2014). Nonnative audiovisual speech perception in noise: Dissociable effects of the speaker and listener. *PloS one*, *9*(12), e114439.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit- formation. *Journal of comparative neurology*, *18*(5), 459-482.

Yu, R. (2015). Choking under pressure: the neuropsychological mechanisms of incentive-induced performance decrements. *Frontiers in behavioral neuroscience, 9*.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, *134*(3), 2225-2234.

Zhang, X. (2013). Foreign language listening anxiety and listening performance: Conceptualizations and causal relationships. *System*, *41*(1), 164-177.