LISTENING EFFORT ALLOCATION, STIMULUS-DRIVEN, GOAL-DRIVEN, OR BOTH?

by

Min Zhang

Bachelor of Clinical Medicine, Tongji Medical School of Huazhong University of Science and Technology, 2002

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCE

This dissertation was presented

by

Min Zhang

It was defended on

July 6, 2017

and approved by

Michael Dickey, PhD, Associate Professor, Department of Communication Science and Disorders

Malcolm Mcneil, PhD, Distinguished Professor, Department of Communication Science and Disorders

Sheila Pratt, PhD, Professor, Department of Communication Science and Disorders

Charles Perfetti, PhD, Distinguished Professor, Department of Psychology

Greg J. Siegle, PhD, Associate Professor, Departments of Psychiatry

Dissertation Advisor: Catherine Palmer, PhD, Associate Professor, Department of Communication Science and Disorders

LISTENING EFFORT ALLOCATION, STIMULUS-DRIVEN, GOAL-DRIVEN, OR BOTH?

Min Zhang, PhD

University of Pittsburgh, 2017

The research in audiology to date about how people listen has been focused narrowly on the impact of the task demand (e.g., speech complexity) on the effort exerted for listening. Very few studies have examined how intention-associated factors affect listening effort regulation, and little is known about how to characterize the individual quality of effort expenditure in terms of efficiency. This study tested a compensatory control model for effort regulation to investigate how reward would modulate the effect of task demand on listening effort. The secondary aim was to propose a modified computational approach for effort efficiency calculation.

The nonclinical sample was comprised of 40 college volunteer participants with normal hearing. All participants completed the Need for Cognition scale, a speech comprehension task which required cost-benefit decision making, and a self-report strategy use survey. Pupil dilation was measured throughout the speech comprehension task as an indicator of listening effort. Results supported the model in which effort regulation during an intended activity is determined not only by stimulus-driven factors such as task demand, but also by goal-driven factors such as reward. Significant interaction effects emerged. Furthermore, the effort efficiency derived by using goal-oriented performance variables demonstrated superiority in distinguishing individuals compared to the use of a simple performance accuracy equation.

This study contributes to the limited literature on proactive listening effort regulation. Examining further how hearing, cognition, and personality interact neurophysiologically and functionally in individuals with normal hearing and hearing loss can help clinicians and

researchers better understand the underlying mechanism of listening effort control, and facilitate implementing strategies to aid effective listening through audiologic interventions.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation and gratitude to my advisor, Dr.Catherine Palmer. Her advice and mentorship profoundly shaped my academic personality. She has made my time as a Ph.D. student extraordinarily enjoyable and intellectually stimulating. I am deeply grateful for this priceless gift. I want to thank all my committee members Dr.Catherine Palmer, Dr. Sheila Pratt, Dr. Michael Walsh Dickey, Dr. Malcolm McNeil, Dr. Charles Perfetti, and Dr. Greg Siegle for all their selfless help, support and confidence in me. I am fortunate that these distinguished scholars were willing to work with me on this exciting interdisciplinary endeavor. I benefited tremendously from their substantive expertise and from their guidance as I started my professional career. Especially, I want to thank Dr. Sheila Pratt for providing me the opportunity of working in her fabulous lab all these years which has helped me establish good manners of conducting serious research. I thank Dr. Michael Dickey for generously lending the eyetracking equipment to me so that my dissertation was able to advance without being held back by any technology issue. I thank Dr. Malcolm McNeil for always encouraging me and making me think critically. I thank Dr. Charles Perfetti for being available for me whenever I need his advice on my dissertation project. I thank Dr. Greg Siegle for opening my eyes to new knowledge and his patient guidance on my data analysis.

I also want to thank the many other faculty members (especially Dr. Elaine Mormer) and my fellow doctoral students (especially Reem Muller) at Pitt who have enriched me with countless hours of constructive discussions, with their honest critical feedback, and by serving as role models and examples. I hope I will become a worthy member of this wonderful academic community.

I dedicate this dissertation to my family, especially my mother Weilan Wu and my sister Li Zhang, whose continued support has allowed me to pursue this goal and encouraged me in difficult times. They came all the way from China to Pittsburgh to rescue on many occasions since I had my daughter. I also want to dedicate this dissertation to my husband Jiangpeng Xu, who has trained me to become stronger and more efficient over the years. I would like to give special thanks to my brilliant daughter, Hailing, for continually providing the requisite breaks from philosophy and the motivation to finish my degree with expediency. I also would like to thank my personal friends both from China and the U.S. for their continued mental support and encouragement.

I wouldn’t have gone this far without such amazing teams behind me, so, the greatest gratitude and appreciation to all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Listening is not just a very useful, but also often a very enjoyable activity. A normal auditory system is able to distinguish thousands of sounds effortlessly (Schnupp, Nelken, & King, 2011), and people tend to take this sense for granted until they lose it and realize that listening is actually a skill which requires effort. Distinct from *hearing* as a passive function that provides access to the auditory world via the perception of sound, *listening* is the process of hearing with intention and attention which requires the expenditure of cognitive resources (Kiessling et al., 2003). From the perspective of cognitive hearing science, listening effort is considered to be the result of imbalance between the implicit (bottom-up) and explicit (top-down) processes (Rönnberg, 2003). Consequently, a better understanding of the interaction between cognitive function and hearing related to speech understanding and communication may allow us to improve our ability to repair problems if the system is impaired, for example, in people with hearing loss whose effort must be particularly concentrated (Sweetow, 2005).

The interaction between bottom-up and top-down processes can be manifest across levels from periphery to cortex, across modalities from auditory-only to multimodality (Rönnberg, Rudner, & Lunner, 2011), and across stages from stimulus preprocessing through response (Sanders, 1983). This review begins by summarizing the definitions of listening effort. The second component of the literature review focuses on the theoretical foundation of listening effort, including neurophysiological mechanisms, underlying philosophies of listening effort driven by stimuli and goals, and the impact of individual differences. The third component of the literature review is a critical analysis of research on listening effort measurements. Research questions and hypotheses are proposed at the end based upon the literature review.

1.1 DEFINITION OF LISTENING EFFORT

Effort is a popular and abstract concept in the fields of cognition as well as in information processing. In the dictionary, effort is defined as exertion of physical or mental power (The Random House College Dictionary, 1980, p.287) and conscious exertion of power (Merriam-Webster's Collegiate Dictionary, 1997, p.368). However, it is often used interchangeably with the terms arousal, attention, capacity, resource, and energy in the literature. Early definitions of effort mainly centered on the attentional aspect (Hicks & Tharpe, 2002; Kahneman, 1973; Pribram & McGuinness, 1975), however, in recent definitions, scientists have added the intention and other cognitive resources (i.e., memory) (Gosselin & Gagné, 2011b; McCoy et al., 2005; Picou, Ricketts, & Hornsby, 2011; Tyler, Hertel, McCallum, & Ellis, 1979). Table 1 lists the definitions from the literature.

Table 1. Effort Related Definitions

Reference	**Term**	**Definition**
Hicks and Tharpe (2002)	Listening effort	The attentional requirements necessary to understand speech
Gosselin and Gagné (2011a)	Listening effort	The attentional and cognitive resources required to understand speech
Picou et al. (2011)	Listening effort	The cognitive resources allocated for speech recognition
Kahneman (1973)	Effort	An intensive form of attention and is involved whenever a human engages in performing mental tasks. (Implied)
Pribram and McGuinness (1975)	Effort	The measure of attention "paid" to increase or maintain the efficiency of a communication channel by reducing its equivocation (enhancing competency).
Yates and Kulick (1977)	Effort	The available resources that are actually applied to the task. The available resources refer to external resources such as tools and energy source, and internal resources such as mental capacity and metabolic energy sources.
Sanders (1983)	Effort	Effort is a co-ordinating process, adjusting the balance of input and output operations, and mediating high level feedback from response outcomes.
Humphreys and Revelle (1984)	Effort	The motivational state commonly understood to mean trying hard or being involved in a task.
Tyler et al. (1979)	Cognitive effort	The engaged proportion of limited-capacity central processing
Piolat, Barbier, and Roussey (2008)	Cognitive effort	A measure of the fraction of attentional resources allocated to a process at a given moment.
Tuovinen, Gerven, Paas, and Tabbers (2003)	Mental effort	The aspect of cognitive load that refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task.
Pichora-Fuller, et.al.(2016)	Mental effort	The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task.

The term resource is widely used in the definition of effort and listening effort. However, the definition of resource itself is vague with reference to almost any processing capability, energetical as well as structural. Norman and Bobrow (1975) defined resources as “such things as processing effort, the various forms of memory capacity and communication channels. Resources are always limited” (p.45). Rabbitt (1979) defined that resources are “acquired information about the structure of particular tasks and about the external world which are used by the subject in order to actively control their momentary perceptual selectivity and their choice of responses” (p. 129). When the cognitive capacity is defined as the mental resources available for storage and processing of information (Rudner, Lunner, Behrens, Thoren, & Rönnberg, 2012), the three terms can hardly be specified because they explain each other.

There is a tendency in the literature to distinguish effort from the other constructs such as arousal, attention, mental load, and working memory (WM). Kahneman (1973) proposed that effort is a special case of arousal, and the difference between effort and other varieties of arousal, such as produced by drugs or by loud noises, depends on whether voluntary processing is involved. Arousal is defined in terms of phasic physiological or psychological responses to input, and is said to occur when an input change produces a measurable incrementing of a physiological or behavioral indicator over a baseline (Pribram & McGuinness, 1975). Although arousal and effort co-occur under many circumstances, which have made it possible to measure indices of effort, the dissociation between arousal and effort has been evidenced by independent manipulations of stimulus set and response set (Broadbent, 1971). Kahneman (1973) also claimed that the voluntary attention is an exertion of effort in activities which are selected by current plans and intentions, and the purposeful aspect of attention corresponds to effort. Later definitions of effort continued to use this concept (Gosselin & Gagné, 2011b; Hicks & Tharpe, 2002; Piolat et al., 2008). Attention and effort are naturally related. We rarely see effort expenditure without attention involved, and there is no evidence indicating that effort can vary independently from attention, although attention might vary independently from effort under certain conditions. Mental load and mental effort are considered as two distinct aspects of cognitive load (Tuovinen et al., 2003). Mental load represents the expected cognitive capacity demands of a particular task by an individual, whereas the mental effort reflects the actual cognitive capacity that the individual allocates to the task. The mental load and effort are impacted by the current knowledge about the task as well as subject characteristics (e.g., age, skill level), hence, they usually are not equal to each other.

Pribram and McGuinness (1975) specified that effort accompanies only those attentional processes that result in a change in the representational organization of the information processing mechanism. Whereas, some researchers suggest that working memory ability, believed to be closely related but separated from attention, plays an important role when effort occurs. In the specific example of speech perception, when the signal is degraded by background noise or hearing loss, the cognitive (explicit) working memory processes help individuals fill in the missing information and compensate for speech perception difficulties (Rönnberg, Rudner, Foo, & Lunner, 2008). The explicit cognitive processes include verbal inference-making (Lyxell & Rönnberg, 1989), word decoding (Lyxell & Rönnberg, 1991), and phonological and lexical access (Rönnberg et al., 1998). These processes are required to retrospectively resolve ambiguities of previous speech elements during a dialogue and also to construct expectations of prospective exchanges in the dialogue. Working memory is often defined as the mental workspace where important information is kept in a highly active state, available for a variety of other cognitive processes (Baddeley & Hitch, 1974). It focuses on dual storage and processing and includes the processes that encode, store, and manipulate information (Baddeley, 2000; Rönnberg, Danielsson, et al., 2011). Whenever a task demands extra storage and processing, the task becomes difficult and is associated with more effort (Zekveld, Kramer, & Festen, 2010).

As can be seen, effort has more overlaps with other cognitive constructs rather than less. The one unique characteristic of effort might be the ability of aggregating all kinds of cognitive resources and processes, and integrating those intensive components into a strategy in order to achieve a goal. Viewing effort from this aerial perspective is helpful because a consensus on the precise definition of the related constructs as well as the differentiation from each other is not required. However, in order to use the term effort appropriately, the context in which effort is used should demonstrate the unique comprehensive aspect of the concept, and a valid measure of effort should be available to quantify it.

Effort could manifest in physiological ways, such as pupil dilation (Globerson, 1983; Kahneman, 1973; Zekveld et al., 2010), increased skin conductance (Mackersie & Cones, 2011; Mead & Lapidus, 1989), heart rate or fast pulse (Mackersie & Cones, 2011), cortisol concentration (Hicks & Tharpe, 2002), evoked potential response (Bernarding, Corona-Strauss, Latzel, & Strauss, 2010; Strauss et al., 2009a) and blood flow in the brain (Croxson, Walton, O'Reilly, Behrens, & Rushworth, 2009). Effort also could manifest in terms of cognitive process, for instance, attention shifting and switching (Fraser, Gagné, Alepins, & Dubois, 2010; Howard, Munro, & Plack, 2010; Piolat et al., 2008; Sarampalis, Kalluri, Edwards, & Hafter, 2009); explicit processing capacity, which may involve executive processes such as shifting, updating, and inhibition to contribute to inference-making (i.e., inferring missing information) (Miyake et al., 2000; Rönnberg et al., 2008); working memory (Rönnberg, Danielsson, et al., 2011; Rudner, Rönnberg, & Lunner, 2011); and response selection (Hockey, 1997; Sanders, 1983).

In addition to the physiological and mental manifestations, the physical effort expended by people with hearing impairment should be included in the scope of the listening effort definition as well. The individual's attempts to compensate and to communicate optimally require a constant effort to hear, to pay attention and to respond appropriately (Demorest & Erdman, 1986). Those who employ beneficial repair strategies, such as using amplification, setting up the acoustic environment as favorably as possible, determining the general content of the conversation, shortening the distance from the speaker, asking the speaker to rephrase or slow down rather than repeat, seeking for visual cues, dominating conversations instead of taking turns with other speakers, pursuing and adjusting their assistive listening devices and so forth, are expending extra effort on listening (Edwards, 2007). It should be noted that the research on listening effort originally started based on vague emotional complaints about effort by people with hearing impairment in everyday life and work, unfortunately, the emotional component and physical component of listening effort are absent in the above definitions, which are usually attached to the concept in a real life environment.

Furthermore, the momentary effort that a task demands should be distinguished from the total amount of effort that is required to complete that task. The momentary effort at a specific instant will be quite different if one can complete a task at his/her comfortable pace compared to the situation when one completes a task within a strict time limit. Therefore, time-pressure has been suggested to be an important determinant of effort (Kahneman, 1973).

The intention of listening in the definition of listening effort is essential in guiding the research direction in this field. Picou et al. (2011) focused on speech recognition, whereas

Gosselin et al. (2011b) and Hicks et al. (2002) emphasized speech comprehension. The ultimate goals of the effort expended on listening could be to comprehend the speech, to participate in a conversation, and/or to enjoy an interesting talk. Speech perception and recognition are necessary but not sufficient to achieve these goals, however, speech comprehension is. In addition, speech comprehension is considered a broader term than speech perception and includes both implicit (bottom-up) and explicit (top-down) processes (Rönnberg, 2003). Therefore, defining listening effort with intension of speech comprehension is more appropriate.

Effort is difficult to define because its mobilizations, motivations and effects vary with each individual. Although difficult, it has been accepted that the effort that a subject deploys at any one time corresponds to what he is doing, rather than to what is happening to him (Kahneman, 1973). Hence, listening effort should be considered as what a subject is doing while listening. In this context, listening effort should be described in terms of the individual’s physiological, mental, behavioral and emotional compensation driven by specific intentions (e.g., speech comprehension).

Based on the above review, a suggestive definition of listening effort could be, any form of energy or resource (e.g., cognitive, physical, emotional) and process that is intentionally allocated (or distributed) in listening for the purpose of speech comprehension and/or pleasure. It is restricted by but independent of an individual's capacity. What effort consists of in a specific listening task is determined by the environment in which the listening activity takes place and how much compensatory strategy use in that environment is allowed.

In summary, despite interest in the topic for the past few decades, there is still no clearly defined, universally accepted definition of listening effort. What we know so far is that listening effort is a mental concept closely related to but also distinct from other cognitive constructs. It has characteristics of activeness, voluntariness, control or top-down processing motivated by intentions (goal-directed), and characteristics of variability as the individual's intention changes. Similar to other mental constructs, listening effort cannot be directly observed; it must be inferred from observation of overt behavior and measurement of psychological and physiological processes. Listening effort can be assessed in different domains, such as time (momentary effort versus total effort), manifestation (physiological, mental and physical), quantity and quality (e.g., efficiency).

1.2 THEORETICAL FRAMEWORK

Compelling support for the essential role that the limbic system, especially the anterior cingulate cortex (ACC) area, plays in effort mobilization comes from neurophysiology and neuroimaging studies on animals and humans. The majority of research on listening effort is based on work of Kahneman (1973) and Broadbent (1971) because they present the earliest attempts to integrate energetical mechanisms, with components such as arousal, activation and effort (Freeman, 1948), into an information processing model including the main stages of encoding and feature extraction, motor adjustment and response preparation, and computational control for central decisions (Sternberg, 1969). Both models implied that the energetic resources may be allocated and controlled. More recent models have postulated the mechanisms that trigger effortful processing or top-down processing, and started to emphasize the influence of active control and individual differences on effort exertion.

This section is divided into three main parts. The first concerns neurophysiological basis of effort, followed by theories and models that relate to effort control and speech understanding.

The third and final part focuses on the association between individual differences and listening effort.

1.2.1 Neurophysiological basis of listening effort

The cognitive neurodynamics of listening effort are not fully understood yet, and research related to the neural correlates is still in its infancy. Pribram and McGuinness (1975) sought to define the major neuroanatomical structures of effort, which are hypothesized to be one of three separate but interacting neural systems in the control of attention. The effort mechanism in their model acted as a coordinating and organizing principle to coordinate the activity of arousal (amygdala circuit) and activation (basal ganglia circuit), and was involved in problem-solving and decision-making.

Effort was said to be a function of the limbic system in a circuit involving the hippocampus, cingulate cortex, septal nuclei, posterior hypothalamus and the anterior thalamic nucleus, centering on the hippocampus (Pribram & McGuinness, 1975). The hippocampus is an area of the primitive cortex, hidden within the medial temporal lobe. The causal relation between the hippocampal circuit and effort was investigated by illustrating the abnormal habituation of hippocampal lesioned animal subjects in discrimination reversal tasks (Douglas, Barrett, Pribram, & Cerny, 1969), and by precise electrical stimulations of selected parts of the hippocampus circuit with micro- and macro-electrodes. Bland and Vandenwolf (1972) observed the hippocampal electrical activity measured by theta rhythm, an oscillatory pattern (i.e., 4-8Hz) in the electroencephalo-graph (EEG) signals recorded either from inside the brain or from electrodes glued to the scalp, from freely moving rats. They found that the theta activity occurred almost exclusively when the rats were making voluntary movements defined as flexible and modifiable responses. Based on the observed long duration characteristics of hippocampal neurons, which indicate a long period of summation preceding hippocampal activity, Pribram and McGuinness (1975) proposed that the hippocampus constitutes part of an error (mismatch) evaluating mechanism which was conceived to process the perturbations resulting from the mismatch among inputs.

More recent neurophysiological and neuroimaging investigations assign particular primacy to the anterior cingulate cortex (ACC) area, which is adjacent to the corpus callosum on the medial surface of the frontal lobe and has both connections to prefrontal and subcortical limbic structures (He, Dum, & Strick, 1995). The working hypothesis is that the ACC encodes information about effort (Rushworth, Walton, Kennerley, & Bannerman, 2004; Walton, Bannerman, Alterescu, & Rushworth, 2003; Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006).

The anterior cingulate cortex (ACC) has been suggested as the neurobiological substrate for decisions engaging effort exertion and is believed to play a role in cognitive control (Fellows & Farah, 2005; Løvstad et al., 2012; Posner & DiGirolamo, 1998). One influential theory postulates that the ACC is involved in response conflict monitoring and serves as a regulator signaling to other executive regions such as the dorsolateral prefrontal cortex (DLPFC) whether executive attention has to be reinforced or alleviated and whether mental effort is demanded (Botvinick, 2008; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004). This was reflected in a number of neuroimaging studies. Carter et al. (1998) observed an increase in blood flow to the ACC when response competition arises as a result of the elicited prepotent but inappropriate response tendency (e.g., the Continuous Performance Test). Buckner et al. (1995) found that when underdetermining response was needed (i.e., override tasks in which conflict or interference takes place among equally permissible lexical representations in working memory), the blood flow to the ACC increased. Braver, Barch, Gray, Molfese, and Snyder (2001) demonstrated increased blood flow in ACC during speeded tasks in which conflict was present between the executed incorrect response and related correct response. In contrast, under circumstances of less challenging auditory discriminations involving little decision-making and simple movement, there was a reduction of blood flow in the anterior cingulate cortex (Cohen et al., 1988). Therefore, the conflict-monitoring theory of ACC function focuses on the initiating mechanism to facilitate the deployment of cognitive control, with an increase in the strength of top-down control following experienced response conflict.

It appears that the hippocampus and the ACC share similar function of conflict-monitoring. However, the approach to address the function of each brain area differs dramatically. The hippocampal desynchronization occurs only when a mismatch between an input and the neuronal model is present. A neuronal model refers to some competence or a patterned memory trace developed in the brain as a representation of the experienced stimulus configuration (Pribram & McGuinness, 1975). This hippocampal desynchronization facilitates the new reticular formation (registration) and activates the arousal function of the amygdala circuit. Discrimination tasks and reasoning tasks examining the contingent negative variation (CNV) and/or recording hippocampal theta rhythm often are used in experiments to test this hypothesis. The main issue with the hippocampal theta rhythm is that observing a response is much more difficult in humans and other primates than in mammals like rats, dogs or cats. The issue with CNV relates to its nonspecificity and contamination by spontaneous activities. In contrast, the Stroop-like tasks are often used in studying conflict-monitoring function of the ACC in combination with functional imaging technique. The characteristics of the conflicts in ACC studies emphasize the complex and response end of the information processing structure, whereas the conflicts investigated in hippocampus studies focused more on the input side of the processing. It is possible that both hippocampus and anterior cingulate cortex constitute the conflict monitoring system but act at different stages.

There are several other theories about relationships between the ACC and effortful behavior. Brown and Braver (2005) demonstrated by a computational model that the ACC might not detect conflict or errors per se, but rather adaptively predict the error-likelihood with continued exposure to task environments. They differentiated the error-likelihood detection from the conflict detection by manipulating response conflict and error probability separately in a change-signal task in which participants had to make a left-right reversed button press response from that indicated by the go-signal after seeing a change-signal. The variable change-signal delay (CSD) represented different levels of error-likelihood with shorter CSD corresponding to a 4% low-error likelihood condition and longer CSD corresponding to a 50% high-error likelihood condition. These two conditions were not explained to the participants prior to the experiment, but they were presented by two colors. The participants demonstrated the ability to learn the meaning of the color cues over time as a result of negative reinforcement in this case, and to adaptively modify their behaviors through this reinforcement learning (RL). A critical finding was that the ACC activity as measured by error-related negativity (ERN/Ne) and fMRI was significantly greater in high-error trials than in low-error trials even with absence of response conflict. The ACC activity occurring in proportion to the error-likelihood therefore determined the amount of effort or cognitive control recruitment.

Although the dominant theories have implied a correlation between the ACC and effortful behavior mediated by a conflict-resolution mechanism, a cognitive-control mechanism, or a reinforcement learning mechanism, as those processes all require effort, they are still vague on effort motivation itself. A growing body of lesion and neurophysiological studies has yielded inconsistent or contradictory results for both the conflict (Nachev, 2011) and reinforcement learning (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006) hypotheses, because ACC lesions in human typically result in global slowing and increased response variability rather than inflexibility of control or lack of the ability to learn from feedback.

The most recent theory proposed by Holroyd and Yeung (2012) reconciles diverse theories of the function of the ACC and suggests a role for ACC not in decision-making and deployment of cognitive control, but rather in motivating effortful behavior. In their hierarchical reinforcement learning (HRL) theory, behavior is organized hierarchically; interrelated states and actions are grouped together to form higher-level behavior options. Each option comprises structured sequences of actions directed toward a specified subgoal along with the initiation associations to the subgoal, and the ultimate goal is the combination of several subgoals. The simple reinforcement learning (better-than-expected outcomes are reinforced) occurs simultaneously within and between levels in the structure so that a goal can be achieved parsimoniously. Mapping the cognitive concepts to the brain, the ACC and basal ganglia operate in parallel but at different levels of the hierarchical organization. The ACC selects, maintains and learns about high-level options, whereas the basal ganglia are concerned with low-level actions. The ACC integrates the individual action values into a general task value to determine whether to sustain optimal performance. The authors suggest that the ACC is responsible for optimal option selection by associating and comparing predictive values (an estimate of the long-term reward by ventral striatum) with different options, and for maintenance (performance monitoring) for the chosen option. Meanwhile, the ACC determines the level of effort to be applied toward executing the policy based on the predictive value. High value corresponds to vigorous top-down control exerted by the dorsolateral prefrontal cortex (DLPFC) directed from the ACC. The DLPFC and motor structures in the dorsal striatum execute those options by signaling basal ganglia.

This hierarchical reinforcement learning theory has incorporated key elements of previous theories and simplified the learning process. In addition, the core ACC functions of option selection, option maintenance, and cost-effort integration allow the theory to accommodate the neurophysiological and neuropsychological evidence that ACC lesions result in effort aversion and even loss of the awareness of effort.

Several investigations have suggested that the ACC is necessary for optimally allocating effort expenditure. Laboratory animals with permanent or transient AAC lesions do not maximally exploit high-effort-high-return options and fail to allocate their effort towards strategies that maximize return (Amiez, Joseph, & Procyk, 2006; Kennerley et al., 2006; Schweimer, Saft, & Hauber, 2005; Walton et al., 2003). Rats were tested after 6-hydroxydopamine infusions into the ACC (Schweimer et al., 2005) or infusing quinolinic acid into the ACC (Walton et al., 2003) in cost-benefit tasks. In the tasks, they could choose from climbing a barrier to obtain a high reward or obtaining a low reward without the barrier. Results demonstrated that the lesioned rats exhibited a reduced preference for the high-cost-high-reward response option when given the choice of obtaining a low reward with little effort, whereas the control rats with intact ACC preferred to select the high-cost-high-reward option. This indicated that the ACC is important when evaluating how much effort to expend for a specific reward.

Kennerley (2006) demonstrated that ACC’s critical role in reinforcement-guided behavior is neither in detecting nor in correcting errors, but in guiding voluntary choices based on the history of action and outcomes. The author trained three of nine rhesus monkeys with selective ACC lesions (removal by surgery) on tasks assessing error- and reward-guided action selection. Monkeys were free to choose between two actions. In one experiment, the response outcomes were categorized as either correct or incorrect, and the monkeys had to use the positive reinforcement information to sustain correct performance because only the correct responses received reward. In the other experiment, the two actions were rewarded according to unequally assigned probabilities, which meant that monkeys had to use another strategy rather than "win-stay, lose-switch" to obtain optimal return. The results showed that the ACC lesioned monkeys did not perform significantly worse than control monkeys on the trials that followed errors, and both groups took approximately three trials to notice that the reinforcement was being increasingly delayed in the first experiment. However, the lesioned monkeys were less likely to repeat a response that had been rewarded and were generally much slower than the control group to approach the optimum ratio threshold of action choices in the second experiment. Thus, the pattern of impairment suggested that the ACC area is vital for sustaining rewarded action selection.

In another experiment, Hillman and Bilkey (2012) argued that heightened ACC activity did not solely result from the level or volume of effort demand of conditions, instead it encodes information as to which course of action provides the best effort-outcome ratio. They recorded AAC neurons in freely moving rats as they performed a competitive two-choice decision-making task. The rats chose whether to physically compete with a peer for food reward under a safe and limited physical competition scenario created by using wire mesh and two peer rats. In this study, the food reward configuration and the dominance of peer rats were manipulated to represent the degree of effort required. They found that the ACC neurons responded to competitive effort costs, increasing their firing rates to indicate the course of action that appeared to provide the best net utility (potential gain minus incurred costs). In addition, the ACC neurons were found sensitive to subtle changes in competitor's strength, it was suggested that the neurons registered the dominant or highly-food motivated competitor as high-effort, low-return options. More interestingly, the heightened ACC activity was not consistently observed for all competitive conditions that required effort. When reward size and dominance of competitor were separately manipulated in isolated experiments, Hillman and Bilkey found a consistent pattern of higher firing rates for the choice that appeared to be the optimal effort-outcome ratio.

Correspondingly, imaging studies in humans have shown that the greater the effort-based cost-benefit net value, the higher the resulting ACC activity (Croxson et al., 2009; Forstmann, Brass, Koch, & von Cramon, 2006; Mars et al., 2005; Walton, Devlin, & Rushworth, 2004; Yoshida & Ishii, 2006). Croxson et al. (2009) scanned young normal participants' brains while they performed a sequence of effortful tasks with four effort levels to obtain secondary reinforcers associated with either high or low money reward. The location of the intersection of a vertical line and horizontal line in a circle visual field served as a cue indicating the levels of reward and of effort to be expected at the beginning of each trial, that is, the net cost-benefit value of the upcoming task. The authors demonstrated several regions, including insula/posterior orbitofrontal cortex (OFT), ventral striatum, dopaminergic midbrain and ACC, which responded at different points during multistep response processing toward the reward. The fMRI Blood Oxygen Level Dependent (BOLD) signal changes of insula/posterior OFT were found to relate only to reward expectation and not to effort expectation, whereas the BOLD signal changes of putamen were only determined by effort expectation alone. Although the cue-locked activities in ventral striatum, midbrain and ACC were all significantly modulated by the net value of the course of action, only ACC activity increased as participants worked through the effort period toward reward, indicating an important role of the ACC in continuously computing the net value and maintaining the optimal goal. In a following behavioral confirmation experiment, participants consistently chose the visual cue corresponding to the higher net value option, which agreed with the fMRI results. The data from Croxson's (2009) study and Kennerley's (2006) substantially supported the hypothetical goal selection and maintenance function of the ACC in the hierarchical reinforcement learning (HRL) theory (Holroyd & Yeung, 2012).

Alterations in ACC structure or function have been associated with clinical conditions of effort misappropriation in humans and cognitive misperception of effort as well. Devinsky (1995) observed that patients with elevated anterior cingulate cortex activity often display tics and obsessive-compulsive behaviors, sometimes even psychopathic or sociopathic behaviors. Conversely, reduced anterior cingulate cortex activity following surgery or infarcts can contribute to behavioral disorders such as diminished self-awareness, depression, impaired motor initiation, reduced responses to pain and aberrant social behavior. Naccache et al. (2005) performed a case study using a Stroop paradigm on a patient with a left mesio-frontal cortex lesion including the left ACC and a group of comparison subjects. The authors initially expected to see control impairments likely to be found in such patient, unexpectedly however, the control abilities of the patient evaluated in various versions of the Stroop tasks were amazingly preserved across response modalities. They accidentally discovered that the patient was totally unable to experience and report a normally associated feeling of mental effort during those tasks. The deficit of sensation of effort could not be simply explained by episodic memory impairment or a failure to understand the task because the patient could perform much better than chance-level when asked to recall the congruity of the stimulus presented on the preceding trials, and she was able to identify the difficulty level of a trial. Naccache et al. thus hypothesized that this deficit could be related to an inappropriate generation of somatic signals (e.g., heart rate, pupil dilation, skin conductance) from mesio-frontal structures, including the ACC. Clinical neuropsychology and brain imaging studies converge to attribute a crucial role to the frontal cortical systems in the generation of those somatic signals, and in the coupling of bodily reactions to a cognitive evaluation (Critchley et al., 2003).

In another Stroop experiment, Naccache et al. (2005) increased the duration of the Stroop task in order to record the skin conductance responses (SCRs). The patient showed behavioral evidence of Stroop interference, but no variation of SCR was observed from congruent trials to incongruent trials, or vice versa. In contrast, each of 10 normal comparison subjects showed a visible event-related SCR on a trial-by-trial basis, and the SCR amplitude correlated with effort measured by asking participants to verbalize which one of the two paired Stroop trials was subjectively felt as the more difficult. Importantly, the patient remained able to generate large SCRs and to report feelings of emotion when presented an emotional picture. Taken together, the series of systematic experiments have led the authors to tentatively suggest that the dissociation between preserved cognitive aspects of control and impaired consciously reportable feelings of mental effort may be due to an inability to translate actual periods of mental effort into physiological emotional signals and eventually into a conscious feeling of having made a mental effort. The preserved control might be the result of the spared right ACC because its event-related potential (ERP) still varied with the requirements for effort.

Several investigations emphasized the importance of dopamine in mediating effort-related behavior (Holroyd & Coles, 2002; Holroyd & Yeung, 2012; Schweimer et al., 2005; Walton et al., 2003). The midbrain dopamine system has been proposed to carry and transfer the reward prediction error signals (positive error indicating better than expected reward and negative error indicating worse than expected reward) to the pyramidal cells in ACC by dopamine neuron activity (Holroyd & Coles, 2002). Transient or phasic dopamine release caused by dopamine neuron firing codes for positive error, and the brief pauses in the dopamine neuron firing codes for negative error (Holroyd & Yeung, 2012). The changes (i.e. dips) in dopamine when participants make errors in cognitive tasks consequently elicit an event-related potential (ERP), referred to as error-related negativity (ERN). The ERN is not only an error or conflict detection mechanism, but its relative magnitude is associated with the extent of learning to avoid maladaptive responses. This hypothesis is supported by Frank et al. (2005) using a modified version of a reinforcement learning paradigm previously shown to be sensitive to dopaminergic manipulation. The data from recording response-locked ERPs suggested that the error-related negativity (ERN) was larger for participants who tended to learn more from the negative consequences of their decisions. By contrast, the tonic dopamine release is said to be responsible for motivating effortful behavior by coding for the average reward probability over time as effortful behavior was dramatically impaired by dopamine-induced ACC lesions (Schweimer et al., 2005; Walton et al., 2006).

Therefore, within the hierarchical reinforcement learning (HRL) framework, the phasic dopamine signal is associated with optimal option selection, whereas the tonic signal is associated with optimal option maintenance. It is the important role of dopamine in effortful behavior that has made some researchers suggest an alternative center of the effort circuit, the ventral striatum, which is a primary target of the midbrain dopamine system (Carlson, Foti, Mujica-Parodi, Harmon-Jones, & Hajcak, 2011). Schmidt et al. (2012) proposed that the ventral striatum is a common motivational center that drives both cognitive and motor effort systems by investigating the correlations between the fMRI activation of ventral striatum and cognitive and mental effort exertion. No evidence was shown in the study regarding the relationship between the ventral striatum and the ACC to either support or to reject the notion that the ACC is superior to the ventral striatum in optimal effort-based cost-benefit option selection and maintenance as indicated in the hierarchical reinforcement learning (HRL) model.

Taken together, convergent evidence has shown that effortful behavior is driven and controlled by the limbic system centering on the anterior cingulate cortex (ACC). The ACC might monitor conflict as an index of task difficulty and/or the mental effort it demands, translating this into cost-benefit analyses underlying action or strategy selection. It then integrates the choice context and the expected and obtained outcomes across time, and determines which actions are worth making and maintaining to achieve goals. The ACC is strongly interconnected with the midbrain dopamine system and receives reward prediction error (RPE) signals carried by phasic and tonic dopamine release from the ventral striatum. The reward-related scalp potentials and fMRI BOLD activation reflect the impact of dopamine reward predicted error (RPE) signals on ACC for learning option-specific values.

Although a growing literature elucidates the crucial contribution of the ACC in motivating effortful behaviors, whether those theories can be used to account for the effortful behavioral change introduced by peripheral sensory disorders (e.g., hearing impairment) is unclear, because neuroanatomically, there is a lack of evidence for direct sensory inputs to ACC. Yet such inputs may arrive either via high-level perceptual pathways or frontal areas that show stimulus specificity (Boussaoud & Wise, 1993). Neuroanatomical evidence supports the view that people with hearing loss use different processes compared to individuals with normal hearing to maintain a comparable level of speech understanding performance (Grady, 2000;

Wingfield & Grossman, 2006). Relative to people with normal hearing, people with hearing loss show a more widespread cortical activation pattern involving the areas associated with working memory and executive function, which reflect reliance on the context (i.e., top-down processing) as a strategy to compensate for difficult listening situations (Cabeza, Anderson, Locantore, & McIntosh, 2002; Pichora-Fuller & Singh, 2006). The reorganization of the neural representation from the brainstem through the auditory cortex, and/or across sensory modalities is often seen in cases with significant attenuation or reduction of peripheral input such as cochlear damage (Gordon et al., 2011; Sweetow, 2005). Because signals from the outside are not carried by these damaged sensory neurons and nerves to the brain, the brain's auditory part begins to weaken and deteriorate from the lack of stimulation and resulting disuse.

A bridge between the general cortical effort framework and the specific auditory pathway is needed to answer questions such as what brain structure in the effort circuit is responsible for the detection of errors or conflicts between peripheral auditory inputs or responses. How are the short-term and long-term statistical probabilities of experienced reward via listening effort calculated to generate effective listening strategies? What is the mediating mechanism of effort control in the connection of the cortical and subcortical auditory area and frontal area of the brain that underlies the strategic or compensatory behavior? Does the stress and constant feeling of effort caused by longstanding hearing impairment adversely impact the function of the dopamine system, consequently offset the motivation to learn and mobilize effort? To what extent, is the effort circuit flexible to compensate for the decrement of hearing?

1.2.2 Theories and models related to effort

Investigation of listening effort in humans with impaired hearing has a rather short history. Several modern theoretical perspectives on this topic are most relevant. These include the unitary-resource theory by Kahneman (1973) , ease of language understanding (ELU) model by Rönnberg (2008) and compensatory control model by Hockey (1997). Compared to the rapidly growing understanding about general effort expenditure in cognitive behaviors (e.g., multiple cognitive tasks), relatively less is known in particular about driving mechanisms of strategic selection and allocation of effort in stages of the speech understanding process.

1.2.2.1 Unitary-resource model of Kahneman (1973)

The unitary-resource model of Kahneman (1973) is widely accepted and serves as the theoretical foundation of many studies of listening effort. In this model, effort is considered equivalent to attentional capacity and attentional resource. The theory is based on the following four assumptions: 1) Humans have a limited attentional capacity that may be exerted when dealing with a task; 2) The amount of attentional capacity exerted at any moment varies depending primarily on the demands of current activities; 3) The attentional capacity is divisible and controllable, it can be allocated to facilitate the processing of selected perceptual units or the execution of selected units of performance; the policy of allocation reflects enduring dispositions and momentary intentions, and the allocation of attention is a matter of degree but more nearly unitary at high levels of task load; 4) Physiological indices of arousal provide a measure that is correlated to the momentary capacity.

One of the central elements of the model is the attentional capacity allocation policy. Expending effort depends on the enduring dispositions and the momentary intentions, and how much effort to be expended on the selected task governed by the evaluation of demands, which is the other central element of the model. According to this model, three factors are required for successful performance on a task: a sufficient total attentional capacity for the task demands, a sufficient momentary supply of the required attentional capacity, and an appropriate allocation policy of the attentional capacity. Kahneman (1973) proposed two kinds of effort – involuntary effort and voluntary effort. Involuntary effort is driven by enduring dispositions, for example, a novel, significant, complex, surprising, or incongruous stimulus spontaneously attracts attention and demands greater processing effort than a familiar stimulus without those properties. In contrast, voluntary effort is determined by current plans and intentions, for example, search for a target object, listen to the sound on the left side, solve a math problem, or comprehend an utterance. The surge of both kinds of mental effort can be reflected in manifestations of arousal, such as dilation of the eye pupil or the electrodermal response.

In a given mental task, the nature of input stimuli and task demands at each moment determines the required total attentional capacity (or effort) and momentary effort respectively. A basic rule of the allocation policy appears to be that activities which demand much capacity are favored over less demanding activities. Kahneman's (1973) theory predicts both single-task performance and multiple-task performance in terms of mental effort. In the single-task condition, the exerted effort increases with the increase of the task complexity until a state of effort overload is approached, at which the exerted effort begins to drop, producing an inverted-U shape. In the multiple-task condition, particularly when two tasks are performed simultaneously, the total capacity varies with the demands of undertaken tasks to some extent, and the transient variation in the effort that one investigates in one activity determines his/her ability to do the other task at the same time. The spare capacity exerted in the secondary task decreases with the increase of the involvement in the primary task. In reference to this model, a task performance can break down due to the insufficient input such as a degraded signal, or because of the inadequate total capacity to meet its demands (e.g., an extremely difficult task), or as a result of allocation policy (e.g., the available effort has been channeled to other activities).

Considerable empirical evidence of listening effort supports Kahneman's (1973) theory. Downs (1978) conducted a study to access effort during auditory learning under degraded listening conditions. A paired-associate spondee word learning task was performed by six groups of randomly assigned young adults with normal hearing, and the effort were accessed by introducing the visual probe reaction-time task as a simultaneous measurement. The six listening conditions of the stimuli presentation included three levels of intensity (50dB SL, 35dB SL, and 20dB SL) in quiet and at +6 dB signal-to-noise ratio, respectively. The author found no effect of presentation intensity or babble noise competition on learning performance measured by the number of trials required to obtain 100% accuracy in a single trial given 100% intelligibility for all stimuli. However, the results showed significantly less effort in terms of secondary task's reaction time in quiet versus competition conditions averaged across intensity levels, and no statistical difference on effort among the stimulus intensity levels, which demonstrated that adding competing noise rather than reducing signal intensity results in extra effort expenditure.

Howard et al. (2010) used a dual-task interference paradigm to investigate the listening effort at typical classroom signal-to-noise ratios on normal hearing children. The primary task was repetition of monosyllabic words presented in a background of children's chatter at quiet, +4, 0, -4 dB SNR; the secondary task was to simultaneously rehearsing sets of five digits for a later recall. The listening effort was indicated by the reduced performance on the secondary task. The authors found that the listening effort increased with the decrease of the SNR while the primary listening task performance was maintained.

More recently, Winn, Edwards, and Litovsky (2015) measured the impact of auditory spectral resolution on listening effort using the pupillometry technology. Results showed that the pupil dilation systematically grew with each successive degradation in spectral resolution of sentences, and the effect of spectral resolution on listening effort persisted even when only analyzing trials in which responses were 100% correct.

There have been several attempts to extend the scope of listening task in studying listening effort under the same theoretical framework. The manipulation of the difficulty level of listening tasks include the sentence intelligibility level (Ohlenforst et al., 2017; Zekveld et al., 2010), type of masker noises (Desjardins & Doherty, 2012; Francis, MacPherson, Chandrasekaran, & Alvar, 2016; Larsby, Hallgren, Lyxell, & Arlinger, 2005), the predictability of the words in the sentences (McAuliffe, Wilding, Rickard, & O'Beirne, 2012; Sarampalis et al., 2009), signal-to-noise ratio (Bernarding et al., 2010; Fraser et al., 2010; Hicks & Tharpe, 2002; Howard et al., 2010; Kramer, Kapteyn, Festen, & Kuik, 1997; Sarampalis et al., 2009), phonetic complexity of words (Bernarding, Latzel, Strauss, & Corona-Strauss, 2011; Strauss, Corona-Strauss, & Froehlich, 2008), syntactic complexity of sentences (Stewart & Wingfield, 2009), combination of visual and auditory modalities (Fraser et al., 2010; Picou et al., 2011), relatedness of paired word list (Tun, McCoy, & Wingfield, 2009), length of auditory digit test (Mackersie & Cones, 2011), order of approximation in auditory recall task (McCoy et al., 2005), uncertainty of speech location and talker (Koelewijn, de Kluiver, Shinn-Cunningham, Zekveld, & Kramer, 2015), accent interference (Iverson, Brint, Song, & Wu, 2016) and first language vs. second language (Mackersie & Cones, 2011). The literature has covered a broad range of populations, including young and old adults, and individuals with normal hearing and hearing loss. A few studies accessed listening effort of children with normal hearing (Globerson, 1983; Hicks & Tharpe, 2002; Howard et al., 2010; Stelmachowicz, Lewis, Choi, & Hoover, 2007).

The convergent finding from the above mentioned studies shows that, as the processing load of the task increases, listening effort increases with or without compromised performance accuracy; persons with hearing loss require more listening effort to solve the task than their counterparts with normal hearing; and older adults expend more listening effort than young adults recognizing and understanding speech in noise. Despite consistence with the predictions of Kahneman's (1973) model, no consensus exists regarding the common unit of effort in order to compare between tasks and subjects. More importantly, tasks used in those empirical studies failed to provide compensatory options to show the strategy pattern that participants would have used in an adverse listening environment.

There are four major limits in applying Kahneman's (1973) theory to listening effort research. First, there is no clear distinction between effort and attention in this model. Kehneman identified effort with attention and capacity as different names for the same construct, and used attention and effort interchangeably. As a result, effort was constrained to the attentional domain, which is treated in the present study as only part of effort according to the previous review of effort definitions. If effort is equivalent to attention, one would argue that any attention model, such as Broadbent's (1958) filter theory, Navon and Gopher's (1979b) multiple-resource theory, and Wickens's (2002) four-dimensional multiple resource model could be considered as an effort model; the likelihood of agreement on that is probably low though. Moreover, the proposed involuntary and voluntary attention does not mesh well with the basic concept of effort. Effort is treated as a voluntary, energy-consuming process, whereby processing resources are voluntarily allocated to a particular task or activity with or without the cost of other tasks and activities (Nocetti, 2005). In this sense, the part of "voluntary effort" determined by momentary intention in this model should be a focus when studying listening effort.

Second, the intention- or goal-driven rule of attention allocation in Kahneman's (1973) model has not received the same weight as the stimulus- or task-driven allocation policy has. According to the model, the allocation of attention is determined principally by three sets of factors: the enduring dispositions which control involuntary attention, the momentary task intentions which control voluntary attention, and the evaluation of task demands. In the experimental work, the momentary task intentions are often set by instructions regardless of single task or dual-task paradigm; for example, task instructions may dictate which parts of a task should be performed first and what subgoals have relatively high or low priority. However, the selection and maintenance of performance stability under demanding conditions in real life is usually an active process under the control of the individual, requiring the strategic management of effort. Evidence has shown difference in task performance between assigned versus participant-set goal conditions (Erez, Gopher, & Arzi, 1990; Strickland & Galimba, 2001), in which participants who self-set goals achieved higher overall performance and reported less task-related cognitive interference compared to those who were assigned goals on dual tasks. Further, the type of goal-setting has an impact on subsequent development of task strategy. Earley, Connolly and Ekegren (1989) demonstrated that participants given specific, difficult goals were more likely to develop plans for idea generation than participants given a general "do your best" goal. Effort allocation is a decision-making type of behavior; therefore experimental manipulation in listening effort studies should allow participants to optimally allocate their divisible but limited effort based on their self-set subgoals on multiple tasks when they are instructed to execute specific responses to auditory stimuli.

Third, the individual differences in effort allocation are not emphasized in this model and neither have been well addressed in the listening effort literature. The original supporting evidence of Kahneman's (1973) model is primarily based on evidence from within-subject designs, which have been carried over to the listening effort literature. Although several studies investigated between-subject factors such as hearing status and age in groups, the variability of the strategic effort allocation among individuals has been overlooked. The impact of task difficulty on exerted effort varies among individuals. Pascual-Leone (1970) found that individual differences, in particular the cognitive style of field-dependence/independence, play a central role in the way subjects manage their effort. Humphreys and Revelle (1984) postulated that two dimensions of personality, impulsivity and anxiety, are closely related to mental effort allocation. Personality accounts for considerable variance of self-reported handicap and effort due to hearing impairment in survey studies (Cox, Alexander, & Xu, 2009). The relationship between individual differences and effort allocation is discussed in more detail in a separate section.

Lastly, there is no mention about the impact of self-perceived or received feedback of performance on effort expenditure. The perception of success or failure may influence the mental effort investment. The motivational intensity theory by Brehm and Self (1989) proposes that the willingness to invest effort into the task is a function of perceived: 1) task difficulty, 2) ability, and 3) likelihood that successful performance on the task will achieve a desired goal (e.g. momentary incentive, pleasant emotion). Effort is invested into performance if individuals believe that they can succeed in achieving a tangible goal. However, if the task is perceived as either too difficult or not worthwhile, then effort will be withdrawn. Contrary to this prediction, Venables and Fairclough (2009) found that the performance feedback of failure produced adverse changes in mood/motivation, but had no significant influence on willingness to invest effort. The authors interpret the results as an agreement with Hockey's (1997) cognitive-energetical model that participants in the failure group chose to invest effort in order to compensate for perceived failure. As stated in the study, the influence of performance feedback on effort investment remains complicated due to some experimental design issues.

A recently proposed framework, called Framework for Understanding Effortful Listening (FUEL) (Pichora-Fuller et al., 2016), adapted Kahneman's (1973) model and interpreted its core components in relation to listening effort. This framework addresses the importance of the intentional factor such as motivation in listening effort research and provides a hypothetical additive interaction effect of motivation and task demand on effort. This framework makes the connection between the Kahneman's model and listening effort research more explicit; however, it does not distinguish effort from attention, and has not yet reached a stage where the strategic effort allocation could be quantitatively predicted beyond what Kahneman's model has been capable of.

In summary, Kahneman's (1973) model has been successfully used by hearing scientists in predicting the allocation of the attentional aspect of effort by listening task demand. What have not been successfully applied are the impact of intention and individual differences on the effort allocation. Although Kahneman's (1973) model covers the majority of the critical elements required to explain the effort regulation behavior, there are lack of clear discrimination between construct terms, the details of the interaction among the determinants of strategic effort regulation, and the role of other important factors such as performance feedback played in effort regulation. Those missing parts implies a necessary collaboration of multiple theoretical models in studying effort allocation.

1.2.2.2 Ease of Language Understanding (ELU) model

A major goal of listening is speech comprehension, and it entails processing a sequence of symbols that is produced and perceived over time, which requires a critical role of working memory not only in performing the symbolic computations and thereby generating the intermediate and final products from the stream of successive words in a context or spoken discourse, but also in storing them (Just & Carpenter, 1992). Speech comprehension involves both attention and working memory, although the relationship between the two constructs is complex (Awh, Vogel, & Oh, 2006; Engle, 2002; Fougnie, 2009), the concept of effort is purposely used to incorporate them rather than differentiate them. While Kahneman's (1973) model contributes the listening effort literature from the attentional perspective of effort, the more recently developed Ease of Language Understanding (ELU) model has provided important supplements from the angle of working memory in the listening effort investigation.

The ELU model is not a model of language understanding per se, but about the ease with which the processing of multimodal language input is accomplished at sentence and discourse level (Rönnberg, 2003). The model was assumed to apply to the conditions where poorly specified language information is delivered through input channels, such as hearing impairment and external noise. A growing amount of studies on listening effort use this model, assuming the ease of language understanding to be reciprocal to listening effort.

The ELU model has five basic assumptions: (1) There is a rapid binding of the speech-related information delivered in different sensory modalities, which takes place in an episodic buffer (Baddeley, 2000). This process is denoted as Rapid Automatic Multimodal Binding of

PHOnology (RAMBPHO) (Rönnberg et al., 2008) and based on research of audio-visual speech integration in the nervous system (Campbell, 2008; McGurk & MacDonald, 1976) and neurophysiology study of hearing native signers in working memory tests (Rudner, Fransson, Ingvar, Nyberg, & Rönnberg, 2007). (2) The syllabic-like phonological representations of language are stored in long-term memory. The precision of the phonological representations and the long-term memory access speed constrain the efficiency of perceptual decoding and lexical and semantic access (Pichora-Fuller, 2003). (3) The single or multi-modality information extracted by RAMBPHO is compared to a corresponding phonological representation in long-term memory through a general mismatch mechanism. The mismatch negativity (MMN), an electrophysiological response, observed during manipulation of the magnitude of stimulus change in auditory memory experiments has served as evidence for the mismatch mechanism (Näätänen, 2008; Näätänen & Alho, 1997). (4) There is a threshold in the mismatch mechanism that determines the initiation of explicit processing because of the constraints on the implicit cognitive functions in humans such as input signal decoding and long-term memory access speed. This assumption has not been systematically tested. (5) The ease of explicit processing depends on the individual's working memory capacity. The higher the WM capacity (i.e., faster processing and bigger storage) an individual has, the easier the inference-making (i.e., inferring missing information) and disambiguation (i.e., repairing misunderstandings) are for this person in language understanding. This assumption was supported by evidence that people with high WM capacity perform better on speech understanding in noise (Pichora-Fuller, Schneider, & Daneman, 1995; Rönnberg, Rudner, Lunner, & Zekveld, 2010), and the observation that people with hearing impairment or deafness who have excellent communication skills generally have higher WM capacity than their average skilled counterparts (Rönnberg et al., 1998). This assumption is also in accordance with the capacity theory of comprehension proposed by Just and Carpenter (1992) which suggests that both processing and storage elements of WM constrain the quality and quantity of the individual's language comprehension in several aspects such as flexible interaction among syntactic and pragmatic information and maintaining multiple interpretations.

As indicated in the ELU model (Rönnberg, 2003; Rönnberg et al., 1998; Rönnberg et al., 2008), when a linguistic signal enters, RAMBPHO serves as an episodic buffer and rapidly and automatically integrates the single or multi-modal linguistic information into a stream of phonological information, which is then compared to the phonological representations in long-term memory. If the incoming signal is well specified, no mismatch arises, and the implicit cognitive processing involved in language understanding under advantageous conditions is sufficient for achieving comprehension. In contrast, when the input is poorly specified (e.g., speech in background noise, speech with strong accent, distorted speech from hearing aids or cochlear implants), this poorly perceived incoming information results in a mismatch between the incoming information and the phonological representation in long-term memory and the implicit processing is no longer sufficient. Consequently, language comprehension processing becomes explicit.

Alternatively, the ELU model can be illustrated by a general formulation (Rönnberg, 2003) as follows:

$$\mathrm{ELU} = \frac{fp(P)fs(S)}{fe(E)fc(C)}$$

The four elements correspond to the assumptions. The parameter fp(P) is determined by the accuracy of phonological representations in long-term memory, and fs(S) is determined by the long-term memory access speed. The fp(P)fs(S) constitute the implicit part of the formula. The parameter fe(E) stands for the explicit process, i.e., predictive inference-making and retrospective disambiguation, and fc(C) represents working memory capacity including both storage and processing. The denominator represents the explicit process. The relationship between the relative weights of implicit and explicit processes and ease of language understanding is that ease of language understanding is positively correlated with the accuracy of the phonological representations in long-term memory and the speed access, and negatively correlated with the amount of explicit process and the WM capacity invoked as a consequence of mismatch. The higher degree of mismatch results in higher weight of explicit process, thus the less ease of language processing.

The ELU model contributes to the listening effort literature by providing a theoretical framework of when, how and where the explicit processing is initiated in language understanding in the working memory domain. The implications of the model have increased the understanding of the phenomena related to hearing loss and listening effort. One of the main implications is that mismatch occurs more frequently for people with post-lingual hearing impairment than their counterparts with normal hearing even after compensating for audibility. It is partially due to the fact that current hearing instruments are still unable to fully restore hearing functions even with advanced technologies (Edwards, 2007), and partially because the phonological representations have been altered to some extent by lack of reinforcement due to longstanding hearing loss. Andersson and Lyxell (1999) studied the contributions of deficits in the phonological representations (in LTM) and deficits in executed processing (in WM) to the phonological deterioration in individuals with acquired severe hearing impairment by using rhyme-judgment task paradigms and lexical decision tasks. Sixteen participants with severe hearing impairment and sixteen age-matched normal-hearing controls were given a cue word, a cue non-word or a cue picture at a time displayed on a computer screen, each of which was followed by a series of test items. They were asked to indicate as quickly and accurately as possible whether the presented item rhymed with the cue word or not by pushing the predefined buttons. In the lexical decision task, given a string of letters on the screen, the participants indicated whether they constitute a real word or not by pushing corresponding buttons. The rhyme judgment task requires participants to develop a phonological representation for each cue, and match the test items with it, while the lexical decision task concerns the ability of an individual to access the meaning of words through the phonological route. The results showed that the group with hearing loss performed significantly less accurately than the group with normal hearing in the rhyme-judgment task; however, no significant difference was found between groups in the lexical decision task, indicating degraded phonological presentations in long-term memory but intact working memory in the group with hearing loss. In the case of the population that is prelingually deaf, their phonological representations of sound in LTM reflect what they hear (i.e., unaided, aided with hearing aid or cochlear implant, etc.), and they are learning a different coding system from peers with normal hearing. As long as their own phonological representations have been established, mismatch should not occur more frequently than people with normal hearing, and their depressed language performance should be primarily due to the degraded quality of phonological representations in LTM. Lyxell et al. (1998) examined the pre-operative cognitive performances of 15 adult users with cochlear implants and their post-operative speech understanding performance in order to investigate the relationship between them. They found that, among the 15 participants, eight of them, who had comparable performance on a rhyme-judgment task and lexical decision-making task with controls with normal hearing, also had high functional communication ability 12 months after cochlear implantation. The remainders had significantly worse performance on cognition tasks than controls with normal hearing and less ability in communication even with 12 months experience with cochlear implants. These results suggested that the phonological representations built through cochlear implant use during 12 months did not have high enough quality to achieve communication for those who did not have intact phonological representations prior to implantation. The intact phonological representation in long-term memory is therefore crucial for people to accomplish the phonological matching and consequent language understanding. Taken together, in both prelingual and postlingual situations, the weight of the implicit process (i.e., the numerator of the ELU formula) decreases, and the language understanding becomes difficult and effortful.

The second implication is that, frequent mismatches would produce a relative disuse of episodic LTM associated with verbal information as the verbal episodes built on degraded signals are hardly successfully encoded into the LTM. The mismatches does not necessarily affect the WM and semantic LTM though because theoretically they are pushed into constant use in order to implement inference-making and disambiguation. However, the semantic LTM might be indirectly affected due to its close relationship with episodic LTM in that the efficient encoding, storage and retrieval of verbal information from episodic LTM depends on the quality of phonological-lexical representations in semantic LTM (Rönnberg et al. 2011). Rönnberg et al. (2011) investigated the relationship between hearing acuity and cognitive functions in 160 hearing aid users with heterogeneous ages and without dementia. The cognitive functions evaluated in this study targeted memory in particular, including STM (measured by free word and sentence recall tasks), semantic LTM (measured by word fluency and vocabulary tasks), and episodic LTM (measured by free recall of subject-performed tasks). The authors found significant negative correlation between hearing loss and episodic LTM performance after adjusting for age. A similar link was found between hearing loss and semantic LTM performance, but a non-significant link between hearing loss and STM performance, which has supported the disuse hypothesis derived from the ELU model. This work, along with other research investigating the association between hearing impairment and cognitive functions, such as memory (Andersson & Lyxell, 1999; Andersson, 2002; Foo, Rudner, Ronnberg, & Lunner, 2007), selective and divided attention (Desjardins & Doherty, 2012; Hauthal, Neumann, & Schweinberger, 2012; Humes, 2007; Rakerd, Seitz, & Whearty, 1996; Shinn-Cunningham & Best, 2008), and executive function (Beer, Kronenberger, & Pisoni, 2011), might expand the researchers' view of the nature of listening effort deployment by people with hearing loss based on their remaining cognitive functions, and help to define and locate the sources of listening effort since the top-down compensatory for language understanding is unlikely limited to a single cognitive domain. The dramatic variability of the findings on the access of cognitive function resulting from hearing impairment raises caution that each case should be considered individually.

The third critical implication from the ELU model is that, when the explicit process (i.e., the denominator of the ELU formula) becomes dominant, the language understanding performance is dependent on the individual's working memory capacity. In other words, once the explicit processing is triggered by mismatch, ease of language understanding perceived by an individuals to achieve a certain level of performance is dependent on his/her working memory capacity. In a study by Foo (2007), participants with hearing loss were tested with unfamiliar hearing aid settings that differed from the ones they were used to in order to produce a phonological mismatch. Foo et al. observed that WM capacity (measured by reading span) was significantly correlated with speech understanding performance and adaptation to a specific compression setting in hearing aids. Rudner et al. (2012) studied the relationship between subjective rating of listening effort and working memory capacity in persons with hearing loss during a speech recognition test at various SNRs corresponding to equated performance levels (i.e., 50%, 80%, 95%, 95%+) with either steady-state noise or modulated noise. The subjective listening effort rating was measured by a visual analog scale, and working memory capacity was measured by a letter-monitoring and a reading span task. The listening effort rating was negatively associated with working memory capacity at all SNR levels in a steady-state noise condition, but the negative relationship was only observed at the two lowest SNRs in a modulated noise condition. These results suggested that adding steady-state noise to speech has triggered an explicit process even at high performance levels, such that participants had to expend more effort at those levels than when the noise was modulated. The modulated noise provided some benefit over the steady-state noise so that the threshold of explicit process initiation (a key assumption of ELU model) shifted toward the more challenging situations. Findings from Desjardins and Doherty (2012) also support the listening effort and working memory association hypothesis. Three groups of participants, fifteen young adults with normal hearing, fifteen older adults with normal hearing, and sixteen older adults with hearing loss, performed a sentence repetition task at a fixed performance level (76% correct) in various types of background noise, and listening effort measurement was administrated using a dual-task paradigm and subjective rating scale. Results revealed that, both groups of older participants expended higher listening effort than the young normal hearing group across all types of background noise conditions, but there was no difference between the two older groups. The listening effort measured by a dual-task paradigm was negatively correlated with working memory for all groups.

From a potential listening effort framework prospective, the ELU model (Rönnberg, 2003; Rönnberg et al., 1998; Rönnberg et al., 2008) supplements Kahneman's (1997) model in several aspects. First, the ELU model brings what is thought to be a different cognitive domain from attention by majority of researchers, working memory, into the listening effort study field. Second, the ELU model was established specifically for challenging situations where the inputs are degraded, while Kahneman's model concerns ideal communication situations and does not explicitly describe how the model accommodates the adverse situations. Third, the ELU model directly bridges the individual difference of cognitive functions and ease of language understanding, especially working memory capacity; whereas the relationship between the individual differences and effort allocation is less obvious in the structure of Kahneman's model. These supplements from the ELU model have made the picture of potential listening effort framework more specific than using Kahneman's model alone. Moreover, the ELU model makes a remarkable contribution to cognitive hearing science by focusing cognitive functions to the population with hearing loss with which audiologists and hearing scientists are most interested.

There are still some crucial questions relative to listening effort left unanswered. First of all, the definition of listening effort is not implied by the ELU model. Ease of language understanding is not identical or reciprocal to listening effort. Desjardins and Doherty (2012) assessed listening effort with a dual-task paradigm with a sentence in noise recognition test as the primary task and a digital visual pursuit rotor tracking test as the secondary task. In addition, ease of listening was measured by asking participants to subjectively rate how easy it was to listen to each sentence on a scale from 0 (very difficult) to 100 (very easy). Although no statistical analysis was conducted to compare the two by the authors, dramatically distinct patterns were shown as a function of difficult listening conditions. The subjective ease of listening rating systematically decreased as the difficulty level of the noise condition increased in all experimental groups: illustrated as a downward slope. However, the listening effort of elderly groups showed a U shape, indicating that they expended less effort under medium difficulty condition than under easy and difficulty conditions, whereas the young normal group displayed an inverted U shape with lower overall values. Therefore, it seems that ease of listening or language understanding represents individual's subjective judgment on the stimulus or input, whereas listening effort represents the invoked output from a person (i.e., what the person is doing) by a stimulus. High ease of listening does not necessarily associate with less listening effort, and vice versa. Secondly, the ELU model does not provide further information about effort allocation than Kahneman's model does, since how the degree of engagement of explicit processing in working memory associates with attention allocation is not addressed. Thirdly, the key assumption of threshold in the mismatch system can be used to equate the difficulty level of task materials across individuals, but how to systematically measure the individual threshold is not indicated in this model. The ELU formula allows quantifying the ease of language understanding of a given speech material for a specific person since the elements in the formula are measurable in some sense. If the individual threshold is known, more valid results can be produced due to better control in the experimental conditions. A within subject trading function demonstration using the concurrent task difficulty manipulation method, a systematic controlled dual-task paradigm adopted from Slansky and McNeil (1997) and McNeil et al. (2004), can potentially serve as a solution.

Most importantly, the explicit compensation cannot be presumed to be passive or stimulus-driven, as a person's will or goal also plays a critical role in effort allocation. Moreover, the influence of performance feedback on effort allocation is still unclear. These two issues consequently raise a need to bring the compensatory control model into the potential listening effort framework.

1.2.2.3 Compensatory control model

In light of Mulder (1986), there are two conceptualizations of effort, one related to the difficulty of the task and the other related to the executive control of state (i.e., compensatory control). The former is mainly determined by the intrinsic attentional demand of tasks, which is in common with Kahneman's (1973) conception of mental effort. The latter concept of effort concerns the active involvement of individuals in performance of mental tasks by changing the current energetical resource state (e.g., arousal level, effort level, activation level) to meet target state. It is primarily affected by cognitive manipulations such as instructions, stressors, incentives, importance, knowledge of performance, achievement motivation, and personality factors. Hockey (1997) presented a two-level compensatory control mechanism of dynamical effort allocation which incorporates both conceptualizations of effort.

The compensatory control model assumes that: 1) Human behavior is essentially driven by goals, including both long-term and short-term goals. 2) Control of goal states is normally a self-regulatory process. 3) Individuals' internally-maintained states depend on the goals, and directly determine the performance criteria for a task (e.g., accuracy, speed, order). The target performance is subject to modification based on the perceived costs and benefits of alternative states and actions. 4) There is a performance monitoring mechanism called action monitor in the lower-level of compensatory control (i.e., routine regulation), which calculates the difference between the actual performance and the target value; and there is an effort monitor in the upper-level of compensatory control (i.e., effort-based regulation), which is responsible for detecting whether a further effort deployment is demanded. 5) The routine regulation is a bottom-up process, while the effort-based regulation is a top-down process. 6) The routine regulation has an effort budget associated with anticipated task demands, and the effort-based regulation has a maximum effort limit mainly associated with individual differences in cognitive capacity, judgment, tolerance and affective state, etc. The difference between the budget and the limit provides a reserve effort to deal with additional demands. 7) Effort may be allocated and controlled rather than stimulus-driven, and subject to strategic resource-management decisions. 8) Regulatory activity attracts costs to other parts of the system in terms of sympathetic activation, such as pupil dilation and increased excretion of cortisol.

According to Hockey's (1997) two-level compensatory control system, when an individual performs a task well within his/her learned skills, a target value for the output (i.e., performance goal) is established first, and then the current performance is continuously compared with the target output value through the action monitor during the task course. The deviation of current activity from the target value can be corrected automatically without active regulation or effort if the deviation is small, given that the processing requirement of the task is well below the individual's functional limit of effort expenditure. However, when the difference is too dramatic for the low-level correction to bring performance to the target range due to the increased complexity of the task and/or increased goal standard, the upper-level effort regulation is needed. This concept is similar to the mismatch mechanism in the ELU model (Rönnberg et al., 2008), except that the ELU model deals with the stimulus end of information processing, comparing the input signal with the phonological representation in semantic LTM; whereas the compensatory control model (Hockey, 1997) is concerned with the response end of the information processing, comparing the actual performance with the target. In addition, the target (i.e., phonological representations in LTM) used for comparison in the ELU model is relatively stable for a given individual, while the performance target in the compensatory control model may vary moment to moment depending upon individual's instantaneous cost-benefit decisions about use of effort (Granholm, Verney, Perivoliotis, & Miura, 2007; Haggard, Cockburn, Cock, Fordham, & Wade, 2000; Segerstrom & Nes, 2006) .

When the effort regulation is shifted from the lower-level to the upper-level because of the performance mismatch, effort is not assumed to automatically increase to meet the new elevated task demands as suggested in Kahneman's (1973) model, rather, a structure in the upper-level control system called the supervisory controller determines what effort allocation strategy is adopted (Hockey, 1997). For example, if the activity is important, a large reserve effort may be set to maintain stable effectiveness level of performance by actively involving working memory or executive control. In this case, the high priority task goals are maintained, and a subjective feeling of effort is perceived with elevated physiological manifestations. This strategy is referred to as active coping. Alternatively, a small reserve effort is likely set when the task value is low to the individual. In this case, the performance target is adjusted downward, which indicates an acceptance of overt performance decrement with little increase in costs. This strategy is referred to as passive coping, and can be adopted by paying less attention, less use of working memory, or reducing the required levels of accuracy or speed. Occasionally, an extreme strategy of complete disengagement from the pursuit of task goals can be observed in some laboratory tasks when the tasks are too stressful to be effectively fulfilled even with further effort expenditure, particularly in sports and education literature (Gendolla, Richter, & Silvia, 2008; Hatzigeorgiadis, 2006).

The above mentioned strategies have been documented in various areas. Hockey and Earle (2006) simulated an office-work environment and asked two groups of participants to conduct a number of standard office tasks including a set of high-priority and urgent jobs (primary tasks) and a set of necessary, low-priority and non-urgent jobs (secondary tasks) within a time limit of 2 hours. The secondary tasks should not be accomplished at the expense of the primary tasks. One group was allowed to establish their own schedule (i.e., high-control condition), while the other group had to follow a fixed schedule (i.e., low-control condition). As predicted by the compensatory control model (Hockey, 1997), different adjustment patterns of trade-offs among performance, effort and costs were observed regardless of the difficulty level of the tasks. The high-control group chose to complete the primary tasks first and allocated sufficient time on them to achieve high performance quality, and then used the remaining time to complete as many secondary tasks as they could with high performance quality as well. In contrast, the low-control group chose to compromise the performance of secondary tasks in order to ensure the completion of the primary tasks within the time limit. The performance of the primary tasks in low-control group was comparable to the high-control group, however, they completed less secondary tasks and made significantly more errors than the high-control group participants. Interestingly, participants in the high-control group rated the perceived demands of tasks and effort higher than low-control group. The authors argued that making time management decision consumes extra effort.

Granholm et al. (2007) investigated effortful processing resource allocation in people with chronic schizophrenia by measuring pupil dilation while performing a span of apprehension

(SOA) task. The pupillary response was considered as an indicator of the extent of active voluntary effort deployment. On the SOA task, participants had to detect which of two target letters was among a letter array (3-letter and 10-letter) shortly flashed on the computer screen. One subgroup of 29 participants showed impaired SOA task performance but normal pupillary responses compared to normal controls under the 10-letter condition. The other subgroup of 29 participants displayed significant decrement in both SOA performance and pupillary response in the 10-letter condition (i.e., passive coping). The result supported the notion that people with schizophrenia have deficits in specific cognitive functions rather than reduced overall resource capacity or effort, and the allocation of effort under high task load is subject to the individual's strategic decision. Ahern and Beatty (1979) studied the coping strategy of healthy individuals in solving arithmetic problems, and found that some participants with low-ability expended more effort measured by pupil dilation in a task to meet increasing difficulty challenges than people with high-ability (i.e., active compensatory coping), however, others with low-ability accepted low performance standards and adopted a passive coping strategy indicated by significantly lower pupil dilation than the high-ability controls.

Studies concerning listening effort, on the other hand, have failed to demonstrate the variability of individuals' active effort allocation. The major finding in those studies was the simple positive correlation between listening effort and the difficulty level of tasks with or without impairment of performance (Mackersie & Cones, 2011; McCoy et al., 2005; Sarampalis et al., 2009; Strauss et al., 2009a; Zekveld et al., 2010). The goal of tasks and the priority of each test item were usually pre-determined by the instructions in those experiments, and no compensatory options (e.g., replay the stimulus, taking notes) were provided; therefore, the measured effort in those studies likely reflects the passive response to changing task demands, rather than any coping strategies postulated by the compensatory control model (Hockey, 1997).

The role of performance feedback is briefly addressed in this model, although not emphasized. Regardless of internal feedback (self-generated) or external feedback (received from others), the function of knowledge of performance results is suggested to inform the individual whether the amount of effort required for the performance of a particular task is allocated or not. Therefore, it appears that the degree of discrepancy between the perceived performance and the target, rather than the origin of the feedback, determines the involvement of the high-level control and subsequently influences the strategy adoption.

Taken together, Hockey's compensatory model provides an active control-as-moderator mechanism underlying the complex relationship between mental effort, task demand and task performance. This is an extremely important supplement to and distinction from Kahneman's (1973) and Rönnberg's (2008) models in that the concept of self-regulatory control allows researchers to not only study the quantity of effort, but also the diversity of effort allocation patterns. Additionally, research based on this model has brought some insights of methodology used to investigate human's motivational regulation of effort. Unfortunately, there have been no direct experimental studies of the compensatory control hypothesis in relation to listening effort.

Besides Kahneman's (1973) unitary-resource model, Rönnberg's (2008) ease of langue understanding model, and Hockey's (1997) compensatory control model, there are some other theoretical frameworks related to effort, such as Yates's (1977) effort and performance model, Sanders's (1983) cognitive-energetical linear stage model of human information processing and stress, Pribram and McGuinness's (1975) model of energetical mechanisms, and Siegrist's (1998) effort-reward imbalance (ERI) model. These models have pronounced differences in centers of interests, assumptions and methodology from the three models mentioned here; however, considerable convergences in explanation of effort exist between them. For example, Yates's (1977) model of effort and performance focuses on the role of individual's subjective judgments (e.g., success probability, difficulty level of a task, effort-performance relationship, potential costs and benefits) in the operation of the effort control system. Those subjective judgments claim individual motivations which in consequence instruct the effort expenditure policy. The model frame is similar to Kahneman's (1973) unitary-resource model while the components more accord with the view of Hockey's (1997) compensatory control model. As Yates's (1977) model was based solely on subjective estimates on hypothetic tasks rather than truly measured outcomes, the reliability and validity tests of this model are needed before it can be applied in other fields. Table 2 outlines a comparison of addressed effort-related questions among the models emphasized in this document from a potential effort-centered framework perspective.

Table 2. Comparison of theoretical models related to effort

	Hierarchical Reinforcement Learning model (Holroyd & Yeung, 2012)	Unitary-resource model (Kahneman, 1973)	Ease of Language Understanding (ELU) model (Rönnberg et al., 2008)	Compensatory control model (Hockey, 1997)
Definition of effort		●		
Neurophysiological basis of effort	●			
Account for degraded input			●	
Origins of effort	●	●	●	●
Effort initiation mechanism	●		●	●
Data-driven effort exertion	●	●	●	●
Goal-driven effort exertion		●		●
Volume of effort		●	●	●
Effort allocation pattern (strategy)				●
Effectiveness of effort		●	●	●
Efficiency of effort		●		●
Prediction of performance	●	●	●	●
Interference among concurrent tasks		●		●
Interference among successive tasks				●
Individual difference			●	●
Active compensatory control of effort				●
Time pressure		●		●
Effort measurement	●	●		●
Upper limit of effort		●		●
Cost/benefit decision	●			●
Effect of performance feedback on effort				●
Potential computational formula of effort			●	
Used in listening effort studies		●	●	

1.2.2.4 Speech comprehension and effort

Speech comprehension is one of the most complex human activities, and the comprehension processes occur at multiple levels across units of language. Wingfield and Tun (2007) proposed a framework of operations involved in speech comprehension at the word, sentence, and discourse levels, as well as cognitive supports and constraints in spoken language comprehension. The model comprises three basic systems: sensory system, perceptual system, and cognitive system, constrained by an overall limited processing resource.

The operations involved in the sensory system include sensory input detection and source discrimination that are carried out by the peripheral and low-level central auditory systems. The acoustical features of the input signal, such as intensity, spectrum, metrical pattern, temporal structure, and spatial localization, are collected at this stage. The status of the auditory system directly determines the quality of the information to be encoded. The second system, the perceptual system, begins with aggregation and/or segregation of the received acoustical elements, and this also is where cognitive attention enters. An attentional filter is engaged to determine which sound to focus on and to minimize the interference as a result of energetic masking and informational masking, so that one can isolate the target speech from the background. The perceptual operations following the attentional filter are phonological analysis and lexical identification, which identify lexical elements, such as nouns, verbs, adjectives, and adverbs that the phonemes represent. The perceptual system extends thus far, not only cognitive attention plays an essential role in perception, memory also contributes to this process by serving as a real time mapping system similar to the matching mechanism mentioned in the ELU model (Rönnberg, 2003). However, understanding of the meaning of the input signal is not required at the perceptual stage, whose main purpose is to make the linguistic elements necessary for comprehension available for further analysis. Speech perception capability is usually examined by speech recognition tests that require listeners to repeat what they heard. Listeners do not have to understand the materials in order to perform those tests. The stimuli can be either speech sound or non-speech sound. The third system, the cognitive system, consists of several interactive operations such as syntactic resolution, thematic role assignment, coherence structure, and discourse comprehension. In other words, given all the necessary linguistics elements from the perceptual stage, listeners must identify the meaning of the individual words, their relationships with other words, and the characteristics by which one can determine what part of speech the word belongs to, and/or how the words are associated with the theme of a certain sentence. At the discourse level, sentences are connected via coherence logical structures; therefore, determination of the propositional content of each sentence and integration of the content within and across sentences are needed to achieve discourse comprehension.

More detailed discussion of language comprehension, at high-level in particular, was supplemented by Perfetti and Adlof’s (2012) schematized model of comprehension components. Although the framework is based on reading comprehension, it can be readily applied to oral language comprehension as the key processes are largely shared between written and spoken language. According to Perfetti and Adlof’s (2012) framework, word-level comprehension is basically a phonetic and lexical process, and sentence-level comprehension is a syntactic and semantic process in addition to phonetic and lexical analysis and closely tied to the grammatical structure of the sentence. The most important supplement to Wingfield and Tun’s (2007) cognitive system was at discourse-level comprehension, which includes high-level comprehension components such as inference-making, comprehension monitoring, and comprehension strategy usage. Inference-making occurs in routine comprehension and helps the comprehender build a coherent mental representation of the discourse. Because inferences are usually triggered by missing or inexplicit elements of the context, the availability and accessibility of the prior knowledge become critical determinants of successful inferences drawing. Comprehension monitoring is a metacognitive skill to monitor or judge the quality of one's understanding (Pitts, 1983), i.e., the ability to be aware, while reading or listening, whether a text is making sense or not. In addition to verifying understanding, comprehension monitoring also plays an important role in allowing the comprehender to make repairs where this understanding fails. As is true for inference-making, retrieval of knowledge from memory is necessary for comprehension monitoring. Comprehenders commonly use strategies to enhance comprehension. For example, strategies enhance reading comprehension (comprehension monitoring, cooperative learning, use of graphic and semantic organizers, question answering, question generation, story structure, summarization, etc.) (National Reading Panel, 2000), and reception strategies enhance listening comprehension (global reprise, specific reprise, hypothesis testing, kinesics, uptakes, faking) (Rost & Ross, 1991; Vandergrift, 1997).

Although described in a sequential fashion, many of the comprehension components and operations are necessarily interactive across linguistic levels (Marslen-Wilson, 1975) and across time (Dahan, 2010), with information flow moving in both directions. Marslen-Wilson (1975) presented evidence in a sentence shadowing task that sentence perception is modulated as a parallel rather than serial process among the four descriptive levels – phonetic, lexical, syntactic and semantic. In the experiment, participants were asked to shadow sentences and rapidly repeat back speech as they heard it. The stimulus materials were constructed such that, 120 pairs of sentences were randomly assigned to 3 context groups – normal group, semantic group and syntactic group. The tri-syllabic target-word in the second sentence of each pair was replaced by a semantically anomalous new word in semantic group and by a semantically and syntactically anomalous new word in syntactic group. Each context group was further divided into 4 subgroups of 10 pairs each in order to produce lexical disruption. In one subgroup in each context group, the target-word was left unchanged, and in the other three subgroups, the first, second or third syllable of the tri-syllabic target-word was changed so as to make it into a nonsense word, for example, *tomorrow* was changed into *tomorrane* (the third syllable change). The word restoration (i.e., the restoration of disrupted words to their original form), context restoration (i.e., reinstatements of the original word that had been replaced by a contextually anomalous word), and repetition latency (i.e., duration between the onset of the target-word in the input and the onset of the word in output) were measured to demonstrate the interaction effect of context disruption and word disruption. The results showed that the word restoration occurred most when the disrupted word was consistent with the preceding semantic and syntactic context (normal group) and when its first one or two syllables were not disrupted; however, this pattern was not seen in the semantic group and syntactic group. In addition, the repetition latencies for context restoration were shorter than those for word restoration. Both findings were contrary to the serial models of sentence processing which would expect the word restoration frequency to be independent of context variables, and the influence of the contextual information to be less effective at short rather than long repetition latency. Therefore, the authors proposed that the listener analyzes the incoming material at all available levels of analysis, such that the information at each level can constrain and guide simultaneous processing at other levels.

In the time domain, the traditional view of comprehension of spoken language over time is that the perceptual interpretation of continuous speech takes place in real time (i.e., a sequential or “left to right” analysis), with linguistic analysis and decision-making accomplished at the rate at which information reaches the senses (Frazier, 1987). The core assumption of this theory is that the lexical processing completes in a serial way with discrete candidate sets and discrete points in time narrowing down as new information adds in to identify a single best-fitting lexical candidate, and there is a brief but measureable temporal delay of the influence of contextual information. However, the recent emerging evidence supports an alternative view that mapping from the input to meaning is a continuous process with merging representations that are continuously updated using multiple information sources such as contextual and phonetic sources, similar to the idea from Bayesian models that multiple sources of information are evaluated simultaneously and in a probabilistic manner to achieve the optimal interpretation of the signal. The propagation of constraints takes place continuously, so that the total support for each alternative is continuously updated (Dahan, 2010; Dahan & Tanenhaus, 2004a). Dahan and Tanenhaus (2004b) designed two eye-tracking experiments to test this theory. Participants were asked to select the object they heard in a sentence from 4 objects displayed on the screen, and the proportion of their fixations on each object was measured. The four displayed objects corresponded to the target referent, a cohort competitor which overlapped with the onset of the referent’s name, a semantic competitor which semantically related to the target, and an irrelevant distractor. In experiment 1, two types of verb-based semantic constraint conditions were created, one with the main verb (i.e., semantic context) preceding the target noun, e.g., literally translated from Dutch “Today crawls the *baby* a bit further”; and the other with an auxiliary verb preceding the target word and the main verb followed the target word, e.g., literally translated from Dutch “Today has the *baby* crawled a bit further”. The authors referred to these two conditions as the constraining-verb condition and neutral-verb condition, respectively. The neutral-verb condition was designed to represent the typical “left-to-right” analysis where the early arriving information has immediate influence on decision making so that the decision space (i.e., set of possibilities) becomes smaller as new pieces of information are added, eventually, a single best-fitting candidate is selected, taking into account goodness of fit with the input and the context. The constraining-verb condition was the critical condition which demonstrated the continuous and simultaneous integration of multiple sources of information, and the modulation of initial hypotheses by later-arriving information (referred to as right context effect by the authors). The eye-tracking results in terms of proportion of fixations on the given words over a time window from 200ms to 500ms after target word onset (i.e., mean duration of the target word) were evaluated. In the neutral-verb condition, participants were equally likely to fixate on the cohort competitor and the target, the fixations to the target and cohort competitor increased with a similar slope from 200ms until around 350ms after target-word onset. Fixations to the target then continued to rise while fixations to the cohort competitor began to drop. The rise and fall in cohort competitor fixations occurred while the target word was heard and processed, illustrating initially consistent and subsequently inconsistent effect of the cohort competitor. A strikingly different pattern was seen in the constraining-verb condition, where the fixations to the target and cohort competitor diverged very early, and the semantic competitor was not activated in either condition. The results indicated an immediate integration of contextual and phonetic information, and fixations were determined by the combination of the two resources. There was no delay of contextual influence on the speech comprehension as assumed by sequential processing models. In experiment 2, the authors additionally manipulated the fragment of the vowel (second formant) of the target word in each sentence such that it anticipated the following consonant, which led to two different words and meanings, either the target referent or the cohort competitor. This manipulation helped to demonstrate that the spectral cues in the vowel mitigated the left-to-right contextual effect. In both neutral-verb and constraining-verb conditions, when the prediction of the word by vowel was incongruent with the verbal contextual information, the fixations on the target decreased and the fixations on the cohort competitor increased; in addition, the point in time of the spectral cues' impact was quite similar between two conditions. This shift of fixation occurred even when the contextual and phonetic information had converged towards a single best-fitting lexical candidate which indicated that the spectral cues in the spoken word's vowel temporarily modulated the interpretation. The above evidence supports that percepts during speech comprehension emerge from both anticipation of upcoming information and integration over a larger temporal window, and the ultimate decision is reached after more time has passed and more information has accumulated.

Taken together, there are enormously sophisticated operations in comprehension of spoken language, especially at discourse-level in the cognitive system specified in Wingfield and Tun's model (2007). The operations embedded in the complete comprehension processing are highly active and interactive, even the audition in the sensory system which was traditionally seen as a passive filter or frequency analyzer based on the Fast Fourier transform (Klatt, 1989), has now been discovered to be an active process (Greenberg, 1996), in that different perceptual strategies are automatically applied in different listening environments. For example, instead of detailed spectral portraiture, the auditory system extracts invariant spectra-temporal representations (syllable-like units) of the speech signal through the computation of the low frequency modulation spectrum in the auditory cortex in order to obtain the same basic meaning across the diverse acoustic conditions such as reverberation, background noise, change of speaking rate, speaker, or style (Greenberg, 1996). High-order operations (e.g., inference-making, comprehension monitoring) not only depend on the accuracy of the relatively low-order operations (e.g., building phonological representations, context-appropriate meaning retrieval from memory), but also have influence on them. Therefore, bottom-up and top-down processing coexist in natural speech comprehension. Corruption at any point in the process will introduce difficulty in understanding speech and require extra effort to fix it in order to achieve successful comprehension.

According to Wingfield and Tun (2007) and Perfetti and Adlof (2012), the cognitive recources draw on a single capacity-limited resource pool, including attention, memory and executive function, are widely involved in operations within speech recognition and speech comprehension. The purpose for that is to select, temporarily store and process the phrases and clauses of syntactically complex sentences in order to determine the correct sentence meaning. Compared to speech recognition, comprehension generally requires more complex operations accompanied by more effort. Dahan and Tanenhaus (2004a) argued that the goal of lexical processing is to make lexically specific information available for ongoing computations about comprehension, rather than mere what words have been said in the sentence (recognition). However, in the literature of listening effort, little attention has been paid to comprehension; instead, researchers tend to focus listening effort at the speech recognition level. The listening goal can modulate the pattern of effort allocation. Fallon et al. (2006) examined the effect of task goal (recall vs. comprehension) on normal-hearing adults' self-paced listening time using an auditory moving window (AMW) task, in which the participants controlled the flow of information in a word-by-word or clause-by-clause fashion at their own rate by pressing a key when they were ready to hear the next segment of the speech message. After presenting each sentence, participants were asked to either repeat the sentence (recall condition) or respond "true" or “false” to a comprehension probe (comprehension condition). The pause duration between segments was recorded to demonstrate the pattern of time expended on the segments, implicitly illustrating the pattern of cognitive effort allocation. The results showed that the pause durations at the major clause boundary and completion of a sentence were significantly longer when the listening goal was to prepare oneself to answer a comprehension question than when the goal was to merely recall what had been heard, and the pause durations at other words in the sentence were at similar low levels. This result indicates that the type of task demand is a highly relevant determinant of the effort exertion. However, the selection of test materials for this type of study should be done carefully because a comprehension task does not always demand more effort than a recognition task. It depends on various factors including the length of the sentence, the comprehension probe type, the memory function of participants, etc. All the sentences used in Fallon (2006) were nine words in length, and none of the materials used in listening effort studies that were asked to be recalled have exceeded this length. Research relative to the effect of listening goals on listening effort is still in its infancy. To what extent the listening goals impact the allocation of listening effort in individuals with hearing loss remains an open question.

In summary, comprehension of spoken speech comprises highly inter-correlated perceptual and cognitive operations. Speech processing is considered more as a parallel and continuous integration of all available levels of analysis across time than a serial process. Cognitive effort in terms of attention, memory and executive function is involved in multiple levels of speech processing. As the general goal of listening is to understand the speech rather than mere recognition, studies on how people manage their effort while listening should take the comprehension-related operations into account in order to simulate the real life situation as well as to reveal the full pattern of effort deployment.

1.2.3 Individual difference and listening effort

Effort has been used in many theories of human behavior to account for individual differences in intellectual performance in terms of differences in the availability of mental resources and in motivation (Humphreys & Revelle, 1984; Mulder, 1986; Strickland & Galimba, 2001). Hockey et al. (1986) suggested that the analysis of individual differences on mental effort not only can provide information on the range of processing options available to the individual, but also on the degree of flexibility in terms of coping strategies that of which the system is capable. The effort exerted by an individual usually relates to his or her physical and mental state. If attempting to characterize the listening effort allocation, it is necessary to know the available physical and mental resources of each individual in order to explain his or her decision on effort allocation in a given task.

The majority of listening effort studies are within-subject designs with manipulation of task difficulty level. Only a few studies have looked at between-group factors such as age (Bernarding, Latzel, et al., 2011; Bernarding, Strauss, et al., 2011; Desjardins & Doherty, 2012; Gosselin & Gagné, 2011a, 2011b; Stewart & Wingfield, 2009; Tun et al., 2009) and hearing status (Bernarding, Strauss, et al., 2011; Hicks & Tharpe, 2002; Kramer et al., 1997; Stewart & Wingfield, 2009; Tun et al., 2009). There is little discussion about the individually different pattern of effort allocation and its underlying components, which might be more informative than focusing on group averages so that the conclusions of the research work on listening effort are not overgeneralized.

There are many dimensions of individual differences that might be related to listening effort allocation, including general factors such as age and intelligence. However, they are not the main focus of this review. The current section discusses three specific potential perspectives in which hearing-impaired individuals may differ from each other: auditory ability, cognition ability, and personality in terms of need for cognition.

1.2.3.1 Auditory ability

Studies concerning listening effort of participants with hearing loss usually administer the common audiometric measures such as pure-tone air-conduction and bone-conduction thresholds, tympanogram and acoustic reflex. The most popular sample in these studies has been participants with mild-to-moderate high frequency sloping sensorineural hearing loss.

One issue with focusing only on this specific sample is that generalization of their effort-related behavior observed in those studies may be limited because other degrees or types of hearing loss have not been investigated in the same manner. One cannot simply assume that mildly hearing-impaired individuals would exert the same amount and pattern of effort in a listening task as individuals with profound hearing loss. On the other hand, severe impairment or handicap does not guarantee more effort from the individual. Hence, holding these assumptions might essentially introduce bias to their task performance and even rehabilitation solution.

Although not studied in hearing impairment, research in other disorders has provided some reference. In an experiment examining the effect of severity of traumatic brain injury (TBI) on performance on the Wisconsin Card Sorting Test (WCST), a total of 176 TBI cases were included in the study, and they were classified as mild or moderate to severe according to initial injury characteristics based on a thorough review of medical records (Ord, Greve, Bianchini, & Aguerrevere, 2010). In order to account for the biases introduced by the various effort levels across participants, researchers administrated four commonly used cognitive performance validity measures to identify participants' effort, which included the Portland Digit Recognition Test (PDRT; Binder, 1993), Test of Memory Malingering (TOMM; Tombaugh, 1996), Word

Memory Test (WMT; Green, Allen, & Astner, 1996), and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994). Sixty-seven participants with mild TBI showed good effort to perform the WCST, and 42 participants with mild TBI showed poor effort. Forty-six participants with moderate-to-severe TBI showed good effort, while 21 participants with moderate-to-severe TBI showed poor effort. Results demonstrated that effort during testing had a larger impact on WCST performance than did the severity of TBI, which is consistent with the reports of the influence of individually different effort on neuropsychological test performance studied by Lange et al. (2012) and Iverson (2010). Green et al. (2001) also reported that effort during testing accounted for 53% of variance on neuropsychological measures. The common comment was that those who fail effort testing are likely to be misdiagnosed as having severe cognitive impairment, and their symptom reporting is likely to be inaccurate. The absence of the direct correlation between the severity of hearing impairment and effort might also exist; thus, research including a variety of samples is needed to better represent the continuum of the population with hearing loss.

Another issue with the focused sample average is that individuals with similar degrees and configurations of hearing impairment identified by audiogram may have large variability in other auditory abilities such as frequency selectivity due to outer hair cell (OHC) damage (Ruggero & Rich, 1991), temporal resolution resulting from damage to inner hair cells (IHCs) and auditory-nerve fibers (Moore, 1993; Moore, 2007), and binaural hearing relying on the central auditory system (Moore, 1991). Evidence shows that considerable variation of deterioration in temporal resolution and/or frequency selectivity exists among patients with diverse inner-ear disorders and among patients with the same diagnosis, regardless of the degree of their hearing loss (Schorn & Zwicker, 1990). Santurette and Dau (2012) evaluated the individual auditory profile of 8 normal-hearing and 14 hearing-impaired listeners to investigate the relationship between specific deficits and binaural pitch perception performance. The auditory profile included measures of hearing thresholds, frequency selectivity, loudness perception, binaural processing, temporal fine structure processing and cognitive ability. The authors found that the loudness-growth slopes varied in a wide range from normal (approximate 0.3 cu/dB) to over 1.5 cu/dB in hearing-impaired listeners. Large variability on frequency selectivity and temporal resolution also were observed in this group. Interestingly, they were both highly correlated with hearing thresholds; however, there was no correlation between frequency and temporal selectivity, suggesting that the ability to process the temporal fine structure is somehow independent of frequency selectivity. The absence of relationship between the function of temporal fine structure processing and frequency selectivity in homogeneous listeners with hearing loss in terms of audibility also was found by Strelcyk and Dau (2009). Those studies are consistent with Moore's (1985) notion that temporal resolution is impaired in most but not all cases of sensorineural hearing loss. The common audiometric measures of hearing ability seem only to provide gross information about listeners' hearing status. If investigating listening effort using tests such as speech perception in background noise or compressed speech, whose performance depends on individual's fine auditory profile (e.g., temporal fine structure processing) (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006; Pichora-Fuller, Schneider, MacDonald, Brown, & Pass, 2007), ignoring the individual differences may lead to errors in interpretation, especially when listening effort is indicated by task performance such as in a dual-task paradigm. For example, a group average may indicate a great amount of effort and poor performance, but some individuals with fairly normal temporal fine structure processing ability may actually achieve normal outcomes with little effort.

1.2.3.2 Cognitive ability

Similar to hearing ability, the correlation between individually different cognitive function and listening effort has seldom been directly and systematically tested. Although there are studies showing the hearing impairment-related declines in working memory (Mackersie, Boothroyd, & Prida, 2000), and focusing selective attention (Gatehouse & Akeroyd, 2006), it has been argued that many of the cognitive deficits observed in people with hearing loss are in fact downstream consequences of degraded hearing ability because the normal processing stages are built on one another (McCoy et al., 2005; Shinn-Cunningham & Best, 2008). Growing evidence suggests that individuals with hearing loss actually differ in cognitive functions as dramatically as the normal-hearing population. People with hearing loss may have impaired phonological representations but intact working memory (Andersson & Lyxell, 1999). A crucial finding on speechreading skills of participants with hearing loss from Rönnberg (1998) is that individuals with hearing loss as a group did not demonstrate superior compensation in speechreading compared to matched normal-hearing controls. Participants with hearing loss who had excellent speechreading skills were those with high working memory function. This indicates that variation in visually-based communication skill is impacted by individual cognitive differences rather than adaptive compensation or neural change. Santurette and Dau (2012) measured cognitive function using lexical decision task and reading span task in a group of participants with hearing loss. They found that there was no correlation between cognitive function and hearing loss. The average accuracy and response time of lexical decision was similar between the two groups with small variations; however, large variability was found in the reading span test with no group difference between the participants.

Just and Carpenter (1992) studied individual differences in working memory capacity for language comprehension, and found that working memory varies among normal young adults in terms of capacity and allocation scheme. In their capacity theory of comprehension, computational processes in combination with storage resources constitute working memory, and there is a trading relationship between the two components. This means that the ability for rapidly and accurately processing information will be compromised if storage of information for later recall also is required. This trading relation between processing and storage in working memory not only suggests a common pool of language resources, but also represents allocation schemes under the condition that the resource demands of the task exceed the available supply, and it is only in this condition that capacity limitations would affect performance. This theory was based on empirical evidence of the relationship between individual working memory capacity and language comprehension performance in various facets of language processing such as lexical access, syntactic parsing, semantic analysis and referential processing studied in a group of college students by means of percent correct, reaction time, reading time per word and gaze duration measurements. Additionally, the human data have been successfully replicated in a proposed simulation model of sentences processing constrained by a single resource capacity. Just and Carpenter (1992) postulated three basic resource allocation schemes under the capacity constraint: evenhanded scheme in which both processing and storage of working memory share the limited capacity equally, processing-favored scheme which ensures fast processing but compromises the accuracy, and storage-favored scheme which maximizes the accuracy at the cost of processing speed.

It appears that the working memory system described in Just and Carpenter (1992) can be readily applied as a listening effort model since it accounts for both quantity (capacity) and quality (allocation strategy) aspects of effort, and working memory is known to widely engage in the operations in spoken language comprehension. However, working memory is only one source of effort among all kinds of cognitive resources, and the working memory resource allocation schemes are difficult to explain individual's favoring some processes or even tasks over the others compared to explaining individual performance strategy in a specific process or task. Moreover, these resource allocation schemes have not been verified in people with hearing loss. To what extent the schemes represent their individual differences is unknown. Just and Carpenter (1992) mentioned but did not emphasized a concept of intensity dimension of thought which represents how much an individual feels himself/herself being engaged in his/her thought processes, which might vary independently throughout the performance of a given task. One can objectively assess it through physiological measurements such as pupillometry (Beatty, 1982) and glucose metabolism (Haier et al., 1988). Nevertheless, the relationship between the intensity dimension of thought and individual working memory capacity or individual performance differences was implied neither in human data nor in the simulation model.

In the language comprehension domain of cognition, the individual differences of language processing ability have been indicated in Rönnberg's (2008) ELU model to determine the ease of language understanding. The implicit part of the ELU formula consists of accuracy of phonological representations in long-term memory and the long-term memory access speed. The precision of the phonological representation is an important parameter related to interindividual differencea in phonological processing (Andersson & Lyxell, 1999). Studies on phonological representation in people with hearing loss support large variability among individuals regardless of hearing loss (Foo et al., 2007; Lyxell et al., 1998; Lyxell et al., 1996), and a wide range of long-term memory retrieval speed also was evidenced in people with hearing loss (Foo et al., 2007; Rönnberg, 1990; Rönnberg et al., 1998; Wingfield, 1996). The time window of lexical access may last up to 300-400ms, and its duration is an important feature of the ELU model for predicting individual difference (Poeppel, Idsardi, & van Wassenhove, 2008). The individual difference of working memory is included in the ELU model as an explicit part, distinct from Just and Carpenter's (1992) capacity model in that it is used to explain ease of language understanding rather than performance of language comprehension.

Theoretically, the higher cognitive functioning one has, the easier a listening task should be for the person, and hence the less effortful one might feel the task to be. However, evidence has proved that is not always the case. Picou et al. (2011) investigated the prediction of listening effort by individual lipreading ability and working memory capacity in a paired-associates recall task under two modality conditions (audio-only and audio-visual). They found no significant relationship between working memory capacity and objective (recall performance) or subjective (self-rating) measures of listening effort when providing visual cues. Participants with high working memory and lipreading skills did not show a release of effort, which was similar to those with low abilities. Fraser et al. (2010) also failed to observe the listening effort decrement with the addition of visual cues in a dual-paradigm task (a speech recognition primary task and a tactile secondary task), but individual cognitive function was not investigated in this study. Picou et al. (2011) argued that integrating visual speech cues and auditory cues might demand extra cognitive resources, and this extra demand overweighs the visual cue benefit. These observations also might result from the different experimental design from other studies which found visual cues reduce listening effort, or might indicate there is something else independent of cognitive abilities that controls listening effort.

Taken together, there are numerous interindividual differences in the population with hearing loss in terms of auditory and cognitive functions. Those differences determine the amount of available internal resources that one can possibly use as effort in performing listening tasks. Various working memory resource allocation strategies can be observed under high task demand or time pressure conditions, and the strategies are constrained by resource capacity. Although researchers have started to cast their attention to individual differences in the hearing-impaired population, studies are limited to the investigation of the relationship between the individual differences and task performance without looking at the listening effort allocation pattern of each individual. Consequently, the contribution or constraint of individual auditory and cognitive characteristics to the individually different listening effort allocation remains unclear.

Auditory ability and cognitive function are critical sources of listening effort; however, they are inadequate to account for all the variability of listening effort among individuals with hearing loss because those abilities are relatively stable within an individual, whereas effort has the nature of fluctuating moment by moment and might affect performance regardless of the absolute level of raw cognitive or sensory ability (Wingfield & Tun, 2007). This is hard to explain without including the psychological dimension of individual differences.

1.2.3.3 Personality

According to the two-level effort system in Hockey's (1997) compensatory control model, the lower level computational effort is associated with task demands, while the higher level computational effort by contrast depends more on individual motivation in origin and is more variable. In the psychology literature, the extent to which effort will be invested in a task is determined by subjective judgment of probabilities of successful performance and anticipated performance consequences (West, Thorn, & Bagwell, 2003; Yates & Kulick, 1977). Humphreys and Revelle (1984) addressed the effect of personality dimensions (introversion-extraversion, achievement motivation and anxiety) on motivational construct of effort, moderated by situational factors such as success, failure, time pressure, incentives, time of day and stimulant drugs. Introversion-extraversion is believed to relate only to arousal level; introverts are more aroused than extraverts. Achievement motivation relates only to effort, and anxiety relates to both arousal and effort. Although effort exertion explained by achievement motivation is quite straightforward in that high motivation corresponds to more effort, it is difficult to quantify the achievement motivation construct.

Two concepts have been introduced as psychology dimensions of individual differences to study human behavior in effortful activities such as perceiving, thinking, problem-solving, decision-making, effort-managing, learning and so forth. One is the need for cognition, and the other is cognitive style.

Need for cognition

Cohen, Stotland, and Wolfe (1955) described the need for cognition (NC) as “a need to structure relevant situation in meaningful, integrated way. It is a need to understand and make reasonable the experiential world” (p. 291). They proposed that the resultant tension would lead to “active efforts to structure the situation and increase understanding” (p. 291). In contemporary literature, need for cognition refers to the tendency for people to vary in the extent to which they engage in and enjoy effortful cognitive endeavors (Cacioppo & Petty, 1982; Cacioppo, Petty, Feinstein, & Jarvis, 1996; Petty et al., 2009). People with high NC constantly engage in and enjoy cognitively challenging activities, whereas people with low NC find thinking to be a chore that is engaged in mostly when some incentive or reason is present.

Need for cognition has been examined in a wide variety of areas. For example, in the domain of survey research, individuals with high NC provide more thoughtful responses and are less likely satisfied with their answers than individuals with low NC (Krosnick, 1991, 1999). In the study of attitudes and persuasion, researchers have found that people with high NC tend to form attitude on the basis of an effortful analysis of the quality of the relevant information in a persuasive message; in contrast, the attitude of people with low NC are based more on simple peripheral cues inherent in the message (Chang, 2007; Haugtvedt, Petty, & Cacioppo, 1992). Individuals high in NC are also more likely to think about their thoughts (i.e., engage in metacognition) than individuals low in NC (Petty, Briñol, Tormala, & Wegener, 2007). At the most basic level, need for cognition has been shown to affect the amount of thought that goes into a decision. This has been demonstrated in the studies of false memory effect, Halo effect and priming effect. In Graham's (2007) study, participants were presented with word lists composed of associates which were semantically related to those contained in the studied lists; on a subsequent recognition test, high need for cognition participants falsely recognized a greater proportion of the non-presented words than low need for cognition participants. Because the high NC individuals elaborated each list item and had stronger interconnections in memory, they were more likely to think about and access the semantically related non-presented items and therefore showed greater false memory. Individuals with low NC have shown to be more susceptible to the Halo effect, a phenomenon in which people rate attractive or likable others as superior on a variety of other trait dimensions (Feingold, 1992), compared to individuals with high NC because the Halo effect can occur when people rely on their stereotypes of attractive others alone to judge a novel target (Perlini & Hansen, 2001). Petty, DeMarree, Briñol, Horcajo, and Strathman (2008) reported that the need for cognition as an individual variable has opposite implications for priming effects, depending on prime blatancy. They found that as need for cognition increases, the magnitude of the priming effect increases with a subtle prime but decreases for a blatant prime.

The need for cognition also is related to efficiency of task performance. Butler, Scherer, and Reiter-Palmon (2003) empirically examined the effectiveness of a solution elicitation technique based on the presentation of problem objectives and also examined the relationship between need for cognition and creative problem solving. They found that need for cognition was positively related to the efficiency of solution elicitation (i.e., the proportion of high-quality solutions) when no objectives were presented; participants with high NC generated higher proportion of high-quality solutions than participants with low NC, but the relationship was absent with the presentation of objectives. Kearney, Gebert, and Voelpel (2009) studied 83 teams from eight different German organizations engaged in various industrial sectors including software development, pharmaceuticals, insurance, telecommunications, manufacturing, media and entertainment, food and energy. They investigated team need for cognition as a moderator of the relationships between the team diversity (both age and educational specialization) and team performance in terms of efficiency, quality of innovations, productivity and overall achievement. Their results showed that both types of diversity were significantly positively related to the team performance only when team need for cognition is high

The previous research of need for cognition in the above areas of social psychology have shown that need for cognition is a stable individual difference in intrinsic motivation of mental effort across a wide range of domain, and individual differences in need for cognition are relevant to understanding not only how people process information, but also how they spend effort. As need for cognition increases, people prefer allocate more effort to cognitive tasks. This preference generalizes to all cognitive tasks regardless of the importance or consequentiality of the task (Betsch & Haberstroh, 2014).

Despite the research pertaining to the need for cognition, relatively little is understood about the relationship between need for cognition and the decision about the effort allocation while listening in challenging situations (e.g., fast speech). Specifically, whether high need for cognition is associated with the increased listening effort is still unknown.

Cognitive style

Cognitive style has been studied as dimensions of individual differences in psychology since the early 1950s, and influences how people look at their environment for information and how they use the information to guide their actions (Hayes & Allinson, 1998). Cognitive styles represent individuals' heuristics or preferences of processing information and allocating cognitive resources, and have both characteristics of stability within individuals over time and adaptability in response to specific environmental circumstances (e.g., profession or education) (Kozhevnikov, 2007). As a result, cognitive styles also are considered as social interactions regulating people's beliefs and value systems (Witkin & Goodenough, 1981).

There are numerous ways of describing cognitive styles. The field dependence/independence (FDI) construct is one of the most cited cognitive styles, and it has become a sort of general theory of perception, intellect and personality. According to Witkin's (1977) definition, field independence (FI) is "the extent to which a person perceives part of a field as discrete from the surrounding field as a whole, rather than embedded in the field; or the extent to which the person perceives analytically". People who exhibit field dependence (FD) tend to rely on information provided by the outer world, and their cognition is based on this overall field rather than embedded parts. They are holistic and socially aware. In contrast, field-independent people perceive a particular relevant item or factor in a field and tend to depend on their inner knowledge to analyze problems without reference to the field. FDI reflects an individual's perceptual and processing characteristics which influence the preferences and strategies one uses to perceive process, store and recall information (Chinien & Boutin, 1993).

Same-age individuals who differ in their cognitive style (i.e., FDI) have the same mental capacity limit but function differently in terms of momentary effort exertion and/or effort allocation strategy employment (Globerson, 1983, 1985; Guisande, Páramo, Tinajero, & Almeida, 2007; Pascual-Leone, 1970). In the study exploring the relationship between Pascual-Leone's (1970) mental capacity model and Kahneman's (1973) unitary-resource mental effort model, Globerson (1983) used Witkin's cognitive style of field-dependence/independence to test the hypothesis that an inverted U-shaped function between measures of momentary mental effort and measures of FDI would be observed as was observed between the functional capacity and cognitive style in Pascual-Leone's model. According to Pascual-Leone (1970) the field-medium individuals normally have highly functional capacity because they are in the middle of the one-dimension continuous scale of FDI, solving problems based on both inner knowledge and outer information. On the contrary, individuals in the two extremes of the scale, either depend solely on inner knowledge or depend solely on outer information, and therefore expend less effort due to less interference.

Regarding the differences in strategy of using the limited mental resources, FD individuals are characterized by less effective resource control compared to immediate and FI individuals. Guisande et al. (2007) investigated cognitive functioning in 149 children (8-11 years olds) with different FDI cognitive styles (field-dependent, intermediate, and field-independent), including capacity to focus, shift, and maintain attention, capacity for sustained attention, storage capacity, and verbal working memory. The authors found that field-independent children did not show better storage capacity measured by digits forward test than field-dependent children, however, they performed significantly better than intermediate and FD children in all other tests, indicating that with the same available mental capacity, FI children are more effective in the use of control strategies and allocation of attentional resources because those tests require resource allocation strategies to maximize the outcomes to different extents (e.g., in the digits backward test, resource allocation between storage and processing needs to be managed efficiently).

There are very limited studies investigating populations of individuals with hearing loss with respect to cognitive styles. Blanton and Nunnally (1964) found male children who were deaf to be more field-independent than male children with normal hearing while female children who were deaf or normally hearing did not differ in performance on a cognitive style test. Fiebert (1967) assessed the effect of sex and the developmental differences on the cognitive style performance in children who were deaf. A total of 90 children at age levels of 12, 15 and 18 were tested. The cognitive styles were measured through 3 tests, Rod and Frame Test (RFT), the Children's Embedded Figures Test (CEFT) and the Poppelreuter Test (P-T). The cognitive style index was the sum of T scores on the 3 tests. Consistent with earlier finding in children with normal hearing, boys who were deaf were significantly more field independent than girls who were deaf. The developmental differences of increasing field independency with age was evident in boys but not in girls, the authors interpreted it as a phenomenon of identity crisis in the late adolescent girls who were deaf. Parasnis and Long (1979) evaluated the relationship between communication skills and field independency in 144 first-year deaf students with mean age of 20, and found that as a group, students who were deaf were more field-dependent compared to students with normal hearing. The communication skills were positively related to field-independency. Florence et al. (2012) carried out a study concerning the effects of teaching techniques and cognitive styles on students' achievement in science concepts in 65 secondary school students with moderate-to-severe hearing loss. The authors found that field independent students with hearing loss gained significantly more benefit from the problem-based learning method of teaching than the field neutral and field dependent participants. These studies have focused on a young population and none of them have assessed the association between the field dependence/independence dimensions of cognitive styles and the listening effort expenditure of adults with normal hearing and adults with hearing loss.

Comparing between the constructs of need for cognition and cognitive style (FDI) in terms of effort exertion, the former seems to be a both quantity and quality dimension in nature (i.e., the volume of effort and the pattern of effort allocation) whereas the latter is mainly a strategic dimension in nature. Cacioppo and Petty (1982) conducted a study to determine whether the Need for Cognition Scale was tapping a construct that was although related to, nevertheless distinguishable from cognitive style or field dependence. They administered both the Need for Cognition Scale and the Embedded Figures Test of field dependence to the same 419 subjects; as expected, a significant but small correlation between the two measures was found.

The previous listening effort literature has not included individual differences of personality in either dimension due to the limitation of the experimental design. For example, the participants usually performed the listening tasks with full motivation as instructed by the experimenters. Moreover, the laboratory experimental listening tasks that most studies used did not involve external environmental variations in addition to the task difficulty manipulation to evoke the individual's strategic effort allocation pattern. The present study will extend the extant literature and allow participants to mobilize effort based on their will, which makes it possible to investigate the relationship between the individual's motivation to expend cognitive effort and the actual effort expenditure when processing informational stimuli. For this purpose, the need for cognition is chosen to serve as the individual difference variable in this study.

1.2.4 Summary

In this section, the neurophysiological basis of effort and human behavioral models related to listening effort were reviewed. The individual differences that might account for listening effort variability were also briefly discussed. Research has reported a strong association between the anterior cingulate cortex (ACC) and effort exertion in both animals and humans, however, the functional connection between the ACC and auditory cortex in explaining listening effort control has not been intensively studied. The tasks used in effort-related ACC research were mainly cost-benefit decision-making tests with the assumption that effort is invested basically according to the judgment of the cost-benefit value. In contrast, the studies of effort-related human behavior in the cognitive hearing science field used speech tests such as speech recognition, holding the assumption that effort expenditure is generally determined by levels of task demands. Hockey's (1997) two-level compensatory effort control model accommodates the external stimulus-driven and subjective goal-driven facets of the effort regulatory system, and has pointed out the value of measuring the strategies that individuals adopt for completing a task. However, what is really lacking in the listening effort literature are studies in which task demand and environmental conditions are independently varied and alternative task strategies are provided to evoke effort allocation strategies and individual differences.

The purpose of listening in real life is to understand speech for successful communication, thus the process involves more than just the auditory functions of the periphery. More complex processes such as selectively attending to sound sources, storing and retrieving information from memory, using context information to improve understanding, resolving ambiguities, and generating appropriate responses quickly (Sarampalis et al., 2009) are also widely engaged and require effort. This implies a consideration of adopting a broader range of speech tasks in the study of listening effort.

In the area of listening effort research, individual differences have not been considered to any great extent. Studying individuals' auditory abilities, cognitive abilities and personality should allow the identification of the sources of individual differences in listening effort.

1.3 LISTENING EFFORT MEASUREMENT

Listening effort is considered an extra dimension that accounts for the disadvantages experienced by persons with hearing loss in daily life (Kramer et al., 1997). This non-audiologic factor is not assessable with traditional audiometric or psychoacoustic tests; however, it may be measured in physiological, behavioral, and subjective domains. In this section, the current techniques of listening effort measurement are reviewed in terms of their rationales, reliability and validity. An overview of the strengths and limitations of the different measurement techniques is outlined.

1.3.1 Physiological measurement

Physiological techniques are based on the assumption that changes in cognitive functioning are reflected by physiological variables (Mulder, 1986). Andreassi (2000) indicated that a normal subject's body is the field of various detectable reactions affecting breathing rate, blood pressure, heart rate, skin conductance or pupil diameter whenever she/he experiences mental effort. The techniques used in listening effort measures include dilation of pupil (Globerson, 1983; Kahneman, 1973; Kramer et al., 1997; Zekveld et al., 2010), skin conductance (Mackersie & Cones, 2011), skin temperature (Mackersie & Cones, 2011), heart rate (Mackersie & Cones, 2011), electromyography (EMG) (Mackersie & Cones, 2011), saliva cortisol concentration (Hicks & Tharpe, 2002), and auditory evoked potential (Bernarding et al., 2010; Bernarding, Latzel, et al., 2011; Bernarding, Strauss, et al., 2011; Okusa, Shiraishi, Kubo, & Nageishi, 1999; Strauss et al., 2010b). Physiological measurement has a particular advantage over the behavioral and subjective rating measurement in that it is measured "online" and continuously so that the fluctuation of expended effort while participants are listening to the stimuli can be demonstrated.

Kahneman (1973) proposed three criteria for any physiological indicator of mental effort: sensitive to within-task variation, between-task variation and between-individual differences. Among all the physiological measures of mental effort, pupil dilation appears to be the best indicator.

1.3.1.1 Pupil dilation

That the pupil of the eye dilates during mental activity has long been known in neurophysiology. Variation of the pupil diameter is a sensitive measure of the invested mental effort to a task (Beatty, 1982). Luria (1973) explained the connection between cognitive activities and autonomic physiological activities by pointing out that the structures maintaining the optimal levels of cortical waking state lie in reticular formation of the brainstem rather than the cortex itself. The efferent fibers leaving reticular structures typically bifurcate, one branch sends upward to higher nervous structures such as the thalamus, caudate body, and terminates in the neocortex; the other branch sends downward to synapse on a wide variety of motor nuclei such as Edinger-Westphal parasympathetic motor nuclei (Brodal, 1981). The higher nervous structures play a critical role in the formation of intensions and plans. Through the extensive corticoreticular connections, those structures are able to modulate the activities of the autonomic periphery systems of the reticular formation (e.g., pupil dilation) (Beatty, 1982).

There are two iris muscles controlling the pupil size. The pupillary sphincter results in pupil contraction and is regulated directly by the parasympathetic nervous system, whereas the pupillary dilator controls pupil’s dilation and ties directly into the sympathetic nervous system (Janisse, 1977). The size of the pupil can be influenced either by the activation of the ocular sympathetic system stimulating the radial dilator muscles, causing enlargement of the pupil, or by inhibiting the oculomoter parasympathetic system. The activation of the parasympathetic system will cause a decrease in pupil size. The size of the pupils at any given time reflects the balance of the sympathetic and parasympathetic systems (Loewenfeld, 1999). The basic responsibility of the sympathetic nervous system is making the task-related physiological adjustments that support performance, but more importantly, sympathetic activity increases especially when active coping and high task engagement are present during the cognitive tasks (Iani, Gopher, & Lavie, 2004).

The pupil size can not only reflect the arousal state like other autonomic systems; for example, the pupil dilation was observed to increase with increasing pure tone intensity level

(Nunnally, Knott, Duchnowski, & Parker, 1967) and broadband noise intensity level (Antikainen & Niemi, 1983), but can also index the cognitive and emotional processes such as mental effort (Beatty, 1982; Granholm & Steinhauer, 2004; Kahneman, 1973) and pleasure (Bradley, Miccoli, Escrig, & Lang, 2008). Beatty (1982) and Karatekin, Couperus, and Marcus (2004) suggested that the tonic change in pupil size is influenced by general factors such as arousal state, anxiety and stress, while the phasic changes in pupil response are time-locked to the onset of stimuli for cognitive processing, and they are independent of tonic changes.

The infrared video-based eye-tracking technology provides accurate assessment of pupil dilation. The pupil diameter can range from 1.5 to more than 9 millimeters in man (Goldwater, 1972). Systems commonly employed today can resolve better than .025 mm on diameter on individual measurement at a rate of up to 240 Hz (Granholm & Steinhauer, 2004; Granholm et al., 2007). The phasic task-evoked dilations normally onset between 100 and 200 milliseconds, peak at 2 to 3 seconds after stimulus onset and terminate rapidly after the completion of processing. Because of the delay of the pupillary reflexes, examining an extended time course is critical (Kuchinke, Võ, Hofmann, & Jacobs, 2007). As in evoked potential research, signal averaging and component extraction techniques are usually used to cancel out the background variations and to specify the event-related responses. Beatty and Lucero-Wagoner (2000) suggested reporting the mean pupil dilation, peak dilation and latency-to-peak for each time interval of interest, and they recommended reporting the pupil dilation in relation to the baseline in absolute values rather than in percentages because when the baseline pupil size is small, the percentage measures inflate the actual changes. It is important to distinguish between factors affecting the task-evoked pupillary response and those that affect the tonic or baseline pupillary diameter, such as light reflex, and between the phasic pupillary responses evoked by cognitive and by emotional processes (Goldwater, 1972). Those confounds can be minimized by using neutral signals, controlling the luminance of the room, requiring participants to maintain fixation and presenting the stimuli aurally.

Task-evoked pupillary responses have been validated in some within-task, between-task and between-individual listening effort experiments. In those studies, pupil size was found to vary as a function of processing demands. For example, Kramer et al. (1997) investigated the relationship between the pupillary response and the difficulty level of speech recognition in a noise task. A list of 13 everyday Dutch sentences was presented in fluctuating noise at various SNR levels referenced to individual's speech reception threshold, and participants were instructed to repeat each sentence as accurately as possible. The pupil diameter was measured during the course of listening. The results demonstrated that listening to the speech with high signal-to-noise ratio (SNR) results in a decrease in pupil dilation; however, the degree of decrement in favorable SNR conditions was more significant in normal-hearing than hearing-impaired listeners. The speech recognition task accuracy was at a similar level across the various SNR conditions between the two groups. The authors concluded that the hearing-impaired listeners need to expend more listening effort in the adverse SNR conditions than in the easy conditions and extra effort is required to achieve equal performance as their hearing peers at the same SNR level. Zekveld et al. (2010) used similar Dutch sentences and applied an adaptive procedure to estimate the SNR required for 50%, 71% and 84% intelligibility for each of 38 normal-hearing young adult listeners in a speech repetition task, during which the pupillary responses were measured. The results were consistent with Kramer et al. (1997) in that the pupil diameter enlarged and the latency to peak dilation prolonged with decreasing the speech intelligibility and the SNR of the speech in noise. In a later study, Zekveld, Kramer, and Festen

(2011) examined speech recognition in noise in 38 middle-aged normal-hearing adults and 36 middle-aged hearing-impaired adults, and confirmed that the pupil response decreased with increasing sentence intelligibility in hearing-impaired listeners but to a lesser degree than in normally-hearing individuals.

Koelewijn, Zekveld, Festen, and Kramer (2011) also found a similar association between the task-related pupil dilation and the listening effort in speech reception with various types of background noise in normal-hearing listeners. The speech with a single-talker masker resulted in larger pupil dilation compared to other types of noises due to the informational masking. A recent listening effort-related pupillometry study was conducted by Kuchinsky et al. (2013) in older hearing-impaired adult participants. The authors manipulated the listening difficulty in terms of acoustic features (i.e., easy SNR or difficult SNR) and word lexical competition (i.e., competitors with or without phonological overlap with the targets). The participants' task was to identify the word that they heard among the four orthographic options displayed on the computer screen in various conditions. The percent correct word identification, reaction time and the pupil size were measured. The results showed that the largest average pupil size, the most delayed peak and the most sustained pupil size following the peak was observed in the condition with difficult SNR and lexical competition. However, the word identification scores seemed more sensitive to the difficulty level of lexical competition only while the reaction time was more sensitive to the SNR manipulation. The authors suggested that the pupil response provides additive measures of task difficulty compared to behavioral measures alone.

Despite the evidence that the task-evoked pupil diameter systematically reflects the task-processing load in listening as a function of the difficulty level of a task, more evidence is needed to demonstrate that the pupil response not only can index the resource allocation directed by changes in task parameters (stimulus-driven), but also can signal the individual subjective goal-directed decision of listening effort investment to tasks, because the complexity of the stimulus is not the only determinant of listening effort, the individual's current performance level and self-appraisal of stress and comfort contribute to the regulation of effort as well (Fairclough, 2001; Hockey, 1997; Venables & Fairclough, 2009). Moreover, the researchers intentionally avoided too easy and too difficult task levels in order to maximize the correlation between the pupil response and the task demand, however, it actually limits the range to investigate the strategic listening effort allocation. Hence, the manipulation of other determinant factors (e.g., goals and performance feedback) of effort in addition to task difficulty and the manipulation range as dynamic as possible should be included in the study design to examine what the pupil response is most sensitive to.

1.3.1.2 Skin conductance, temperature, Heart rate, Cortisol level, EMG

Under the same general principle that the active engagement in effortful cognitive tasks typically results in increased activity in the endocrine system and the sympathetic branch of the autonomic nervous system (ANS) accompanied by decreased activity in the parasympathetic nervous system (Staal, 2004), some other measurable physiological changes in addition to the pupil dilation index the mental effort, including the skin conductance response (SCR), skin temperature, cardiac activity, cortisol level in saliva, electromyography (EMG) and evoked response potentials (ERP) (Andreassi, 2000). However, they are not as robust as pupil dilation measurement in the study of listening effort.

Skin conductance

The skin conductance refers to how well the skin conducts electricity when an external direct current of constant voltage is applied. It is one form of electrodermal activity and is associated with the activity of eccrine sweat glands innervated by sympathetic nerves (Figner & Murphy, 2010). Not only the hypothalamus and the brainstem (i.e., central origins of autonomic nervous system), but also the limbic networks are involved in the control of eccrine sweating (Dawson, Schell, & Filion, 2000). Skin conductance is often used in judgment and decision making research as an indicator of the involvement of affective and emotional processes (Mackersie & Cones, 2011; Naccache et al., 2005; Venables & Fairclough, 2009). For example, Naccache et al. (2005) investigated the subjects' conscious feeling of mental effort in the Stroop task and the Iowa gambling task using the skin conductance measurement.

The skin conductance response (SCR) is typically measured from the volar surfaces of the fingers, the palms of the hands, or soles and inner sides of the feet using a small constant voltage. Skin conductivity is revealed by the amount of current that passes between the electrodes. The SCR is measured in microsiemens (μS), and can be divided into tonic and phasic phenomena like pupil response. The phasic SCR consists of many discrete sharp peaks of skin conductance response which usually relate to a stimulus, whereas the tonic skin conductance refers to the spontaneous responses unrelated to a specific event. The most common measures of SCR are the onset latency (typically 1-3 seconds), the rise time (typically 1-3 seconds), and the peak amplitude. More recent studies indicate that the measure of area under the curve is a more valid indicator than any aspect alone (Naqvi & Bechara, 2006).

Cardiac activity

Cardiac activity has been used as a physiological measure to indicate changes in mental effort (Mulder, 1986). Cardiac activity can be measured in terms of heart rate, heart rate variability, and blood pressure. Heart rate is influenced by both the sympathetic and parasympathetic nervous system. When the sympathetic activation is dominant, the heart rate is found to respond to effort mobilization (Berntson, Cacioppo, & Quigley, 1993). Heart rate variability is the fluctuation in the time interval between consecutive heartbeats and usually illustrated as the power spectral density (PSD) as a function of frequency (Akselrod et al., 1981). Due to the relatively slow reaction of the sympathetic nervous system, mental effort is reflected mainly in the low frequency band (i.e., 0.04~0.15 Hz) of the heart rate variability which associates with short-term blood pressure regulation (Mulder, Van Roon, Veldman, Elgersma, & Mulder, 1995). In contrast, the high frequency band (i.e., 0.15~0.40 Hz) mainly reflects respiratory activity (Akselrod et al., 1981). Previous research showed that an increase in invested mental effort is related to an increase of heart rate (Gellatly & Meyer, 1992), and reported a strong reverse relationship between mental effort and heart rate variability in that as mental effort increases, the heart rate variability decreases in power within the low frequency band (Capa, Audiffren, & Ragot, 2008). Mukherjee, Yadav, Yung, Zajdel, and Oken (2011) examined the test-retest reliability of heart rate variability in measuring mental effort of 40 healthy seniors during the visual working memory tasks with various levels of difficulty. The authors found that a large number of heart rate variability parameters (time domain and frequency domain) were sensitive and reliable indices of mental effort.

Skin temperature

There is a growing investigation of the change of human facial skin temperature in relation to mental workload (Or & Duffy, 2007). Researchers have found that the skin temperature drops after the exposure to the task load during arithmetic tests and tracking tasks (Genno, Ishikawa, Kanbara, & Kikumoto, 1997; Ohsuga, Shimono, & Genno, 2001; Or & Duffy, 2007). According to Wallin (1981), the underlying mechanism is that mental load or negative emotion can result in activation of the sympathetic nervous system, and consequently induce peripheral metabolic responses such as vasoconstriction. The reduction of blood flow in the peripheral capillary vessels causes the decrease of skin temperature.

Skin temperature can be measured easily in the areas of nose, fingers and toes by thermistors or thermocouplers with physical contact and infrared camera (Genno et al., 1997). However, the problem with this measure is that it can be affected by the changes of the environment temperature. As a result, the difference between nose and forehead temperature is proposed as an index of mental strain because the temperature of the forehead skin is independent of taskload (Genno et al., 1997).

Electromyographic (EMG) activity

Another effort-related physiological index is the facial electromyographic (EMG) activity. The increase in facial EMG activity (e.g., the frontalis, the corrugator supercilii) is often manifest when exposed to physical or mental effort demanding tasks (Veldhuizen, Gaillard, & de Vries, 2003). It is considered as an expression of effort compensating for the decrement in performance efficiency caused by habituation, boredom, and fatigue (Van Boxtel & Jessurun, 1993; Waterink & van Boxtel, 1994). Waterink and van Boxtel (1994) asked 21 healthy college students to perform a visual two-choice serial reaction task (i.e., press the right bottom on green light and press the left bottom on red) with externally paced (EP) signal presentation rate, and 23 students with self-paced signal presentation rate. The EMG activity of frontalis, corrugator supercilii, orbicularis oculi, zygomaticus major, and anterior temporalis muscles were measured during the performance of the task. A group of participants showed a decline in performance with time and their corresponding EMG amplitudes displayed an inverted U shape pattern in which an initial EMG increase transferred into a decrease over time. However, the EMG amplitudes of the other group of participants who maintained stable performance until the end of the experiment increased uninterruptedly. This result suggested that the EMG responses of specific facial muscles are reliable and sensitive to mental effort mobilization.

There has been few applications of skin conductance, cardiac activity, skin temperature and EMG in listening effort research. Mackersie and Cones (2011) included these four physiological measures in their listening effort study to determine the physiological and emotional cost of maintaining performance near ceiling level (i.e., higher than 96% correct for all listening tasks) in normal-hearing adults, and the relations between psychophysiological measures, performance measures and subjective measures of listening effort as well. Participants were asked to repeat all of the digits they heard in a Dichotic Digits Test with various demand level. The information of heart rate, skin conductance, skin temperature, electromyography, percentage correct and subjective rating of effort was collected. Results showed that the mean recognition scores were close to 100% as expected, and the subjective ratings of listening effort systematically increased as the task difficulty increased. The significant task demand effect was found only in EMG activity and skin conductance with a monotonic increase in the skin conductance and EMG activity as the task became more difficult, however, the heart rate and skin temperature remained little changed across test conditions. The authors pointed out the importance of individual differences as participants varied in their patterns of ANS reactivity. When averaged across task sessions (medium- and high-demand), no significant change in EMG activity was found compared to the baseline reference (i.e., low-demand), whereas 60% of the participants showed an increase in skin conductance. It was concluded that among the four physiological measures, the skin conductance was most sensitive to detect the task demands and the individual changes.

Cortisol level in saliva

The change of cortisol in saliva is part of a physiological stress reaction. Increased cortisol level is associated with perceived stress, whereas low cortisol level is correlated with fatigue or burnout (Hellhammer, Wüst, & Kudielka, 2009; Pruessner, Hellhammer, & Kirschbaum, 1999). The cortisol level is controlled by the hypothalamic-pituitary-adrenocorticol (HPA) axis. The hypothalamus produces corticotropin-releasing factor which leads to the production of adrenocorticotropic hormone (ACTH) by the pituitary gland. It is the ACTH that causes the adrenal glands to release cortisol (Nemeroff, 1998). In high-stress situations, stressors trigger this hormone chain, resulting in detectable changes of cortisol level in saliva. The salivary cortisol is a useful physiological assessment in stress literature, however, the reliability and validity are still under investigation (Hellhammer et al., 2009)

Hicks and Tharpe (2002) have looked for fatigue and listening effort in children with normal-hearing and hearing loss by sampling cortisol concentrations in saliva and dual-task paradigm, respectively. The saliva was sampled twice a day (morning and afternoon) for two days from 14 normal-hearing children and 14 children with mild-to-moderate hearing loss. The results showed no significant difference on the cortisol level between two groups. The cortisol level was higher in the morning samples than in the afternoon samples for both groups. In the listening effort experiment, the same groups of children were asked to perform a speech-recognition test in various signal-to-noise ratios (primary task) and respond to a random probe light simultaneously. They were motivated to pay primary attention to the word recognition task. For all SNR conditions, children with hearing loss had greater reaction time change from baseline in the secondary task than children with normal hearing, indicating that children with hearing loss expended extra effort in listening. This study failed to demonstrate a greater decrease in cortisol level in hearing-impaired children throughout a school day than normal-hearing children as an indicator of greater fatigue or stress experienced by children with hearing loss. The authors argued that the salivary cortisol levels might not be sensitive enough to capture changes in listening-related fatigue or stress. As the listening effort was not directly measured by salivary cortisol level in this study, the correlation between the cortisol level and listening effort remains unknown.

1.3.1.3 Event-related Potentials (ERPs)

The electrophysiological evaluation of high level auditory processing as indices of listening effort is receiving growing interest. The most frequently measured components of late ERPs are N100, P200, MMN, N2b and P300. The N100 and P200 peaks are normally evoked by both target and non-target stimuli in the oddball paradigm, whereas the MMN, N2b and P300 only occur following the target stimuli (Okusa et al., 1999). The MMN is regarded as a fully automatic cerebral response to a deviant stimulus, and the N2b is assumed to correlate with the controlled mismatch detection process or effortful processing (Näätänen & Alho, 1997). These auditory late responses can be visually identified within the following latency ranges: N100 as the most negative peak in the time interval of 80-160ms; P200 as the most positive peak in 140-230ms; MMN, 150-200ms; N2b, 190-350ms, and P3 as the most positive peak in the range of 290-520ms (Bernarding et al., 2010; Okusa et al., 1999).

Okusa et al. (1999) recorded the cognitive event-related potentials (ERPs) from participants with cochlear implantation to determine the extent of the listening effort required in a 2-tone discrimination task with various levels of difficulty. The participants were instructed to lift the thumb whenever they heard the target tone. The ERP recordings showed that the peak latency and amplitude of N100 and P200 did not change as a function of task difficulty. However, the N2b and P300 latencies became significantly longer as the target stimuli become closer to the non-target stimuli, and the P300 amplitude also became significantly smaller. The MMN was also delayed when the task became more difficult, but the MMN was absent in the most difficult condition. These results indicated that the late components of ERPs such as N2b and P300 can reflect task demand and might be used to assess the listening effort in hearing-impaired population.

Strauss et al. (2010b) proposed a computational transformation of the auditory late responses entitled Wavelet Phase Synchronization Stability (WPSS) in order to extract information associated with effortful endogenous modulation from the raw individual ERP sweeps, and suggested that the increased listening effort can be reflected in an increased WPSS of auditory late response sequences. The WPSS trace generated from the data sets of difficult listening conditions was larger than from the data sets of easy listening conditions for both people with normal hearing and hearing loss, especially in the N100 and P200 wave temporal range (Bernarding et al., 2010; Bernarding, Strauss, et al., 2011; Strauss et al., 2010a). Despite the robustness of the WPSS in illustrating listening effort within task (various demanding) and between tasks (tonal and syllable paradigms), the evidence between subjects is lacking. The complexity of the computational equation of WPSS has made this measure of listening effort less feasible in the clinical settings. In addition, the critical requirement of the response mode has limited the experimental design in listening effort studies, and the visual inspection of the WPSS trace for final analysis compromises the objectivity of the listening effort measurement just as much as the raw ERP data do.

Polich (2007) described the P300 as a fundamental attribute of CNS reactivity associated with attention and memory operations. The P300 amplitude and latency is sensitive to task demand and individual differences in cognitive capacity. The greater amount of attentional resources that a task requires, the P300 amplitude becomes smaller and the peak latency becomes longer (Kok, 2001; Okusa et al., 1999; Polich, 2007). Other late ERPs such as N400 and P600 has been studied to show high association to language comprehension, in which the N400 has been correlated with lexico-semantic processes while the P600 was assimilated with syntactic reanalysis or integration under a sentence correctness judgment task (Kotz, von Cramon, & Friederici, 2005). Although there are only a few ERP studies investigating listening effort, given the relationship between the auditory late responses and language comprehension, it seems plausible that the late components of evoked response potentials may be used to indicate listening effort.

1.3.1.4 Brain imaging technique

Among the popular neuroimaging techniques today, such as Functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) and Magnetoencephalography (MEG), the fMRI is most widely used. This technique detects the increase in blood oxygen levels when fresh blood is brought to a particular area of the brain due to the neuronal activity change following an event. The event-related activation is usually determined in terms of the average blood-oxygenation-level-dependent (BOLD) signal intensity (i.e., the percentage change in the image signal between the stimulus and silence blocks) (Barth & Poser, 2011; Hwang, Li, Wu, Chen, & Liu, 2007). The fMRI technique allows for detailed maps of brain areas underlying human mental activities in health and disease, and has the advantages of high spatial resolution and noninvasiveness. However, the most serious challenge faced in auditory fMRI studies is the interference of the acoustic noise of the gradient coil oscillation caused by rapid switching during echo-planar imaging with the presentation of auditory stimuli (Ravicz, Melcher, & Kiang, 2000).

There have been some attempts to minimize the impact of the acoustic noise associated with echo-planar imaging. Hall et al. (1999; 2000) suggested a sparse imaging procedure in which the repetition time of a sequence is longer than its acquisition time, allowing the auditory stimuli to be presented during the silent period between the acquisitions of consecutive volumes. The disadvantage is that within a limited scanning time, fewer images are acquired in the sparse imaging paradigm compared to the conventional continuous paradigm. Another approach to reduce the scanner noise is to change its qualitative nature by implementing a quasi-continuous gradient switching pattern in order to emit a continuous noise rather than a pulsed noises during scanning, which has less effective stimulation to the auditory cortex (Seifritz et al., 2006). However, this technique leaves the issue of energetic masking. Schmitter et al. (2008) developed a quiet echo-planar imaging sequence in which a constant phase encoding gradient and a sinusoidal readout echo train are used to change the gradient switching frequency to one that is associated with low acoustic response, therefore reducing the acoustic noise by approximately 20dBA.

Peelle, Eason, Schmitter, Schwarzbauer, and Davis (2010) investigated the effect of echo-planar imaging sequences (standard, sparse and quiet) on brain activity in a sentence listening task. The authors hypothesized that an additional neural activity would be observed in the most adverse condition (standard sequence). Six normal-hearing young adults were asked to judge if a probe word was semantically related to the sentence they heard. The fMRI results showed significantly more speech-related activity in left temporal cortex as well as left inferior parietal cortex when using the standard echo-planar imaging sequence than when using the sparse and quiet sequence while maintaining high level of performance accuracy ($\geq$ 93%), which indicated an extra effort invested in listening in the low signal-to-ratio situation. Davis and

Johnsrude (2003) implemented three different acoustic manipulations to distort sentence stimuli at various intelligibility levels and collected imaging data from 12 healthy young adults during a speech repetition task. They found that elevated activity in large portions of left temporal, prefrontal and premotor cortices was associated with listening to the more degraded sentences, providing evidence that listeners expended increased effort to extract information from degraded signal.

It is notable that the measures of listening effort in audiology imaging studies are different from that in other fields such as psychology (as reviewed in the earlier section). The former illustrate effort in terms of the increased activation in the sensory-related brain areas as a function of task demand and define effort as quantitative signal processing. However, the latter literature demonstrates effort in terms of activity in the specific effort-related area (i.e., ACC) and considers effort as a general decision-making process regardless of modality. Further research is needed to validate the application of fMRI in the study of listening effort.

Table 3.The strengths and limitations of psychophysiological measurements in listening effort assessment.

	Variables	Strengths	Limitations
Pupil dilation	Mean pupil dilation Peak dilation Latency to the peak Response duration	• Noninvasive • Fast • Easy and convenient to operate • Reliable • Not interfere with mental tasks • Able to reflect the variations in the cognitive demands evoked by different tasks and variations in the perceived taskload between individuals • Highly sensitive to subtle fluctuations in listening effort • Can be measured continuously • Less expensive than fMRI and EEG.	➢ Delay due to the time lag involved in autonomic responses ➢ May be confounded by illumination and eye movement (involuntary looking or blinking) ➢ Exclude some data due to distortion ➢ Subjective judgment bias when measure peak and baseline in traces
Skin conductance	Peak amplitude Onset latency Rise time Recovery half time Spontaneous SCRs per time unit	• Noninvasive • Relatively cheap • Can be measured continuously • Can be measured unobtrusively and reliably	➢ Slow and time-lagged ➢ Need a number of repetitions ➢ May prolong the experiment due to using relatively long interstimulus intervals
Skin temperature	Temperature in degrees Fahrenheit	• Noninvasive • Easy • Relatively cheap	➢ Slow ➢ Confounded by environment temperature ➢ Relatively low reliability in effort measurement
Cardiac activity	Heart rate Heart rate variability	• Noninvasive • Easy • Fast • Reliable	➢ Potential effect of breathing pattern on heart rate variability ➢ Relatively low reliability in effort measurement

Table 3 (continued)

EMG activity	Peak amplitude Onset latency	• Noninvasive • Fast • Reliable • Can be measured continuously	➢ Artifact ➢ Bias of visual inspection
Salivary cortisol	Cortisol concentration	• Noninvasive • Easy • Fast • Purely reflect energy mobilization, relatively independent of behavioral or cognitive conditioning • Relatively cheap	➢ Mainly used as indices of fatigue rather than effort ➢ May be affected by the presence of food or drink in the mouth ➢ Need to consider variables such as estrogens (gender, menstrual cycle, oral contraceptives) or medical conditions could affect cortisol binding and hypothalamic-pituitary-adrenal (HPA) responsivity.
Event-related response (ERP)	Amplitude and latency of N100, P200, N2b, MMN, P300, N400, P600	• Noninvasive • Easy • Fast	• Artifact • Bias of visual inspection • Require complex computation on raw data to obtain reliable effects
fMRI	The blood oxygenation level dependent (BOLD)	• The brain areas can be identified in a functional manner • Reliable	➢ Slow and time-lagged ➢ Significant acoustic noise which interferes with the presentation of auditory stimuli ➢ Weak temporal resolution ➢ The activity as measured by fMRI simply is not sufficiently fine-grained to pick up certain effects ➢ Insensitive to changes in the low or medium level of invested effort

1.3.2 Task- and performance-based behavioral measurement

Listening effort has been studied using single and dual-task procedures. This research typically finds decrements in individuals' performance in the most difficult conditions. The performance is usually presented in terms of percent correct and/or response time, indicating how accurate and/or how fast participants accomplish the task.

1.3.2.1 Single-task paradigm

To measure listening effort using a single task, researchers tend to choose a task that involves extra operations in addition to speech perception, such as encoding speech into memory. This is based on the limited resource capacity theory of Kahneman (1973) which suggested that multiple tasks performed concurrently or in close sequence compete for limited resources. The more resources demanded by a particular task, the fewer resources will be available for the other tasks, resulting in a compromised performance in the latter. McCoy et al. (2005) compared the performance accuracy of 12 normal-hearing and 12 hearing-impaired older adults in a running memory task, which required participants to recall the final three words of the presented sentences with various orders of word contextual constraint (i.e., 0-order means none of the word is constrained by the remaining words in the sentence, 1^{st}-order means each word is constrained by the prior word, and 2^{nd}- and higher order means the word is constrained by at least two prior words in the sentence). The two groups of listeners were matched for age, education, verbal ability and recall performance for the last word of three-word recall sets. The results showed that both groups reached near ceiling recall for the first two words of three-word recall sets of 2^{nd}- and higher order of contextual constraint sentences; however, a significant difference on the recall performance between groups was observed in 0- and 1^{st}-order sentences, with the hearing-impaired group having much lower percent correct score than the normal-hearing group. The authors concluded that the hearing-impaired listeners must expend extra effort on perception that might otherwise be available for encoding speech to memory, and the effect of contextual constraint indicated that hearing-impaired listeners used top-down processing to compensate for the degraded input.

The similar performance decline as an indicator of listening effort also were observed in other studies. Rabbitt (1990) reported decreased accuracy in a word list recall task performed by hearing-impaired participants when noise was added. Comparable results were found in normal-hearing listeners as well when the listening environment became more challenging (Sarampalis et al., 2009). Stewart and Wingfield (2009) found that participants with hearing loss displayed a compromised speech comprehension at difficult levels of syntactic complexity compared to those with normal hearing due to the extra listening effort expended on perceiving the sentences.

However, there is a theoretical issue with interpretation of the observed decrements in task performance, because they could be caused by either data limitation or resource limitation. The term data limitation refers to changes in task difficulty that cannot be compensated for by increased effort, while the resource limitation refers to deficits due to insufficient effort investment that may result from low motivation or concurrent involvement with other tasks (Navon & Gopher, 1979a; Norman & Bobrow, 1975). These two types of constraints exist despite the fact that participants supposedly invest all their effort in task performance. Gopher (1994) suggested that studies of difficulty manipulations and changes in task demands that use exclusively task performance measures are unable to distinguish between these two causes of performance decrements. Alternatively, the combined use of psychophysiological measures and performance measures may allow researchers to distinguish between these two determinants of performance decrements (Iani et al., 2004).

There is also evidence that the addition of response-time measures may be more sensitive in reflecting effort than accuracy measures alone (Apoux, Crouzet, & Lorenzi, 2001; Gatehouse & Gordon, 1990; Kuchinsky et al., 2013; Piolat et al., 2008). For example, Apoux et al. (2001) found that the envelope expansion in vowel-consonant-vowel syllables did not improve the identification scores in groups with normal hearing and hearing loss; however, it significantly decreased the response time in both groups, suggesting an ease of listening improvement. However, the response time measure is also subject to the data and resource constraints, and the interpretation of individual data may be complicated by time-order effects (Larsby, Hallgren, & Lyxell, 2008; Mackersie, Neuman, & Levitt, 1999). For example, Mackersie et al. (1999) reported systematic changes in response time (gradually increasing or decreasing) across a test session for some participants.

Self-paced listening technique described byFerreira, Anes, and Horine (1996), also referred to as auditory moving window technique, has been used to explore allocation of attentional resources while listening. This technique allows listeners to control the flow of the information at their own rate by pressing a key when they are ready to hear the next segment (e.g., word, clause) of the speech message. It is presumed that if a listener requires more time (i.e., more effort) to process a particular segment, he or she will exhibit a prolonged pause before initiating the next segment. A recall request or a comprehension question often follows the presentation of the speech message. Piquado, Benichov, Brownell, and Wingfield (2012) studied the narrative recall accuracy of listeners with normal hearing and hearing loss under a conventional continuous speech presentation and a self-paced presentation condition. They found that listeners with hearing loss recalled the narratives significantly less accurately than normal-hearing controls in the continuous presentation condition, however, the difference was eliminated in the self-paced condition. The results suggested that processing time can be a valid variable to index listening effort. Although the self-paced listening technique is mainly applied in the second language learning literature and rarely used in audiology research, it is worth incorporating the technique to investigate listening effort in the population with hearing loss with respect to the quantity and the individual allocation strategy of effort in terms of the absolute processing time and the pattern of the pause duration across segments, respectively. It is also feasible and valuable to obtain the information about efficiency of listening effort through the use of the self-paced technique.

1.3.2.2 Dual-task paradigm

The application of the dual-task paradigm or double stimulation procedures in the audiology field is based on Kahneman's (1973) limited capacity theory. The classic dual-task design holds the following assumptions: ⑴ Humans have a certain limited capacity to process information; ⑵ When asked to divide their attention in a dual-task paradigm, individuals presumably use the majority of mental capacity in the primary task and use the spare mental capacity in the secondary task; ⑶ When the primary task is made more difficult, less spare capacity remains for completion of the secondary task, thus hindering performance on the secondary task; ⑷ Decrements in the performance of one task are taken as an indicator of processing load incurred by a second, concurrently performed task.

Table 4. Studies of listening effort using dual-task paradigm

<table>
<tr><th colspan="2" rowspan="3"></th><th colspan="7">Secondary task</th></tr>
<tr><th colspan="4">Visual</th><th>Auditory</th><th>Tactile</th><th>Combination</th></tr>
<tr><th>Arrow_number matching</th><th>Probe light</th><th>Serial digit recall</th><th>Target tracking</th><th>Random probe</th><th>Pattern recognition test</th><th>Dot-to-dot game</th></tr>
<tr><th rowspan="6">Primary task</th><td>Notetaking and writing in L1 and L2</td><td></td><td></td><td></td><td></td><td>Piolat(2008)</td><td></td><td></td></tr>
<tr><td>Audio Sentence recognition</td><td></td><td></td><td></td><td>Desjardins (2011)
Desjardins (2014)</td><td></td><td>Gosselin,& Gagné(2011)
Fraser (2010)
Fraser (2007)</td><td></td></tr>
<tr><td>Audiovisual speech recognition in noise</td><td></td><td></td><td></td><td></td><td></td><td>Gosselin & Gagne (2011)
Fraser (2010)
Fraser (2007)</td><td></td></tr>
<tr><td>Word recognition</td><td>Sarampalis (2009)</td><td>Hicks (2002)
Downs (1982)
Downs (1978)
Feuerstein (1992)</td><td>Howard (2010)
Choi (2008)
Stelmachowicz (2007)</td><td>Broadbent (1958)
Picou (2013)</td><td></td><td></td><td>McFadden & Pittman (2008)</td></tr>
<tr><td>Word list recall</td><td></td><td></td><td></td><td>Tun (2009)</td><td></td><td></td><td></td></tr>
<tr><td>Speech comprehension</td><td></td><td></td><td>Rakerd (1996)</td><td></td><td></td><td></td><td></td></tr>
</table>

There are three broad categories of dual-task procedures in the literature (McNeil et al., 2004). The simplest method is the single-to-dual task comparison, which requires that the task(s) of interest be performed in isolation and concurrently with a secondary task. This is the method that has been applied in the listening effort research. The primary task is typically speech recognition, speech comprehension or word recall, combined with a variety of tests across modalities as a secondary task (see Table 4). Participants are asked to perform two tasks alone first and then perform both of them concurrently. Presumably, in single task conditions, all attention can be devoted to either task, whereas in dual task conditions, attention must be divided between the two tasks. The priority is assigned to the primary task in a dual-task paradigm by instructing participants to shadow the primary task or using tangible payoffs with larger rewards for the primary task. Participants receive reward for the secondary task only if they have performed the primary task adequately. Listening effort is usually computed using the formula based on the performance of the secondary task $Listening\ effort = (Single_{2nd} - Dual\ task_{2nd})/Single_{2nd}$ (Desjardins & Doherty, 2012; Kemper, Schmalzried, Hoffman, & Herman, 2010), or $Listening\ effort = Single_{2nd} - Dual\ task_{2nd}$ (Tun et al., 2009). The second dual-task procedure is the voluntary effort allocation method in which participants perform two concurrent tasks with explicit instructions to vary their allocation of effort between them according to the experimental condition (Slansky & McNeil, 1997). The third procedure is called concurrent task difficulty manipulation method, in which participants perform two concurrent tasks and the difficulty (or some other parameter) of each task is systematically and independently manipulated (McNeil et al., 2004; Wickens, 2002)

Although the dual-task paradigm is accepted to be a relatively sensitive and objective way to reveal the hidden behavioral dimension of effort in theory (Hafter, 2010) and might potentially permit the comparison of different tasks in common units by devising a standard subsidiary-task, researchers must apply it with caution. Firstly, according to Kahneman (1973), there are two types of interference between concurrently performed tasks, capacity interference and structural interference. Capacity interference arises as a function of the attentional demands of competing activities, whereas the structural interference occurs when the activities occupy the same mechanisms. The interference could take place not only when the total taskload exceeds one's capacity, but also when it is far below the total capacity, and even when the stimuli presentation modes (or response modes) are not in the same sensory modality (Pashler, 1992). The extent of interference will depend in part on the load which each of the activities imposes (Kahneman, 1973). The structural interference appears to be a confounding factor in studies that attempted to measure capacity interference which is the case in listening effort studies.

Secondly, the premise of using the secondary task performance to index listening effort is that the primary task performance should maintain at a certain level across experimental conditions, because listeners presumably shift their effort from the secondary task to the primary task to prevent a decrease in the primary task performance (Kerr, 1973). Some studies reported the stability of the primary task performance when the difficulty of the primary task varied (Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008; Desjardins & Doherty, 2012; Downs, 1982; Downs & Crum, 1978; Sarampalis et al., 2009; Tun et al., 2009), while others did not (Feuerstein, 1992; Hicks & Tharpe, 2002; McFadden & Pittman, 2008). It is also necessary to look at the accuracy of the secondary task when measuring response time. Researchers normally included only the data from correctly performed trials into analysis, which might lose some important information about effort exertion strategy. Variation of performance in the primary task and/or secondary task will suffer from the uncertainty between data limitation and resource limitation in performance results interpretation (Norman & Bobrow, 1975).

Thirdly, investigating effort allocation strategy is challenging in the dual-task paradigm. The invested effort is constrained to the two experimental simultaneous tasks with assigned priority order, which does not commonly happen in real life situations. Individuals' own decision of listening effort employment and the pattern of effort allocation among concurrent and/or sequential tasks with various levels of difficulty are topics that researchers are most interested in but cannot easily analyze in the dual-task paradigm.

Lastly, the dual-task paradigm requires high cooperation from subjects which might results in a disparity between the performance of children and that of adults (McFadden & Pittman, 2008). The nature of the crossmodal design in this paradigm to some extent moves the focus away from the auditory modality (Strauss et al., 2009b).

1.3.3 Subjective rating

Another approach to quantify listening effort is subjective rating. Rating scale techniques are based on the assumption that people are able to introspect on their cognitive processes and to give a numerical indication of mental effort expended. The rating scale can be independent unidimensional scales which only address specific questions about listening effort. It also can be a portion of some multidimensional questionnaires which assess groups of associated variables, such as mental effort, task difficulty, work load and stress. Various self-assessment scales, such as Speech, Spatial and Qualities of Hearing Scale (SSQ) (Gatehouse & Noble, 2004), the NASA Task Load Index (Hart & Staveland, 1988) and the Visual Analog Scale (Rudner et al., 2012), have been used experimentally. Participants generally rate their perceived listening effort on a scale immediately after the completion of a trial, a condition or the whole task. Such scales are sensitive to relatively small differences in cognitive load and are reliable. More details about the rating scales applied in the listening effort literature are displayed in Table 5.

Table 5. Self-assessment instrumentations for listening effort measurement

	Measured construct	Sample question and scale
Fraser et al. (2010)	Listening effort	"How effortful was it to identify the components of the sentence" 0% (no effort) ~ 100% (very effortful)
Downs and Crum (1978)	Task difficulty	Rate the difficulty of the learning task 1 (Very easy) ~ 7 (Very difficult)
Larsby et al. (2005)	Perceived effort	Rate the degree of effort perceived during the listening task 0 (None at all) ~ 10 (Extremely great)
Gosselin and Gagné (2011b)	Listening effort	"How much effort was required for you to identify the components of the sentence?" 0 (Negligible amount of effort) ~100 (High degree of effort)
Picou et al. (2011)	Listening effort	Verbally rate listening effort 0 (No effort) ~ 10 (Most effort)
Zekveld et al. (2010)	Listening effort	Rate listening effort 1 (No effort) ~ 9 (Very high effort)
Rudner et al. (2012)	Listening effort	Rate listening effort via a Visual Analog Scale (VAS) 0cm (No effort) ~ 11.7cm (Maximum possible effort)
Brons, Houben, and Dreschler (2012)	Listening effort	Rate listening effort (5-point) 1 (Extremely high effort) ~ 9 (No effort)
Mackersie and Cones (2011)	Listening effort	"How hard did you have to work to accomplish your level of performance?" 0 (Lowest effort) ~ 10 (Highest effort)
Luts et al. (2010)	Listening effort	Rate listening effort (13-point) 0 (No effort) ~6 (Extreme effort)
McAuliffe et al. (2012)	Listening effort	Rate listening effort via a Visual Analog Scale (VAS) 0cm (Minimal effort) ~ 10cm (Maximum effort)

Although a self-report measure of effort scale gives insight into the experienced listening effort, there are at least three limitations to this approach. First, the subjective rating measurement appears to be affected by both task difficulty (Gopher, 1994) and compensatory control (Hockey, 1997). However, the experimental design rarely takes the compensatory control factor into account, thus it is uncertain whether the subjective indication of effort truly reflects the availability or demand on processing resources (Wickens, 1992). Second, there are inter-individual differences in discriminating between questionnaire item dimensions (e.g., distinguishing task demand from effort investment) and in decision criteria for rating, which makes the interpretation of subjective effort ratings complex (Recarte, Pérez, Conchillo, & Nunes, 2008; Yeh & Wickens, 1984; Zekveld et al., 2010). Third, the subjective rating is usually conducted after task completion. It is likely that individuals may not accurately recall the perceived effort during the performance. As a result, their responses will reflect the average of perceived effort across many trials rather than momentary effort (Kuchinsky et al., 2013; Kuchinsky, Eckert, & Dubno, 2011). In addition, some common factors such as age, cognitive abilities and IQ influence the subjective effort ratings (Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004; Zekveld, Kramer, Kessens, Vlaming, & Houtgast, 2009).

1.3.4 Summary

Researchers have been exploring techniques to answer the question of what sort of test should be included to capture multiple dimensions of listening effort. Three categories of measurements have been proposed: physiological measurement, performance-based measurement and self-reported measurement. Only a few studies have combined two or three techniques of the categories in a single experiment (Fraser et al., 2010; Gosselin & Gagné, 2011a, 2011b; Hicks &

Tharpe, 2002; Mackersie & Cones, 2011; McAuliffe et al., 2012; Picou et al., 2011; Zekveld et al., 2010), among which little evidence supports the strong relationship between the measurements. Only two studies found significant correlation. Fraser et al. (2010) reported significant negative correlation between the effort rating score and the actual accuracy of the secondary tactile task in audio-visual (AV) modality, and Mackersie and Cones (2011) found weak but statistically significant relation between physiologic changes in skin conductance and subjective ratings of perceived effort measured by NASA-TLX.

The lack of relationship between these measurements supports that the autonomic responses of body, the speed and accuracy of behavioral performance and the self-reported perception do not yield redundant information about listening effort. It also can be viewed as evidence that the human is not merely a passive difficulty detector. Hence, a well-designed study should incorporate these complementary sources and reveal a full picture of listening effort. It should be able to measure both the volume (i.e., amount) and pattern (i.e., distribution or strategy) of effort allocation, the effectiveness (i.e., quality of performance) and the efficiency (i.e., the relation between the quality of performance and the effort invested in it) of effort, the momentary (i.e., effort at any instant) and total (e.g., overall in terms of area under curve) effort. The combination of pupil dilation assessment, self-paced listening test with some compensatory options and a detailed subjective assessment of listening effort (i.e., tackle multidimensional listening effort and differentiate the perceived effort in real life environment from lab environment) would address the various dimensions of effort.

1.4 RESEARCH QUESTIONS AND HYPOTHESES

Based on the literature review of listening effort in audiology and related fields, mounting evidence from neurophysiological and behavioral studies compel audiology researchers to reach out further and systematically investigate the interaction effect between external stimulus-driven and subjective goal-driven factors on listening effort exertion. The present study was the first to systematically examine the effect of the two factors. Prior to applying the compensatory control model to potential populations of interest (e.g., people with hearing loss), it was important to establish performance and quantify the interaction effect between external stimulus-driven and subjective goal-driven factors on listening effort exertion in a young, normal hearing population.

The primary aim of the current study was to investigate whether listening effort is driven by task demand (speech rate) only, goal (reward points) only, or by both task demand and goal. The specific research question was, to what extent is the stimulus-driven pattern of listening effort allocation in speech comprehension modulated by reward condition under time pressure? It was hypothesized that the listening effort allocation pattern among the speech rate levels would be modulated by the reward condition. Individuals with normal hearing were hypothesized to vary their listening effort based on the taskload/reward net value of the stimuli, meaning that people tend to spend more listening effort on the low demand-high benefit items.

The secondary aim of the proposed study was exploratory in nature and was to investigate the efficiency of listeners' effort allocation in a complex cognitive task within limited time. The effort efficiency is defined in the present study as the ability to achieve the target reward using a minimum amount of effort with the same consumption of time. The aim was to compare two computational approaches of efficiency calculation in characterizing individual differences. One approach was proposed by Paas and Van Merriënboer (1993), $Efficiency =$

$\frac{Z(performance)-Z(mental\ effort)}{\sqrt{2}}$, and the other was proposed by the current study, $Efficiency = \frac{Z(reward)-Z(mental\ effort)}{Z(time)+K}$. As the two equations incorperate different elements, the specific question was, which one of the two computational approaches characterizes the effort efficiency and differentiates individuals better? It was hypothesized that the new equation would be superior to the original equation with the inclusion of the reward and time elements.

Answering these questions will specify some critical components that should be accounted for in the theoretical frameworks of listening effort, as well as in the experimental design, especially when studying listening effort in a population with hearing loss. The results of this study may help to explain various degrees of perceived fatigue from listening. Furthermore, the previous research has not established how the effort efficiency should be characterized. The detailed information about listening effort efficiency may help in tailoring individualized amplification solutions and auditory rehabilitation programs, which would aim at reducing the demands of the environment and maximizing the abilities of the individual, respectively.

2.0 METHOD

2.1 EXPERIMENTAL DESIGN

This study involved a 5×5 within-subjects Latin Square experimental design, with an independent variable of speech rate (5 levels) and an independent variable of reward point (5 levels). In order to control for both order effect and carryover effect, a digram balanced Latin square table was used to create the test conditions of the speech rate and reward point combination. As the level of independent variables was an odd number, two Latin square tables were generated so that each reward point level was followed by a different point level with the same probability, as shown in Table 7. Both tables needed to be used in the experiment, therefore, participants were randomly assigned to receive the testing condition sequence following one of the two tables.

Table 6. The diagram balanced latin square design.

A						B					
	Slow	**Normal**	**Slightly fast**	**Fast**	**Extremely fast**		**Slow**	**Normal**	**Slightly fast**	**Fast**	**Extremely fast**
Block 1	1 point	9 points	3 points	7 points	5 points	**Block 1**	5 points	7 points	3 points	9 points	1 point
Block 2	3 points	1 point	5 points	9 points	7 points	**Block 2**	7 points	9 points	5 points	1 point	3 points
Block 3	5 points	3 points	7 points	1 point	9 points	**Block 3**	9 points	1 point	7 points	3 points	5 points
Block 4	7 points	5 points	9 points	3 points	1 point	**Block 4**	1 point	3 points	9 points	5 points	7 points
Block 5	9 points	7 points	1 point	5 points	3 points	**Block 5**	3 points	5 points	1 point	7 points	9 points

The 25 combinations of speech rate and reward point were presented in a block manner. Each block consisted of all 5 levels of speech rate and the different reward points assigned to every speech rate level. There were 10 speech stimuli in each combination and therefore 50 speech stimuli in one block. Breaks were given upon the completion of every block. The speech stimuli were questions about the spatial relationship between 3 objects. Participants were asked to answer where object X is in relation to the object Y by pressing the appropriate key (left, right, front, or back) on a keypad. Participants must answer the question correctly in order to get the corresponding incentive points, and their task goal was to earn the full reward payment upon the completion of the experiment within a limited time period. Their pupil dilation were tracked as an indicator of listening effort throughout the whole experiment except during breaks. It was hypothesized that participants will manage their listening effort based on the cost-benefit decision.

2.2 PARTICIPANTS

Forty healthy young native American English speaking individuals (33 females) aged 19 to 32 (M=22.43, SD=2.39 years) were recruited from the campus at the University of Pittsburgh. They signed a statement of informed consent that had been approved by the University of Pittsburgh's Institutional Review Board and were given a copy of the signed statement. All participants were screened for hearing, vision and auditory processing abilities.

Hearing was screened following a typical clinic audiometry procedure (ASHA, 2005). All participants met the normal bilateral hearing sensitivity criteria defined as having pure-tone air-conduction thresholds less than 20 dB HL at audiometric test frequencies 250 Hz through 8000 Hz in both ears (ANSI, 2004). All participants had a visual acuity of 20/20 or better when tested under the binocular condition without correction using the standard Snellen chart (Bailey & Lovie, 1980). Participants' auditory processing capabilities were documented by the Computerized Revised Token Test (CRTT) (Eberwein et al., 2007; Heilman, 2008; McNeil et al., 2009; Turkyılmaz & Belgin, 2012). The CRTT is a diagnostic test used to evaluate auditory processing and comprehension abilities, auditory attention, auditory memory and temporal processing (McNeil et al., 2015). In the test, participants are required to follow auditory commands varying in sentence length and complexity and to identify and manipulate objects of standardized shapes (circles and squares), colors (black, red, blue, white and green) and sizes (big and small) displayed on a computer screen by either using a computer mouse or touchscreen. Studies have shown no significant difference between the two response modes (Heilman, 2008; McNeil et al., 2009). The CRTT subtest VII and VIII were used in this study. The CRTT speech commands were prerecorded using natural voice with average rate of 3.0–3.5 syllables per second (approximately 178 wpm)(McNeil et al., 2015). These subtests were chosen because the command type is directional prepositional and the sentences are relatively long, which is most similar to the speech test in the present study. An example of speech commands in the CRTT subtests is “Put the little green circle to the left of the big red square”, participants are expected to select and drag one object to the other by using a computer mouse. A score of 15 is the maximum score that a person can achieve on any part of the test, Participants in this study had a mean score of 13.72 (range, 12.49-14.58) for subtest VII and a mean score of 14.42 (range, 13.39-14.90) for subtest VIII.

Participants were monetarily compensated for their participation due to the incorporation of motivation in terms of incentives (i.e., certain reward points redeem a certain amount of money) in the study design, This was a one-visit experiment, with each participant performing five 20-minute sessions with 5-minute breaks in between. Including screening and instructions, total time was approximately 2.5 hours.

2.3 STIMULI

Recording

A set of 432 speech sentences were developed for the study. They were variants of a reasoning question about the spatial relationship between two of three objects (i.e., Stone, Kite, Ball). The three objects were selected from word lists of the Northwestern University Auditory Test No. 6 (NU-6) (Wilson, Coley, Haenel, & Browning, 1976), and they were phonetically different from each other. The sentences contained two parts, one was the description of the spatial relationship among the three objects, and the other was a question of one object’s location in relation to another object (e.g., “The kite is on the left of the ball, the stone is on the right of the ball, where is the kite in relation to the stone?”).

There were 6 different permutations of the three objects array in the left-right direction, and 6 different descriptions of each array, and 6 combinations of questions about the spatial relationship between 2 objects out of 3. This generated 216 sentences ($6 \times 6 \times 6$). And another 216 sentences were formed in the same way in the front-back direction. An example of one permutation of the left-right array and front-back array is given in Figure 1.

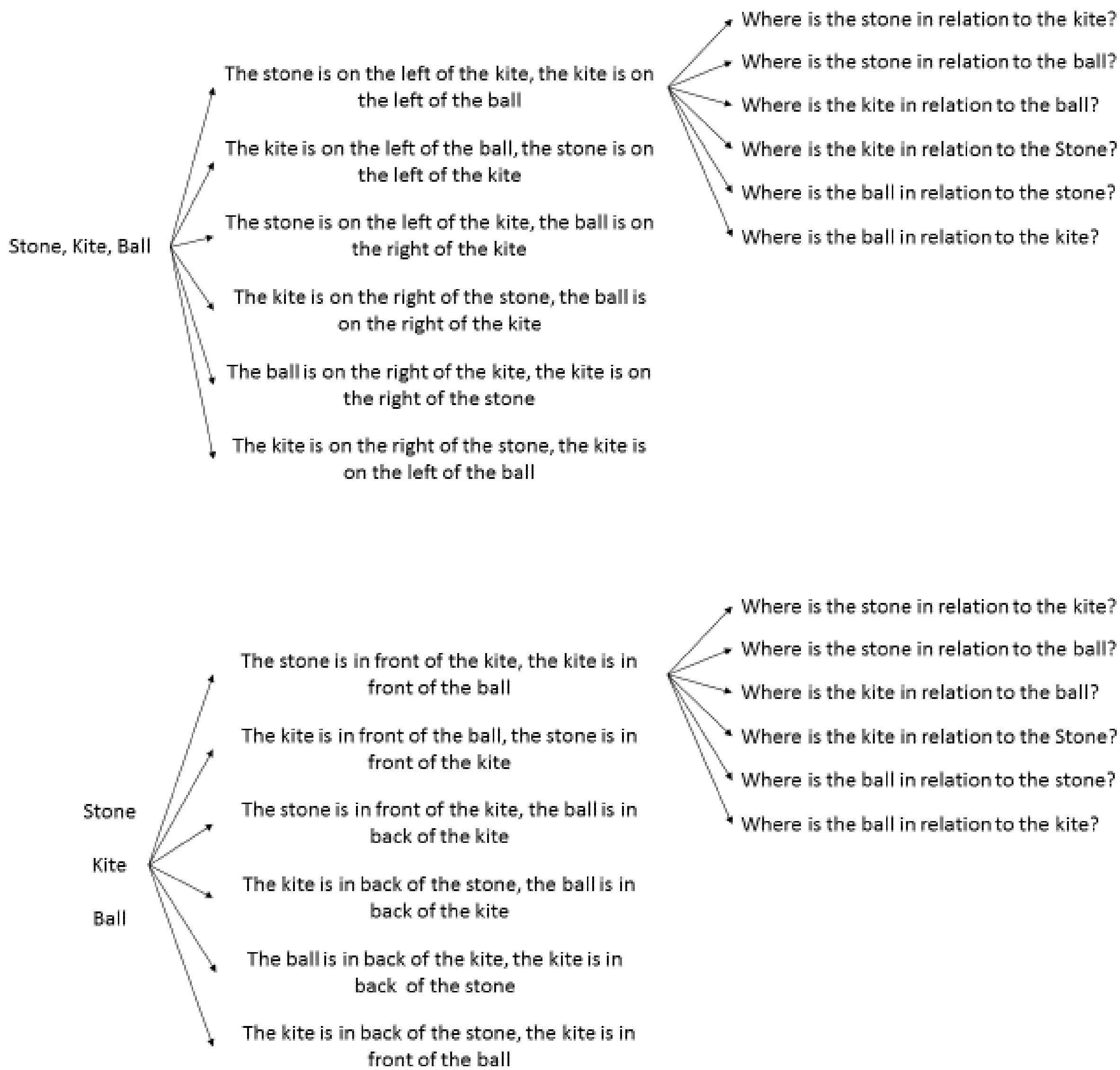


Figure 1. An example of speech sentence stimuli.

The advantages of using these sentences are, first, they have identical lexical and syntactic structure. The only component that varies is the relative location of one object with respect to the others. The 6 different descriptions of one object array creates differences in iconicity (i.e., the conceived similarity or analogy between the sequential order of the objects and that described in the speech chain) (Haiman, 1980), as well as differences in answer priming (i.e., whether the answer is repeated in the sentence); which result in various difficulty levels of speech understanding. This had been resolved by equating the sentence distribution across test conditions. Second, these sentences require a fairly high level of cognitive processing even at a slow rate due to their low predictability. Participants had to listen through the end of the sentence to be able to solve the question because each question was random, which facilitates participants' engagement in the task. Third, these sentences allow participants to respond by pushing the appropriate key on the keypad instead of repeating what they hear, which evaluates listening comprehension. In addition, participants' performance (e.g., correct rate, error rate, response time) can be accurately recorded, minimizing the errors made by the raters commonly required in speech repetition tasks.

A male talker and a female talker were asked to read the 432 sentences three times producing maximum intelligibility, at a very slow rate that might be directed toward the hard of hearing or a non-native English speaker, at their normal conversational speech rate and as fast as they could respectively. The purpose of recording the original speech at 3 rates was explained in the following speech rate manipulation section. Both talkers were asked to read the sentences without exaggerated emphasis on rate, pitch and intensity, even though the last part of the sentence is a question.

The speech stimuli were recorded at 44.1 kHz sampling rate through stereo channels in a sound-treated chamber on a Marantz solid state digital recorder (PMD 670) with the Shure SM48 dynamic microphone. The ambient noise in the sound-treated room was less than 25 dB A. The talkers were required to read at an intensity level that passed the midline of the VU meter on the recorder display, and the microphone was set at 25 cm from the talker's lips. A separate recording for the cue words which indicate the speech rate and reward points (i.e., slow, normal, slightly fast, fast, extremely fast, 1 points, 3 points, 5 points, 7 points, 9 points) were obtained under the same setting.

A cue phrase (i.e., speech rate + points) proceeded each sentence in the experiment, indicating the net value of a specific sentence. The net value was defined as the level of reward divided by the difficulty level of the task, which in the present study is the speech rate ($Net\ value = \frac{Reward_{(points)}}{Speech\ rate_{(words\ per\ minute)}}$). This net value computation has been used in neurophysiology and neuropsychology studies of effort in both human and animals (Apps & Ramnani, 2014; Botvinick, Huffstetler, & McGuire, 2009; Croxson et al., 2009; Kennerley, Dahmubed, Lara, & Wallis, 2009; Kennerley et al., 2006).

The full-version of a speech stimulus token comprised a net value cue phrase, a 2.5 seconds decision time period, a 0.5 second 1000Hz pure tone signaling the onset of the sentence and the sentence. An example of a speech stimulus token is shown in Figure 2.

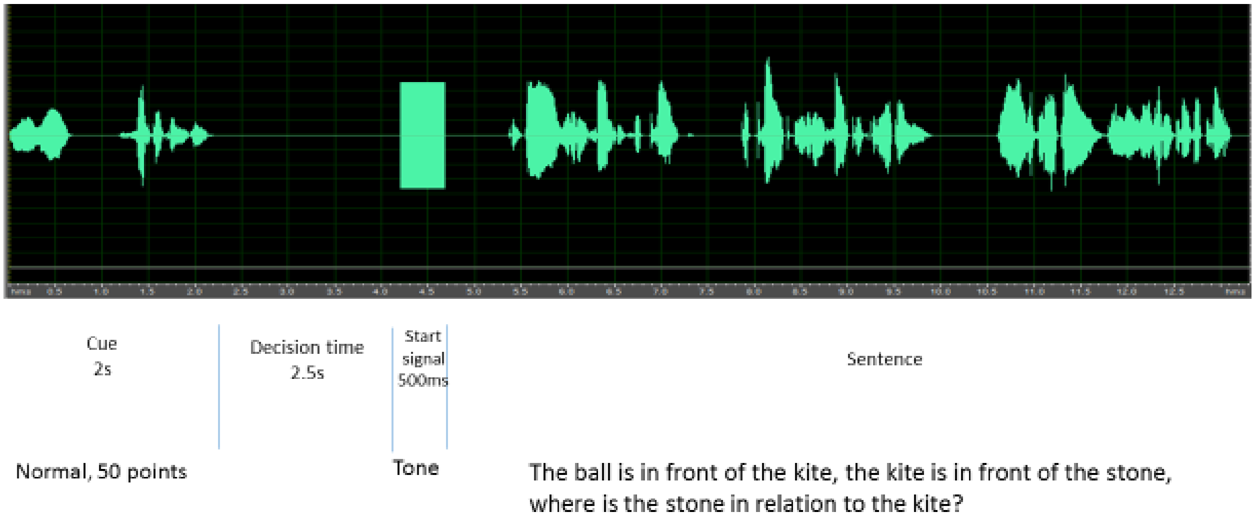


Figure 2. An example of a complete speech token used in the experiment.

Speech rate manipulation

The waveforms and spectrums of the two talkers' recordings were compared in a digital audio editing program, Adobe Audition 3.0. The speech intelligibility and the quality of neutral tone maintenance were judged by a group of 10 native English speakers. The criteria for speech intelligibility was that all the sentences could accurately be repeated by all the raters. The criteria for neutral tone was that every word in the sentence had the same amount of stress and each sentence was perceived as non-emotional. These restrictions minimize the possible confounds to the task demand manipulation. The speech from the two talkers was highly intelligible, however, the male talker kept the neutral tone significantly better than the female talker. Additionally, the speech rates and the speech-to-pause ratio at 3 levels (i.e., slow, conversational and the fastest) were calculated, see Table 7. The speech rate was computed as the number of words in the sentence divided by the duration of the sentence in the unit of words per minute (wpm). Some studies have used syllables per second (syll/s) as the unit of speech rate; however, it doesn't make a difference in the present study since all sentences consist of identical words.

Table 7. Speech rate and speech-to-pause ratio comparison between two talkers

		Slow	Conversational	Fastest
Male talker	Speech rate (wpm)	163	236	324
	Speech-to-pause ratio	4.23	4.18	6.463
Female talker	Speech rate (wpm)	172	219	352
	Speech-to-pause ratio	4.75	5.93	11.82

According to the literature, speech rates vary tremendously among normal speakers, but in general, adults produce an average of 270 wpm during conversational speech and 160-180 wpm during oral reading (Calvert & Silverman, 1983). The speaking rate of normal adults ranges from 220-410 wpm during uninterrupted discourse (Weiner, 1984). When speaking fast, a talker unintentionally changes relative attributes of his/her speech such as pause durations, consonant-vowel duration, etc. The speech-to-pause ratio is defined as the duration of speech segments divided by that of pause segments in the present study. The pause boundaries were identified automatically by the audio editing program. As the distinction between hesitation or resting pauses and phonetic events (e.g., voiceless stops, articulation shifts) can be difficult even in healthy speakers, researchers have typically adopted a pause threshold around 250ms as proposed by Eisler (1968) and Niimi and Nishio (2001). Therefore, a 250ms pause duration threshold was used in this study.

When linearly regressing the speech-to-pause ratio against speech rate for the speech of the two talkers, as shown in Figure 3, the relationship between the speech-to-pause ratio and speech rate of the female talker seemed to follow the linear path better than the male did. Considering the speech intelligibility, neutral tone control and speech rate/speech-to-pause control, the male talker's voice was used but the female's linear equation (i.e., $y = 0.0403x - 2.4758$) was applied to generate the target speech rates variables for the experiment.

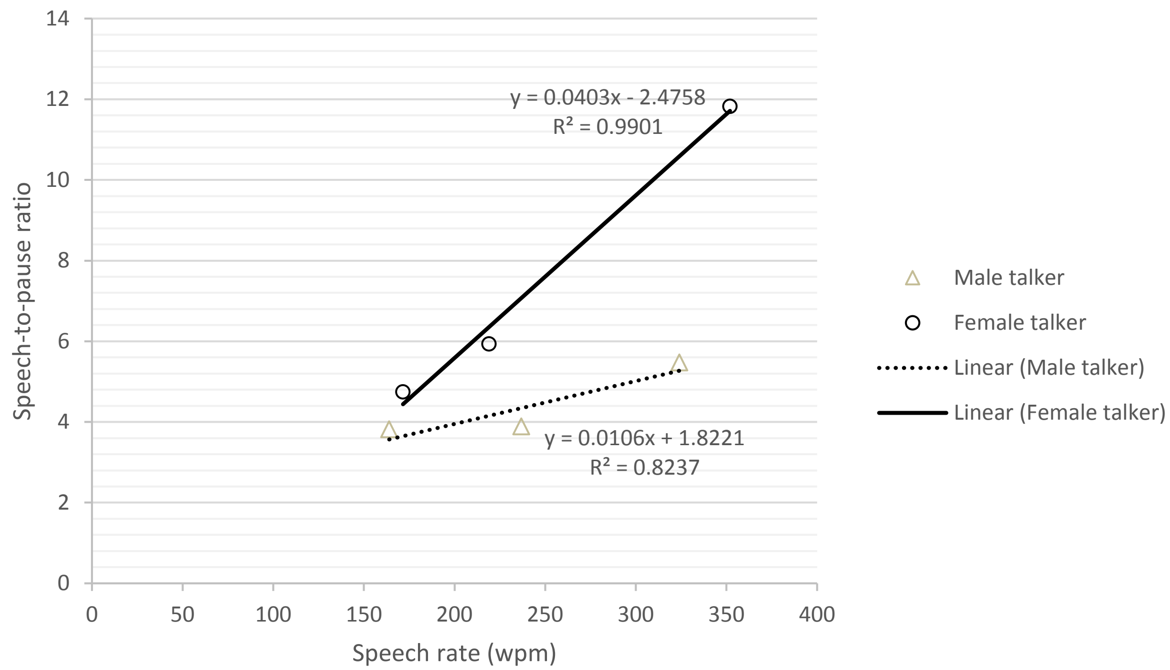


Figure 3. Speech-to-pause ratio as a function of speech rate

The range of the speech rate variables used in the present study were intended to cover the normal speech rate range indicated in the literature and a level beyond the upper limit of the normal range. The purpose for a wide range was to investigate participants' decision on their listening effort exertion not only when task demands were within their capability, but also when task demands exceed their ability. A preliminary experiment was conducted to determine the exact levels of speech rate that were used in the main experiment.

In the preliminary experiment, seven levels of speech rate were generated from the original recording, they were 130wpm, 180wpm, 230wpm, 280wpm, 330wpm, 380 wpm and 430wpm. According to the literature, as the time-compression ratio of the speech signal increases, there is a degradation of speech recognition performance (Wilson, Preece, Salamon, Sperry, & Bornstein, 1994). The time-compression ratio specifies the percentage reduction in the total duration of the original speech sample. For example, in the case of an entire sentence, 40% time compression results in a sentence in which the duration is 40% less than the duration of the original sentence. Wilson found that compression ratios above 50% substantially affected speech perception. These findings are consistent with earlier findings that recognition performance is most affected by compression ratios above 60% (Beasley & Freeman, 1977; Beasley & Maki, 1976). In order to minimize the adverse impact of time compression on the speech intelligibility, samples of 130wpm and 180wpm were generated from the original slow recording, sample of 230wpm and 280wpm were generated from the original conversational recording, and sample of 330wpm, 380wpm and 430wpm were generated from the original fastest recording. As a result, none of the time-compression ratios exceeded 50%, see Table 9.

Table 8. Target speech rates and their time-compression ratios for the preliminary experiment.

	Slow		Conversational		Fastest		
Original recording	163 wpm		236 wpm		324 wpm		
Target speech rate	130 wpm	180 wpm	230 wpm	280 wpm	330 wpm	380 wpm	430 wpm
Compression ratio	25.4%	9.4%	2.6%	15.7%	1.8%	14.7%	24.7%

To manipulate speech rate, the durations of the 7 target speech rate levels were calculated first. It was computed as $\text{Duration (second)} = {Number\ of\ words}/{Speech\ rate\ (wpm)} \times 60$. The sentences regarding the left-right direction consist of 27 words, and the sentences regarding the front-back direction consist of 25 words, therefore, the sentence duration were slightly different between the two sets of 216 sentences. The correspondent 7 target speech-to-pause ratios were identified from the regression line of the female talker's speech, as displayed in Figure 4. Given the appropriate durations and speech-to-pause ratios (see Table 10), the 7 levels of speech rate were then manipulated through the audio editing program (i.e., Adobe Audition 3.0).

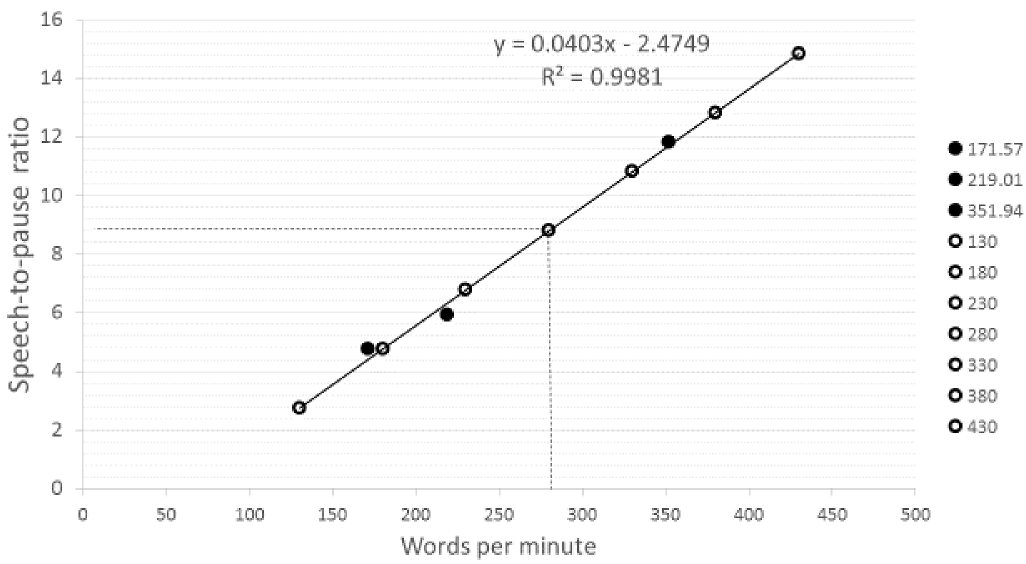


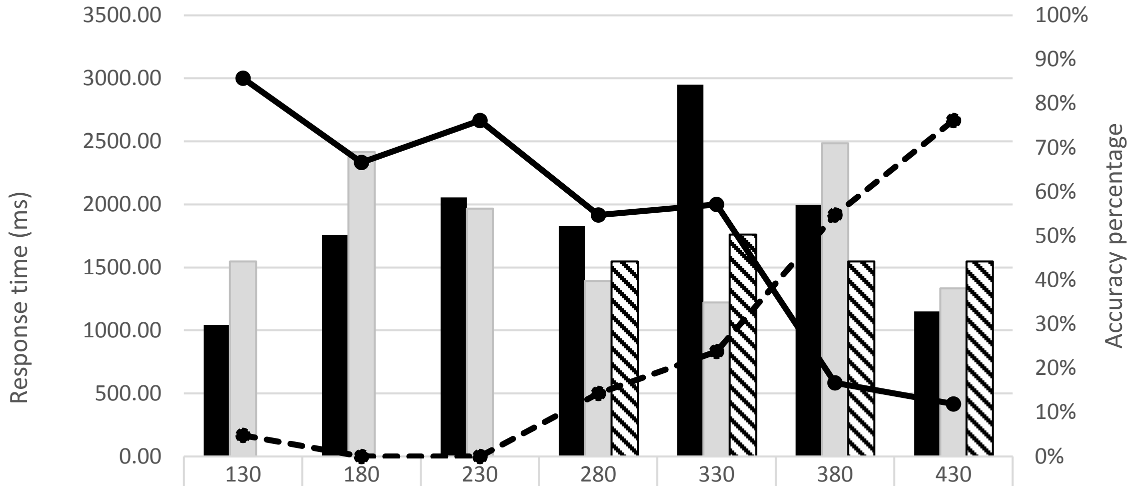
Figure 4. The 7 speech-to-pause ratios corresponding to the 7 target speech rates.

Table 9. Target sentence durations and speech-to-pause ratios

	130 wpm	**180 wpm**	**230 wpm**	**280 wpm**	**330 wpm**	**380 wpm**	**430 wpm**
27 words duration (Left-right)	12.46s	9.00s	7.04s	5.79s	4.91s	4.26s	3.77s
25 words duration (Front-back)	11.54s	8.33s	6.52s	5.36s	4.55s	3.95s	3.49s
Speech-to-Pause ratio (from linear function)	2.76	4.78	6.79	8.81	10.82	12.84	14.85

In order to minimize the contamination to the speech caused by processing artifacts, the pauses in the sentences were adjusted in duration to achieve the target speech-to-pause ratios before Adobe Audition uniformly stretched or compressed the whole sentence. A pause (or silence) was defined as a segment in the waveform that had an intensity lower than 10dB for at least 250ms, and could be automatically marked by Adobe Audition. Based on the measured speech duration and pause duration within a sentence, precisely calculated silent segments were either added to or removed from the original pause segments so that the speech-to-pause ratio reached the target value. Once the speech-to-pause ratio was set, the processed sentence was then time-compressed or time-expanded by Adobe Audition without changing the pitch. The processing algorithm was based on the pitch synchronous overlap and add method (PSOLA) (Kawahara, Masuda-Katsuse, & De Cheveigne, 1999; Moulines & Laroche, 1995). First, the input waveform was decomposed into a stream of short-time signals based on pitchsynchronous marks. Second, the pitch-synchronous short-time signal was either eliminated or duplicated based on the predefined stretch factor. Third, the modified short-time signal was added to synthesize the stretched and compressed stimulus. The original pitch was preserved during processing and the duration of each voiced or silent segment in the speech was uniformly changed (Liu & Zeng, 2006). The intensity of each sentence was finally equated to 65dB SPL in root mean square (RMS) across all the speech stimuli.

A total of 42 sentences (6 sentences in each speech rate level) randomly selected from the 432 sentences pool were used in the preliminary experiment. The reward condition was not included in the preliminary experiment because the purpose was to determine which five levels of speech rate would to be used in the main experiment. Eight young adults were asked to answer the questions as accurately and fast as possible by pushing the appropriate key on the keypad. They were allowed to push the middle key on the keypad to skip to the next sentence if it was too difficult for them to answer. The results of accuracy, skip rate, response time for correct items, incorrect items and skipped items are presented in Figure 5. The bar graph reads the response time on the left Y-axis, and the curve graph reads percentage correct on the right Y-axis. The accuracy decreased as the speech rate increased, and there was a rapid decrease in accuracy at 380wpm, this was also where the response time for the correct items started to decrease, and the skipping rate at 380 wpm was fairly high. All these factors made the level of 380 wpm valuable to be included in the main experiment. The highest rate 430 wpm was clearly an extreme condition, with the lowest accuracy, the highest skipping rate and the shortest response time for all items. The rate of 430 wpm was included to serve as a condition where no listening effort is expected to be allocated to regardless of reward points. The remaining 5 levels of speech rate seemed to be within participants' processing capability. The response time for the correct items increased with the increase of the speech rate, reaching the maximum at 330 wpm. The accuracy declined with the increase of the speech rate but had been maintained at a fairly good level, Skipping items occurred at 280 wpm and 330wpm with low percentage. These performance patterns seen in the range of 130 ~330 wpm were consistent with the prediction of Kahneman's unitary resource theory (1973) and the results of numerous studies investigating listening effort within cognitive capabilities. Three out of the 5 levels were therefore selected to represent the pattern, and these are 130 wpm, 230 wpm and 330wpm.



	130	180	230	280	330	380	430
RT	1044.18	1759.35	2054.97	1826.43	2949.40	1994.71	1150.38
RT (incorrect)	1546.36	2415.88	1965.79	1390.96	1223.02	2484.87	1334.14
RT (skip)	0	0	0	1548	1760	1548	1547.29
Accuracy	86%	67%	76%	55%	57%	17%	12%
% of skip	5%	0%	0%	14%	24%	55%	76%

Speech rate (wpm)

Figure 5. Results of the preliminary experiment for speech rate range determination purpose.

2.4 PROCEDURE

2.4.1 Screening and descriptive tests

Participants first completed a brief medical and audiologic history questionnaire (see Appendix A). A vision test then was conducted using the Snellen chart, and the eligibility criteria was a visual acuity of 20/20 or above without correction tested binocularly. An otoscopic examination was performed prior to the audiometric evaluation to ensure that there was no occluding cerumen preventing visualization of the eardrum. Individuals who had occluding cerumen were referred to an audiologist. Participants' hearing sensitivity was screened using a Beltone 2000 clinical audiometer following the ASHA guidelines (ASHA, 2005). Participants were eligible for the study if they could respond to signals at 20dB HL at octave frequencies from 250 Hz-8000 Hz in both ears tested separately (ANSI, 2004). The subtests VII and VIII of the Computerized Revised Token Test (CRTT) were administrated through the experimental computer. Participants listened to the CRTT speech commands through a loudspeaker at 65dB SPL calibrated at seating position.

The short form of the Need for Cognitive (NC) test was administered to the eligible participants. Although the NC scale was originally developed as a 34-item inventory, the most commonly used version contains 18 statements that people rate on 5-point scales to reflect how characteristic the statement is of themselves (Cacioppo & Petty, 1984). The scale has high internal consistency and test-retest reliability. Participants were instructed to indicate, on a 5-point Likert-type scale, the degree to which each of the 18 items characterized them (see Appendix B).

.

2.4.2 Main experiment

All of the procedures of the main experiment took place in a 1.5x2.2m double-walled sound-treated booth that meets specifications for maximum permissible ambient noise levels (ANSI, 2003). The booth was illuminated with medium lighting with a background luminance of 300lx.

The experiment setup was presented in Figure 6. Participants were seated at a table in front of the computer monitor (size of the display: 19 in., display resolution: 1024×768) at a distance of approximately 32 inches (81.28cm) with a five-key response keypad in front of them. The experiment instructions and the fixation image (for the purpose of pupil dilation measurement) were displayed on this monitor. Speech stimulus presentation was controlled by a computer running the experimental control software, SuperLab 5 (Cedrus, Phoenix, Arizona). Speech signals were played out via a SoundBlaster soundcard on the experimental computer, and subsequently fed to the Beltone audiometer. Participants received the signals via the audiometer's ER-3A insertion earphones. The calibration speech-shaped noise for the experimental stimuli was generated by Praat program and matches the Long Term Average Speech Spectrum (LTASS) of the speech produced by the male speaker whose speech was recorded for this experiment. The sound intensity level calibration was performed using a Larson-Davis 824 sound level meter to ensure that the output from the left and right ER-3A insert earphones was 65 dB SPL.

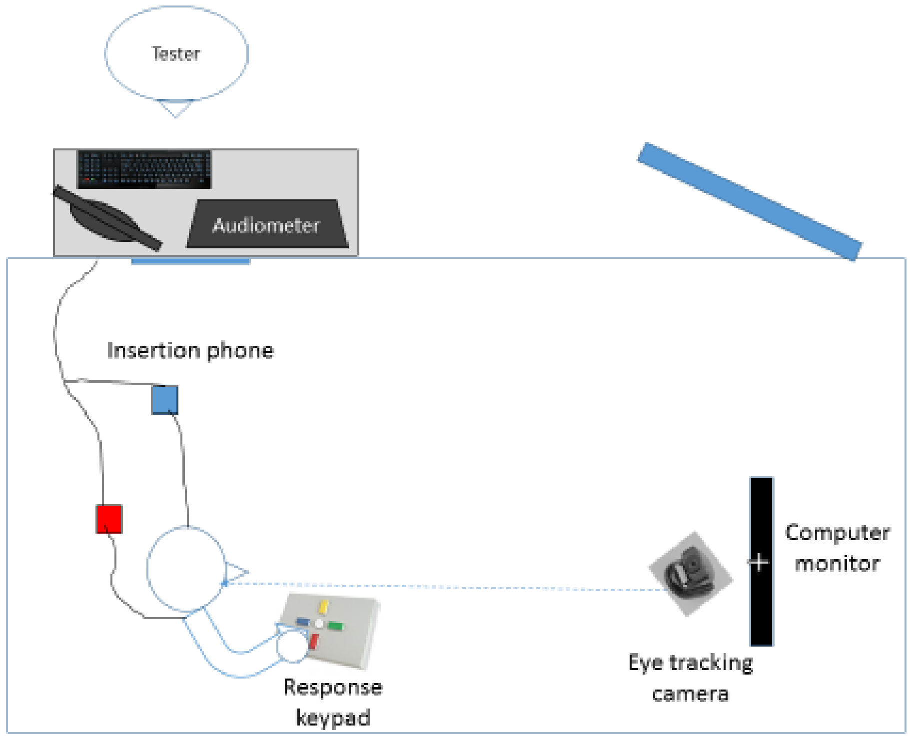


Figure 6. Experiment setup.

Prior to the experiment, participants performed a sequence of practice trials to ensure they were familiarized with the response keypad and the speech stimuli, and that they understood the task and directions. A list of practice trials along with the specific instructions that were both read to the participants and displayed on the screen is shown in Table 11. Fifty-five sentences randomly selected from the 432 speech stimuli developed for this study were used in the practice trials. Participants were allowed to repeat the trials until they correctly followed all of the commands in the practice trials.

Table 10. Practice trials objectives, contents and instruction to participants.

Trial ID	Objective	Content	Instruction
1	• Be familiar with the response keypad.	Commands in text on the screen (e.g., "Please press the left key".)	Please follow the commands by pressing the appropriate key on the keypad without looking at it.
2	• Be familiar with the speech stimuli. • Ensure that the speech stimuli are intelligibility to the participants. • Be used to fixate their gazes on the cross display on the screen while listening.	A total of 15 speech stimuli (3 for each of 5 speech rates) were given through the ER-3 insertion phone in an order from slow to fast. (e.g., The stone is in front of the kite, the kite is in front of the ball, where is the kite in relation to the ball?)	You will hear some sentences with a questions at the end, and the speech rate will vary from slow, normal, slightly fast, and fast to extremely fast. You don't need to answer the question, only repeat the question part of the sentence. Please look at the cross image on the screen all the time.
3	• Ensure that participants understand the speech • Be familiar with the "skipping key" on the keypad	A different set of 15 speech stimuli are given, similar to the trial 2.	You will hear similar sentences to the ones you heard before, instead of repeating, please answer the question by pressing the appropriate key on the keypad. You can press the middle key (namely "skipping key") on the keypad when you decide not to answer the question. The next sentence will appear automatically. Please look at the cross image on the screen all the time.

Table 10. (continued)

4	• Be familiar with the reward conditions. • Be familiar with the full-version structure of the stimuli.	A list of 25 full-version stimuli are given in a random order. (e.g., "Slow, 7 points (0.5 seconds) + pause (2.5 seconds) +1000Hz tone (0.5 seconds) + stimuli sentence")	This is a simulation of the real experiment, in which each correctly answered question will earn you certain points, and each point is worth 5 cents, your goal is to earn as many points as possible. In each full-version stimuli, you will be told how fast the upcoming sentence is, and how many points you can earn if corrected answered, then you'll have 2.5 seconds to make decision of how much effort you are going to put in it. A tone will occur at the end of the pause signaling the beginning of the sentence. After the sentence offset, an answer prompt (a low frequency tone) will indicate that you can push the appropriate button to respond. Again, this is only a practice and is meant to prepare yourself for the real experiment. Don't forget to look at the cross image on the screen all the time.

In the formal experiment, a total of 250 sentences (5 speech rate levels × 5 reward levels × 10) were selected from the remaining 377 speech stimuli (excluding the 55 used in the practice trials) such that each speech rate×reward combination had 10 sentences with the same distribution of the sentence structure variants. The 250 stimuli were presented in 5 blocks of 50 sentences (see Table 7) with breaks in-between blocks. The order of the 5 speech rate×reward combinations presentation within each block was random, whereas the presentation order of the 10 stimuli within each combination condition remained fixed. Participants were randomly assigned to receive one of the testing condition sequences.

Participants were informed both verbally and in writing (shown on the computer screen) that:" ***There are 5 blocks in total (the last practice was half block). Please remember, each correctly answered question will earn you certain points (1, 3, 5, 7 or 9), and each point is worth 5 cents. The maximum points you can possibly get is 3700, but you only need 1000 points to get the full payment ($50). Your goal is to EARN the FULL PAYMENT within a limited time. A clock ticking sound will appear when you have spent half of your time in each block. Your response won't be registered until the sentence presentation is complete. Don't forget to look at the cross image on the screen all the time.***"

Participants were told that they only had a limited time to finish each block, and the time limit was set individually based on the amount of time they spent on the 4th practice trial. As the 4th practice trial has half the number of the sentences in one block, and participants performed it without time pressure, the time limit for a single block was set at double the practice time minus 5 minutes. For example, if one participant finished the 4th practice trial in 20 minutes, he/she was told that 35 minutes (20×2 – 5) was the time limit to complete each block. The performance feedback in terms of percent correct and the total gained reward points for the 4th practice trial was provided so that each individual could self-calibrate and prepare a strategy for the experiment. The countdown signal was a clock ticking sounds presented at halfway through a block. Because no missing data were desired, there was no actual time limit for the experiment. Participants were told there was a time limit in order to impact their decision making strategies (Verplanken, 1993). Any devices that indicated time in the laboratory were taken away during the experiment. Participants were asked to turn off their phones and remove watches and leave them in a safe container in the lab during the experiment.

Participants were required to fixate on a cross-hair presented in the middle of the screen while listening to the speech stimuli and responding. They could either answer the question by pressing one of the four direction buttons on the keypad or skip to the next stimulus by pressing the middle button, depending on their effort management strategies. The stimulus presentation was programed to avoid a response occurring before the sentence finishes. The fixation image remained on the screen throughout the block.

2.4.3 Pupillometry

Listening effort was indicated by pupil dilation in the present study. Participants' pupil diameters were monitored using an ASL Eye-Trac 6 system (Applied Science Laboratories, Bedford, MA), which consists of a video camera and an infrared light source pointed at the participant's right eye. The ASL Eye-Trac 6 system allows free movement of the head with a magnetic sensor, attached to a headband, tracked and adjusted for head movement. However, using a head and chin rest has been shown to be very effective in minimizing head movements, keeping the participant in focal range of the video camera, and ensuring a consistent distance from the screen (Raney, Campbell, & Bovee, 2014). A chin-rest was mounted at the end of the table, and the height was adjusted so that participants were able to comfortably maintain the position for 30 minutes at a time (approximate one block). The computer monitor (size of the display: 19 in., display resolution: 1024×768) was located approximately 32 inches (81.28cm) from the participants. The center of the monitor was at the same level as the participant's eye. The recording video camera was located at the same vertical plane with the monitor, and at 0 azimuth to the measured eye. The distance of camera to eye was approximately 24 inches (60.96 cm). The ASL Eye-Trac 6 system employs a Pupil-Centre Corneal Reflection (PCCR) method (Mason, 1969) to calculate pupil size and to track the eye diameter and eye gaze location at 60 Hz (i.e., every 16.7 ms). The spatial resolution of the pupillometer is 0.1mm.

The luminance of the visual field was controlled to avoid the floor and ceiling of the range of pupil size, which is affected relatively more strongly by the light reflex than by cognitive load (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Zekveld & Kramer, 2014). For each participant, the brightness of the computer screen was adjusted from black to white in successive shades of gray to elicit the range of pupil sizes attributable to the light reflex. The luminance required to elicit an intermediate pupil size (midway between the minimum and maximum measured sizes) was calculated, and the corresponding shade of gray was used as screen background color for the rest of the experiment. This calibration process is consistent with Winn and Edwards (2013) and Zekveld et al. (2010). The standard 9 point calibration was used to teach the ASL system each individual's pupil/CR/scene relationship prior to collecting eye diameter data.

As the ASL program was run by a separate computer, a PCI-DIO 24 board manufactured by Measurement Computing was used to connect it to the experimental computer which runs SuperLab software. The pupil tracking by the ASL system therefore was able to synchronize with the stimulus presentation, having markers of the onset and offset of each trial in the pupil dilation traces.

2.4.4 Data selection, cleaning, and reduction

2.4.4.1 Pupillary data

The pupillary data were analyzed to answer the primary research question of listening effort allocation.

Data selection and cleaning

The output of the eye tracker data files include the major axis length (width) of the pupil and the minor axis length (height) at each time sample (i.e., every 16.67 ms, 60Hz sampling rate). The minor axis becomes small as the eye rotates away from the camera; however, the major axis does not, which indicates that the width is a more accurate pupil size measurement during eye movements (Kuchinsky et al., 2013). Therefore, the width of the pupil was used to calculate the pupil dilation indices in this study.

Data were cleaned using the procedures suggested by Siegle, Ichikawa, and Steinhauer (2008). The data cleaning process included blink identification and interpolation, smoothing, and artifactual trial removal. Correctness and reaction time for each trial were not part of the data selection criteria because the experiment allowed subjects to guess, which was considered one of the effort allocation strategies.

As noted in Siegle et al (2008), blinks were identified as large changes in pupil dilation occurring too rapidly to signify actual dilation or contraction. Specifically, blinks were coded as samples with estimated pupil diameter meeting any of the following criteria: 1) difference between the raw and a smoothed (3 point moving average applied twice) version greater than 1mm, 2) below 1mm; 3) below the minimum diameter in a subject's waveform +0.1mm; 4) below the median diameter minus 4mm; 5) below two times the interquartile range below the 25^{th} percentile (i.e., the Tukey extreme outlier hinge). Because intervals between blinks of less than 10 samples (0.16s) were unlikely to represent periods of clear vision, when an interval of less than 10 samples separated two blinks, both blinks and the interval between them were judged to be part of the same single blink.

Linear interpolations beginning 4 samples before and ending 9 samples after a blink replaced blinks throughout the data set. This technique prevented interpolation to poor pre- and post- blink pupil estimates due to partial lid closure. The average pupil diameter in the 10 samples (0.17 sec) preceding the onset of the tone was subtracted from pupil diameter after tone onset to produce pupil dilation difference score indices.

Data were selected using methodology similar to that described by Granholm et al. (1996). Specifically, trials comprised of over 50% blinks were removed from consideration. Data selection and cleaning procedures resulted in the exclusion of 5 participants, and elimination of 26.6 trials on average per participant (SD=12.5, Median=24 trials). After removing the 5 participants, the actual sample size of 35 still met the requirement by the power analysis.

Determination of windows of significant differences.

As the effect of speech rate and reward factors on pupil dilation across the entire duration of stimulus presentation and response period was of interest in this study, the pupillary data at each sample point were included in the analysis. Contrasts on pupil dilation were examined via linear regression at each time point along pupil dilation. To control type 1 error for this large number of tests, Guthrie and Buchwald's (1991) technique was used. This method was developed to combine graphical presentation with statistical theory. Briefly, this technique involves using Monte Carlo simulations to estimate the number of consecutive significant differences long enough to be judged to not have occurred by chance with $p<.05$ given the temporal auto-correlation of the data. Thus, contiguous sample-by-sample tests are considered replications. This technique has been successfully applied in previous publications on pupil dilation data sets (Siegle et al., 2008; Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003; Siegle, Steinhauer, & Thase, 2004).

The autocorrelation was calculated to be 0.98 in the data. The Monte Carlo simulations suggested that a window of 85 (i.e., 1.41 seconds given the 60Hz sampling rate) consecutive regression tests on pupil dilation along the time course significant at $p<.1$ would yield a window of differences significant at $p<.05$. When results are reported for an entire time window, they represent tests of the mean regression coefficients in a window of consecutive significant differences.

Wavelet analysis

Due to the temporal property of the speech rate variable, the pupillary time-courses between the beginning and the end of the sentence presentation were unaligned in time, with slow rate conditions having much longer durations than fast rate conditions (see Figure 7). As a result, Guthrie and Buchwald's (1991) approach is not appropriate for analyzing the pupillary data within the sentence presentation time window. The analysis solution aligned the trial segments of the pupil data to the beginning of the sentence (circle marker) first, and then to the end of the sentence (triangle marker). The temporally aligned cue and pause window (pre) and response window (post) adopted the massively univariate approach proposed by Guthrie and Buchwald (1991), whereas the wavelet analysis was performed for the sentence presentation window (mid).

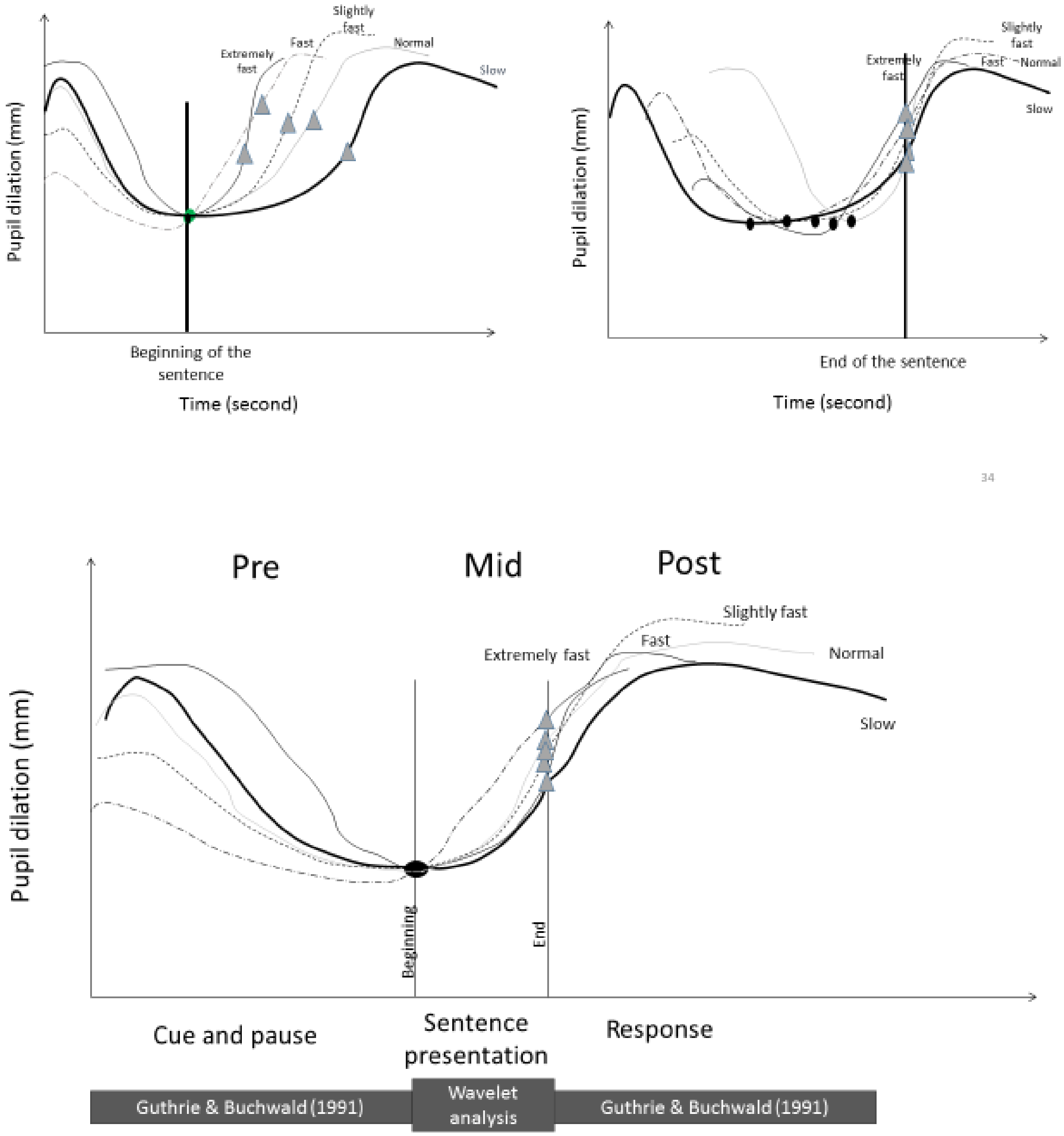


Figure 7. Illustration of the alignment issue of the pupillary data and the analysis solution.

Wavelet analysis is very useful for processing nonstationary signals, such as the pupillary data in this study. Like Fourier analysis, wavelet analysis deals with expansion of functions in terms of a set of basic functions. Unlike Fourier analysis, wavelet analysis expands functions not in terms of trigonometric polynomials but in terms of wavelets that are generated in the form of translations and dilations of a fixed function called the mother wavelet (Lee & Yamamoto, 1994). The Morlet wavelet was chosen to be the mother wavelet for the analysis. It is a wavelet composed of a complex exponential (carrier) multiplied by a Gaussian window (envelope), and it is closely related to human perception, both hearing (Tognola, Grandori, & Ravazzani, 1998) and vision (Daugman, 1985).

Briefly, the Morlet wavelet with a specific width, e.g., a width of 13.5 seconds for a slow rate condition, was compared to a section at the start of the original continuous-time pupil dilation signal, and a correlation coefficient was calculated. This single number gives a measure of the projection of this wave packet on the data during the first 13.5 second period, i.e. how much (amplitude) does the data's 13.5-second period resemble a sine wave of this width (frequency). By shifting this wavelet along the time series, a new time series of the projection amplitude versus time was constructed. The process was repeated many times, varying the scale of the wavelet by changing its width each time, until the defined scale range was covered. The result was a collection of time-scale representations of the signal, all with different resolutions (see Figure 8). The duration of the sentences in the study ranged from 4.8 seconds (extremely fast) to 13.5 seconds (slow), so the scale range for the wavelet analysis was set to be 0.02~0.2 Hz.

The wavelet power value for each of the valid trials was calculated for each participant. Finally, 25 condition mean wavelet power values were computed and served as the dependent variable in the significance test for the sentence presentation time window.

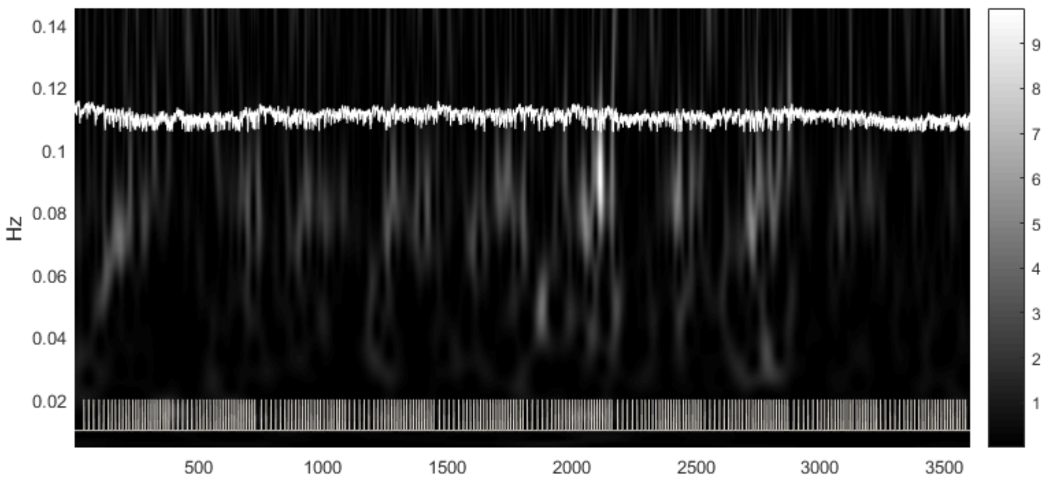


Figure 8. One subject wavelet data. The pupillary data (top white curve) and the trial onset markers (bottom white vertical lines) are shown in the wavelet spectrogram.

Two-way within-subject ANOVA

Additional two-way within-subject ANOVA was performed on the peak pupil dilation in comparison to Guthrie and Buchwald's (1991) approach. The purpose was to demonstrate the non-redundant information provided by using Guthrie and Buchwald's (1991) approach. The traditional method of analyzing the physiological data such as EEG and pupillary data is to conduct statistics on the peak, the peak latency and the mean of the response waveforms. However, the peak latency and mean dilation were not appropriate dependent variables due to the nature of the experimental design in this study (i.e., the trial duration was not equal among the speech rate conditions); therefore, only the peak dilation was included in the analysis.

2.4.4.2 Behavioral data

The behavioral data were analyzed to answer the exploratory research question about the quality of the effort regulation, denoted by effort efficiency. The individual's overall performance accuracy, total reward achievement and the total response time of the listening task were obtained and applied in the effort efficiency calculation. Each individual was scored on the Need for Cognition questionnaire, and the eighteen questions were averaged to produce a scale score. The Pearson correlation analysis was performed between the two efficiency scores and the score of need for cognition.

In the formula proposed by Paas and Van Merriënboer (1993), $Efficiency = \frac{Z(performance)-Z(mental\ effort)}{\sqrt{2}}$, the Z score of performance accuracy and the Z score of the total peak pupil dilation (i.e., the overall effort spent) were used, which are the same variables that Paas and Van Merriënboer used. In contrast, the variables used in the new formula proposed by the current study, $Efficiency = \frac{Z(reward)-Z(mental\ effort)}{Z(time)+K}$, hwere the Z score of earned reward in points, the Z score of the total peak pupil dilation, and the Z score of the total response time spent plus a constant K ($K > |\text{minimum } Z_{time}|$).

Additionally, the group's performance, error type, and self-reported strategy use were described. Some individuals' data also were reported.

3.0 RESULTS

3.1 PRIMARY RESEARCH QUESTION

To what extent is the stimulus-driven pattern of listening effort allocation in speech comprehension modulated by reward condition under time pressure?

3.1.1 Guthrie and Buchwald's (1991) approach

To illustrate the extent to which time windows differed in the effects of speech rate and reward, condition-related differences were evaluated at each time point along the grand-mean waveforms via multiple regression tests and the subsequent t-tests on the regression coefficients. Figure 9 shows the grand mean waveforms and wavelet powers for each speech rate (upper panel) and for each reward (lower panel) across all participants.

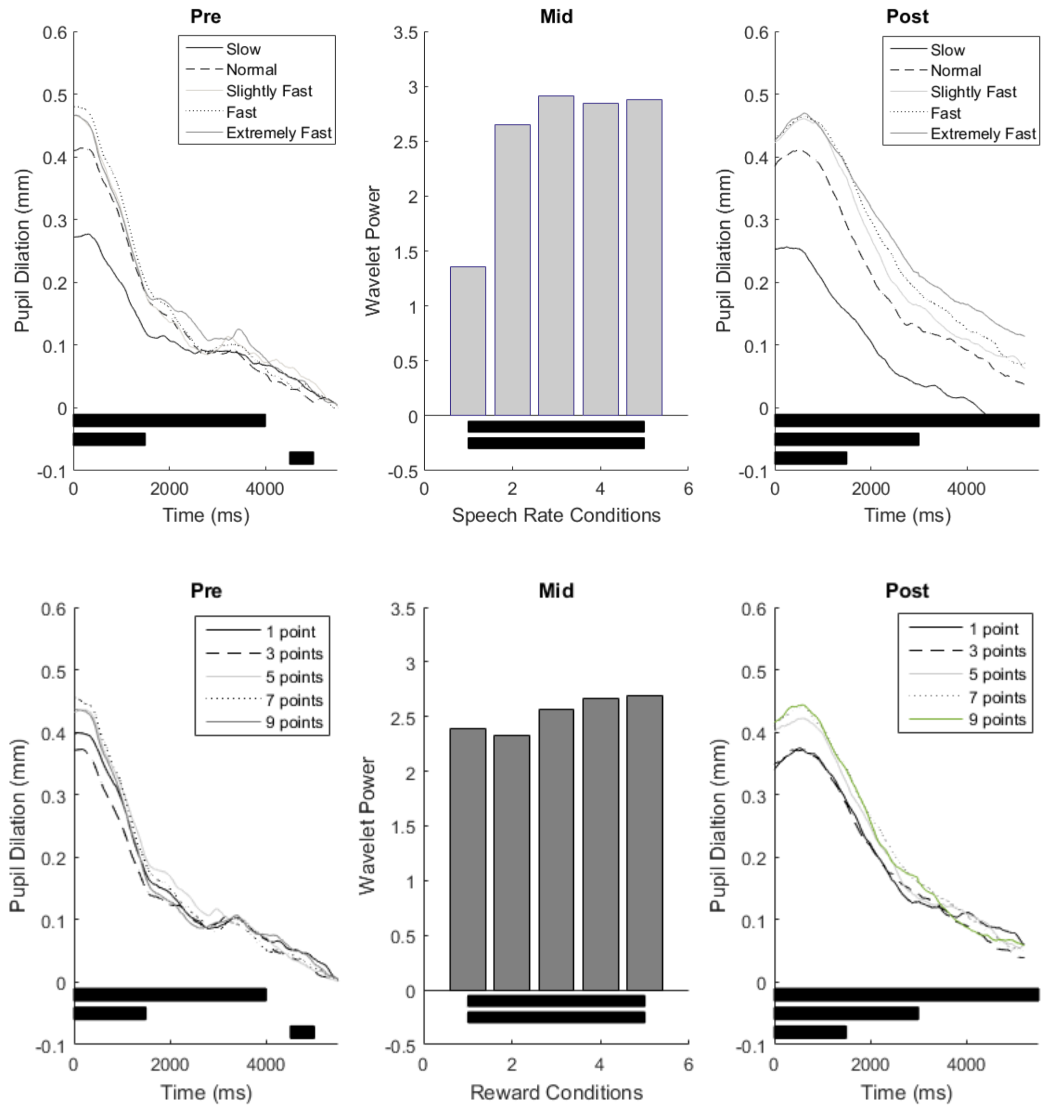


Figure 9. Grand mean pupil waveforms and the results of significance tests. The main effect of speech rate (upper panel) and the main effect of reward (lower panel). Regions of significant differences (at *p*<0.1) in regression test are highlighted along the X-axis by black bars: significant main effect of speech rate (the top horizontal bar), main effect of reward (the middle horizontal bar) and the interaction effect of speech rate and reward (the bottom bar).

During the pre-sentence time period, there was one significantly long window of main effect of speech rate extending from 0.02 to 4.00 seconds after the beginning of a trial (sentence onset occurred at 5 seconds) as determined using Guthrie and Buchwald (1991) method. The speech rate positively predicted the pupil dilation after adjusting for the reward, $B = 0.02$, t (34) = 6.48, $p < 0.01$, $d = 1.10$. There also was a significant window of positive prediction on the pupil dilation by reward after adjusting for the speech rate, from 0.02 to 1.92 s, $B = 0.01$, t (34) = 2.91, $p = 0.01$, d=0.49. The observed interval of the interaction effect of speech rate and reward closely preceding the onset of the sentence presentation was not long enough to be considered as statistically significant.

The same multiple regression and subsequent t-test was performed on the condition-related wavelet power values for the sentence presentation period. The results showed that the speech rate significantly predicted the pupil dilation with 0.005 mm increase in dilation for every 1 wpm increase in speech rate, t (34) = 4.48, $p < 0.01$, $d = 0.76$. The prediction by the reward condition also was significant, with 0.05 mm raise in dilation for every 1 point increase in reward, t (34) = 3.10, $p < 0.01$, $d = 0.52$. There was no significant interaction effect of speech rate and reward, $B < 0.001$, t (34) = 0.47, $p = 0.639$, $d = 0.08$.

In the post-sentence time period, a significant interaction effect of speech rate and reward was observed within the window from -0.95 to 1.22 seconds relative to the offset of the sentence presentation, $B < 0.005$, t (34) = -2.40, $p = 0.02$, $d = -0.40$. As shown in Figure 10 (left panel), the pattern of pupil dilation change as a function of speech rate was different at different reward levels. At low reward level (i.e., 1 point), the pupil dilation increased with the increase of the speech rate, at median reward level (i.e., 5 point), the pupil dilation did not differ among the speech rate levels, whereas at high reward level (i.e., 9 point), the pupil dilation declined as the speech became fast. The maximum dilation occurred in the slow-rate/9-points condition, and the minimum dilation took place in the slow-rate/1-point condition. There was significant main effect of speech rate extending from 0 to 5 s, $B = 0.04$, t (34) = 8.69, $p < 0.01$, $d = 1.47$, as well as significant main effect of reward extending from 0 to 3.28 s, *B* = 0.01, *t* (34) = 3.99, *p* < 0.01, *d* = 0.67, similar to the main effects in the previous time periods.

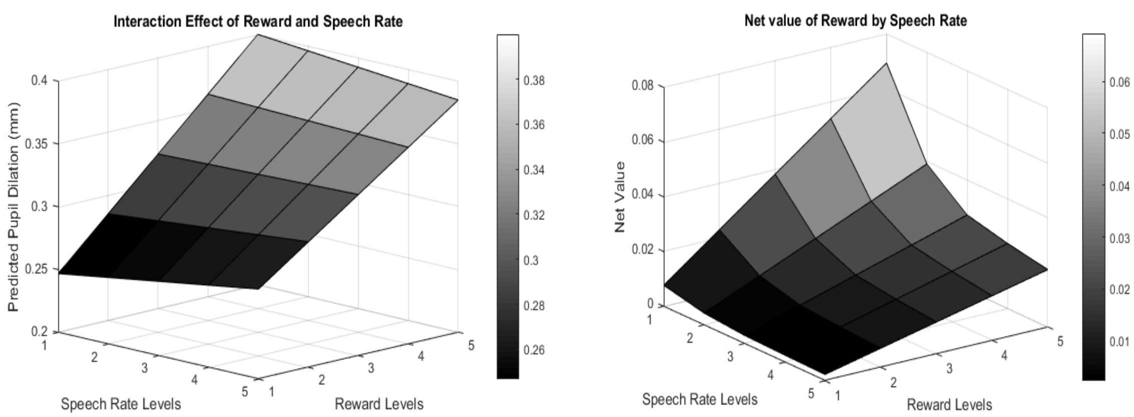


Figure 10. Interaction effect of reward and speech rate during the post-sentence time period.

The above data support the hypothesized effect of increasing listening effort as the speech rate and reward increase. However, the interaction effect was different from the findings in the effort literature which have shown that the effort varies as a function of the cost-benefit net value (Croxson et al., 2009; Forstmann et al., 2006; Mars et al., 2005; Walton et al., 2004; Yoshida & Ishii, 2006). The pattern of effort allocation among the 25 speech rate/reward combinations would resemble the pattern in the right panel of Figure 10, where the minimum effort would be spent under the extremely-fast/1-point condition, and the most effort would be spent under the slow/9-point condition.

3.1.2 Two-way within-subject ANOVA

As a comparison to Guthrie and Buchwald's (1991) approach on analyzing the pupillary data, a traditional 5×5 within-subjects analysis of variance was performed on the peak dilation as a function of speech rate (130wpm, 230wpm, 330wpm, 380wpm and 430wpm) and reward point (1 point, 3 points, 5 points, 7 points and 9 points). As the assumption of sphericity was violated, the statistics were reported with Huynh-Feldt correction. As displayed in Figure 11, there was no interaction effect of speech rate and reward on the peak dilation, $F(6.075, 206.543) = 1.713$, $p = .118$, $\eta_p^2 = .048$, indicating that the pattern of the peak dilation among the speech rate levels was not significantly different under different reward conditions. However, there was a significant difference on the peak dilation among the levels of speech rate averaged across reward, $F(3.032, 103.075) = 32.419$, $p < .001$, $\eta_p^2 = .488$. The peak dilation was significantly smaller under the slow rate condition (*M*=.373, *SE*=.028) than under all the other faster rate conditions (*M*=.536, *SE*=.039), $F(1, 34) = 77.669$, $p < .001$, $\eta_p^2 = .696$. The peak dilation at the normal rate condition (*M*=.502, *SE*=.037) was significantly smaller than at the slightly fast condition (*M*=.556, *SE*=.044), $F(1, 34) = 11.349$, $p = .002$, $\eta_p^2 = .250$. There was no significant difference on peak dilation between any other speech rate conditions.

A significant difference on the peak dilation among the 5 levels of reward averaged across speech rate also was observed, $F(3.126, 106.301) = 5.201$, $p = .002$, $\eta_p^2 = .050$. The peak dilation under 7 points condition (*M*=.540, *SE*=.043) and 9 points condition (*M*=.515, *SE*=.036) were significantly higher than that under 3 points condition (*M*=.466, *SE*=.031), $F(1, 34) = 10.754$, $p = .002$, $\eta_p^2 = .240$ and $F(1, 34) = 13.495$, $p = .001$, $\eta_p^2 = .284$, respectively. No significant difference existed between any other reward conditions.

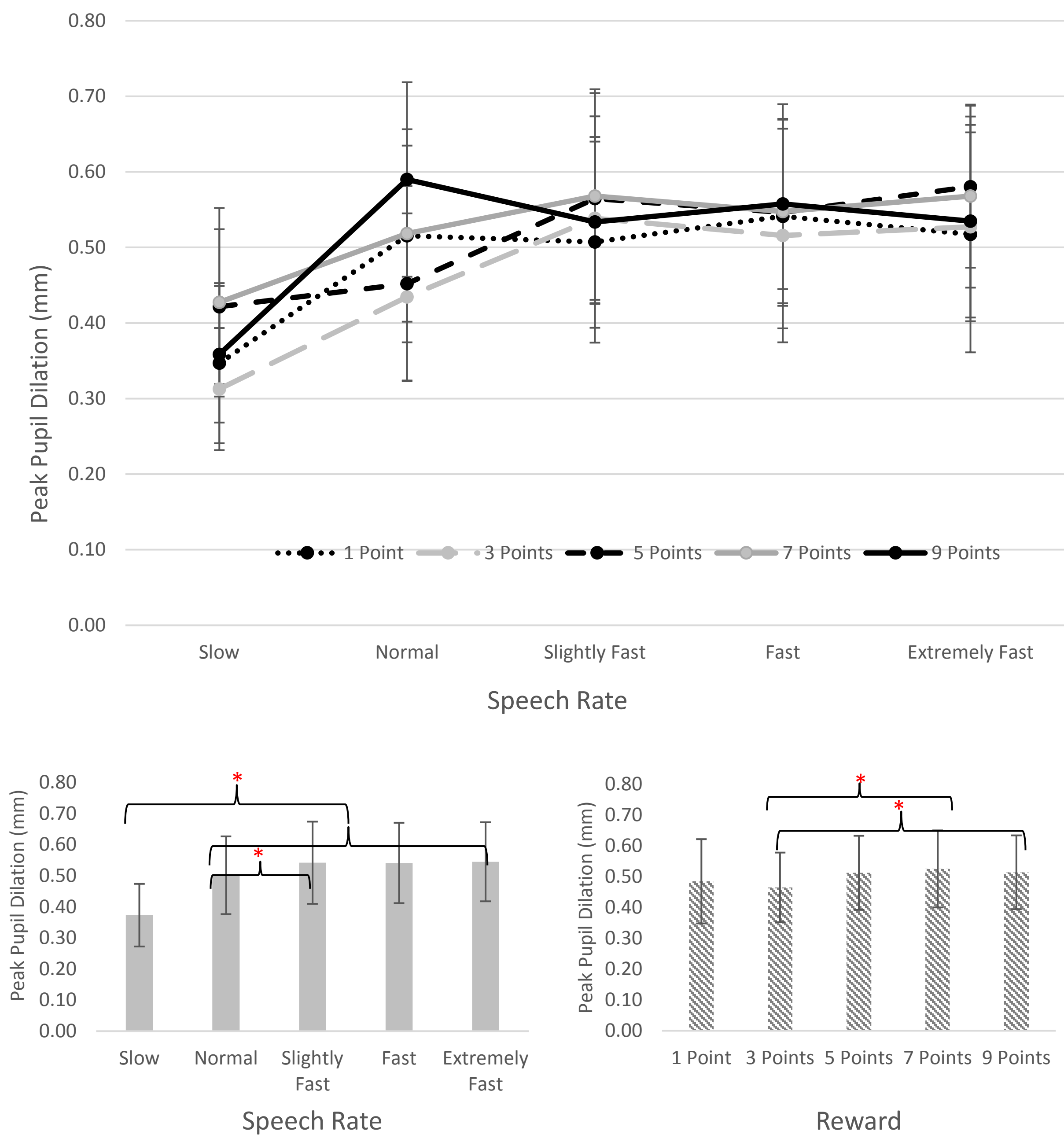


Figure 11. Results of the traditional analysis on peak dilation. Top: Mean peak pupil dilation in the 25 speech rate/reward combination conditions. There was no interaction effect of speech rate and reward. Bottom: Significant main effect of speech rate and reward. * $p < .05$.

3.2 SECONDARY RESEARCH QUESTION

Does the inclusion of the reward achievement, pupillary indices of effort and time component in the effort efficiency computational equation improve characterizing the quality of effort regulation?

Two efficiency scores were computed for each individual based on the equations in the previous section to assess the relationship between the effort efficiency and the Need for Cognition score. As shown in Figure 12, there was a significant negative correlation between the Need for Cognition score and the effort efficiency calculated by the proposed new equation, r = -.406, n =37, p =.013. The Need for Cognition score was not significantly correlated with the efficiency calculated by the original equation, r = -.187, n =37, p =.267, although the two effort efficiency scores were significantly correlated, r = .478, n =37, p =.003. There was no significant association between the performance accuracy and the Need for Cognition score, r = -0.036, n =37, p =.830.

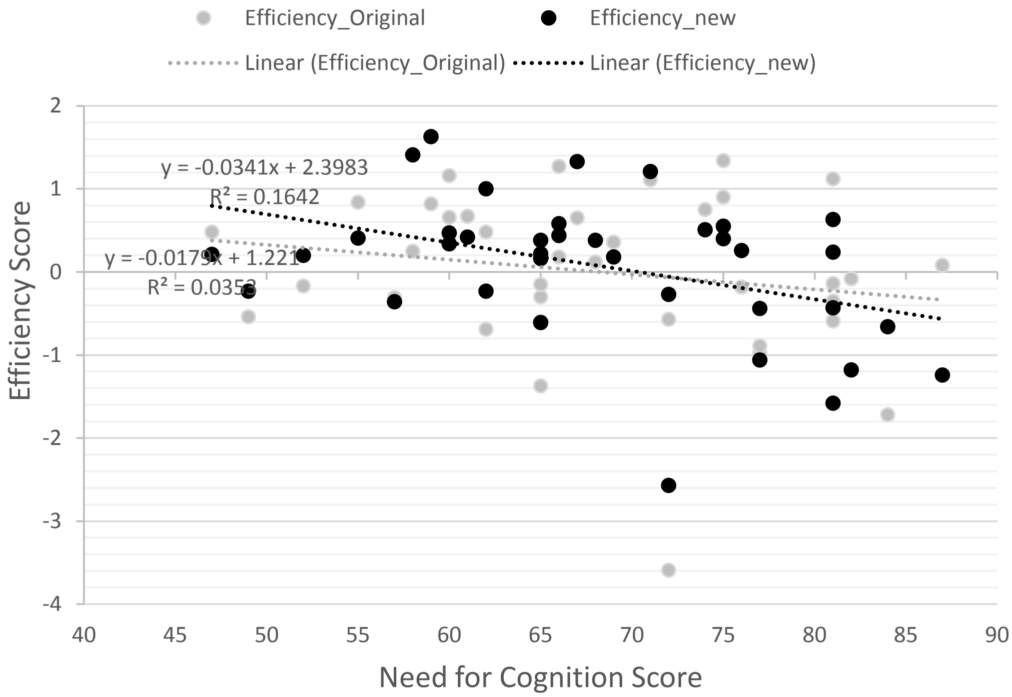


Figure 12. Correlation between the effort efficiency and the Need for Cognition score.

Further support of the advantage of using the new equation over the old one was illustrated by the following individual case comparisons. An example pair of participants, #14 and #31, were identified as equally efficient by the old equation, however, they were differentiated by the new equation in terms of effort efficiency (Table 11).

Table 11. Individual case comparison between the two efficiency equations.

	Participant #14	Participant #31
	Accuracy: 89% Total Effort: 10.99 mm Reward: $50 Time: 376.26 s	Accuracy: 80% Total Effort: 6.89 mm Reward: $48.8 Time: 138.90 s
$Efficiency = \frac{Z(performance) - Z(mental\ effort)}{\sqrt{2}}$	0.48	0.48
$Efficiency = \frac{Z(reward) - Z(mental\ effort)}{Z(time) + K}$	0.20	1.00

The most and least efficient participants identified by the two equations were different (see Table 12). According to the efficiency score computed by the original equation, participant #40 was most efficient. In contrast, the efficiency score computed by the new equation indicated that participant #15 was most efficient. Participant #36 was identified as the least efficient by both equations.

Table 12. Top 5 most efficient participants identified by the two equations.

Rank	Old Equation	New Equation
1st	Participant #40	Participant #15
2nd	Participant #20	Participant #32
3rd	Participant #12	Participant #16
……	……	……
35th	Participant #7	Participant #38
36th	Participant #34	Participant #21
37th	Participant #36	Participant #36

Table 13. Individual efficiency score comparison between the most and least efficient participants.

	Most efficient		2nd Least efficient	
	Participant #40	Participant #15	Participant #34	Participant #21
	Accuracy: 97% Effort: 8.56 mm Reward: $50 Time: 244.16 s	Accuracy: 81% Effort: 4.93 mm Reward: $50 Time: 134.64s	Accuracy: 61% Effort: 13.79mm Reward: $38 Time: 519.97 s	Accuracy: 60% Effort: 5.38 mm Reward: $37 Time: 109.52 s
$\boldsymbol{Efficiency} = \frac{\boldsymbol{Z(performance) - Z(mental\ effort)}}{\sqrt{\boldsymbol{2}}}$	1.34	0.82	-1.72	-3.59
$\boldsymbol{Efficiency} = \frac{\boldsymbol{Z(reward) - Z(mental\ effort)}}{\boldsymbol{Z(time) + K}}$	0.54	1.62	-0.64	-2.55

Participant #40 (the most efficient) and participant #38 (the least efficient) had the exact the same performance accuracy, both achieved nearly 100% correct on the task. They gained the same amount of reward; however, participant #40 spent less effort on the task than participant #38, although participant #40 needed longer to complete the experiment.

In order to investigate whether the effort allocation was based on the net value in the most efficient participants as originally hypothesized, a regression analysis was performed on the pupil dilation predicted by net value for both participant #40 and participant #15 (Figure 13). There was a significant prediction of pupil dilation by the net value for participant #15, $F(1, 23) = 5.125$, $p = 0.033$, $\eta^2 = .182$, the pupil dilation wavelet power decreased by .185 with every .01 increase in net value. When the slow condition data were removed to eliminate its possible confounding impact on effort (i.e., the slow condition was too easy for them to spend more than required effort on even with the highest reward), the prediction became not significant. There was no significant prediction of pupil dilation by the net value for participant #40, $F(1, 23) = 1.621$, $p = 0.216$, $\eta^2 = .066$, regardless of the inclusion of the slow conditions (Figure 13). These results suggest that even for the most efficient individuals, the listening effort allocation was not necessarily based on the cost-benefit judgement.

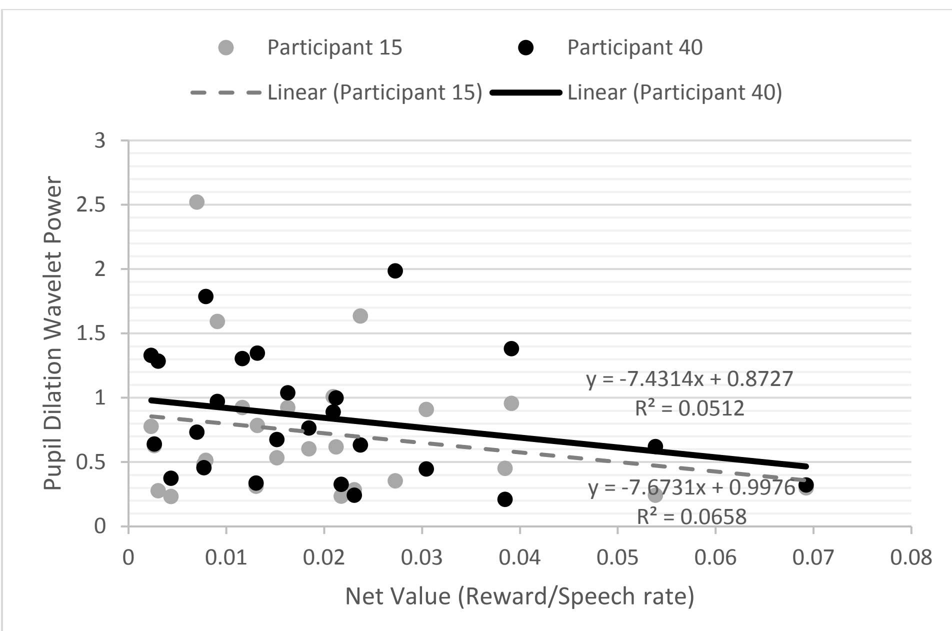


Figure 13. Effort as a function of reward/speech rate net value for the two most effort efficient participants.

To summarize, compared to the original efficiency computation equation with only performance accuracy and effort elements, the inclusion of the three components, reward, effort and time, has shown the advantage of differentiating individuals and characterizing the quality of effort allocation, which implies that the resource is optimally allocated to achieve the goal in the best way while minimizing the time spent.

3.3 DESCRIPTIVE DATA

3.3.1 Performance and response time

There were a total of 29 participants (74.4%) who gained enough reward points to receive the full $50 payment. The accuracy range among the 29 participants was 20% (78%~98%). Thirteen participants had an accuracy above 90%.

On average, the accuracy decreased as the speech rate increased, and the response time (range from 0.7 to 1.1 seconds) increased with the increase of speech rate as displayed in Figure 16. This pattern was consistent across all 5 reward levels; however, it differed from the inverted U response time pattern shown in the preliminary data collected without giving the participants any reward (see Figure 5), where the accuracy (range from 12% to 86%) decreased as the speech rate increased, and the response time (range from 1.0 to 2.9 seconds) increased as the speech rate went from slow, normal to slightly fast, then decreased as the speech rate became extremely fast.

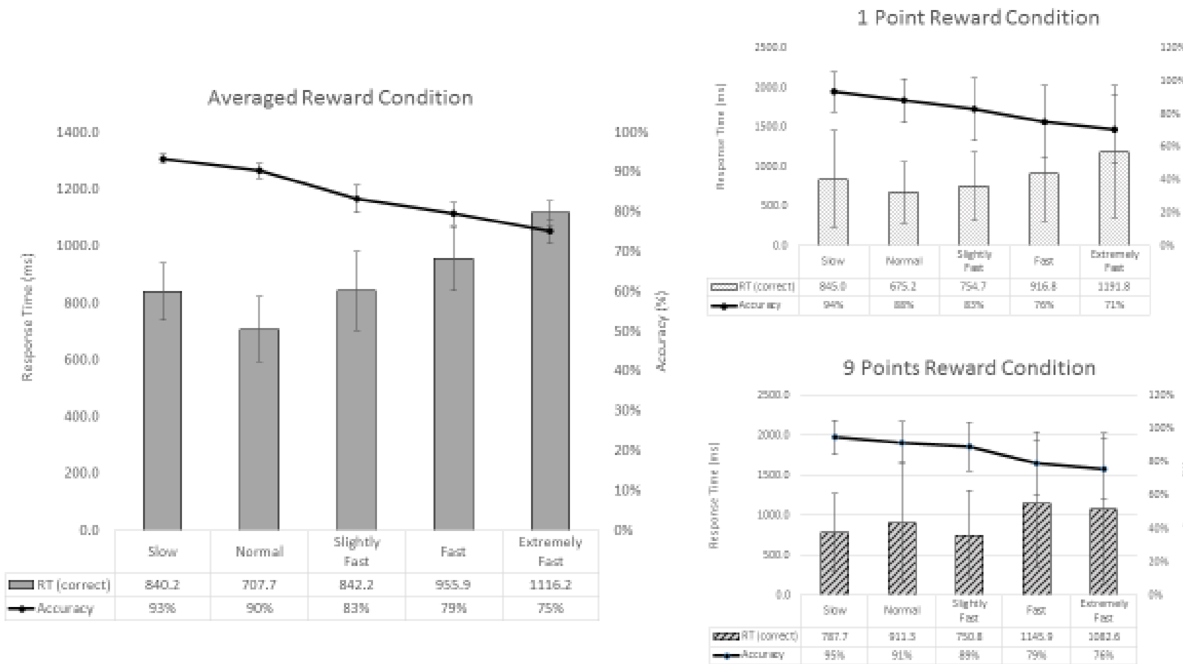


Figure 14. Averaged performance accuracy and response time.

All participants successfully passed the hierarchically organized practice trials prior to the main experiment. There were 19 out of 39 participants who chose to use the skip key when they did not know the answer to the spatial questions even if there was no punishment for any incorrect answer. There were 9 participants who pressed the Left/Right key when answering the Front/Back question, and 5 participants who pressed the Front/Back key when answering the Left/Right questions, but they only made the former error less than 10 times, and only once for the latter error (see Figure 15). These errors are meaningful in the perspective of individual's overall effort regulation and its subsequent efficiency, therefore, the effort allocated on those error items was included in the above data analysis.

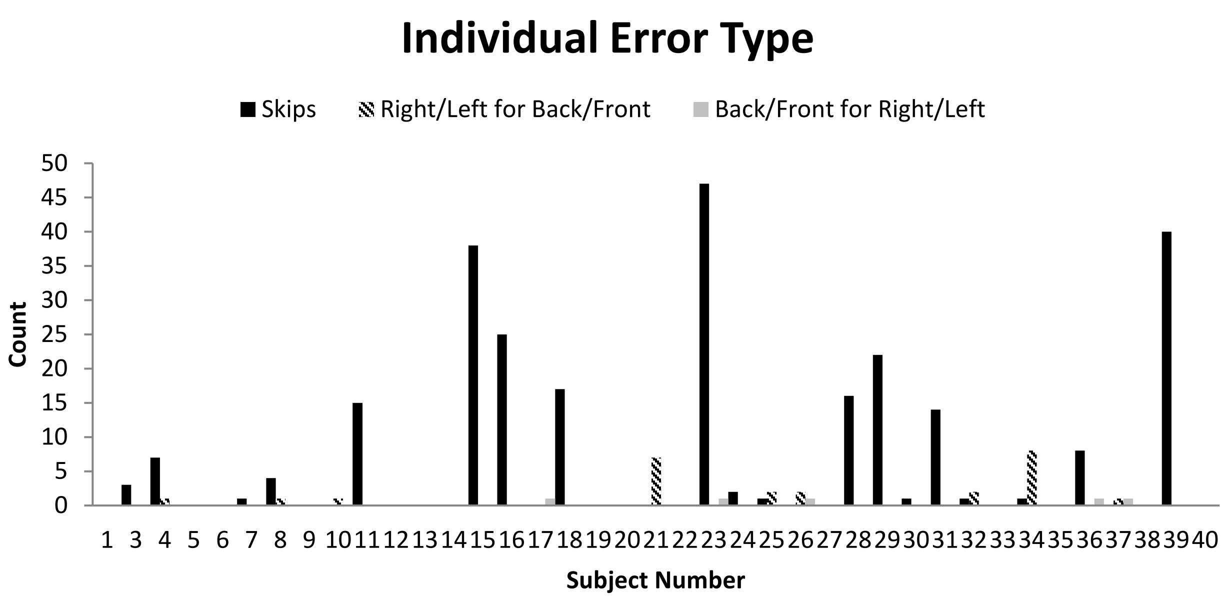


Figure 15. Histogram of the performance error type.

3.3.2 Self-report strategy use

All subjects responded to two open questions about strategy by typing their answers on the computer immediately after completing the main experiment and before they were told how much reward they had earned. The first question was: “What strategy did you use to answer the object spatial relationship questions?” Nineteen participants reported that they used visual cues the most, for instance, a participant described that she tried to use the circle in the fixation image as the ball, and visually placed the kite and stone around it. The remaining 21 participants reported using finger strategies. For example, a participant described that she assigned locations of the objects to a space on the table or control pad with her fingers. She oriented her fingers to the commands; if the commands were talking about left and right her fingers were positioned going left to right; if the commands were about front to back, she rotated her hand so that her fingers went front to back.

In answering the second strategy question, “How did you distribute your effort among the items?” the majority of participants reported that they exerted more effort on the high-reward items. A typical response was: “I felt as though I tried to be hyper-attentive when I knew that a subset of commands was worth higher value of points. I noticed that when the fast and super-fast conditions were worth 1 point only, I was less attentive. If I felt that I did poorly on even the heavy-weighted commands, I would try to compensate on the 1-3 point commands despite the fact that they were weighted differently to try to make up for my mistakes”. Only 3 participants intentionally spent more effort on the fast rate items than on the slow rate items. Eight participants claimed that they used the same amount of effort across the items. There was one subject who felt that the slow condition was difficult and consumed her additional attentional effort, see Figure 16.

The results of self-reported strategy use suggested an impact of reward on effort regulation. When given both task demand and reward information, the reward became more salient than the task demand in driving the effort allocation.

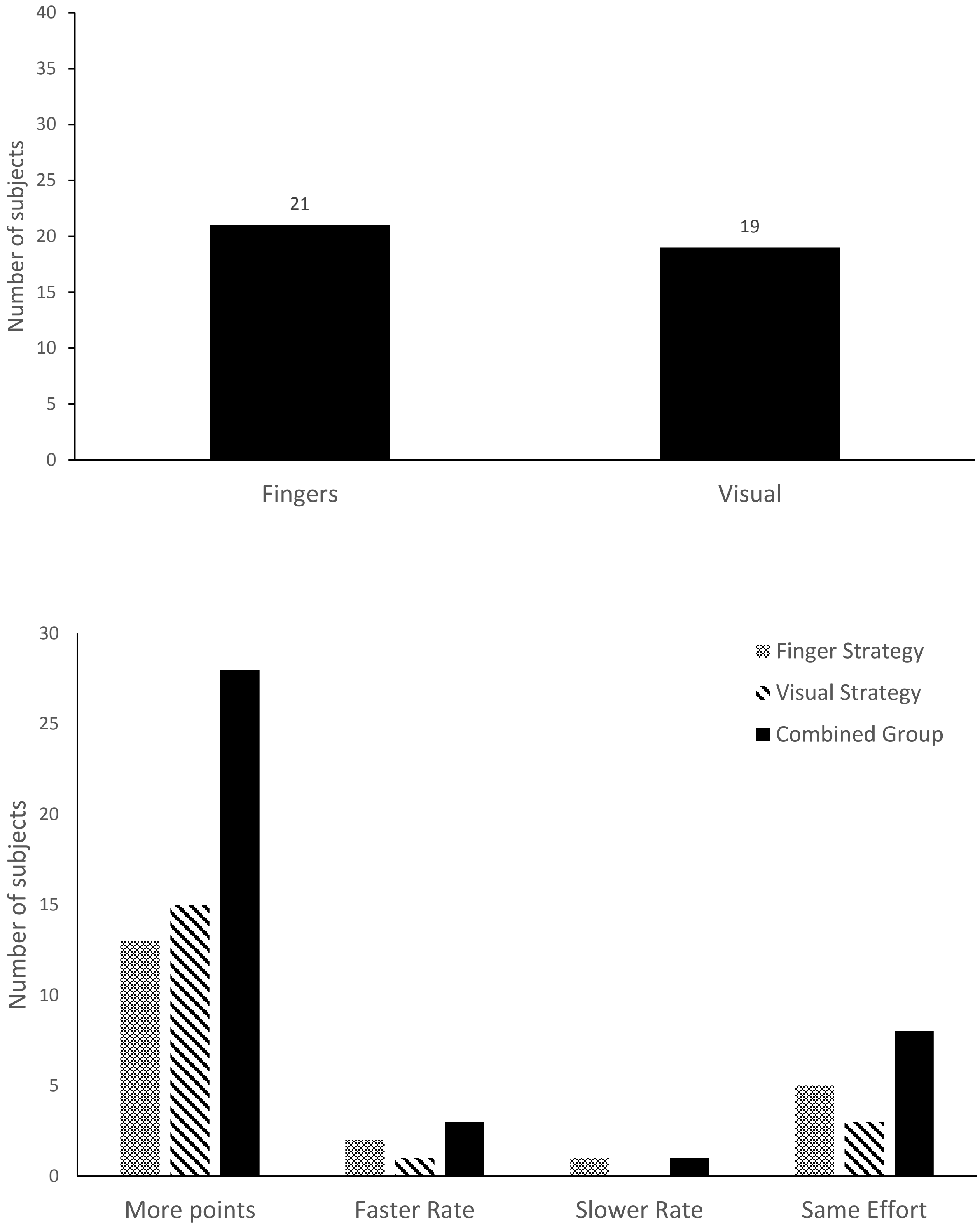


Figure 16. Self-reported strategy use.

4.0 DISCUSSION

4.1 GENERAL DISCUSSION

The goal of the current study was to evaluate the effect of speech rate and reward on listening effort allocation in young normal hearing adults and to explore a method to quantify the quality of effort allocation, which refers to how efficiently the effort is spent. The main hypothesis was that listening effort allocation would be driven by both factors, and the inclusion of the obtained reward points, total effort and time spent information in the efficiency computation equation would properly differentiate individuals and quantify the effort quality.

The present study used Guthrie and Buchwald's (1991) statistical-graphical method to evaluate the statistical significance of difference pupil dilation, which provided a new way of viewing the significant effects of speech rate and reward on effort at each cognitive processing stage along the continuous time-course during the whole speech stimulus trial. From the beginning to the end of each trial, the following cognitive processes were expected to happen to lead to the completion of the task. In the first 5.5 seconds (i.e., pre-sentence presentation period), participants auditorily received the information about speech rate and reward weight of the upcoming sentence, and had 2.5 seconds silence to decide how much effort to invest. The effort at this stage included attention, auditory perception and comprehension, value evaluation, and effort-based decision making. As evidenced by the significant main effects of speech rate and reward on the pupil dilation in this time frame (see Figure 9), the effort was driven by both speech rate and reward information, and the impact of speech rate lasted longer than the reward. The sustained effect of speech rate is a possible result of the extra cognitive effort required to evaluate the less straightforward speech rate cue compared to the straightforward reward cue. At the end of the decision-making silent period, the pupil diameter was expected to return to baseline after the stimulation of the auditory cues to get prepared to hear the speech sentences; therefore, neither factor should impact the pupil dilation right at that particular time. This hypothesis was supported by the data.

During the sentence presentation period (4.8~13.5 seconds depending on the speech rate condition), participants used whatever strategy they chose to solve the spatial question while listening to the speech sentence. The effort at this stage included various levels of attention depending on their earlier decision, auditory perception and comprehension, storage and encoding in working memory, and physical effort if using finger strategies. As supported by the results of significant main effects of speech rate and reward on the pupil dilation wavelet powers in this time window (see Figure 9), the effort spent on listening to the sentence stimuli for the purpose of comprehension, solving the problem and obtaining the reward points was driven by both speech rate and reward. As expected, the effort significantly increased as the speech rate became faster and the designated reward point was higher. At this stage, the pattern of effort allocation among the speech rate levels was not expected to significantly vary as a function of reward levels due to the difficulty in suppressing the salient auditory stimuli, therefore, it is not surprising that no significant interaction effect was observed.

At the response stage (i.e., post-sentence presentation period), participants were supposed to answer the spatial question as quickly as possible given the time pressure and the overall goal of gaining the full payment. The effort in this period included the attention to the answer of the question, physical response action and task goal. As hypothesized, in addition to the main effects of speech rate and reward similarly seen in the previous two stages, the reward significantly interacted with speech rate on the effort allocation. At low reward level, effort was driven by speech rate in the positive direction, representing a stimulus-driven pattern; whereas at high reward level, effort was driven by speech rate in the opposite direction, representing a goal-driven pattern.

Although significant, the modulation of speech rate and reward on each other was not in line with the effort-based cost-benefit decision making literature (Croxson et al., 2009; Hardy, 1982; Hillman & Bilkey, 2012), which claimed that the deployment of effort on a task is based on the cost-benefit net value judgment. According to the results, participants as a group did not show a clear relationship between effort and net value; even the two participants identified to be most efficient did not distribute their effort based on the speech rate-reward net value. This may have been due to the many levels of speech rate and reward yet small variability of the net value. Both Figure 10 and 15 reveal that the net value clustered between 0.002 and 0.030. It is possible that participants were not able to distinguish the subtle value difference between normal/3 points (net value=0.0130) and fast/5 points (net value=0.0132). It may be that because performance feedback was not provided until the completion of the whole experiment, that some participants spent effort on the low value items unnecessarily to compensate for their perceived failure on high value items, as described in the self-reports of strategy use.

The pupillary data in the present study demonstrated an impact of stimulus factor (speech rate) and goal factor (reward point) on the listening effort distribution, which implies that human effort allocation during listening is driven by both. The interaction effect at the response stage suggested complex strategic resource-management decisions. These all fit the general predictions of Hockey's (1997) compensatory control model of the individual difference, in terms of cognitive capacity, judgment, tolerance, and affective state, as a critical influential factor for high level effort control. More specifically, when individuals are actively involved in performance of effortful mental tasks by changing their current energetical resource state (e.g., effort level) to meet target state, effort is not automatically increased to meet the new elevated task demand. Rather, the involvement is affected by the subjective judgment on the cognitive manipulations such as incentives, importance, knowledge of performance, and achievement motivation.

Compared to the conventional analysis approach using mean, latency and peak pupil dilation within a discrete window proximal to response prompt, Gutherie and Buchwald's (1991) method of testing the significance on time-series data has shown some unique advantages. First, the graphical display of the cross-sectional significance levels provided a clear picture of the range of significant differences along the entire cognitive processing time period; second, requiring significant intervals of cross-section significance provides experiment-wise protection of type I error rate without invoking the very conservative multiple comparison methods; third, the method is sensitive to the presence of intervals of significant difference physiological activities, such as evoked potentials and pupil responses, in long periods rather than brief periods. All the above justified the proper application of this analysis in the present study. This data analysis verified that Gutherie and Buchwald's (1991) approach revealed the interaction effect of speech rate and reward which the conventional ANOVA approach performed on the peak dilation failed to capture.

Given the actual allocation pattern of the effort regulation, a logical subsequent question of interest was about the quality of the effort exertion, i.e., how efficiently participants had spent their effort. This question was answered by computing efficiency scores using two mathematical equations with different performance elements and correlating the efficiency scores to the individual differences on the Need for Cognition test. The hypothesis of the superiority of using the new equation proposed by the present study was borne out. The efficiency score generated by the new equation showed larger variability among individuals and was significantly correlated with the individual Need for Cognition score. Interestingly, the correlation was negative, indicating that high Need for Cognition scores corresponded to low effort efficiency. A possible explanation was that, individuals high in need for cognition are more likely to think about and elaborate cognitively on issue-relevant information when forming attitudes but individuals low in need for cognition might tend to be driven more by the task difficulty than by the overall task goal (Petty, Cacioppo, Kao, & Rodrigues, 1986). In other words, individuals high in need for cognition are inclined to take the challenge of difficult tasks regardless of the performance outcome. This was demonstrated by the non-significant correlation between the performance accuracy and the need for cognition score shown in the present study. The low efficiency of the participants high in need for cognition implied that they wasted their effort on tasks beyond their capacity.

In considering the optimal description of individual effort efficiency, the goal-related performance, mental effort and time are reasonable and necessary elements to be included in the efficiency computation. In a traditional point of view, performance accuracy and mental effort are sufficient to characterize relative effort efficiency during complex cognitive tasks (Camp, Paas, Rikers, & van Merrienboer, 2001; Paas & Van Merriënboer, 1993). However, as the effort allocation is significantly driven by goal, the total reward achievement was considered more appropriate to represent the task performance than the accuracy in the context of this study. The addition of total response time spent on task was also important to capture individual differences because time is another form of resource. The mental effort measured by pupillometry provided a wide range of variability in the study sample, which might have played a role in the differentiation among individuals. The new effort efficiency equation was able to reflect individual differences in effort use preferences and in avoiding effort waste.

Unlike typical cognitive psychology studies, which usually focus on the accuracy and speed measures of performance and only include correct trials in the analysis due to the difficulty of identifying participants' strategy in speed-accuracy tradeoff and their guess process (Glickman, Gray, & Morales, 2005), the present study purposely pooled the correct and incorrect trials into the data analysis for two reasons. One reason was that, because the study aimed to investigate participants' strategy of effort allocation, the task goal for each individual was not to respond as quickly and accurately as possible to every item, but to selectively answer questions to gain the full payment within a time limit, which only required an accuracy of 50%~80% depending on which items were selected. Participants were allowed, and to some extent encouraged by the time pressure to disengage in some of the speech comprehension processes (i.e., use the skip button to indicate a giving up) while having the overall task goal in mind. The focus of the pupil dilation analysis was on how much effort was actually spent on each test item, and therefore the effort spent on the correct and incorrect items both were meaningful in this scenario.

Second, based on the distribution of performance accuracy and reward achievement (shown in Figure 17), only 4 participants were suspected of losing the task goal during the experiment. Their accuracy scores were 58%, 60%, 61% and 64%, corresponding to the reward achievement of $36, $37.15, $37.15 and $38.65 respectively. The identical statistical analysis procedure on the pupillary data was performed without the 4 participants, and the same results as shown in Figure 9 emerged.

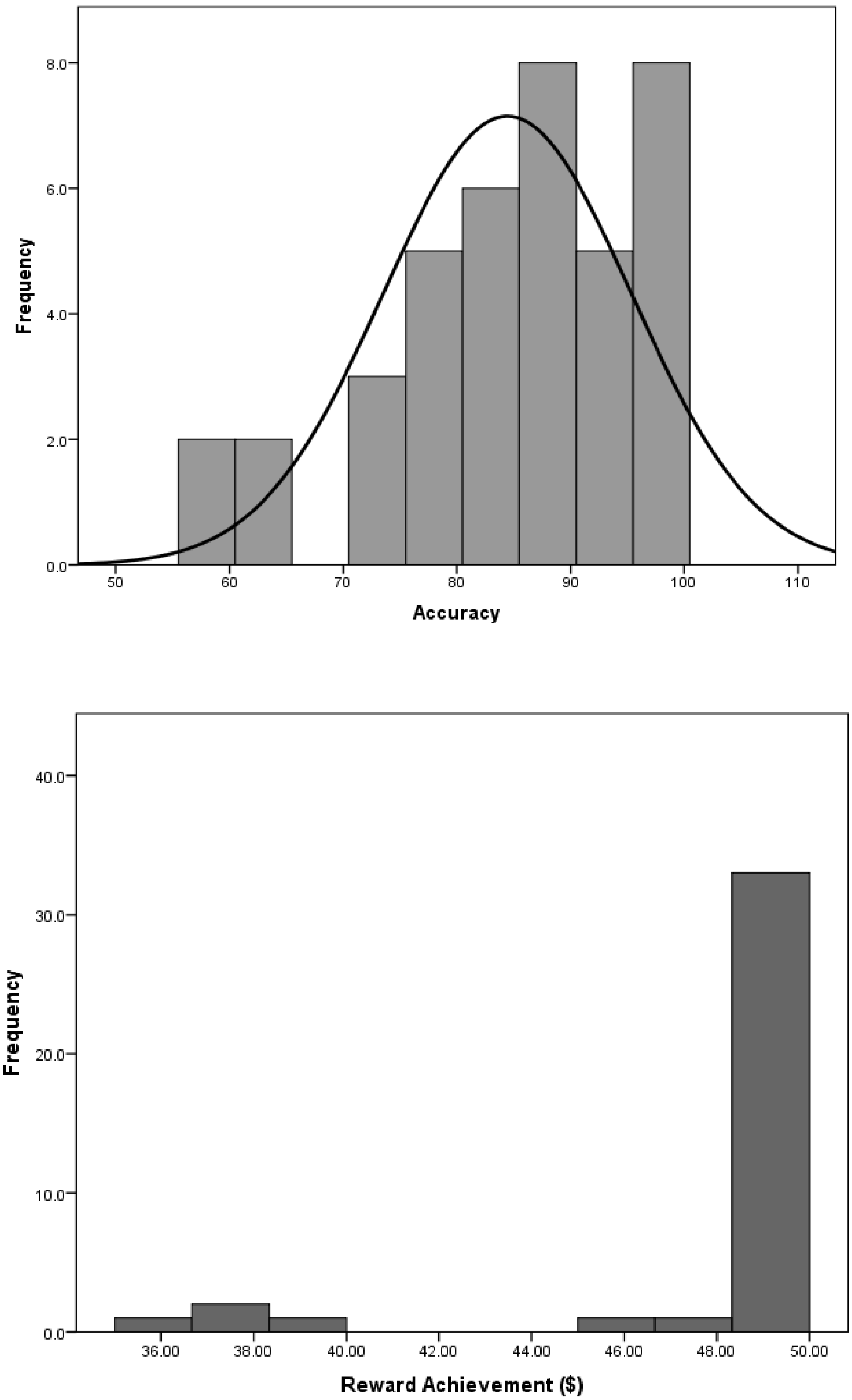


Figure 17 Histogram of performance accuracy and reward achievement.

The third reason for including the error trials was that the efficiency characterizes individuals' effort expenditure preferences as well as the ability to avoid wasting resources. Only with the total effort and time spent on both correct and incorrect trials could a clear strategic pattern of used and wasted effort be drawn, which illustrated an effective proportion of effort and time spent to achieve the overall task goal.

Typical errors observed in a task include skill-based errors (e.g., slip of action, wandering attention) and knowledge-based mistakes (e.g., unable to solve the problem, guess). Not all errors were the result of a strategy. However, there was a special strategic error type in the present study. During the experiment, participants were given an option to push the skip button if they decided not to answer the question for whatever reason. Intuitively using the skip button would not be a good strategy to gain the reward since guessing had a 50% probability of producing a correct response; however, this option saved time to complete the whole experiment. For example, participant 15, the second most efficient participant, selectively used the skip option and guessing strategy, as described in her report: "The points were a strong motivation factor, but the slower the speed the easier. I think overall I concentrated harder on the sentences worth the most points. If the slow speed was only worth 1 point, I still attempted an answer because slow speed required a lot less effort. If the extremely fast speed was 9 points, I tried to follow to get the points, but since it required a lot more effort, sometimes I missed the relations completely and just hit skip. If the extremely fast speed was only worth 1 point, I (more often than not) skipped it. For the middle range (both speed and points) sometimes I hit skip if I missed the first relation, but other times I figured it couldn't hurt to just take a guess". This participant obtained 81% accuracy but received enough points to get the full payment.

Although the task performance accuracy itself was not a focus in the current study, the accuracy results conveyed additional information that supports the important role that the goal-driven factor played in an effortful task. Specifically, the speech rate range used in the study, 130~430 wpm, was determined by a preliminary experiment in which the reward was not given. The preliminary data showed that the performance accuracy was 90% at slow rate, around 80% at normal rate, close to chance (57%) at slightly fast rate, and fell below 20% when the rate was fast and extremely fast, which suggested an adequate range, from easy (requiring only a small amount of effort) to very difficult (make participants quit). However, the chosen speech rate range based on these preliminary data failed to accurately cover the actual full cognitive capability of the participants, because the performance accuracy varied between 75% and 93% across the speech rate levels in the formal experiment, and even higher under the 9 points reward condition. These findings suggested that, goal-driven factors, such as the incentive motivation in this study, not only regulate the effort allocation among items with different difficulty levels, but also facilitated strategy use and shifted the position of the task relative to the full cognitive capability scale.

The majority of participants reported that reward points drove their effort allocation, few participants reported allocating more effort on difficult items (i.e., faster rate conditions) than on easy items, although the pupillary data indicated that they actually did. When the task demand and reward information are both available as in the present study, the task demand appears to become a much less salient driving factor perceived by participants. It is unclear whether the absent self-report of stimulus-driven pattern is due to participants' choice of not explicitly expressing their effort preference on the difficult items or due to the deployment of effort in the stimulus-driven fashion without participants' awareness. The latter raises an interesting question about whether the bottom-up computation of task demand takes place regardless of the current task instruction and is simply added to the effort weight associated with top-down value assessment. This might explain why the subjective measure of listening effort is often not strongly related to the objective measure (Capa et al., 2008), as the self-report measures of effort seem to tap what appear to be separate processes.

4.2 RESEARCH AND PRACTICAL IMPLICATIONS

4.2.1 Research implications

Beyond the immediate context of the study, this research offers important and closely linked research implications. First, this is, to the best of our knowledge, the first study that objectively measures listening effort with systematic manipulation of both a goal-driven factor and a stimulus-driven factor in the context of the listening effort and the general cognitive effort control literature. This study provides direct support to the two major effort theories. Kahneman's (1973) unitary resource model is a dominant and influential theoretical foundation for the listening effort studies, yet, the role of intention in resource allocation regulation proposed in this model has not received adequate attention. Hockey's (1997) effort compensatory control model incorporates both goal-driven factors and stimulus-factors in explaining how effort is controlled during a cognitive task; however, this model has not been adopted in the listening effort research. The lack of nuanced and direct empirical support in the field of Audiology for goal-driven effort allocation indicated in both models represented an opportunity for this research. Confirmation of the impact of reward on listening effort allocation is a first, yet important step for motivation-based listening research direction, and it improves our understanding of the underlying mechanism of listening effort allocation.

Second, the study brings the dimension of quality of effort allocation into the listening effort literature and proposes a method of calculating the effort efficiency. The new equation reflects goal-driven effort control and characterizes the individual differences in how well the limited resources are distributed in a cognitive task such as listening. By comparing two different efficiency computation approaches, the importance of using the appropriate performance variables based on task goal to evaluate the quality of effort expenditure has been raised. The correlation between the efficiency score and the Need for Cognition score supports an often ignored notion that more effort does not guarantee better outcomes. It also sheds some light on the role the individual's dispositional factor played in effort regulation, which is part of the high level effort control loop in Hockey's (1997) compensatory control model.

Other additions of the present study to the listening effort literature include: the introduction of the innovative pupillary data analysis approach to allow viewing the pattern of effort allocation in a continuous temporal domain; the empirical evidence of effort change as a function of speech rate which is a real life listening difficulty factor (Pichora-Fuller, 2003); the introduction of the Need for Cognition scale that assesses individual dispositional differences in cognitive motivation; and the direct subjective measure of the driving factors of effort allocation by open questions about strategy used.

4.2.2 Practical implications

Currently, the main objective of the audiological intervention in clinic is to take full advantage of patients' residual auditory function to improve hearing. This is usually accomplished by using the diversity of acoustical signal processing technologies. However, listening effort evaluation has not become a part of clinic practice thus far. When hearing is improved, listening effort is normally reduced to some extent. Even when speech perception and comprehension are not significantly improved by the technologies, the release of listening effort can take place (Sarampalis et al., 2009). Hence, assessing this cognitive benefit clinically has become as critical as the speech performance in order to evaluate the full dimension of intervention outcomes. In the past two decades, researchers have been devoting effort to develop an objective tool that can be readily implemented clinically to measure listening effort. However, little attention has been paid to in which context listening effort should be measured.

In daily life, listening activities are accompanied with various types of value (e.g., incentive reward, emotional reward), and listeners make decisions frequently on how much effort to spend on those activities. Their decision-making performance is important for their communication success. Since human decision making is essentially driven by goals, it is not surprising that listening effort mobilization is determined by more than just task demand. Findings from the current study suggest the need to not only focus the on stimulus-driven pattern of listening effort allocation, but to consider both jointly and independently the implication of goal-directed factors and individual dispositional differences.

Specifically, this research provides evidence that when listeners are motivated with a reward, they could discover their potential capacity and find coping strategies to improve their performance. Adding a motivational factor such as reward in the performance outcome measurement can provide a clearer picture of what a given auditory and cognitive system is capable of to both audiologists and patients. This information, in combination with patients' own value judgement system, will help them establish a realistic individualized matrix of cost-benefit for listening during the counselling and aural rehabilitation sessions. The audiologists can then direct patients to efficiently use their effort. Although we still do not have an easily implemented clinical measurement tool of listening effort so that clinicians can quantify the quality of listening, and use the calculated effort efficiency as an indicator of listening improvement rather than hearing improvement by intervention, the present study nevertheless recommends a rough estimate of effort efficiency by assessing individuals' Need for Cognition. As shown in the study results, people high in Need for Cognition tend to have relatively low efficiency, thus need more advice on strategic effort use compared to their high efficiency peers.

4.2.3 Limitations and directions for future research

Limitations

Several limitations should be considered when evaluating the current results. First, due to a practical reason, 80% of the participants were females, which limits the generalizability of the results to the female population. Repeating the experiment in a population in which males account for the majority may reveal some new insights. Second, the speech rate range was not wide enough for some participants with high cognition function to provide a difficulty level beyond their capability. Having items at such levels is preferred to demonstrate the subjective control on selecting items to which to respond. Since all levels of speech rate were manageable for some participants, the interaction effect of speech rate and reward might have been compromised to some extent. An individualized task demand range should be used in future studies. Third, the task demand-reward net value contrast between items was not explicitly distinguishable enough, this could be the reason why effort did not seem to be allocated as a function of the speech rate-reward net value. Fewer independent variable levels and/or larger increments between levels than used in this study might be desired. Fourth, the experiment did not collect the quantitative data of self-report effort allocation mainly due to the built-in time restrictions and the large amount of trials. This has limited our ability to compare the objective and subjective measure of effort allocation. Future studies could implement a self-report at the end of each block, for example, ask participants to assign a retrospective estimated effort expenditure to each of the speech rate-reward conditions in that block.

Directions for future research

The present study has made the first step testing the theories of active effort control in the context of audiology, and evidenced that young normal hearing adults allocate their listening effort depending on both stimulus-driven factors and goal-driven factors. There are four directions for future research based on the findings in this study.

First, assessing whether different populations will demonstrate similar patterns of listening effort allocation using the same task demand/reward paradigm. Those populations include people with different levels of hearing sensitivity, different age, different culture background, etc. Abundant evidence in the literature has shown the relationship between the task demand and listening effort in populations with different age and hearing sensitivity, however, the impact of goal-driven factors on listening effort control has not been empirically documented. Given the potential negative consequences of excessive and sustained effort in people with hearing loss to cognition, general health, well-being and quality of life in the long-term (Pichora-Fuller, 2016), it is of particular importance to know whether the hearing loss actually alters listening effort allocation behavior, if this is the case, it is of interest to know, whether it is due to the changes in their cost-benefit judgement and whether their judgement can be recalibrated to normal. Furthermore, it is worth studying how the intrinsic motivation and reward value judgement may be influenced by additional factors such as age and ethnic background which are thought to relate to performance and choice of activities (Deci, Koestner, & Ryan, 1999; Iyengar & Lepper, 1999; Wigfield & Eccles, 2000).

Second, investigating the neuronal mechanism of listening effort regulation is of interest. So far, little is known regarding the functional brain circuits that are selectively engaged during motivated listening. The reward-modulated listening process should represent a highly coordinated and interactive neural network among the auditory, reward and affect systems rather than a single-directional pathway from the auditory cortex to other brain regions as documented in the previous neuroimaging studies using auditory stimulation. Future research in building the neurocognitive framework of listening effort regulation is necessary.

Third, examining the effects of performance feedback and/or affective feedback on listening effort regulation is warranted. The present study has tapped an important determinant of high level effort control, yet has not looked in depth into how exactly the effort control is executed. According to Hockey's (1997) model, the effort control is achieved by comparing target output values with current activity (in this model through an action monitor), and changing the output until the discrepancy is removed (or kept within acceptable limits of error tolerance). It implies that the knowledge of results on task performance and/or the perception of success on a task might affect the choice of coping strategy, but it has not been studied in the audiology field. It is important to examine whether the population with hearing loss relies on and benefits from the performance and affective feedback more than their normal hearing peers in order to achieve effective communication.

Fourth, developing a clinical instrument to predict listening effort efficiency would be an important contribution. In parallel to the effort of establishing a clinic-friendly objective test of listening effort to meet the need for precise characterization of effort allocation and its efficiency, exploring the association between the available effort-related survey tests and effort efficiency could help to develop an alternative way of assessing listening effort. For example, researchers can construct a battery of survey tests to estimate the quality of effort expenditure. The topics of surveys that are relevant in the scope of effort would include cognitive resource allocation, reward attitude, intrinsic and extrinsic motivation, value judgement, coping strategy, etc. More ideally, an integrated survey test specific to listening behavior could be created to directly measure the multi-dimensional attributes of effort in the audiology clinic.

4.3 CONCLUSION

Findings of the current study enhance our understanding of the strategic resource allocation behavior during effortful listening among normal hearing young adults, and have provided support to the theory of effort regulation that highlights the interactive roles of task demand and intention as important to the key process of effort monitoring and control (Hockey, 1997).

Results have suggested that it is critical to account for the effect of goal-driven factors such as motivation when studying listening effort.

Redefining effort efficiency in this study raises the attention of future investigation on the quality of listening effort allocation and of implementing the efficiency concept in the audiologic intervention. Results of this study suggest that adding the goal-oriented performance variables in the effort efficiency calculation has the advantage of differentiating individuals in terms of how well the effort is spent in a task.

Evidence from the descriptive data illustrated in this study provide insights into the role of motivation in assessing individuals' capacity range, the role of task errors in demonstrating the strategic behavior in a cost-benefit decision making listening process, and the possible underlying cause of the disassociation between the objective and subjective measure of listening effort observed in previous literature.

In summary, the current study advances theoretical and methodological considerations for listening effort research. Although the results of these experiment are limited to listeners with normal hearing, the findings are encouraging and warrant further investigation of seeking to integrate current models of speech comprehension, resource allocation, and effort regulation to help optimize listening effort for individuals with hearing loss.

APPENDIX A

CASE HISTORY FORM

University of Pittsburgh

Case history form

Date: ____☐___/________/_______.

Last Name:________________________First Name:____________________M.I.___________.

Sex: Male ☐Female Dominant hand: ☐ Left ☐Right

Race/ethnic group: ☐ Caucasion, Non-Hispanic ☐Hispanic ☐ African-American ☐ Asian

☐ Native American ☐Other

General health: ☐ Good ☐Fair ☐ Poor

Birth Date:_______/______/________. Age:__________.Native language:______________.

Phone number:_______________________________. Email:_________________________.

Address:____.

Otologic history

Surgery:____;

Most recent otologic exam:____.

Ophthalmologic history

Surgery:____;

Most recent ophthalmologic exam:____.
Better side: ☐ Left ☐Right

APPENDIX B

THE NEED FOR COGNITION SCALE

For each of the statements below, please indicate whether or not the statement is characteristic of you or of what you believe. For example, if the statement is extremely uncharacteristic of you or of what you believe about yourself (not at all like you) please place a "1" on the line to the left of the statement. If the statement is extremely characteristic of you or of what you believe about yourself (very much like you) please place a "5" on the line to the left of the statement. You should use the following scale as you rate each of the statements below.

1. I prefer complex to simple problems.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

2. I like to have the responsibility of handling a situation that requires a lot of thinking.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

3. Thinking is not my idea of fun.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

5. I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

6. I find satisfaction in deliberating hard and for long hours.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

7. I only think as hard as I have to.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

8. I prefer to think about small daily projects to long term ones.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

9. I like tasks that require little thought once I’ve learned them.

1	2	3	4	5
extremely uncharacteristic of me	somewhat uncharacteristic of me	uncertain	somewhat characteristic of me	extremely characteristic of me

10. The idea of relying on thought to make my way to the top appeals to me.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

11. I really enjoy a task that involves coming up with new solutions to problems.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

12. Learning new ways to think doesn't excite me very much.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

13. I prefer my life to be filled with puzzles I must solve.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

14. The notion of thinking abstractly is appealing to me.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

16. I feel relief rather than satisfaction after completing a task that requires a lot of mental effort.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

17. It's enough for me that something gets the job done; I don't care how or why it works.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

18. I usually end up deliberating about issues even when they do not affect me personally.

1 extremely uncharacteristic of me	2 somewhat uncharacteristic of me	3 uncertain	4 somewhat characteristic of me	5 extremely characteristic of me

BIBLIOGRAPHY

Ahern, SK., & Beatty, J. (1979). Physiological signs of information processing vary with intelligence. *Science, 205*, 1289-1292.

Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Berger, A. C., & Cohen, R. J. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science (New York, N.Y.), 213*(4504), 220-222.

Amiez, C., Joseph, J. P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex, 16*(7), 1040-1055.

Andersson, U., & Lyxell, B. (1999). Phonological deterioration in adults with an acquired severe hearing impairment: a deterioration in long-term memory or working memory? *Scandinavian Audiology, 28*(4), 241-247.

Andersson, Ulf. (2002). Deterioration of the phonological processing skills in adults with an acquired severe hearing loss. *European Journal of Cognitive Psychology, 14*(3), 335-352.

Andreassi, J. L. (2000). *Psychophysiology human behavior and physiological response*. New Jersey: Lawrence Erlbaum Associates.

ANSI, ANSI. (2004). S3. 6 Specification for Audiometers. *Society of America, New York, NY, USA.*

Antikainen, J., & Niemi, P. (1983). Neuroticism and the pupillary response to a brief exposure to noise. *Biological Psychology, 17*(2), 131-135.

Apoux, F., Crouzet, O., & Lorenzi, C. (2001). Temporal envelope expansion of speech in noise for normal-hearing and hearing-impaired listeners: effects on identification performance and response times. *Hearing Research, 153*(1-2), 123-131.

Apps, M. A., & Ramnani, N. (2014). The Anterior Cingulate Gyrus Signals the Net Value of Others' Rewards. *The Journal of Neuroscience, 34*(18), 6190-6200.

ASHA. (2005). Guidelines for manual pure-tone threshold audiometry [Guidelines]. *Available from www.asha.org/policy.*

Awh, E., Vogel, E.K., & Oh, S.H. (2006). Interactions between attention and working memory. *Neuroscience, 139*(1), 201-208.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation, 8*, 47-89.

Bailey, I. L., & Lovie, JE. (1980). The design and use of a new near-vision chart. *American Journal of Optometry and Physiological Optics, 57*(6), 378-387.

Barth, M., & Poser, B. A. (2011). Advances in High-Field BOLD fMRI. *Materials, 4*(12), 1941-1955.

Beasley, D.S., & Freeman, B.A. (1977). Time altered speech as a measure of central auditory processing. *Central Auditory Dysfunction*, 223-256.

Beasley, D.S., & Maki, J.E. (1976). Time-and frequency-altered speech. *Contemporary issues in experimental phonetics*, *12*, 419-458.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276-292.

Beatty, J., & Lucero-Wagoner, B. (2000). The Pupillary System. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 142-162). Cambridge: Cambridge University Press.

Beer, J., Kronenberger, W. G., & Pisoni, D. B. (2011). Executive function in everyday life: implications for young cochlear implant users. *Cochlear Implants International*, *12 Suppl 1*, S89-91.

Bernarding, C., Corona-Strauss, F. I., Latzel, M., & Strauss, D. J. (2010). Auditory streaming and listening effort: an event related potential study. *Conf Proc IEEE Eng Med Biol Soc, 2010*, 6817-6820. doi: 10.1109/iembs.2010.5625957

Bernarding, C., Latzel, M., Strauss, D. J., & Corona-Strauss, F. I. (2011). On the Objective Electrophysiological estimation of Listening Effort in Different Age Group. *Conf Proc IEEE Eng Med Biol Soc, 2011*, 671-674.

Bernarding, C., Strauss, D. J., Hannemann, R., Latzel, M., Seidler, H., Jobst, U., . . . Corona-Strauss, F. I. (2011). The effects of age and hearing impairment on the extraction of listening effort correlates. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, 2011*, 2323-2326.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Cardiac Psychophysiology and Autonomic Space in Humans: Empirical Perspectives and Conceptual Implications. *Psychological Bulletin*, *114*(2), 296-322.

Betsch, T., & Haberstroh, S. (2014). *The routines of decision making*: Psychology Press.

Binder, L.M. (1993). *Portland Digit Recognition Test Manual* (2 ed.). Portland, OR: Author.

Bland, B.H., & Vandenwolf, C.H. (1972). DiencepWc and hippocampal mechanisms of motor activity in the rat: Effects of posterior hypothalamic stimulation on behavior and hippocampal slow wave activity. *Brain Research, 43*, 67-88.

Blanton, Richard L., & Nunnally, Jum C. (1964). Evaluational Language Processes in the Deaf. *Psychological Reports, 15*(3), 891.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(5), 201-208.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539-546.

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1), 16-27.

Boussaoud, D., & Wise, S. P. (1993). Primate frontal cortex: effects of stimulus and movement. *Experimental Brain Research, 95*(1), 28-40.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral cortex (New York, N.Y. : 1991), 11*(9), 825-836.

Brehm, Jack. W. , & Self, Elizabeth. A. . (1989). The Intensity of Motivation. *Annual Review of Psychology, 40*, 109-131.

Broadbent, D. (1958). *Perception and communication*. New York: Pergamon Press.

Broadbent, D. E. (1971). *Decision and stress*. London: Academic Press.

Brodal, A. (1981). *Neurological anatomy in relation to clinical medicine* (3rd ed.). New York: Oxford University Press.

Brons, I., Houben, R., & Dreschler, W. A. (2012). Perceptual Effects of Noise Reduction With Respect to Personal Preference, Speech Intelligibility, and Listening Effort. *Ear and Hearing*.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118-1121.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 15*(1 Pt 1), 12.

Butler, A. B., Scherer, L. L., & Reiter-Palmon, R. (2003). Effects of solution elicitation aids and need for cognition on the generation of solutions to ill-structured problems. *Creativity Research Journal, 15*(2-3), 235-244.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage, 17*(3), 1394-1402.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116.

Cacioppo, J. T., & Petty, R. E. (1984). The need for cognition: Relationship to attitudinal processes. *Social perception in clinical and counseling psychology, 2*, 113-140.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B.G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197.

Calvert, D. R., & Silverman, S. R. (1983). *Speech and deafness: a text for learning and teaching*: Alex Graham Bell Assn for Deaf.

Camp, G., Paas, F., Rikers, R., & van Merrienboer, J. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior, 17*(5), 575-595.

Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 363*(1493), 1001-1010.

Capa, R. L., Audiffren, M., & Ragot, S. (2008). The interactive effect of achievement motivation and task difficulty on mental effort. *International Journal of Psychophysiology, 70*(2), 144-150.

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage, 57*(4), 1608-1616.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749.

Chang, C. (2007). Diagnostic advertising content and individual differences: Testing a resource-matching perspective with a Taiwanese sample. *Journal of Advertising, 36*(3), 75-84.

Chinien, C. A., & Boutin, F. (1993). Cognitive style FD/I: An important learner characteristic for educational technologists. *Journal of Educational Technology Systems, 21*(4), 303-311.

Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language, and Hearing Research, 51*(4), 1042-1054.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology, 51*(2), 291.

Cohen, R. M., Semple, W. E., Gross, M., Holcomb, H. H., Dowling, M.S., & Nordahl, T. E. (1988). Functional Localization of Sustained Attention: Comparison to Sensory Stimulation in the Absence of Instruction. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 1*(1), 3-20.

Cox, R.M., Alexander, G. C., & Xu, J.J. (2009). Development of the Device-Iriented Subjective outcome (DOSO) Scale. from http://www.memphis.edu/csd/harl/downloads/posters/AAS-cox.pdf

Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B. K., . . . Dolan, R. J. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain, 126*(Pt 10), 2139-2152.

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E., & Rushworth, M. F. (2009). Effort-based cost-benefit valuation and the human brain. *Journal of Neuroscience, 29*(14), 4531-4541.

Dahan, D. (2010). The time course of interpretation in speech comprehension. *Current Directions in Psychological Science, 19*(2), 121-126.

Dahan, D., & Tanenhaus, M. K. (2004a). Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 498-513.

Dahan, D., & Tanenhaus, M.K. (2004b). Continous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constrains. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 498-513.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A, 2*(7), 1160-1169.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 23*(8), 3423-3431.

Dawson, M.E., Schell, A.M., & Filion, D.L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed.): Cambridge University Press.

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

Demorest, M. E., & Erdman, S. A. (1986). Scale composition and item analysis of the Communication Profile for the Hearing Impaired. *Journal of Speech and Hearing Research, 29*, 515-535.

Desjardins, J. L., & Doherty, K. A. (2012). Age-Related Changes in Listening Effort for Various Types of Masker Noises. *Ear and Hearing*.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain, 118 (Pt 1)*, 279-306.

Douglas, R. J., Barrett, T. W., Pribram, K. H., & Cerny, M. C. (1969). Limbic lesions and error reduction. *Journal of Comparative and Physiological Psychology, 68*(3), 437-441.

Downs, D. W. (1982). Effects of hearing and use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders, 47*(2), 189-193.

Downs, D. W., & Crum, M. A. (1978). Processing demands during auditory learning under degraded listening conditions. *Journal of Speech and Hearing Research, 21*(4), 702-714.

Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology, 74*, 24-33.

Eberwein, C. A., Pratt, S. R., McNeil, M. R., Fossett, T. RD., Szuminsky, N. J., & Doyle, P. J. (2007). Auditory performance characteristics of the computerized Revised Token Test (CRTT). *Journal of Speech, Language, and Hearing Research, 50*(4), 865-877.

Edwards, B. (2007). The future of hearing aid technology. *Trends in Amplification, 11*(1), 31-45.

Eisler, F. G. (1968). *Psycholinguistics: Experiments in spontaneous speech*: Academic Press.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Erez, M., Gopher, D., & Arzi, N. (1990). Effects of Goal Difficulty, Self-Set Goals, and Monetary Rewards on Dual Task Performance *Organizational Behavior and Human decision Processes*, *47*, 247-269.

Fairclough, S.H. (2001). Mental effort regulation and the functional impairment of the driver. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue*. Mahwah, NJ: L. Erlbaum.

Fallon, M., Peelle, J. E., & Wingfield, A. (2006). Spoken sentence processing in young and older adults modulated by task demands: evidence from self-paced listening. *The journals of gerontology. Series B, Psychological sciences and social sciences, 61*(1), P10-17.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*(2), 304.

Fellows, L. K., & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control? *Brain : a journal of neurology, 128*(Pt 4), 788-796.

Ferreira, F., Anes, M. D., & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, *25*(2), 273-290.

Feuerstein, J. F. (1992). Monaural versus binaural hearing: ease of listening, word recognition, and attentional effort. *Ear and Hearing, 13*(2), 80-86.

Fiebert, M. (1967). Cognitive styles in the deaf. *Perceptual and Motor Skills, 24*, 319-329.

Figner, B., & Murphy, R. O. (2010). Using skin conductance in judgment and decision making research. In A. K. M. Schulte-Mecklenbeck, & R. Ranyard (Ed.), *A handbook of process tracing methods for decision research.* New York, NY: Psychology Press.

Florence, N.N., Liziana N, O., & UCHE N, E. (2012). Effects of problem-based learning and cognitive style on basic science achievement and self-efficacy beliefs of students with hearing impairment. *International Journal of Educational Science and Research, 2*(3), 21-36.

Foo, C., Rudner, M., Ronnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology, 18*(7), 618-631.

Forstmann, B. U., Brass, M., Koch, I., & von Cramon, D. Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience, 18*(3), 388-398.

Fougnie, D. (2009). The relationship between attention and working memory. In I. N. B. Johansen (Ed.), *New research on short-term memory*. New York: Nova Science Publishers.

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., & Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. *Frontiers in Psychology, 7*.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501.

Fraser, S., Gagné, J. P., Alepins, M., & Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research*, *53*(1), 18-33.

Frazier, L. (1987). Sentence processing: A tutorial review. In I. M. Coltheart (Ed.), *Attention and Performance XII* (pp. 559-586). Hillsdale NJ: Lawrence Erlbaum Associates.

Freeman, G.L. (1948). *The energetics of human Behavior*. NY: Cornell University Press.

Gatehouse, S., & Akeroyd, M. (2006). Two-eared listening in dynamic situations. *International Journal of Audiology, 45*(Suppl.1), S120-S124.

Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, *24*(1), 63-68.

Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology, 43*(2), 85-85.

Gellatly, I. R., & Meyer, J. P. (1992). The Effects of Goal Difficulty on Physiological Arousal, Cognition, and Task Performance. *Journal of Applied Psychology, 77*(5), 694-704.

Gendolla, G. H., Richter, M., & Silvia, P. J. (2008). Self-focus and task difficulty effects on effort-related cardiovascular reactivity. *Psychophysiology, 45*(4), 653-662.

Genno, H., Ishikawa, K., Kanbara, O., & Kikumoto, M. (1997). Using facial skin temperature to objectively evaluate sensations. *International Journal of Industrial Ergonomics, 19*, 161-171.

Glickman, M. E., Gray, J. R., & Morales, C. J. (2005). Combining speed and accuracy to assess error-free cognitive processes. *Psychometrika, 70*(3), 405-425.

Globerson, T. (1983). Mental Capacity, Mental Effort, and Cognitive Style. *Developmental Review, 3*, 292-302.

Globerson, T. (1985). Field Dependence/Independence and Mental Capacity: A Developmental Approach. *Developmental Review, 5*, 261-273.

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin, 77*(5), 340-355.

Gopher, D. (1994). Analysis and measurement of mental load. In e. a. G. D'Yderwalle (Ed.), *International perspectives on psychological science, Vol. 2: The state of the art* (pp. 265-291). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gordon, K. A., Wong, D. D., Valero, J., Jewell, S. F., Yoo, P., & Papsin, B. C. (2011). Use It or Lose It? Lessons Learned from the Developing Brains of Children Who are Deaf and Use Cochlear Implants to Hear. *Brain Topogr*.

Gosselin, P. A., & Gagné, J. P. (2011a). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *International Journal of Audiology, 50*(11), 786-792.

Gosselin, P. A., & Gagné, J. P. (2011b). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research, 54*(3), 944-958.

Grady, C. L. (2000). Functional brain imaging and age-related changes in cognition. *Biological Psychology, 54*(1-3), 259-281.

Graham, L.M. (2007). Need for cognition and false memory in the Deese–Roediger–McDermott paradigm. *Personality and Individual Differences, 42*(3), 409-418.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*(4), 457-461.

Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology, 52*(1), 1-6.

Granholm, E., Verney, S. P., Perivoliotis, D., & Miura, T. (2007). Effortful cognitive resource allocation and negative symptom severity in chronic schizophrenia. *Schizophrenia Bulletin, 33*(3), 831-842.

Green, P., Allen, I., & Astner, K. (1996). *Manual for Computerized Word Memory Test*. Durham, NC CogniSyst.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M., 3rd. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*(12), 1045-1060.

Greenberg, S. (1996). *Understanding speech understanding: Towards a unified theory of speech perception.* Paper presented at the ESCA Workshop on the "Auditory Basis of Speech Perception", Keele University, pp 1-8.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesic measures with a large clinical sample. *Psychological Assessment*, *6*, 218–224.

Guisande, M.A., Páramo, M.F., Tinajero, C., & Almeida, L.S. (2007). Field dependence-independence (FDI) cognitive style: An analysis of attentional functioning. *Psicothema, 19*(4), 572-577.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240-244.

Hafter, E. R. (2010). Is there a hearing aid for the thinking person? *Journal of the American Academy of Audiology, 21*(9), 594-600.

Haggard, P., Cockburn, J., Cock, J., Fordham, C., & Wade, D. (2000). Interference between gait and cognitive tasks in a rehabilitating neurological population. *Journal of Neurology, Neurosurgery and Psychiatry, 69*(4), 479-486.

Haier, R. J., Siegel, B. V, Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., . . . Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence, 12*, 199-217.

Haiman, J. (1980). The iconicity of grammar: isomorphism and motivation. *Language*, 515-540.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., . . . Bowtell, R. W. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping*, *7*(3), 213-223.

Hall, D. A., Summerfield, A. Q., Gonçalves, M. S., Foster, J. R., Palmer, A. R., & Bowtell, R. W. (2000). Time-course of the auditory BOLD response to scanner noise. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *43*(4), 601-606.

Hardy, A. P. (1982). The selection of channels when seeking information: Cost/benefit vs least-effort. *Information Processing & Management, 18*(6), 289-293.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press.

Hatzigeorgiadis, A. (2006). Approach and avoidance coping during task performance in young men: the role of goal attainment expectancies. *Journal of Sports Sciences, 24*(3), 299-307.

Haugtvedt, C. P., Petty, R. E., & Cacioppo, J. T. (1992). Need for cognition and advertising: Understanding the role of personality variables in consumer behavior. *Journal of Consumer Psychology, 1*(3), 239-260.

Hauthal, N., Neumann, M. F., & Schweinberger, S. R. (2012). Attentional spread in deaf and hearing participants: face and object distractor processing under perceptual load. *Attention, Perception, & Psychophysics, 74*(6), 1312-1320.

Hayes, J., & Allinson, C.W. (1998). Cognitive style and the theory and practice of individual and collective learning in organizations. *Human Relations, 51,* 847-871.

He, S. Q., Dum, R. P., & Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *Journal of Neuroscience, 15*(5), 3284.

Heilman, L. E. (2008). *An examination of the effects of mode of access on the Computerized Revised Token Test.* University of Pittsburgh.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology, 34*(2), 163-171.

Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research, 45*(3), 573-584.

Hillman, K. L., & Bilkey, D. K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. *Nature Neuroscience, 15*(9), 1290-1297.

Hockey, G. R. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetical framework. *Biological Psychology, 45*(1-3), 73-93.

Hockey, G. R. J., & Earle, F. (2006). Control over the scheduling of simulated office work reduces the impact of workload on mental fatigue and task performance. *Journal of Experimental Psychology: Applied, 12*(1), 50-65.

Hockey, G.R. J., Gaillard, A. W. K., & Coles, M. G. H. (1986). *Energetics and Human Information Processing*. Dordrecht: Martinus Nijhoff.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*(2), 122-128.

Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology, 49*(12), 928-932.

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology, 18*(7), 590-603.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: a theory of the relationship between individual differences and information processing. *Psychological Review, 91*(2), 153-184.

Hwang, J. H., Li, C. W., Wu, C. W., Chen, J. H., & Liu, T. C. (2007). Aging Effects on the Activation of the Auditory Cortex during Binaural Speech Listening in White Noise: An fMRI Study. *Audiology and Neurotology, 12*(5), 285-294.

Iani, C., Gopher, D., & Lavie, P. (2004). Effects of task difficulty and invested mental effort on peripheral vasoconstriction. *Psychophysiology, 41*(5), 789-798.

Iverson, G. L. (2010). Mild traumatic brain injury meta-analyses can obscure individual differences. *Brain Injury, 24*(10), 1246-1255.

Iverson, P., Brint, E., Song, J., & Wu, M. (2016). Listening effort and accent in competing-talker speech recognition by native and non-native listeners: Electrophysiological measures of auditory and lexical processing. *The Journal of the Acoustical Society of America, 140*(4), 3436-3437.

Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: a cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology, 76*(3), 349.

Janisse, M. P. (1977). *Pupillometry. The Psychology of the Pupillary Response*. Washington, D. C.: Hemisphere Publishing Corporation.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122-149.

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

Karatekin, C., Couperus, J.W., & Marcus, D. J. (2004). Attention allocation in the dual‐task paradigm as measured through behavioral and psychophysiological responses. *Psychophysiology, 41*(2), 175-185.

Kawahara, H., Masuda-Katsuse, I., & De Cheveigne, A. (1999). Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech communication, 27*(3), 187-207.

Kearney, E., Gebert, D., & Voelpel, S. C. (2009). When and how diversity benefits teams: The importance of team members' need for cognition. *Academy of Management Journal, 52*(3), 581-598.

Kemper, S., Schmalzried, R., Hoffman, L., & Herman, R. (2010). Aging and the Vulnerability of Speech to Dual Task Demands. *Psychology and Aging, 25*(4), 949-962.

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience, 21*(6), 1162-1178.

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience, 9*(7), 940-947.

Kerr, B. (1973). Processing demands during mental operations. *Memory and Cognition, 1*(4), 401-412.

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., . . . von Wedel, H. (2003). Candidature for and delivery of audiological services: special needs of older people. *International Journal of Audiology, 42 Suppl 2,* 2S92-101.

Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 169-226). Cambridge: MIT Press.

Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B.G., Zekveld, A.A., & Kramer, S. E. (2015). The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing Research, 323,* 81-90.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2011). Pupil Dilation Uncovers Extra Listening Effort in the Presence of a Single-Talker Masker. *Ear and Hearing*.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*(3), 557-577.

Kotz, S. A., von Cramon, Y. D., & Friederici, A. D. (2005). On the role of phonological short-term memory in sentence processing: ERP single case evidence on modality-specific effects. *Cognitive Neuropsychology, 22*(8), 931-958.

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: toward an integrated framework of cognitive style. *Psychological Bulletin, 133*(3), 464-481.

Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *Audiology, 36*(3), 155-164.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*(3), 213-236.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*(1), 537-567.

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *Int J Psychophysiol, 65*(2), 132-140.

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, Jr K. I., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology, 50*(1), 23-34.

Kuchinsky, S. E., Eckert, M. A., & Dubno, J. R. (2011). The eyes are the windows to the ears: pupil size reflects listening effort. *Audiology Today, 23*(1), 56.

Lange, R. T., Pancholi, S., Bhagwat, A., Anderson-Barnes, V., & French, L. M. (2012). Influence of poor effort on neuropsychological test performance in U.S. military personnel following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 34*(5), 453-466.

Larsby, B., Hallgren, M., & Lyxell, B. (2008). The interference of different background noises on speech processing in elderly hearing impaired subjects. *International Journal of Audiology, 47 Suppl 2*, S83-90.

Larsby, B., Hallgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44*(3), 131-143.

Lee, D. TL., & Yamamoto, A. (1994). Wavelet analysis: theory and applications. *Hewlett Packard journal, 45*, 44-44.

Liu, S., & Zeng, F. G. (2006). Temporal properties in clear speech perception. *The Journal of the Acoustical Society of America, 120*(1), 424-432.

Loewenfeld, IE. (1999). *The Pupil: Anatomy, Physiology and Clinical Applications*. Boston: Butterworth Heinemann.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 18866-18869.

Løvstad, M., Solbakk, A. K., Funderud, I., Meling, T., Krämer, U. M., Voytek, B., . . . Knight, R. T. (2012). Anterior cingulate cortex and cognitive control: neuropsychological and electrophysiological findings in two patients with lesions to dorsomedial prefrontal cortex. *Brain and Cognition, 80*(2), 237-249.

Luria, A.R. (1973). *The working brain*. New York: Basic Books.

Luts, H., Froehlich, M., Puder, H., Grimm, G., Hohmann, V., Leijon, A., . . . Ljud- och, B. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *The Journal of the Acoustical Society of America, 127*(3), 1491-1505.

Lyxell, B., Andersson, J., Andersson, U., Arlinger, S., Bredberg, G., & Harder, H. (1998). Phonological representation and speech understanding with cochlear implants in deafened adults. *Scandinavian Journal of Psychology, 39*(3), 175-179.

Lyxell, B., Andersson, J., Arlinger, S., Bredberg, G., Harder, H., & Ronnberg, J. (1996). Verbal information-processing capabilities and cochlear implants: implications for preoperative predictors of speech understanding. *J Deaf Stud Deaf Educ, 1*(3), 190-201.

Lyxell, B., & Rönnberg, J. (1989). Information-processing skill and speech-reading. *British Journal of Audiology, 23*(4), 339-347.

Lyxell, Björn, & Rönnberg, Jerker. (1991). Visual speech processing: Word‐decoding and word‐discrimination related to sentence‐based speechreading and hearing‐impairment. *Scandinavian Journal of Psychology, 32*(1), 9-17.

Mackersie, C. L. , Boothroyd, A., & Prida, T. (2000). Use of a simultaneous sentence perception test to enhance sensitivity to ease of listening. *Journal of Speech, Language, and Hearing Research, 43*, 675-682.

Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology, 22*(2), 113-122.

Mackersie, C., Neuman, A. C., & Levitt, H. (1999). Response time and word recognition using a Modified-Rhyme monitoring task: list equivalency and time-order effects. *Ear and Hearing, 20*(6), 515.

Mars, R. B., Grol, M. J., Toni, I., Nieuwenhuis, S., Holroyd, C. B., Hulstijn, W., & Coles, M. G. H. (2005). Neural dynamics of error processing in medial frontal cortex. *Neuroimage, 28*(4), 1007-1013.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science, 189*(4198), 226-228.

McAuliffe, M. J., Wilding, P. J., Rickard, N. A., & O'Beirne, G. A. (2012). Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss. *Journal of Speech, Language, and Hearing Research*.

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

McFadden, B., & Pittman, A. (2008). Effect of Minimal Hearing Loss on Children's Ability to Multitask in Quiet and in Noise. *Language, Speech, and Hearing Services in Schools, 39*(3), 342-351.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

McNeil, M., Doyle, P., Hula, W., Rubinsky, H., Fossett, T., & Matthews, C. (2004). Using resource allocation theory and dual‐task methods to increase the sensitivity of assessment in aphasia. *Aphasiology*, *18*(5-7), 521-542.

McNeil, M. R., Pratt, S. R., Szuminsky, N., Sung, J. E., Fossett, T. R.D., Fassbinder, W., & Lim, K. Y. (2015). Reliability and validity of the Computerized Revised Token Test (CRTT): Comparison of reading and listening versions in persons with and without aphasia. *Journal of Speech, Language, and Hearing Research*.

McNeil, M.R., Kim, A., Sung, J.E., Pratt, S.R., Szuminsky, N., & Doyle, P.J. (2009). *A comparison of left versus right hand, and mouse versus touchscreen access methods on the Computerized Revised Token Test in normal adults and persons with aphasia.* Paper presented at the Clinical Aphasiology Conference.

Mead, R. A., & Lapidus, L. B. (1989). Psychological differentiation, arousal, and lipreading efficiency in hearing-impaired and normal children. *Journal of Clinical Psychology, 45*(6), 851-859.

Merriam-Webster's Collegiate Dictionary. (Ed.) (1997) (10th ed.). Springfield, MA: Merriam-Webster.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Moore, B. C. (1985). Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *British Journal of Audiology, 19*(3), 189-201.

Moore, B. C. (1993). Temporal analysis in normal and impaired hearing. *Annals of the New York Academy of Sciences, 682*, 119-136.

Moore, B.C.J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues* (2 ed.). Chichester: Wiley.

Moore, D. R. (1991). Anatomy and physiology of binaural hearing. *Audiology, 30*(3), 125-134.
Moulines, E., & Laroche, J. (1995). Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech communication*, *16*(2), 175-205.

Mukherjee, S., Yadav, R., Yung, I., Zajdel, D. P., & Oken, B. S. (2011). Sensitivity to mental effort and test-retest reliability of heart rate variability measures in healthy seniors. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 122*(10), 2059-2066.

Mulder, G. (1986). The concept and measurement of mental effort. In G. R. J. Hockey, A. W. K. Gaillard & M. G. H. Coles (Eds.), *Energetics and Human Information Processing*. Dordrecht: Martinus Nijhoff.

Mulder, L.J.M., Van Roon, A.M., Veldman, J.B.P., Elgersma, A.F., & Mulder, G. (1995). Respiratory pattern, invested effort, and variability in heart rate and blood pressure during the performance of mental tasks. In M. Di Rienzo, G. Mancia, G. Parati, A. Pedotti & A. Zanchetti (Eds.), Computer analysis of Cardiovascular Signals. Amsterdam: IOS Press.

Näätänen, R. (2008). The mismatch negativity (MMN) as an index of auditory memory. *International Journal of Psychophysiology*, *69*, 167.

Näätänen, R., & Alho, K. (1997). Mismatch negativity--the measure for central sound representation accuracy. *Audiology and Neuro-Otology*, *2*(5), 341-353.

Naccache, L., Dehaene, S., Cohen, L., Habert, M. O., Guichart-Gomez, E., Galanaud, D., & Willer, J. C. (2005). Effortless control: executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia*, *43*(9), 1318-1328.

Nachev, P. (2011). The blind executive. *Neuroimage, 57*(2), 312-313.

Naqvi, N. H., & Bechara, A. (2006). Skin conductance: A psychophysiological approach to the study of decision making. In T. R. M. S. G. C. Senior (Ed.), *Methods in mind*: The MIT Press.

Navon, D., & Gopher, D. (1979a). On the economy of the human-processing system. *Psychological Review*, *86*(3), 214-255.

Navon, D., & Gopher, D. (1979b). On the economy of the human processing system. *Psychological Review*, *86*, 214-255.

Nemeroff, C. B. (1998). The neurobiology of depression. *Scientific American, 278*(6), 42-49.

Niimi, M., & Nishio, S. (2001). Speaking rate and its components in dysarthric speakers. *Clinical Linguistics & Phonetics, 15*(4), 309-317.

Nocetti, D. (2005). A Model of Mental Effort and Endogenous Estimation Risk. *Economics Bulletin, 4*(14), 1-10.

Norman, D.A., & Bobrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Nunnally, J.C., Knott, P.D., Duchnowski, A., & Parker, R. (1967). Pupillary response as a general measure of activation. *Perception and Psychophysics, 2*, 149-155.

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., . . . Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing Research*.

Ohsuga, M., Shimono, F., & Genno, H. (2001). Assessment of phasic work stress using autonomic indices. *Int J Psychophysiol, 40*(3), 211-220.

Okusa, M., Shiraishi, T., Kubo, T., & Nageishi, Y. (1999). Effects of discrimination difficulty on cognitive event-related brain potentials in patients with cochlear implants. *Otolaryngology - Head and Neck Surgery, 121*(5), 610-615.

Or, C.K.L, & Duffy, V.G. (2007). Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement *Occupational Ergonomics, 7,* 83-94.

Ord, J. S., Greve, K. W., Bianchini, K. J., & Aguerrevere, L. E. (2010). Executive dysfunction in traumatic brain injury: the effects of injury severity and effort on the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 32*(2), 132-140.

Paas, F. GWC., & Van Merriënboer, J. JG. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 35*(4), 737-743.

Panel, National Reading. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. ((NIH Publication No. 00- 4769)). Washington, D.C.: U.S. Government Printing Office.

Parasnis, I., & Long, G. L. (1979). Relationships among spatial skills, communication skills, and field independence in deaf students. *Perceptual and Motor Skills, 49*(3), 879-887.

Pascual-Leone, J. (1970). A Mathematical Model for the Transition Rule in Piaget's Developmental Stages. *Acta Psychologica, 32*, 301-345.

Pashler, H. (1992). Attentional Limitations in Doing Two Tasks at the Same Time. *Current Directions in Psychological Science, 1*(2), 44-48.

Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., & Davis, M. H. (2010). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *Neuroimage, 52*(4), 1410-1419.

Perfetti, C.A., & Adlof, S.M. (2012). Reading Comprehension: A conceptual framework from word meaning to text meaning. In I. J. Sabatini & E. Albro (Eds.), *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences.* Lanham, MD: Rowman & Littlefield Education.

Perlini, A. H., & Hansen, S. D. (2001). Moderating effects of need for cognition on attractiveness stereotyping. *Social Behavior and Personality: an international journal, 29*(4), 313-321.

Petty, R. E., Briñol, P., Loersch, C., McCaslin, M. J., Leary, M.R., & Hoyle, R.H. (2009). The need for cognition. *Handbook of individual differences in social behavior*, 318-329.

Petty, R. E., Briñol, P., Tormala, Z. L., & Wegener, D. T. (2007). The role of metacognition in social judgment. *Social psychology: Handbook of basic principles, 2*, 254-284.

Petty, R. E., Cacioppo, J.T., Kao, C. F., & Rodrigues, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology, 51*(5), 1032-1043.

Petty, R. E., DeMarree, K. G., Briñol, P., Horcajo, J., & Strathman, A. J. (2008). Need for cognition can magnify or attenuate priming effects in social judgment. *Personality and Social Psychology Bulletin, 34*(7), 900-912.

Pichora-Fuller, M Kathleen. (2016). How Social Psychological Factors May Modulate Auditory and Cognitive Functioning During Listening. *Ear and Hearing, 37,* 92S-100S.

Pichora-Fuller, M Kathleen, Kramer, Sophia E, Eckert, Mark A, Edwards, Brent, Hornsby, Benjamin WY, Humes, Larry E, . . . Mackersie, Carol L. (2016). Hearing impairment and cognitive energy: the Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*, 5S-27S.

Pichora-Fuller, M. K. (2003). Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *International Journal of Audiology, 42 Suppl 1*, S59-67.

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

Pichora-Fuller, M. K., Schneider, B., MacDonald, E., Brown, S., & Pass, H. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research, 223*, 114-121.

Pichora-Fuller, M. K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification, 10*(1), 29-59.

Picou, E. M., Ricketts, T. A., & Hornsby, B. W. (2011). Visual cues and listening effort: individual variability. *Journal of Speech, Language, and Hearing Research, 54*(5), 1416-1430.

Piolat, A., Barbier, M. L., & Roussey, J. Y. (2008). Fluency and Cognitive Effort During First- and Second-Language Notetaking and Writing by Undergraduate Students. *European Psychologist, 13*(2), 114-125.

Piquado, T., Benichov, J. I., Brownell, H., & Wingfield, A. (2012). The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. *International Journal of Audiology, 51*(8), 576-583.

Pitts, M.M. (1983). Comprehension Monitoring: Definition and Practice. *Journal of Reading, 26*(6), 516-523.

Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 363*(1493), 1071-1086.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128-2148.

Posner, M.I., & DiGirolamo, G. . (1998). Executive attention: conflict, target detection and cognitive control. In R. Parasuraman (Ed.), *The attentive brain* (pp. 401-423). Cambridge: MIT Press.

Pribram, K. H., & McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review, 82*(2), 116-149.

Pruessner, J. C., Hellhammer, D. H., & Kirschbaum, C. (1999). Burnout, perceived stress, and cortisol responses to awakening. *Psychosomatic Medicine, 61*(2), 197.

Rabbitt, P. (1990). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Oto-Laryngologica. Supplementum, 476*, 167.

Rabbitt, P.M.A. (1979). Current paradigms and models in human information processing. In V. Hamilton & D. M. Warburt (Eds.), Human stress and cognition (pp. 129). New York: Wiley.

Rakerd, B., Seitz, P. F., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear and Hearing, 17*(2), 97-106.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using Eye Movements to Evaluate the Cognitive Processes Involved in Text Comprehension. *Journal of visualized experiments: JoVE*(83).

Ravicz, M.E., Melcher, J.R., & Kiang, N.Y.S. (2000). Acoustic noise during functional magnetic resonance imaging. *The Journal of the Acoustical Society of America, 108*(4), 1683-1696.

Recarte, M.A., Pérez, E., Conchillo, A., & Nunes, L. M. (2008). Mental workload and visual impairment: differences between pupil, blink, and subjective rating. *The Spanish journal of psychology, 11*(2), 374.

Rönnberg, J. (1990). Cognitive and communicative function: The effects of chronological age and ‘‘handicap age’’. *European Journal of Cogniti6e Psychology, 2*, 253-274.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *International Journal of Audiology, 42 Suppl 1*, S68-76.

Rönnberg, J., Andersson, J., Andersson, U., Johansson, K., Lyxell, B., & Samuelsson, S. (1998). Cognition as a bridge between signal and dialogue: communication in the hearing impaired and deaf. *Scandinavian Audiology, 49*, 101-108.

Rönnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternang, O., Wahlin, A., & Nilsson, L. G. (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *Journal of Speech, Language, and Hearing Research, 54*(2), 705-726.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47 Suppl 2*, S99-105.

Rönnberg, J., Rudner, M., & Lunner, T. (2011). Cognitive Hearing Science: The Legacy of Stuart Gatehouse. *Trends in Amplification*.

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health, 12*(49), 263-269.

Rost, M., & Ross, S. (1991). Learner use of strategies in interaction: Typology and teachability. *Language Learning, 41*(2), 235-268.

Rudner, M., Fransson, P., Ingvar, M., Nyberg, L., & Rönnberg, J. (2007). Neural representation of binding lexical signs and words in the episodic buffer of working memory. *Neuropsychologia, 45*(10), 2258-2276.

Rudner, M., Lunner, T., Behrens, T., Thoren, E. S., & Rönnberg, J. (2012). Working Memory Capacity May Influence Perceived Effort during Aided Speech Recognition in Noise. *Journal of the American Academy of Audiology, 23*(8), 577-589.

Rudner, M., Rönnberg, J., & Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology, 22*(3), 156-167.

Ruggero, M. A., & Rich, N. C. (1991). Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. *Journal of Neuroscience, 11*(4), 1057-1067.

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences, 8*(9), 410-417.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychol (Amst), 53*(1), 61-97.

Santurette, S., & Dau, T. (2012). Relating binaural pitch perception to the individual listener's auditory profile. *Journal of the Acoustical Society of America, 131*(4), 2968-2986.

Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research, 52*(5), 1230-1240.

Schmidt, L., Lebreton, M., Clery-Melin, M. L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology, 10*(2), e1001266.

Schmitter, S., Diesch, E., Amann, M., Kroll, A., Moayer, M., & Schad, L. R. (2008). Silent echo-planar imaging for auditory FMRI. *Magnetic Resonance Materials in Physics, Biology and Medicine, 21*(5), 317-325.

Schnupp, J., Nelken, I., & King, A. J. (2011). *Auditory Neuroscience: Making Sense of Sound*. Cambridge, MA: MIT Press.

Schorn, K., & Zwicker, E. (1990). Frequency selectivity and temporal resolution in patients with various inner ear disorders. *Audiology, 29*(1), 8-20.

Schweimer, J., Saft, S., & Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behavioral Neuroscience, 119*(6), 1687-1692.

Segerstrom, S. C., & Nes, L. S. (2006). When Goals Conflict But People Prosper: The Case of Dispositional Optimism. *Journal of Research in Personality, 40*(5), 675-693.

Seifritz, E., Di Salle, F., Esposito, F., Herdener, M., Neuhoff, J. G., & Scheffler, K. (2006). Enhancing BOLD response in the auditory system by neurophysiologically tuned fMRI sequence. *Neuroimage*, *29*(3), 1013-1022.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283-299.

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*(5), 679-687.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage*, *20*(1), 114-124.

Siegle, G. J., Steinhauer, S. R., & Thase, M. E. (2004). Pupillary assessment and computational modeling of the Stroop task in depression. *International Journal of Psychophysiology*, *52*(1), 63-76.

Siegrist, J. (1998). Adverse health effects of effort–reward imbalance at work:theory, empirical support, and implications for prevention. In I. C. L. Cooper (Ed.), *Theories of organizational stress* (pp. 190-204). Oxford: Ox ford University Press.

Slansky, B. L., & McNeil, M. R. (1997). Resource allocation in auditory processing of emphatically stressed stimuli in aphasia. *Aphasiology*, *11*(4-5), 461-472.

Staal, M. (2004). Stress, cognition, and human performance: A literature review and conceptual framework. from http://humanfactors.arc.nasa.gov/flightcognition/Publications/IH 054 Staal.pdf.

Stelmachowicz, P. G., Lewis, D. E., Choi, S., & Hoover, B. (2007). Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear and Hearing*, *28*(4), 483-494.

Sternberg, S. (1969). The discovery of processing stages: extension of Donders's method. In W. G. Koster (Ed.), *Attention and Performance Ⅱ*. Amsterdam: North Holland.

Stewart, R., & Wingfield, A. (2009). Hearing loss and cognitive effort in older adults' report accuracy for verbal materials. *Journal of the American Academy of Audiology, 20*(2), 147-154.

Strauss, D. J., Corona-Strauss, F. I., Bernarding, C., Reith, W., Latzel, M., & Froehlich, M. (2009a). On the cognitive neurodynamics of listening effort: a phase clustering analysis of large-scale neural correlates. *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, 2009*, 2078-2081.

Strauss, D. J., Corona-Strauss, F. I., Bernarding, C., Reith, W., Latzel, M., & Froehlich, M. (2009b). *On the cognitive neurodynamics of listening effort: A phase clustering analysis of large-scale neural correlates.* Paper presented at the Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE.

Strauss, D. J., Corona-Strauss, F. I., & Froehlich, M. (2008). Objective estimation of the listening effort: towards a neuropsychological and neurophysical model. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, 2008*, 1777-1780.

Strauss, D. J., Corona-Strauss, F. I., Trenado, C., Bernarding, C., Reith, W., Latzel, M., & Froehlich, M. (2010a). Electrophysiological correlates of listening effort: Neurodynamical modeling and measurement. *Cognitive Neurodynamics, 4*(2), 119-131.

Strauss, D. J., Corona-Strauss, F. I., Trenado, C., Bernarding, C., Reith, W., Latzel, M., & Froehlich, M. (2010b). Electrophysiological correlates of listening effort: neurodynamical modeling and measurement. *Cognitive Neurodynamics, 4*(2), 119-131.

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *Journal of the Acoustical Society of America, 125*(5), 3328-3345.

Strickland, O. J., & Galimba, M. (2001). Managing time: the effects of personal goal setting on resource allocation strategy and task performance. *Journal of Psychology, 135*(4), 357-367.

Sweetow, R. (2005). Training the adult brain to listen. *The Hearing Journal, 58*(6).

The Random House College Dictionary. (Ed.) (1980) (Rev ed.). New York: Random House.

Tognola, G., Grandori, F., & Ravazzani, P. (1998). Wavelet analysis of click-evoked otoacoustic emissions. *IEEE Transactions on Biomedical Engineering, 45*(6), 686-697.

Tombaugh, T. (1996). *Test of Memory Malingering manual*. New York: MultiHealth Systems.

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging, 24*(3), 761-766.

Tuovinen, J. E., Gerven, P. W. M. van, Paas, F., & Tabbers, H. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.

Turkyılmaz, M. D., & Belgin, E. (2012). Reliability, Validity, and Adaptation of Computerized Revised Token Test in Normal Subjects. *Journal of International Advanced Otology, 8*(1), p103-112.

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cogntive Effort and Memory. *Journal of Experimental Psychology:Human Learning and Memory, 5*(6), 607-617.

Van Boxtel, A., & Jessurun, M. (1993). Amplitude and bilateral coherency of facial and jawelevator EMG activity as an index of effort during a two-choice serial reaction task. *Psychophysiology, 30*, 589-604.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*(2), 167-174.

Vandergrift, L. (1997). The Cinderella of Communication Strategies: Reception Strategies in Interactive Listening *The Modern Language Journal, 81*(4), 494-505.

Veldhuizen, I. J. T., Gaillard, A. W. K., & de Vries, J. (2003). The influence of mental fatigue on facial EMG activity during a simulated workday. *Biological Psychology, 63*(1), 59-78.

Venables, L. , & Fairclough, S. H. . (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion, 33*(1), 63-74.

Verplanken, B. (1993). Need for cognition and external information search: Responses to time pressure during decision-making. *Journal of Research in Personality, 27*(3), 238-252.

Wallin, B. G. (1981). Sympathetic Nerve Activity Underlying Electrodermal and Cardiovascular Reactions in Man. *Psychophysiology, 18,* 470-476.

Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *Journal of Neuroscience, 23*(16), 6475-6479.

Walton, M. E., Devlin, J. T., & Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience, 7*(11), 1259-1265.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E., & Rushworth, M. F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Network, 19*(8), 1302-1314.

Waterink, W., & van Boxtel, A. (1994). Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology, 37*(3), 183-198.

Weiner, A. E. (1984). Stuttering and syllable stress. *Journal of Fluency Disorders, 9*(4), 301-305.

West, R.L., Thorn, R.M., & Bagwell, D.K. (2003). Memory performance and beliefs as a function of goal setting and aging. *Psychology and Aging, 18,* 111-125.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance*. NY: HarperCollins Publishers Inc.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81.

Wilson, R. H., Coley, K. E., Haenel, J. L., & Browning, K. M. (1976). Northwestern University Auditory Test No. 6: normative and comparative intelligibility functions. *Journal of the American Audiology Society, 1*(5), 221.

Wilson, R. H., Preece, J.P., Salamon, D.L., Sperry, J. L., & Bornstein, S. P. (1994). Effects of time compression and time compression plus reverberation on the intelligibility of Northwestern University Auditory Test No. 6. *Journal of the American Academy of Audiology, 5*, 269-269.

Wingfield, A. (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. *Journal of the American Academy of Audiology, 7*(3), 175-182.

Wingfield, A., & Grossman, M. (2006). Language and the Aging Brain: Patterns of Neural Compensation Revealed by Functional Brain Imaging. *Journal of Neurophysiology, 96*(6), 2830-2839.

Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology, 18*(7), 548-558.

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation. *Ear and Hearing*.

Winn, M., & Edwards, J. R. (2013). The impact of spectral resolution on listening effort revealed by pupil dilation. *The Journal of the Acoustical Society of America, 134*(5), 4233.

Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive style: Essence and origins*. New York: International Universities Press.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field dependent and field independent cognitive styles and their educational implications. . *Review of Educational Research, 47*, 1-64.

Yates, J.F., & Kulick, R.M. (1977). Effort Control and Judgments. *Organizational Behavior and Human Performance, 20*, 54-65.

Yeh, Y.Y., & Wickens, C. D. (1984). *The dissociation of subjective measures of mental workload and performance*. Moffett field, Calif.: National Aeronautics and Space Administration, Ames Research Center.

Yoshida, W., & Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron, 50*(5), 781-789.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology, 51*(3), 277-284.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear and Hearing, 31*(4), 480-490.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing, 32*(4), 498-510.

Zekveld, A. A., Kramer, S. E., Kessens, J. M., Vlaming, M. S., & Houtgast, T. (2009). The influence of age, hearing, and working memory on the speech comprehension benefit derived from an automatic speech recognition system. *Ear and Hearing, 30*(2), 262-272.