RESOURCE AND BOTTLENECK MECHANISMS OF ATTENTION IN LANGUAGE PERFORMANCE

by

William D. Hula

B.A., University of North Carolina, 1994

M.S., University of North Carolina, 1997

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

William Hula

It was defended on

April 27, 2007

and approved by

Thomas Campbell, Professor, Communication Science and Disorders

Patrick J. Doyle, Associate Professor, Communication Science and Disorders

J. Richard Jennings, Professor, Psychiatry

Connie Tompkins, Professor, Communication Science and Disorders

Dissertation Advisor: Malcolm R. McNeil, Distinguished Service Professor and Chair,

Communication Science and Disorders

RESOURCE AND BOTTLENECK MECHANISMS OF ATTENTION IN LANGUAGE PERFORMANCE

William Hula, Ph.D.

University of Pittsburgh, 2007

The view that impairments of attention may constitute an important factor underlying impaired language performance in aphasia has gained support in recent years. Aphasiologists taking this view have generally proceeded from resource allocation models of attention, with little or no attention given to alternative models. One alternative model of dual-task performance is the central bottleneck model, which proposes a single-channel limit at response selection or other central processing stages. The first purpose of the present experiments was to further examine the effects on word production of lexical frequency in the context of the psychological refractory period (PRP) dual-task method. The second purpose was to examine whether the reaction time (RT) patterns obtained under conditions promoting equal task emphasis are more consistent with the central bottleneck or central resource models. Three dual-task experiments were conducted using speeded picture naming and tone identification tasks presented at varying timulus onset asynchronies (SOAs). In experiment 1, lexical frequency affected primary-task naming and secondary-task tone identification RTs approximately equally. In experiment 2, lexical frequency affected secondary-task naming RTs similarly at all levels of SOA, after potentially confounding variables were taken into account. It was concluded that frequency-sensitive lexical processing in picture-naming participates in the central processing stage of the dual-task models under study. In the third experiment, the two tasks were presented in variable order and subjects were instructed to give equal attention to both. On tone-primary trials, tone

RTs increased with decreasing SOA, a result consistent with the central resource model and inconsistent with the central bottleneck model, unless augmented by the assumption that particpants grouped responses on short SOA trials. Also, additional analyses restricted to those participants demonstrating a lexical frequency effect on the secondary naming task found that lexical frequency and SOA interacted on primary-task tone RTs such that tone responses preceding low-frequency naming responses were slower than those preceding high-frequency names. This further suggests that these subjects allocated more central processing capacity to the naming task on low-frequency trials. Comparison of results across the three experiments suggested that participants in Experiment 3 demonstrated less dual-task interference than predicted by either model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I could not have accomplished the work represented by this document without the personal and professional support of many people to whom I owe much. I would first and foremost like to thank Mick McNeil. He is a gifted scientist and clinician with whom I have been proud to study. He was perennially willing to engage in thoughtful discussion of whatever I brought to him, and for that I am grateful. I owe no less a debt of gratitude to Pat Doyle, whose patient and steadfast support was essential, and who, like Mick, consistently encouraged me to tackle new problems, consider alternative points of view, and challenge myself in new ways. Together, they provided me with the opportunity to practice the skills of our profession with a well-balanced combination of independence and supervision.

Connie Tompkins has been an excellent source of intellectual stimulation and encouragement throughout my time in Pittsburgh, and I thank her for that. Tom Campbell and Dick Jennings made valuable contributions as members of my dissertation committee, and I appreciate their efforts on my behalf. Looking back to my undergraduate, master's, and clinical fellowship training, Laura Janda, Robert Peters, John Grose, Celia Hooper, Alex Johnson, and Barbara Jacobson were all important mentors and positive examples for me at various stages of my academic and professional development.

Mike Tombu was a helpful correspondent in the early stages of planning this project and provided valuable comments on an early version of the proposal. I would also like to thank

Jessie VanSwearingen, Paula Leslie, Elizabeth Hile, Connie Tompkins, Pat Doyle, Sheila Pratt, and Catherine Eberbach for assistance with subject recruitment. The work was supported by the VA Pittsburgh Geriatric Research, Education, and Clinical Center.

I would like to thank my colleagues in the VA Pittsburgh Healthcare System Audiology and Speech Pathology Service, Randy Aker, Kim Eichhorn, Katie Davis, and Mike Biel, for their support. I thank Christine Matthews in particular for her selfless collegiality and friendship. Working with such dedicated, bright, and generous colleagues has been an invaluable component to my education, and one that I hope will continue.

Shannon Austermann Hula, my colleague, best friend, and wife, has been as essential to the successful completion of this project as any other single person, myself included. She enthusiastically endured countless discussions of this work, and provided invaluable comments, feedback, and assistance. Her patience and generosity with me are beyond words and I hope that I can be half the friend and colleague to her that she has been to me.

Finally to my mother and father, Pat and Tom Queen, who have had infinite patience with my seemingly neverending education and provided all of the encouragement and support that any son could ask for, thank you and I love you.

1.0 INTRODUCTION

The view that impairments of attention constitute an important factor in impaired language performance in aphasia has attracted increasing interest and support in recent years. Aphasiologists who have taken this view have generally proceeded from *resource allocation or capacity theories* of attention, with little or no consideration of alternative accounts, such as the *central bottleneck model.* Resource theories propose that limitations on cognitive performance are the result of insufficient attentional resources or difficulties in the graded allocation of attention, while bottleneck models hold that such limitations are due to features of the cognitive architecture that permit processing of only one stimulus or task at a time. The primary objective of the present study was to compare a central *resource* model and the central *bottleneck* model in their ability to account for dual-task performance involving language.

> **Specific Aim A: To investigate the temporal locus of cognitive performance limitations resulting from lexical frequency effects.** It is well established that frequently encountered words are produced more quickly in speeded naming tasks than those that are encountered infrequently. To accomplish the objective of comparing the central resource and central bottleneck models, it was first necessary to ascertain whether these lexical frequency effects operate at *central processing stages* associated with response selection or more *peripheral stages* associated with perceptual encoding or

response production. The first two experiments employed a well-established dual-task procedure, the psychological refractory period (PRP) method, in which two temporally offset but overlapping stimulus-response tasks are presented, with one task designated as primary and the other as secondary.

Specific Aim B: To investigate whether dual-task performance limitations in single word production are more consistent with a central resource or central bottleneck model. The *central resource model* proposes that the two tasks will be handled *in parallel* with graded sharing of resources, permitting secondary task factors to influence primary task performance. The *central bottleneck model* predicts that, whenever central processing for two tasks overlaps in time, that they will be handled *serially*, and that factors associated with the second task should have no effect on reaction times to the first task. To address this aim, a third experiment was conducted in which the traditional PRP method was modified to present both tasks in primary and secondary positions.

2.0 BACKGROUND AND SIGNIFICANCE

2.1 BACKGROUND

2.1.1 Attention and Aphasia

The general view that attention constitutes an important explanatory mechanism in aphasia has often been described in opposition to the classical connectionist view of aphasia, exemplified in the writings of Wernicke and Lichtheim (Eggert, 1977), and the modern proponents of the classification systems based largely on their original categories (Geschwind, 1965a; Geschwind, 1965b; Goodglass, Kaplan, & Barresi, 2001; Kertesz, 1979). In general, they have tended to describe aphasia as a loss of representations subserving various specific language functions or the disconnection of serial components in language computation. This view has emphasized the differences among persons with aphasia by dividing them into a variety of behavioral categories arranged along multiple dimensions, and have ascribed to these categories substantial ability to localize brain lesions. In contrast, theories of aphasia that incorporate attention, resources, or processing capacity as important constructs tend to have more in common with accounts that regard aphasia primarily as a unidimensional disorder of performance in which representations are not lost, but access to them is impaired. Such theories have tended to minimize the connections between specific aphasic symptoms and particular lesion sites and have emphasized both the similarities between persons with aphasia and the variability within them according to both internal and external factors. Thus, the writings of Freud (1953), Marie (reviewed in Caplan, 1987), Head (reviewed in Caplan, 1987), Schuell (Schuell, Jenkins, & Jimenez-Pabon, 1964), and Darley (1982) can be viewed as important historical precedents for theories that regard aphasia as a disorder of resource allocation or diminished processing capacity.

Among contemporary aphasiologists operating from a resource allocation or capacity perspective, McNeil and colleagues (Arvedson & McNeil, 1987; Campbell & McNeil, 1985; McNeil, 1988; McNeil et al., 2004; McNeil, Odell, & Tseng, 1991; Slansky & McNeil, 1997; Tseng, McNeil, & Milenkovic, 1993) have chosen to emphasize the general nature of the deficit while eschewing linguistic analyses of aphasic performance as explanatory constructs. They have been among the most fastidious in applying the details of resource theory and methods as proposed by Kahneman (1973) and Navon and Gopher (1979; 1980) to the study of aphasia. McNeil and colleagues (McNeil, 1982; 1988; McNeil & Kimelman, 1986; McNeil et al., 1991) have stressed five arguments against prevailing neoclassical notions of aphasia as a loss of various specific linguistic or psycholinguistic functions and in favor of the view that aphasia is the result of impairments in the ability to allocate attentional or processing resources supporting language. They have noted that (1) aphasia almost universally affects all domains and modalities of language regardless of lesion location (Darley, 1982; Schuell et al., 1964), (2) persons with aphasia are stimulable for correct performance on tasks that they rarely or never perform correctly on their own (Darley, 1976; Duffy & Coelho, 2001), (3) aphasic persons typically demonstrate marked within-person variability in their ability to produce particular responses (Crisman, 1971; Freed, Marshall, & Chulantseff, 1996; Hageman, McNeil, Rucci-Zimmer, & Cariski, 1982; Head, 1926; Howard, Patterson, Franklin, Morton, & Orchard-Lisle, 1984;

Kreindler & Fradis, 1968; McNeil & Hageman, 1979; McNeil, Odell, & Campbell, 1982; Schuell et al., 1964), (4) aphasia can be transient, as in epilepsy or transient ischemic attack (McNeil, 1982; 1988; Valachovic, Smith, Elisevich, Jacobson, & Fisk, 1998), and (5) there is substantial qualitative similarity between normal and aphasic language performance, suggesting that aphasia may be viewed as the low end on a continuum shared with normal language (the so-called "continuity hypothesis") (Bates, Frederici, & Wulfeck, 1987; Blackwell & Bates, 1995; Brookshire & Nicholas, 1984; Brookshire & Nicholas, 1980; Dick et al., 2001; Ernest-Baron, Brookshire, & Nicholas, 1987; Freud, 1953; Hageman, 1980; Kilborn, 1991; Miyake, Carpenter, & Just, 1994; Nicholas & Brookshire, 1986; Schwartz, Saffran, Bloch, & Dell, 1994; Shewan, 1976; Shewan & Canter, 1971; Silkes, McNeil, & Drton, 2004; Wilson, Saygin, Schleicher, Dick, & Bates, 2003). McNeil and colleagues have argued that these observations suggest a disorder of performance or access, rather than one in which linguistic rules or representations are lost, and they have proposed that the broad and variable impairments typically observed in aphasia can only be accounted for by a "superordinate mechanism [that] is shared by linguistic processing units" (McNeil et al., 1991, p. 28).

A number of dual-task studies have been carried out over the past two decades to directly address the issue of whether aphasic and normal language performance are well described by the constructs of resource allocation theory. In most cases, the strategy has been to manipulate the presence, difficulty, and/or priority of two concurrent tasks, one or both of which involves language, and then to observe whether or not performance trades occur. One relatively consistent result has been that performance is slower or less accurate in dual-task compared to single-task conditions, and that the performance decrement is larger for aphasic individuals than for those with normal language (Arvedson, 1986; Erickson, Goldinger, & LaPointe, 1996;

LaPointe & Erickson, 1991; Murray, 2000; Murray, Holland, & Beeson, 1997a; Murray, Holland, & Beeson, 1997b) However, McNeil and colleagues (2004) have argued that interpretation of such single-to-dual-task decrements as evidence for resource sharing is problematic, because of the qualitative differences between single and dual-task conditions. They have further suggested that the effects of task priority manipulations are also suspect, because of methodological difficulties reviewed below, that are inherent in the voluntary effort allocation method. They have argued that the best evidence for the graded sharing of processing resources is provided by the observation of performance trades induced by changes in task difficulty among dual-task conditions that can be supposed to share the same qualitative processing requirements.

Studies following this general strategy have often found, in accordance with predictions of resource theory, that increasing the difficulty of one task induces a decrement in the performance of a competing task, at least where normal individuals are concerned (Blackwell & Bates, 1995; Campbell & McNeil, 1985; McNeil et al., 2004; Tseng et al., 1993). There has been less consistency observed in the effects of dual-task manipulations in persons with aphasia. Some studies have found that persons with aphasia are insensitive to certain dual-task manipulations, such as emphatic stress on a primary task stimulus (Slansky & McNeil, 1997) or probability of target occurrence, that cause concurrent task performance decrements in normals (Tseng et al., 1993). On the other hand, manipulations of presentation rate of auditory commands (Campbell & McNeil, 1985), concreteness of lexical decision stimuli (Arvedson, 1986), and propositional load in sentence comprehension stimuli (Caplan & Waters, 1996) have been shown to cause decrements in competing task performance in groups with disordered language. One way of accounting for such inconsistencies would be to suppose that task manipulations might differentially affect demand for processing resources and task allocation ratio. If one proposes the existence of allocation deficits in aphasia, then one might predict that aphasic dual-task performance would be more sensitive to factors that increase or decrease demand, and relatively less affected by factors that primarily influence allocation ratio in normal individuals. One difficulty in interpreting dual-task studies is that task manipulations may confound the effects of demand and allocation ratio. One would expect that in most cases, increasing the demand of a task for resources might also increase the proportion of resources allocated to it (Tombu & Jolicoeur, 2002a). Historically, dual-task studies motivated by resource theory have required subjects to voluntarily allocate different percentages of their effort between tasks by explicit instruction (Arvedson, 1986; Gopher, Brickner, & Navon, 1982; Matthews & Margetts, 1991; Navon, 1990; Slansky & McNeil, 1997). This method is suspect because it assumes that a subject's perception of his or her own effort is a veridical reflection of resource investment. It has been shown that sense of effort or subjective measures of workload dissociate from performance measures under many conditions (Clark & Robin, 1995; Gopher & Braune, 1984; Vidulich & Wickens, 1986). Also, the voluntary allocation method has been criticized on the grounds that it simply invites subjects to respond according to the experimenter's wishes (Navon, 1984).

Another, broader issue in the interpretation of most dual-task studies of language (and most dual-task studies in general), is that they do not rule out the possibility that performance is achieved by serial back-and-forth switching between the two tasks, rather than by parallel processing that utilizes a limited capacity of shared resources (Pashler, 1994a; Wickens, 1984). It is this position that has been taken by proponents of the central bottleneck model of dual-task performance.

2.1.2 Central Bottleneck Theory

The central bottleneck model as discussed in current literature had its primary inception in the work of Craik (1947; 1948), Telford (1931), and Welford (1952; 1959; 1967). They noted, in the context of both simple reaction time (RT) and manual tracking tasks, that, when responses to two successive stimuli had to be made under time pressure, the shorter the interval between presentation of the stimuli, the greater the delay of the second response. Telford (1931) termed this phenomenon the "psychological refractory period," an analogy with the refractory period observed in the activity of individual neurons. Welford formalized the account into a model that proposed that this delay is due to a bottleneck in the central, response selection stage that permits processing of only one task at a time. Although the explicit analogy with the functioning of individual neurons has since been discarded, the name has stuck.

The current version of the central bottleneck model proposes three serial stages of processing for each task: a perceptual analysis and encoding stage, a central or response selection stage, and a response execution stage (Pashler, 1994a). The central stage admits processing for only one task at a time, causing response selection for any competing task to be delayed until central processing of the first is complete. The perceptual processing and response execution stages for a given task are hypothesized to run concurrently with any stage of a competing task. While it is acknowledged that capacity limits and/or processing bottlenecks may occur in these two stages (perceptual analysis and response execution) under some circumstances, such limits are proposed to be independent of and qualitatively different from the more ubiquitous central bottleneck (Pashler, 1998). Thus, central bottleneck models by definition require serial processing of multiple stimulus-response tasks, always delaying central processing of a concurrent task until the bottleneck stage of any competing task is complete. This requirement of serial processing in the central stages of concurrent sensorimotor tasks is considered a fixed architectural feature of the human cognitive system, rather than a dynamic feature resulting from the interaction of goals, strategies, or task requirements.

Most of the evidence for the central bottleneck theory in its current form has been generated by the PRP method, wherein two simple, discrete stimulus-response tasks are presented at a variety of stimulus onset asynchronies (SOAs) and reaction times to both tasks (RT1 and RT2, respectively) are measured. Typical examples of tasks used in this method include pressing a button to indicate whether a tone is high or low in pitch, naming aloud the highest number in a visually presented array of digits, or pressing a button to indicate the location of a visual stimulus. The stimulus for one task is usually presented 50ms to 1000ms before the other, and this task is likewise usually given priority in the instructions, i.e., subjects are told to respond as quickly and accurately as possible to it and only then to respond to the second stimulus. Figure 1 presents a schematic representation of the central bottleneck model of RT performance in the PRP method at a short SOA, when task 2 central processing is subject to serial postponement by task 1 central processing. The model makes a number of specific predictions about the effects of various task manipulations on RT1, RT2, and their relationships to one another within the PRP method, and these have been presented and discussed extensively (Kahneman, 1973; Navon & Miller, 2002; Pashler, 1984; 1994a; 1998; Tombu & Jolicoeur, 2002a).

One prediction, often referred to as the “PRP effect”, is that RT2 should increase with decreasing SOA. Over the range of the shortest SOAs, any reduction in SOA should be reflected by a millisecond-for-millisecond increase in RT2, resulting in an RT2-SOA function with a slope of -1. As can be seen in Figure 1, this situation is proposed to result from stage b2 (task 2 central processing) having to wait for the completion of stage b1 (task 1 central processing). This pattern of results has been consistently observed in a large number of PRP studies (Carrier & Pashler, 1995; Fagot & Pashler, 1992; Ferreira & Pashler, 2002; Johnston, McCann, & Remington, 1995; McCann & Johnston, 1992; Pashler, 1989; 1990; 1991; 1994b; Pashler & Johnston, 1989; Welford, 1959), with RT2-SOA slopes approaching negative one.

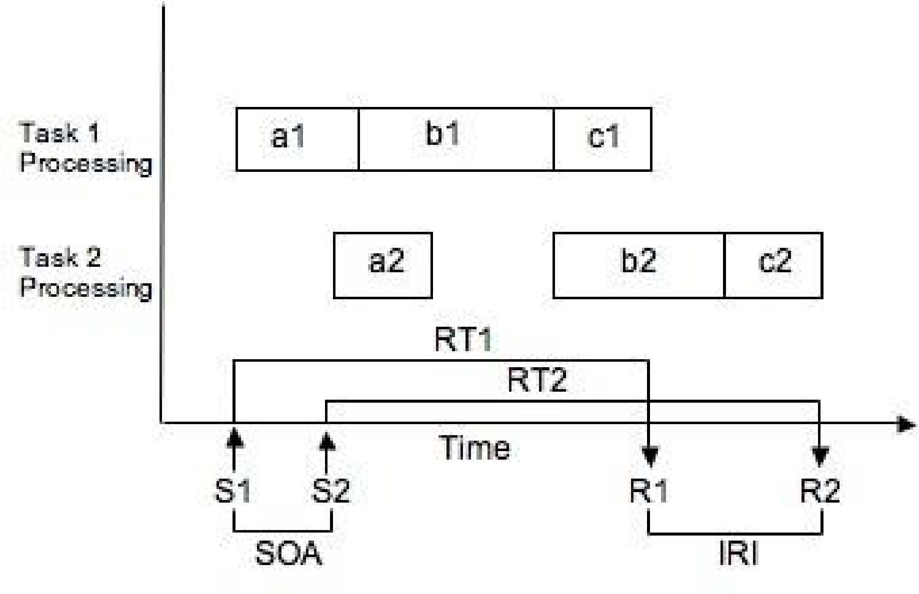

Figure 1. A time diagram of the central bottleneck model of PRP dual-task performance in which task 2 central processing (b2) is postponed. Abbreviations: 1a, 2a = perceptual processing of tasks 1 and 2, respectively; 1b, 2b = central or response selection processing of tasks 1 and 2; 1c, 2c = response execution stages of each task; RT1, RT2 = reaction times for task 1 and 2; S1, S2 = stimulus onset for each task; SOA = stimulus onset asynchrony; R1, R2 = task responses; IRI = inter-response interval (after Navon and Miller, 2002, p. 228)

A second prediction is that the effect of task 2 perceptual encoding difficulty should interact with SOA, such that it has a smaller effect at shorter SOAs. Manipulations of task 2 perceptual encoding are hypothesized to affect the length of stage a2, and it can be seen from

Figure 1 that any lengthening of stage a2 would have little or no effect on RT2 at the short SOA depicted. In contrast, at longer SOAs, as depicted in Figure 2 below, any increases in Task 2 perceptual encoding demands would be reflected in a longer RT2.

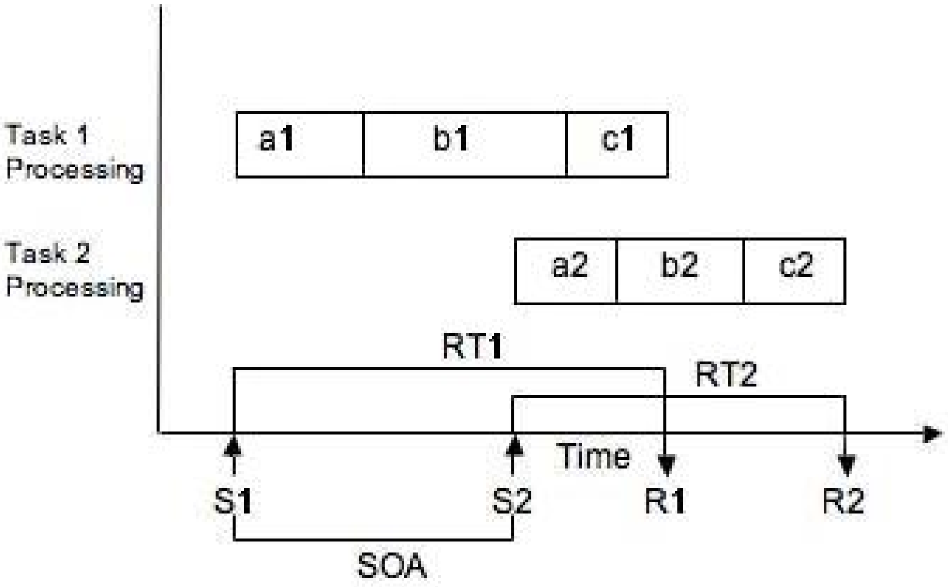

Figure 2. A time diagram for the central bottleneck model of PRP dual-task performance when task 2 central processing (b2) is not postponed by task 1 central processing. Abbreviations are as in Figure 1. (after Navon and Miller, 2002)

This predicted interaction has been observed in many experiments across several papers (Pashler, 1984; 1989; 1991; Pashler & Johnston, 1989), but is has been acknowledged that the lack of any apparent limits on dual-task perceptual processing in these studies likely depends on the simplicity of the stimuli involved (Pashler, 1998). Evidence suggests that limitations on perceptual processing become more apparent when stimuli are even modestly complex, for example, words as opposed to letters, or when visual display size is increased (Fisher, 1984;

Kahneman & Treisman, 1984; Kleiss & Lane, 1986; Ninio & Kahneman, 1974; Puleo & Pastore, 1978; Treisman & Fearnley, 1971).

Central bottleneck model predictions for RT1 in the PRP method are relatively straightforward. First, neither SOA nor any Task 2 factors should affect RT1. This prediction follows directly from the absolute and privileged access to the bottleneck stage that is given to the first task: If the first stimulus to arrive occupies the central bottleneck stage until its response selection processing is complete, then the time at which the second stimulus arrives and its central processing requirements are irrelevant to the first response. These predictions have been upheld in many experiments (e.g., Fagot & Pashler, 1992, exp. 1, 7; Johnston et al., 1995; Pashler, 1989, exp.1, 2, 5; Pashler & Johnston, 1989), but challenged by others (e.g., Carrier & Pashler, 1995; Fagot & Pashler, 1992, exp. 2; McCann & Johnston, 1992, exp. 2; Pashler, 1984; Pashler, 1989, exp. 3, 4; Pashler, 1991, exp. 3, 4; Smith, 1969).

Second, any increases in Task 1 perceptual encoding or central processing difficulty should have the same effect across all levels of SOA and regardless of the dual-task requirement itself. Stated differently, there should be no interaction between factors affecting Task 1 difficulty and SOA, or even the dual-task requirement itself. This null interaction has been observed in at least two studies (Ferreira & Pashler, 2002; Pashler, 1984).

2.1.3 Resource Theory

Resource theory, which has its roots in both human factors engineering (e.g., Knowles, 1963) and psychology (e.g., Moray, 1967), received its most comprehensive early application to human attention and dual-task performance in the work of Kahneman (1973). Kahneman proposed that human cognitive activity is powered by a pool of resources, which he equated with the notion of mental effort and that is flexibly allocated to various processes according to a number of internal and external factors. Resource theory differs from the central and other bottleneck theories primarily in that it assumes that attention is a divisible quantity that is potentially capable of driving multiple operations in parallel across all stages of processing. Specifically, Kahneman proposed that the available supply of resource capacity fluctuates over time and is determined by task demands, overall level of arousal, and other miscellaneous factors. This capacity is then applied to various ongoing activities according to an allocation policy. The allocation policy is determined by some of the same factors that determine the availability of resources, including task demand and arousal level, and also by enduring dispositions (the "rules of involuntary attention", p. 11) and momentary intentions (situation-specific performance criteria and biases).

Kahneman specifically contrasted his single resource model of attention with structural theories, including the central bottleneck model reviewed above. While he admitted that structural interference plays a role in dual-task performance, he concluded that prominent effects of response selection requirements are not due to a ubiquitous bottleneck at that stage of processing. Rather, he argued that response selection causes interference by virtue of its high demand for processing resources (p. 185), which may be greater than the demand imposed by, for example, perceptual recognition (see p. 148-150). It has been noted however, that the serial response selection processor of the central bottleneck model may be considered, at least in some sense, a resource (Navon & Miller, 2002), and that it could behave like a divisible resource provided that it shifts between tasks with enough fluidity and rapidity (Pashler, 1994a; Wickens, 1984). Furthermore, even structural bottleneck theories implicitly include the concept of capacity by virtue of their assumption that more difficult tasks occupy the bottleneck for longer periods of time (Wickens, 1984). If the amount of time the central response selection processor spends on a given task before switching to a concurrent one is susceptible to influence by priorities and strategies, then distinguishing between the central bottleneck and resource models on empirical grounds becomes very difficult indeed, and may in fact be an arbitrary exercise (Wickens, 1984). However, the extent to which and the conditions under which the human cognitive system permits parallel processing is a nontrivial issue.

2.1.4 A Resource Model of PRP Dual-Task Performance

The inability of the central bottleneck model to account for certain results produced by the PRP method were discussed by Kahneman (1973), and more recently by Navon and Miller (2002) and Tombu and Jolicoeur (2002a; 2003). Both Navon and Miller and Tombu and Jolicoeur presented mathematically specified resource models of PRP performance, which were essentially identical and will be referred to here collectively as the central resource model. The two presentations differed primarily in their hypotheses regarding the effects of task difficulty on allocation ratio, as will be discussed below. The central resource model is presented below in Figure 3, which was taken from Navon and Miller (2002, p. 232). A mathematical formulation of the model, based on presentations by Navon and Miller and Tombu and Jolicoeur (2002a; 2003) is provided in Appendix A.

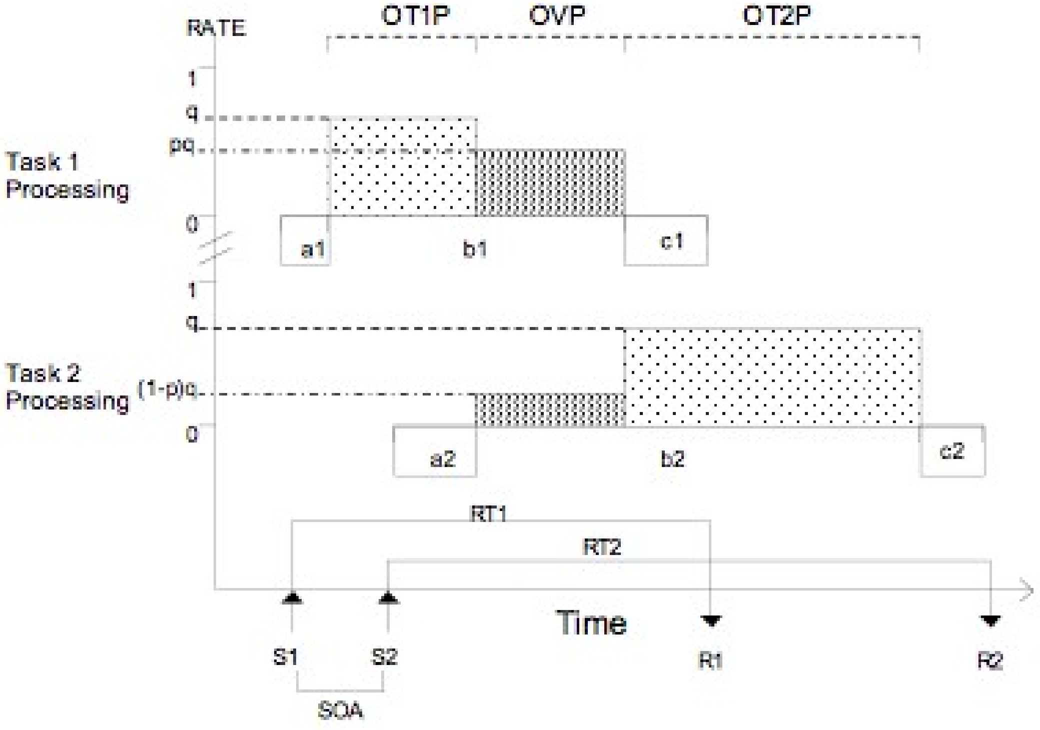

Figure 3. A time diagram of the central resource model. Abbreviations as in Figure 1, with the following additions: 1 = total resource capacity available under ideal conditions; q = resource capacity available on a given trial; p = task 1 allocation ratio; OT1P = period during which the resource-limited stage of only task 1 is processed; OT2P = corresponding period for task 2; OVP = the overlap period during which resource-limited processing for both tasks is carried out (after Navon and Miller, 2002, p. 232)

In this model, resources determine processing rate and their availability fluctuates from trial to trial, up to a maximum (designated '1' in the figure). Central or response selection processing (stage b) for each task imposes a resource demand, which is expressed as the size of an area on the coordinate plane of the diagram. The length of the response selection stage (the length of the stage along the x-axis, equal to time in milliseconds) is determined as the resource demand (the total area on the coordinate plane associated with stage b of a given task) divided by processing rate (the height(s) of the associated area(s) along the y-axis, equal to the available resources multiplied by the allocation ratio, or pq for task 1 and (1-p)q for task 2). Thus, an increase in the availability of resources to a given task will decrease the time needed to carry it out. During the time that only central processing of task 1 (stage b1) is being carried out (OT1P), all of the available capacity (q) is allocated to it. As soon as central processing for task 2 (stage b2) begins, the resources are divided between the two tasks according to the allocation ratio (p: 1-p) for the entire time that b1 and b2 overlap (OVP). Once central processing for task 1 is completed, all available resources are allocated to the completion of stage b2 (OT2P). The model assumes that the central stage of processing is the only one that is resource-limited. Other assumptions include: Both allocation ratio (p) and available capacity (q) remain constant during a trial; available capacity (q) does not vary systematically with experimental factors including SOA; there is no cost of online re-allocation of resources from one task to the other; and the central resource capacity is undifferentiated and may be applied to any task domain with equal efficiency and productivity. Also, under short SOA conditions, it is necessary to assume that central processing of task 1 begins before task 2 central processing begins, that there is overlap in central processing of the two tasks, and that task 1 central processing finishes before task 2 central processing finishes.

It is important to note that under the extreme allocation ratio favoring task 1 (p = 1), that the central resource model reduces to the central bottleneck model and the two make identical predictions for both RT1 and RT2 (Navon & Miller, 2002; Tombu & Jolicoeur, 2003). Furthermore, it is reasonable to assume that such biased allocation ratios have been present in most PRP experiments, given that stimulus order is typically known and quick responses to task 1 are emphasized (Navon & Miller, 2002; Tombu & Jolicoeur, 2003). It is also important to note that the two predictions for RT2 emphasized most often and reviewed above (a –1 slope of the

RT2-SOA curve and a reduced effect of task 2 perceptual demand at low SOAs) are identical for the central bottleneck and central resource models, regardless of allocation ratio (Navon & Miller, 2002; Tombu & Jolicoeur, 2003). The counterintuitive insensitivity of RT2 to the difference between the serial processing of the bottleneck model and the parallel processing of the resource model is explained in the following way: When resource capacity is limited, any concurrent processing will be accomplished at a slowed rate relative to single-task processing. Whenever task 2 gets earlier access to limited central resources, any RT2 advantages will be offset by a lengthening of task 1 central processing and consequent lengthening of the time that the two tasks must share capacity. Conversely, when task 2 waits longer to begin utilizing central resources, more of its processing will be carried out at a faster rate that is possible only when it does not have to share capacity. Given that the central bottleneck and central resource models make similar predictions for RT1 under extreme allocation ratios and identical predictions for RT2 under any allocation ratio, it is observations of RT1 under conditions promoting more equal allocation of processing capacity that have the potential to differentiate the two models (Navon & Miller, 2002).

To date, there are few examples in the literature of PRP experiments that have invited equal emphasis on both tasks. In one such experiment (Pashler, 1991, exp. 3), participants made pushbutton responses to auditory tone and visual letter stimuli. Tone stimuli were presented first, and the task was to classify them as high or low. Letter stimuli were presented at 50 ms, 150 ms, and 650ms SOA and the secondary task was to identify the probed letter in an eight-letter array. Subjects were told to complete both responses as quickly as possible, in the order of presentation. While the tone stimuli were always presented and responded to before the letter stimuli, and this may have biased subjects to devote more capacity to the first task (tone ID), these instructions might also have promoted a more even allocation ratio than typical in PRP experiments. In this particular experiment, decreasing SOA significantly elevated the mean of RT1, in contrast with central bottleneck model predictions, and consistent with the predictions of the central resource model.

Another experiment conducted by Pashler (1994b) may have better promoted equal priority between the two tasks. In this study, the tasks were roughly similar to those described in Pashler (1991), but SOA was varied between 1000ms, 500ms, 0ms, -500ms, and –1000ms, so that the order of the tasks varied. Furthermore, subjects were explicitly instructed to place equal emphasis on both tasks and not to favor a particular stimulus, even if it arrived first. The primary hypotheses of this experiment concerned the distribution of inter-response intervals (IRIs). It was argued that graded sharing of processing resources should produce a broad, unimodal distribution of IRIs centered about zero. The operation of a central processing bottleneck, on the other hand, was predicted to produce either a unimodal "spike" distribution, representing grouped responses clustered tightly around zero with little variance, or a bimodal, "double-ridge" pattern, with the modes located on either side of zero, representing cases in which responses were produced serially. Twenty-three of 24 subjects produced either a spike (n = 6) or a double-ridge (n = 17) distribution, and support was inferred for the bottleneck model. However, one aspect of the data was problematic for this interpretation. Namely, SOA had a significant effect on RT1: RTs for the first task responded to in the 0ms SOA condition were significantly longer than the same RTs in the 1000ms SOA condition. Pashler accommodated this result by reference to stimulus-order uncertainty effects, which have been noted to elevate RTs in a variety of conditions (Pashler, 1990). Additionally, it has been suggested that the predictions of spike and double-ridge patterns are not unique to the bottleneck model (Tombu & Jolicoeur, 2003). The response grouping proposed to cause the spike distribution is equally available under a resource model, and any asymmetry in task allocation ratio would promote the same double ridge distribution predicted by the bottleneck model.

While neither of these experiments included manipulations of response selection difficulty that would allow for examination of predictions other than SOA effects on RT1, a recent study using the PRP method did include such a manipulation (Tombu & Jolicoeur, 2002a). In this study, subjects were required to judge whether a tone was high, medium, or low in pitch and whether two irregular polygons were the same or mirror images of one another. The difficulty of the polygon task was manipulated by varying the size ratio between the two stimuli (1:1, 1:2, 1:4), a factor that had previously been shown to impact central processing in a PRP dual-task experiment (Tombu & Jolicoeur, 2002b). In order to encourage more even allocation of resources, the two tasks were presented in random order with SOAs of 50ms, 250ms, 750ms and 1250ms. The central resource model employed in this study differed significantly from the model presented by Navon and Miller (2002) in that Tombu and Jolicoeur (2002a) proposed that changing the difficulty of the secondary task should influence subjects to allocate more resources to it on a given trial.

There were several specific hypotheses examined by Tombu and Jolicoeur (2002a). First, the authors predicted that decreasing SOA should result in increased task 1 RT. In agreement with the earlier finding discussed above, this prediction was upheld when both the tone and shape tasks were presented and responded to first. The second prediction of this study was that task 1 central processing difficulty should interact with SOA, such that increases in task 1 difficulty should cause larger increases in RT1 at shorter SOAs. This prediction was not upheld: on shape-first trials, increasing the difficulty of the shape had a similar effect on shape RT at all

SOAs. The authors speculated that the analysis might have lacked sufficient statistical power to detect this interaction effect, which would have been subtle under the allocation ratios apparently achieved in this study (see immediately below).

In addition to analysis of variance, Tombu and Jolicoeur (2002a) also employed a model fitting procedure in which they conducted a simulation of their results using both a central bottleneck and a central resource model. In fitting the resource model, they permitted task 2 difficulty to influence allocation ratio by diverting more resources to task 2 for trials with increased task 2 difficulty. The simulation results suggested that, for tone-first trials, the proportion of resources allocated to task 1 (p in Figure 3 above) varied between 0.91 in the easiest task 2 condition and 0.78 in the hardest. A corresponding main effect of task 2 difficulty on RT1 was noted in the observed data. This version of the central resource model also predicts, however, an interaction between task 2 difficulty and SOA on RT1 that was not observed. If increased task 2 difficulty causes capacity to be diverted away from task 1, the effect should be more apparent at shorter SOAs (Tombu & Jolicoeur, 2002a, 2003).

It is interesting to question the assumption that task 2 difficulty influences allocation ratio (and thus RT1) without making the reciprocal prediction for task 1 difficulty. It seems that if arrival of a more difficult secondary task can cause subjects to shift more capacity to that secondary task, then a more difficult primary task should also cause subjects to retain more capacity to this primary task when a secondary task arrives. This could explain the null shape difficulty-by-SOA interaction observed in the real data for the shape-primary RTs. The reasoning is that the shifting of resources toward the primary shape task in the more difficult conditions would partially counteract the effect of the difficulty increase, and mitigate the expected interaction with SOA (Navon & Miller, 2002; Tombu & Jolicoeur, 2005). Depending on the specific allocation ratios involved, the resource model might actually predict a small interaction in the opposite direction between shape difficulty and SOA for shape-primary RTs, while preserving the observed main effects. Thus, RT1 could be predicted to increase at both low SOAs and with difficult size ratios (as observed), but the predicted difficulty effect would be slightly smaller at low SOAs. Ideally, one would parametrically manipulate both task difficulty and allocation ratio, and examine their three-way interaction with SOA, but this would likely be difficult, if not impossible to achieve in a PRP-type experiment. Moreover, the size of this interaction would depend on the balance between the increase in task demand and the increase in resources applied to meet that demand, factors that would be difficult to control.

In a more recent paper, Tombu and Jolicoeur (2005) examined a different prediction of the central resource model, one concerning the effects of manipulating pre-central stages of task 2. When perceptual encoding of the secondary stimulus becomes more difficult under non-extreme allocation ratios, the resource model makes the counterintuitive prediction that RT1 should be *faster* than when encoding of the secondary stimulus is less difficult. This is because the lengthening of the pre-central stage of the secondary task gives the primary task exclusive access to the available central processing capacity for a longer period of time before it must share capacity and begin central processing at a reduced rate. The greater the task overlap, the greater the benefit to RT1 of increased stimulus 2 encoding difficulty. Tombu and Jolicoeur adapted a procedure from Carrier and Pashler (1995), in which the primary task was tone identification and the secondary task was an old-new visual word recognition task. Subjects were presented with lists of words during a study phase, and then during the dual-task trials, were required to classify the words according to whether or not they had been presented during the study phase. The visual contrast of the words was varied across two levels to create a perceptual encoding difficulty manipulation of the secondary word task. Based on previous PRP studies showing underadditive interactions between secondary stimulus visual contrast and SOA, it was hypothesized that the visual contrast manipulation would affect pre-central processing of the word stimulus. Furthermore, it was predicted that low contrast trials would be associated with *shorter* primary-task tone RTs at short SOAs only. As predicted, it was found that the contrast manipulation interacted underadditively with SOA on RT2 such that low contrast words took longer to classify at the longest SOA (75 ms difference), with no difference observed between high and low contrast words at the shortest SOA (5 ms difference). The predicted effect on RT1 was also obtained, that is, contrast and SOA interacted on tone RT such that tone responses associated with low contrast words were significantly faster (by 18 ms) than tone responses associated with high contrast words at the shortest SOA. This RT1 advantage in the low contrast condition diminished and reversed as SOA increased: Tone RT1 was 14 ms faster in the low contrast condition at 200ms SOA, and 6 ms slower at 1100 ms SOA.

Both studies reviewed immediately above demonstrated effects of SOA and task 2 difficulty on RT1 that are consistent with the central resource model and inconsistent with the central bottleneck model. However, the bottleneck model can account for most of these effects if it is augmented with an additional hypothesis that participants engage in response grouping, particularly at short SOAs. According to this account, participants select their response to task one, but postpone the execution of the response until response selection for task 2 is complete, at which time they emit both responses as a set. Pashler and Johnston (1989) manipulated task instructions to encourage response grouping at all SOAs and found the predicted results: RT1 increased dramatically at longer SOAs and task 2 factors had strong effects on RT1. However, in order to account for an increase in RT at short SOAs, the most diagnostic and frequently observed result consistent with the resource model, the response grouping hypothesis must additionally hold that participants group their responses more often at short SOAs and infrequently or not at all at long SOAs (Navon & Miller, 2002; Tombu & Jolicoeur, 2002a). Tombu and Jolicoeur (2002a) examined the distribution of inter-response intervals (IRIs) in their data, and concluded, based on the relatively small number of trials with IRIs close to zero, that the response grouping hypothesis did not account for their data. Tombu and Jolicoeur drew similar conclusions about their subsequent (2005) study, and also noted that their results did not differ appreciably when they excluded from analysis all trials with IRIs less than 200 ms.

Thus, the central resource model has been shown to account for a substantial amount of data generated by PRP experiments, both those motivated by it (Pashler, 1994b; Tombu & Jolicoeur, 2002a; Tombu & Jolicoeur, 2005), and those designed to test particular aspects of the central bottleneck model (e.g., Carrier & Pashler, 1995; Fagot & Pashler, 1992, exp. 2; McCann & Johnston, 1992, exp. 2; Pashler, 1984; Pashler, 1989, exp. 3, 4; Pashler, 1991, exp. 3, 4; Smith, 1969).

2.1.5 Models of Word Production

The PRP method has been infrequently used to investigate language processing in dual-task situations, perhaps in part because it requires relatively simple, discrete stimulus-response sequences. There has, however, been a small handful of studies in recent years that have begun to integrate models of lexical access with the central bottleneck model of PRP dual-task performance. Prior to discussing those studies, current models of word retrieval will be reviewed, with a focus on issues relevant to the dual-task models under consideration.

The current literature on single word production is dominated by spreading activation models that describe lexical access as a multi-stage activity in which earlier stages are dominated by semantic processing and later stages by phonological processing (Cutting & Ferreira, 1999; Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992; Roelofs, 1997; Roelofs, Meyer, & Levelt, 1996). In the specific case of a picture-naming task, first the picture must be recognized. Recognition activates a conceptual representation of the idea conveyed by the picture, and this conceptual representation links to the first properly lexical representation, the lemma. The lemma, originally proposed by Kempen and Huijbers (1983), is semantically and syntactically specified, but has no phonological content. A given word's lemma links directly to its lexeme, which consists of its phonological form and syllable structure. Following selection of the appropriate lexeme, the individual phonemes comprising the word are assembled. At each level, linguistic rules external to the model build frames and the activated representations are selected to fill slots within those frames, and the products of this assembly are fed to the articulatory system for motor planning, programming, and execution.

Word production models that treat semantic and phonological processing in detail describe speech motor processing either not at all (Cutting & Ferreira, 1999; Dell, 1986; Dell et al., 1997) or with much less specificity (Levelt et al., 1999). Current models of motor speech production, on the other hand, tend to under-specify pre-motor linguistic processing (e.g., Guenther & Perkell, 2004; Van der Merwe, 1997). Because the dual-task models under consideration emphasize central, response selection processing, the current review will focus on models motivated by semantic and phonological factors, acknowledging that aspects of motor speech control may also be subject to bottleneck/resource constraints.

Aside from the general features described above, there are substantial differences between models of lexical retrieval. One primary unresolved issue concerns the existence of feedback between the various stages. A second, related question is whether processing proceeds in a discrete manner, with selection of the appropriate representation at one level occurring prior to any activation of nodes at subsequent levels, or in a cascaded fashion that permits for example, partial activation of one or more lexemes prior to lemma selection. One class of models proposes that activation spreading is both cascaded and characterized by feedback via excitatory, bi-directional links between adjacent levels (Cutting & Ferreira, 1999; Dell, 1986; Dell et al., 1997). Such interactive, cascaded models propose that partial activation of nodes at lower levels can influence the activation and subsequent selection of representations at higher levels.

The interactive models proposed by Dell and colleagues in particular had as one of their primary goals the explanation of naturally occurring and induced speech errors. For example word substitution and exchange errors typically involve members of the same syntactic category (Dell & Reich, 1981; Fay & Cutler, 1977; Garrett, 1975; Garrett, 1980; MacKay, 1982). Thus, the exchange error between nouns in example (A) is much more likely to be observed than the exchange of a noun and a determiner in example (B). At the same time, the interacting elements in word-level errors of this type are often phonologically unrelated (Garrett, 1975).

(A) put some *mustard* on this *pretzel* → put some *pretzel* on this *mustard*

(B) put *some* mustard on this *pretzel* → put *pretzel* mustard on this *some*

This category constraint on whole-word errors constitutes one of the primary pieces evidence supporting the inclusion of phonologically empty lemma representations in models of lexical retrieval.

Another phenomenon for which word production models should account is the tendency of sound-level errors to create real-word rather than nonword outcomes (Dell, 1986; 1990; Dell & Reich, 1981). Thus, the phrase “club pays” is more likely to induce an initial consonant exchange, producing “cub plays” than the phrase “cheap table” is to produce an analogous error, “teap chable”. This lexical bias effect is explained within the context of interactive models by the feedback between individual phonemes and the lexeme representations to which they connect. When phonological strings corresponding to real words are being encoded, they send activation back to those lexemes, thus enhancing the probability of selection. On the other hand, nonword strings have no lexeme representations to receive feedback, and thus will be selected less often. Other speech error effects that find similar explanations in interactive models include semantic bias effects, the tendency of sound errors to create words that are semantically related to other words in the linguistic environment, and similarity effects, such as the tendency of lexical errors to involve words that are similar in both sound and meaning (Dell, 1986).

In contrast to cascaded, interactive models discussed above, the models proposed by Levelt and colleagues (Levelt et al., 1999; Roelofs, 1992; Roelofs, 1997) are characterized by discrete spreading of activation from one level to the next and exclusively uni-directional, feed-forward connections between them. The discrete aspect of these models refers to the fact that activation spreads to a lower level only after the corresponding item at the preceding level has been fully selected.

Levelt and colleagues’ (1999) model includes three potential mechanisms to account for the speech error data discussed above. The first concerns their solution to the problem of correctly binding representations across levels of the model. For example, given a conceptual representation of *The teacher told the student* and activation of corresponding lexeme nodes, it is not clear how a speaker avoids producing *The student told the teacher.* Dell's (1986) model solved the problem by timing the model's activation dynamics such that the most activated representation at a specific moment is the correct one. Levelt and colleagues, on the other hand, proposed a binding-by-checking mechanism. This mechanism checks that an activated representation corresponds to an appropriate representation in the preceding level. Occasional failures of this checking procedure result in many of the kind of errors predicted from natural and elicited speech error data. The second mechanism proposed by Levelt and colleagues (1999) is a post-lexical self-monitoring system that feeds the phonetic plan for an utterance through the speech comprehension system. They argue that imperfect self-monitoring should result in both the lexical bias and similarity effects that interactive models explain by feedback. The third potential source of errors discussed by Levelt et al. (1999) is the possibility that the lemma retrieval mechanism in their model may occasionally inadvertently select two lemmas. Given the dynamics of lemma activation, these two competing representations are likely to be semantically related to one another, leading to a semantic similarity effect. Moreover, phonological encoding of an intruder is more likely to proceed to completion when it is phonologically related to the target.

Whereas interactive models have been motivated primarily to explain speech error data, discrete feed-forward models have most of their empirical basis in chronometric studies of word production tasks. One productive method has been speeded picture naming in the presence of auditory or visual word distractors. For example, Schriefers, Meyer, & Levelt (1990) presented subjects with pictures for speeded naming along with auditory word distractors. The distractors were presented at three different SOAs: either simultaneously with, 150 ms before, or 150 ms after the picture stimuli. Also, the distractor words were systematically varied in their relationship to the target word: They could be related by meaning, related by sound, or unrelated. While all distractor conditions resulted in longer picture-naming latencies than a no-distractor condition, semantically related distractors interfered relatively more than unrelated distractors, producing even longer reaction times. On the other hand, phonologically related words were less interfering, eliciting shorter reaction times than words unrelated to the picture target. Crucially for a serial stage model of lexical retrieval, these relatedness effects depended on SOA. Semantic distractors had their relatively interfering effects only when they occurred 150 ms before the picture stimulus, and phonological distractors produced faster reaction times (relative to unrelated distractors) only when presented simultaneously with or 150 ms after the picture stimulus.

Schriefers and colleagues (1990) explained their results as follows: Picture-word interference occurred with semantically related distractors because they activated conceptual representations that partially overlapped with those of the naming target, resulting in increased competition and difficulty in lemma selection. The facilitation observed with phonologically related distractors occurred because these distractors activated the phonemic segments shared with naming target, thereby speeding their selection for encoding of the target word form. The fact that semantic interference occurred only at an earlier point in the picture naming process while phonological facilitation effects were observed only at later points supports the ideas that (1) lemma selection precedes lexeme selection and (2) there is no feedback between levels of the system.

Other evidence cited in support of a discrete, feed-forward account was produced by Levelt and colleagues (1991), who performed a series of experiments involving dual-task picture-naming and auditory lexical decision. The lexical decision stimuli were either nonwords or words that were unrelated, semantically related, phonologically related, or identical to the picture-naming target. One result compatible with both interactive and discrete, feed-forward accounts was the finding that lexical decision reaction times in the semantically and phonologically related conditions were slower than those in the unrelated condition. However, when lexical decision stimuli were phonologically related to semantic associates and alternatives to the naming target, no such interference effects were found. For example, given the picture-naming target *sheep*, lexical decision reaction times to *wool* (semantic associate), *goat* (semantic alternative), and *sheet* (phonologically related) were elevated compared to *car* (unrelated), but lexical decisions to *wood* (phonologically related to *wool*) and *goal* (phonologically related to *goat*) were not affected. Levelt and colleagues argued that these results were incompatible with an interactive account, which they argued predicts spreading of activation from semantically related items to their phonological relatives. This mediated spreading of activation was predicted to interfere with the lexical decisions under an interactive model, and since it did not, Levelt and colleagues inferred support for their discrete stage model.

How then, do interactive models account for the apparent lack of interactivity between semantic and phonological encoding in such speeded naming and lexical decision data? The explanation advanced by Dell and O'Seaghdha (1991) is that graded spreading of activation and feedback connections make such models *locally* interactive, but not *globally* interactive. They argued that three properties of models like Dell's (1986) constrain interactive effects primarily to adjacent levels of the model (e.g., between lexeme selection and phoneme assembly, but not between lemma selection and phoneme assembly). First, they suggested that feedback connections may be relatively weak (compared to the feedforward connections), and still produce the predicted biases in speech error data. Second, Dell and O'Seaghdha appealed to the system of linguistic rules that builds the slots filled by the model. They proposed that the activation delivered into the model by these rules at each successive step causes a particular step to dominate processing at any given moment, and thus limits the reach of both feedforward and feedback interaction effects. Third, to explain Levelt and colleagues' failure to observe activation of phonological neighbors of items semantically related to a naming target, Dell and O'Seaghdha focused on a particular feature of their interactive model. They noted that this sort of mediated priming (activation of *wood* given the target *sheep*, mediated by the semantic associate *wool*) involves spreading of activation by a multiplicative rule as it traverses each connection between the three units. Because each connection has a weight of less than one, the extra step (from *wool* to *wood*) results in relatively less activation of words bearing such a mediated relationship to the target, thus explaining the lack of any observed effect of mediated priming.

In summary, there is reasonable evidence that lexical access in single word production can be productively modeled as a series of stages including access of a semantically and syntactically specified, but phonologically empty representation (lemma), retrieval of the phonological word form (lexeme), and selection of individual phonemes. While evidence from natural and elicited speech errors has been cited in support of the claim that processing proceeds through these levels in a cascaded and interactive fashion, evidence from chronometric studies of word production suggest that the effects of interaction between the stages are relatively small and localized within the system.

2.1.6 Lexical Frequency Effects in Word Production

In addition to the question of interactivity between stages of word production, the locus of the lexical frequency effect within these stages is another issue important for the present discussion. The lexical frequency effect refers to the fact that words that occur more often are retrieved more quickly in speeded naming tasks (Bonin & Fayol, 2002; Griffin & Bock, 1998; Jescheniak & Levelt, 1994; Meschyan & Hernandez, 2002; Wingfield, 1968) and result in fewer speech errors (Dell, 1990). The lexical, as opposed to general, nature of the frequency effect in picture naming is supported by the observation that it is eliminated when an object must only be recognized, but not named (Jescheniak & Levelt, 1994; Wingfield, 1968).

Within the context of the stage models of lexical access discussed above, the phonological word form (lexeme) level is considered by some investigators to be the primary focus of the lexical frequency effect. One interesting line of evidence for this conclusion concerns the representation of homophone words pairs in the mental lexicon and their behavior with respect to frequency of occurrence. Homophone word pairs such as *week* and *weak* have distinct lemma representations by virtue of their semantic and grammatical differences, but share a common lexeme representation because they are composed of identical sounds (Dell, 1990; Jescheniak & Levelt, 1994). Jescheniak and Levelt (1994) found that when speakers produced the lower frequency member of a homophone pair in an English-Dutch translation task, they did so as quickly as would be expected based on the total frequency of occurrence of both members. That is, low frequency homophones benefited from sharing a lexeme representation with their higher frequency sound-alikes. Dell (1990) also observed the homophone frequency inheritance effect in a speech error generating task: Errors were as rare on low frequency members of homophone pairs as they were for their high frequency counterparts.

The existence of the homophone frequency inheritance effect on word production latencies, however, is not without controversy. First, it has been observed only in translation experiments where bilingual speakers are provided with a written stimulus word in their L2 that will elicit a homophone word in their native language. Second, because of the need to control for recognition latencies of the stimulus words, most of the analyses have been performed on reaction time difference scores, the interpretation of which can be problematic. Second, the effect has not been universally replicated by other groups, either in translation, picture naming, or word naming studies. Subsequent attempts to replicate the effect in French, English, and Chinese picture-naming and in Spanish-English translation have failed (Bonin & Fayol, 2002; Caramazza, Costa, Miozzo, & Bi, 2001), and debate over its existence and interpretation continue (Caramazza, Bi, Costa, & Miozzo, 2004; Jescheniak, Meyer, & Levelt, 2003; Miozzo & Caramazza, 2005).

Another study relevant to the locus of lexical frequency effects was one conducted by Griffin and Bock (1998), whose purpose was to produce evidence to distinguish between the predictions of a discrete, feed-forward model of lexical access and a cascade model permitting partial activation of unselected representations. To this end, they employed two different manipulations in a picture-naming task. First, lexical frequency was assumed, largely on the basis of the homophone evidence reviewed above, to affect phonological word-form selection. The other manipulation was designed to affect lemma access. They used a cloze-picture naming task in which the pictures were preceded by rapid serial visual presentation of printed cloze sentences that could be completed with the target name. These sentences differed in the degree of constraint they provided for the target picture naming. For example, the picture naming target “car” was preceded by “George taught his son to drive a” in the highly constraining condition, "The commercial was for a new" in the medium constraint condition, and "Peter saw a drawing of a" in the low constraint condition. This cloze constraint manipulation was assumed to affect the stage of lemma access because it provided more or less contextual redundancy, with the more constraining contexts providing more specification of the lexical concept to be encoded. This is in a way the opposite of the semantic interference effect in picture naming discussed earlier, which can be interpreted as creating uncertainty at this stage.

Griffin and Bock (1998) hypothesized that a discrete, feed-forward model of lexical access, such as Levelt's, should produce additive effects of semantic constraint and lexical frequency because, on that account, a lemma is fully selected before any phonological processing begins, and the two factors have their effects at different stages. On the other hand, under a cascade model such as Dell's, the factors could be expected to interact because of the interdependence of the processing stages. Specifically, assuming that lexeme nodes (1) have a logarithmic activation function and (2) have higher resting levels of activation when they represent high frequency words, one would predict a larger effect of contextual constraint for low frequency items. Such an interaction is indeed what they found: At increasing levels of contextual constraint, the lexical frequency effect was diminished and it disappeared entirely at the highest level of constraint. Griffin and Bock noted that a discrete, feed-forward model could account for their result by simply assuming that the input activation to the lexeme stage is weighted by the "amount of evidence favoring the selected lemma" (p. 331).

An alternative explanation not considered by Griffin and Bock (1998) is the possibility that lexical frequency and contextual constraint both affect lemma selection. According to the same additive factors logic that they employed to generate their predictions for the discrete model, a single, lemma-stage locus of the frequency and constraint effects would produce exactly the same kind of interaction as the one they observed.

Dell (1990) presented other arguments in favor of placing the frequency effect at the stage of lemma access in his model. First, he noted that lexical frequency effects in picture-naming tasks tend to be larger than those observed in written word naming tasks (Levelt, 1989). While the latter permit more direct access to a word's phonological representation, picture-naming requires lemma retrieval in every case. Second, while *target* word frequency had a strong influence on speech error rates, no frequency effect was observed in real word error *outcomes*. Dell argued that this evidence, taken together with the homophone frequency inheritance effect he observed in his error data, was most compatible with coding of word frequency at the lemma. Coding of frequency at the lemma permits a *target* word's frequency to influence processing, because the lemma is recruited early in processing, but also explains why an *outcome* word's frequency would not be a potent factor, because activation of a non-target lemma would come only via a convoluted feedback loop beginning with the target lemma and including the target lexeme, a given target phoneme, a non-target lexeme sharing that phoneme, and finally the non-target lemma. This relatively small amount of activation of the non-target lemma would in turn provide minimal additional activation to its lexeme or associated phonemes. In contrast, if frequency were coded in the lexeme, one might expect both target and outcome frequency to demonstrate an effect. Finally, according to Dell (1990), coding of lexical frequency at the lemma is compatible with the homophone frequency inheritance effect because a lexeme activated by a low frequency homophone lemma can quickly recruit its high frequency counterpart via feedback.

There is one relatively minor, but notable, inconsistency in Dell's (1990) arguments. As mentioned above, he asserted that word reading, as opposed to picture naming, may deemphasize lemma processing and thus attenuate any lexical frequency effects found at that stage of the model. With this in mind, it is somewhat surprising that Dell's method of inducing speech errors, which required subjects to read aloud two-word phrases under time pressure, demonstrated significant effects of lexical frequency.

In summary, the locus of the lexical frequency effect in word production has been placed either at or before the stage of phonological access within the current dominant models of word production. The research in this area is far from conclusive, and other studies not reviewed here have reached conflicting conclusions falling into both the former (Barry, Morrison, & Ellis, 1997; La Heij, Puerta-Melguizo, van Oostrumm, & Starreveld, 1999) and the latter (Bonin & Fayol, 2002; McCann & Besner, 1987; Wheeldon & Monsell, 1992) of the two general categories just described. A relatively small number of PRP dual-task studies employing naming tasks are also relevant to this discussion and they will be reviewed below.

2.1.7 Language Processing and the PRP Method

2.1.7.1 PRP Studies of Lexical Decision

McCann, Remington, and Van Selst (2000) performed a series of experiments using both visual word naming and visual lexical decision to investigate the attentional demands of frequency-sensitive lexical processing. They framed their experimental questions in terms of interactive activation (connectionist) models of word processing, suggesting that the stimulus-driven and self-organizing nature of these models predicts automaticity and thus minimal slowing of frequency-sensitive processing by a competing task. Two experiments employed tone identification as the primary task and visual lexical decision as the secondary task, both with manual responses. In both experiments, they found a null interaction between lexical frequency and SOA on secondary task lexical decision RTs, suggesting that frequency effects operate at or after a bottleneck or resource-limited stage, rather than at some prior stage such as perceptual encoding. Based on these results and findings of four additional experiments using visual naming as the secondary task, they concluded that frequency-sensitive lexical processing is either not as automatic as connectionist accounts suggest or is subject to top-down blocking of activation. For the current purposes, these results are consistent with the hypothesis that lexical frequency effects operate at a central stage that is subject to dual-task processing limitations. For reasons discussed above regarding the indifference of secondary-task reaction times in the traditional PRP method to serial vs. capacity-limited parallel processing, these results are also consistent with both bottleneck and resource accounts of those limitations.

Cleland, Gaskell, Quinlan, and Tamminen (2006) investigated the locus of frequency effects in word recognition using methods similar to the lexical decision experiments reported by McCann and colleagues (2000). In their first experiment, color discrimination was the primary task and auditory lexical decision was secondary. In the second experiment they used a primary tone task and a visual lexical decision task, as did McCann et al. The findings of these two experiments were quite consistent with each other: lexical frequency was *partially* underadditive with SOA. The lexical frequency effect was smaller in 100 and 200 ms SOA conditions than in the 800 ms SOA conditions, but planned simple contrasts revealed that the frequency effects were still significant in the short SOA conditions. Cleland et al. concluded that some portion of frequency-sensitive processing is automatic and/or is not blocked by competing tasks requiring central attention, and they emphasized the consistency of this result with current models of word recognition. They attributed the discrepancy with McCann et al. (who found that frequency effects were additive with SOA) to differences in their stimulus lists, noting that the absolute RT differences elicited by the frequency manipulation in the former study (50-60ms at 800 ms SOA) were much smaller than in their study (~140 ms at 800 ms SOA). Cleland and colleagues' frequency effect at 100 ms SOA was 70-80 ms, suggesting that the manipulation of lexical frequency in the previous study was simply too weak to elicit the interaction.

Cleland and colleagues' (2006) claim in this regard, however, is suspect and difficult to evaluate. Their stimuli were taken from the CELEX database (Baayen, Piepenbrick, & Gulikers, 1995) and their participants were speakers of British English, while McCann et al. (2000) used the Kucera and Francis (1967) frequency norms, and their data were collected from American English speakers. The low frequency lists used in the two studies had similar counts in their respective corpora, but the high frequency lists differed significantly ($p < 0.001$), with the Cleland et al. list having a higher count ($m = 374$, range 100-2000 occurrences per million) in the CELEX database than the McCann et al. list ($m = 236$, range 100-1030) had in the Kucera and Francis count. Even if the two lists were directly comparable on frequency, however, it seems unlikely that the modestly increased strength of the manipulation by Cleland et al. could have more than doubled the size of the frequency effect in terms of absolute RT. Allen, Wallace, and Weber (1995) found no differences in lexical decision RT to a medium-high frequency word list (151-236 occurrences per million) and a very high frequency word list (240-1016 occurrences). Given the considerable difference in lexical frequency effects obtained in the two studies, it seems likely that other factors were operating in addition to lexical frequency. Candidate factors include imageability (Morrison & Ellis, 2000), number of meanings (Ferraro & Hansen, 2002), and orthographic neighborhood size and/or frequency (Ferraro & Hansen, 2002; Morrison &

Ellis, 2000; Sears, Hino, & Lupker, 1995; Siakaluk, Sears, & Lupker, 2002; Ziegler & Perry, 1998), all of which have been shown to interact with lexical frequency and/or affect lexical decision RTs in their own right. Neither McCann et al. nor Cleland et al. explicitly controlled any of these variables, suggesting that some or all of them may have been partly or wholly responsible for the underadditive SOA by frequency interaction found in the latter study.

Allen et al. (2002), in a study of age differences in dual-task performance, found more complete underadditivity of frequency effects with SOA. They performed two PRP experiments using a two-alternative forced choice shape ID task as the primary task and visual lexical decision as the secondary task, both with manual responses using opposite hands. Because the primary aim of the study was to investigate age effects, they included younger and older subjects. Overall, they found that frequency effects on lexical decision RT were on average smaller at shorter SOAs than at longer SOAs, suggesting that frequency-sensitive processing was not subject to the central limitations proposed by either the central bottleneck or central resource models discussed here. In their second experiment, in which more practice was given and where the results were more consistent across SOA conditions, this result was qualified by a three-way Age x SOA x Frequency interaction. The results for their younger group demonstrated a 114 ms frequency effect at 1000 ms SOA vs a 37 ms effect at 250 ms SOA (SOA x Frequency interaction $p = 0.053$). Their older group showed greater underadditivity, with a 265 ms frequency effect at 1000 ms SOA and a 16 ms effect at 250 ms SOA (interaction $p < 0.01$).

A follow-up PRP study (Lien et al., 2006) employing both auditory and visual discrimination tasks in the primary position with visual lexical decision as the secondary task produced similar results. Specifically, older adults demonstrated the same underadditive lexical frequency by SOA interaction that they had shown in the earlier study (98 ms and -2 ms frequency effects at 900 ms and 100 ms SOA, respectively). Younger adults, by contrast demonstrated a lexical frequency effect that was consistent across levels of SOA (71 ms and 92 ms frequency effects at 900 ms and 100 ms SOA, respectively). A second experiment in this study found that both older and younger adults demonstrated additive effects of task difficulty and SOA for a secondary shape discrimination task. The authors concluded that older adults' greater cumulative experience with word reading enable them to perform frequency-sensitive lexical processing in parallel with response selection for a competing task, while the same processing requires central attention in younger adults. To support this hypothesis, they cited correlations of WAIS-R Vocabulary scores with indices of parallel processing derived from task 1 and 2 RTs in long and short SOA conditions. These correlations were moderately strong and statistically significant for the older adults only, suggesting that their ability to perform frequency-sensitive stages of lexical access in parallel with the competing task was associated with their overall verbal ability. These correlations were weak and nonsignificant for younger adults, and for both groups when the secondary task was a nonlexical shape discrimination task.

Both Allen et al. (2002) and Lien et al. (2006) employed stimulus lists with Kucera and Francis (1967) frequency counts that were comparable to McCann et al. (2000), except that their low frequency words were slightly more frequent (10-30 occurrences per million vs. 1-14 occurrences). Their lexical frequency effects were larger on average than those found by McCann et al., but as with the studies discussed above, they did not control for potentially confounding factors. Nevertheless, their young participants demonstrated null or weak interactions between lexical frequency and SOA on secondary task lexical decision RT, in general agreement with McCann et al. Their older participants, on the other hand showed even more complete underadditivity than observed by Cleland et al. (2006).

Tombu (personal communication, 2005) proposed an explanation for the underadditive interactions between lexical frequency and SOA on task 2 lexical decision RTs observed by Allen et al. (2002), based on earlier evidence for the existence of a post-central response *execution* bottleneck when competing tasks require similar responses (De Jong, 1993). Both Allen et al. and Lien et al. (2006) employed manual responses for task 1 and task 2, meaning that a response execution bottleneck may have been operating in addition to the more commonly discussed central bottleneck. Such a response execution bottleneck would lead to an underadditive SOA by lexical frequency interaction regardless of whether the locus of the frequency effect is central or pre-central. This explanation is intriguing but does by not itself account for the observed differences between younger and older adults, as De Jong's (1993) work demonstrated evidence for the manual response execution bottleneck in younger subjects. Thus, the younger groups in both studies described above should have also shown the underadditive SOA by frequency interaction, as should have McCann and colleagues' (2000) participants. However, De Jong (1993) also noted that sufficiently long task 2 central processing would effectively eliminate the response execution bottleneck, thereby restoring the null SOA by task 2 difficulty interaction predicted by the (single) central bottleneck model. In order for a post-central bottleneck to account for Lien and colleauges' (2006) three-way SOA by frequency by age interaction, it would have to be the case that frequency-sensitive processing central processing takes *longer* for younger subjects. This hypothesis, which is similar to (but distinct from) Lien and colleagues' hypothesis of increased parallel processing of frequency-sensitive stages of word recognition, seems unlikely, because older adults had longer overall RTs in their experiments, by 200 to 400 ms depending on the condition. In order to be consistent with the observed data, the post-central bottleneck hypothesis would require that older subjects' post-central, response execution stages be much longer than younger subjects (by more than 200-400 ms) in order to offset their shorter central processing in the context of longer overall RTs. Whereas Lien and colleagues' proposed that older adults are more able to perform lexical processing in parallel with competing tasks than are younger adults, the post-central bottleneck hypothesis described here proposes that aging is associated with more efficient central, frequency-sensitive lexical processing, but drastically slower response initiation and execution. Current data regarding the increase of simple RT and basic choice RT with age provide equivocal support for this hypothesis. Analysis of RT data from a large sample of British adults found average increases of approximately 175 ms between age 20 and 75 for simple RT and approximately 280 ms for a basic 4-alternative choice RT task (Der & Deary, 2006). If the dual-task context exacerbated these differences for older subjects and the extra processing time were required at stages associated with response initiation and execution, the post-central bottleneck explanation might be tenable.

As unlikely as the post-central bottleneck explanation for the age-dependent SOA by frequency interaction appears, it would also explain an apparent contradiction pointed out by Reynolds and Besner (2006) in the results obtained by Cleland et al. (2006), who found that lexical frequency was partially underadditive with SOA in lexical decision. This contradiction concerns the effects of lexical frequency on primary task tone identification RT in their second experiment: Both the main effects of SOA and lexical frequency were significant, such that primary tone RTs were longer at shorter SOAs (by 40ms) and in the low lexical frequency conditions (by 20 ms). Their graphical presentation also suggested that the lexical frequency effect on RT1 was larger at shorter SOAs, but the interaction was not significant. These findings are similar to Tombu and Jolicoeur (Tombu & Jolicoeur, 2002a) in showing effects of SOA and task 2 difficulty on RT1. These results are consistent with the resource model of the PRP dual-task performance, if the task 2 difficulty manipulation has its effects at the central stage of the model and task 2 difficulty influences allocation ratio (Tombu & Jolicoeur, 2002a; 2003). As discussed above, Cleland et al. (2006) found an underadditive SOA x Frequency interaction, which is traditionally considered diagnostic of a *pre-central* locus for at least some of the effect. Tombu and Jolicoeur (2003; 2005) pointed out that a pre-central manipulation of task 2 difficulty should have the opposite effect on RT1 under the central resource model. By delaying the point at which the two tasks must share scarce central resources, a longer task 2 pre-central stage actually *speeds up* RT1 at short SOAs. Cleland et al. found that RT1 was slower in low frequency conditions, suggesting that a substantial portion of their frequency effect was occurring at central stages of the secondary lexical decision task. The post-central bottleneck hypothesis is one way to resolve the seeming contradiction: if the bottleneck (or resource-scarcity) causing the underadditive SOA x frequency interaction on RT2 were post-central, then the frequency effect could plausibly still be operating at the central stage, giving the observed result of longer task 1 tone RTs with low frequency words. An alternative account of the seeming contradiction between task 2 RTs showing underadditivity with SOA and task 1 RTs suggesting a central frequency effect is that the frequency manipulation affected both pre-central and central stages. If this were true, and if the shift in allocation ratio induced by the central portion of the frequency effect were even moderately large, it would overwhelm the RT1 savings due to the pre-central portion of the effect, and the observed result would obtain. Given Cleland and colleagues' lack of control over potentially confounding factors such as imageability and neighborhood effects, among others, it seems likely that the frequency manipulation operationalized in their stimuli had its effects at multiple loci.

In summary, PRP studies using lexical decision as a secondary task have shown that lexical frequency is additive with SOA in some cases (Allen et al., 2002; Lien et al., 2006, McCann et al., 2000) and least partially underadditive in others (Allen et al., 2002; Cleland et al., 2006; Lien et al, 2006). The lack of experimental control over potentially confounding factors, however, makes the differences between the studies difficult to interpret and suggests the need for further study in this area.

2.1.7.2 PRP Studies of Naming

In addition to the lexical decision experiments discussed above, McCann et al. (2000) also investigated frequency effects in naming. In four separate PRP experiments with tone ID as the primary task and visual word naming as secondary, they varied the lexical frequency of the naming targets. Tone responses were manual and naming RTs were collected by voice key. In all four of these experiments they found a null interaction between lexical frequency and SOA on naming RT2, suggesting a central locus for frequency sensitive processing in visual word naming. As noted by Reynolds and Besner (2006), however, in the five comparisons across McCann and colleagues' four naming experiments showing additivity of frequency and SOA, the frequency effect was smaller at the shortest SOA than at the longest SOA in each case. This suggests partial underadditivity that approaches significance by a sign test ($p < 0.0625$).

Reynolds and Besner (2006) performed a series of PRP visual word naming experiments to further investigate the attentional demands of visual word processing. In their first experiment, they examined repetition priming, which has been shown to speed up production of low-frequency words more than high-frequency words in both word naming (Scarborough, Cortese, & Scarborough, 1977) and picture naming (Griffin & Bock, 1998; Oldfield & Wingfield, 1965). They found that repetition priming was significantly underadditive with SOA on secondary-task visual word naming RTs for low-frequency exception words, with a 54 ms priming effect observed at 750 ms SOA and a 12 ms effect at 50 ms SOA. From this result, they concluded that there is an early stage of visual lexical processing that does not require central attention, and that at least a portion of the lexical frequency effect relating to orthographic-lexical processing is underadditive with SOA. They further conjectured that the additive component of the frequency effect relates to accessing the phonological output lexicon. An alternative interpretation is that repetition priming could have both a precentral and a central component, and that it is only the central component of the priming effect that interacts with frequency. In a series of additional experiments, Reynolds and Besner found that nonword letter length, grapheme-phoneme complexity, and orthographic neighborhood density were all additive with SOA. It was concluded that assembled phonological recoding in reading aloud requires central attention.

Ferreira and Pashler (2002) also studied single word production in the context of the central bottleneck model using the PRP method. In contrast to both written word naming PRP studies reviewed above, they used a picture naming task. The goal of the study was to determine whether various stages of word production are subject to central bottleneck effects similar to those observed previously in non-linguistic tasks, or whether those stages are modular and automatic. Word production processes were described in terms of a general three-stage spreading activation model containing lemma selection, lexeme (phonological word form) selection, and phoneme selection stages. One experiment in this study involved a primary cloze-picture naming task and a secondary 3-alternative forced choice tone identification task. Naming RTs were measured by voice key, and the tone responses were given by manual keypress. The cloze-picture naming task was manipulated in the following ways: the cloze sentence contexts were either highly constraining or unconstraining for the naming target, and the naming targets were of either high or low lexical frequency. The cloze constraint manipulation was assumed to affect the difficulty of lemma selection, with lemma selection proceeding more quickly when the sentence context was highly constraining. The lexical frequency manipulation was assumed to affect phonological word form selection, with word form selection proceeding more quickly for high frequency naming targets. The important result for present purposes was that, after accounting for the expected interaction between cloze constraint and lexical frequency, both manipulations significantly affected RTs for naming (RT1) and tone ID (RT2) by approximately the same amount across the low range of SOAs. These results suggest that frequency sensitive processes, along with those affected by conceptual constraint, participate in "central processing mechanisms" (Ferreira and Pashler, 2002, p. 1187) from the perspective of the central bottleneck model. Using the terminology of their general model of word production, their findings place both lemma selection and lexeme selection within the central stage of the central bottleneck model.

In a second PRP experiment, Ferreira and Pashler (2002) varied the difficulty of a primary picture naming task by visually presenting semantically and phonologically related distractor words 0 or 100 ms prior to the picture stimulus. The former were assumed to affect lemma selection, and the latter were assumed to operate at the phoneme selection stage of the spreading activation model. In this experiment, slowing in naming RTs produced by the semantically related distractors was propagated onto secondary tone ID responses, suggesting that lemma selection occurs at or before the central stage of the dual-task model. Additionally, facilitation of naming provided by the phonologically related distractors did *not* significantly affect tone RTs, suggesting that phoneme selection operates after the central bottleneck stage, and may be considered part of response execution processing.

A similar study by Sullivan and Macchi (2002) provides further evidence regarding the picture-word interference effect and the placement of lemma selection in the central bottleneck model. In this study, tone identification was the primary task and naming was secondary. Auditory word distractors that were semantically related or unrelated to the target name were presented with the picture stimuli. Distractor relatedness produced an underadditive interaction with SOA such that semantic distractors significantly slowed naming RTs at 1000 ms SOA but not at 50 ms SOA. This finding was replicated in a recent study that differed meaningfully only in that that the word distractors were presented in the visual modality, centered inside the picture naming stimuli (Dell'Acqua, Job, Peressotti, & Pascali, in press). Considered together with Ferreira and Pashler’s (2002) findings, these results are consistent with a *pre-central* locus for the effects of semantically related distractors, suggesting (in contrast to Ferreira and Pashler’s conclusion) that lemma selection occurs in whole or part prior to the central bottleneck. Furthermore, if lexical frequency is additive with SOA on secondary-task naming RTs, as found by McCann et al. (2000), these results are most consistent with a phonological word form (lexeme) locus for frequency-sensitive processing.

2.2 SUMMARY AND STATEMENT OF PURPOSE

In the context of the central bottleneck and resource models discussed above, the results of Ferreira and Pashler (2002) suggest that lexical frequency affects either the pre-central or central components of picture naming, and their results effectively rule out the post-central response execution stage as a potential locus. By the same token, McCann and colleagues' (2000) results place lexical frequency effects in visual word naming at either the central or post-central stage, and argue against a pre-central locus. Taken together with picture-word interference data suggesting a pre-central locus for semantic processing (Dell'Acqua et al., in press; Sullivan & Macchi, 2002), these results are most consistent with the view that frequency-sensitive processing in lexical retrieval is associated with the phonological word form, and that phonological word form selection participates in the central (bottleneck or resource-limited) stages of the dual-task models under consideration. Reynolds and Besner (2006) found evidence that repetition priming effects on word naming have a pre-central locus, suggesting that a portion of the lexical frequency effect may also be pre-central. However, they further conjectured based on this result that there is a central component of the frequency effect involving access of the phonological output lexicon, in agreement with the arguments advanced immediately above. A small handful of other PRP studies have found evidence that frequency effects occur at pre-central stages, but only in the context of lexical decision tasks (Allen et al., 2002; Cleland et al., 2006; Lien et al., 2006). Furthermore, in one case the diagnostic underadditivity of frequency with SOA was incomplete, suggesting both pre-central and central effects (Cleland et al., 2006), and in both other cases underadditivity was reliably observed only in older adults.

Each of these studies employed the traditional PRP method, which designates a given task as primary and emphasizes the speed of the first response. For this reason, none of them directly addressed the issue of whether lexical frequency effects are better characterized as imposing a serial processing bottleneck or participating in graded sharing of capacity-limited resources, although aspects of Cleland and colleagues' (2006) results suggested resource-sharing. The purposes of this research were threefold: The first purpose was to replicate Ferreira and

Pashler's (2002) results localizing lexical frequency effects in picture-naming to central or pre-central stages of dual-task processing. The second purpose was to replicate and extend to picture naming the results of McCann et al. (2000) with regard to a central locus for frequency-sensitive processing. The third purpose was to replicate and extend to the domain of word production the work of Tombu and Jolicoeur (2002a) by investigating whether, under conditions promoting equal task emphasis, patterns of dual-task interference associated with lexical frequency effects in picture-naming are more consistent with the central resource or central bottleneck model.

All of the proposed experiments followed Ferreira and Pashler (2002) in proceeding from a stage model of spoken word production including lemma, lexeme, and phoneme access. Based on evidence reviewed above, it will be assumed that lexical frequency primarily affects phonological word form (lexeme) selection. Although the question of whether lexical access proceeds in discrete or cascaded steps is important, it has little direct bearing on the proposed work, provided that the relevant stage(s) of the word production model involve a selection operation that requires modular central processing as conceived in the bottleneck and resource models being considered (Ferreira & Pashler, 2002). For similar reasons, the related question of whether the stages of lexical access are strictly feed-forward or interactive is also not crucial in the current studies. If feedback were operating across stages of the dual-task model, for example, from a central lexeme selection stage to a pre-central lemma selection stage, such effects should be small in comparison to those observed on the stage that is itself the source of those effects.

2.3 SIGNIFICANCE

The present research was designed to contribute to the understanding of normal language processing and dual-task performance, and to form the basis for future investigations of language and dual-task performance in aphasia. Resource models thus far represent the major alternative to the theoretical view of aphasia as a disorder of linguistic competence and the associated Wernicke-Lichtheim model. Despite its status as the dominant paradigm in aphasia theory and practice, this disconnection-based model's neuroanatomical underpinnings are tenuous at best (De Bleser, 1988) and the classification system it supports is of limited theoretical and clinical value (Darley, 1982; Marshall, 1986; Schwartz, 1984). At the same time, resource models as applied to aphasia are currently under-specified and alternative models of attention and dual-task performance have not been adequately considered (Shuster, 2004). The present research provides a basis from which to examine these issues. Future studies in this line of work will further define important variables in aphasic language performance and help elucidate how those factors might be used to improve diagnosis, assessment, and treatment of aphasia.

3.0 EXPERIMENT 1

3.1 RATIONALE

Investigating whether the central bottleneck or central resource model better describes dual-task interference related to lexical frequency processing requires assumptions regarding the locus of this frequency-sensitive processing. Based on the prior studies reviewed above, it is proposed that in word production tasks, lexical frequency effects engage central processing related to response selection, as opposed to more peripheral processes related to perceptual encoding or response execution. However, given the small number of studies in this area, it was desirable to replicate the results supporting this conclusion. Experiment 1 attempted to replicate Ferreira and Pashler's (2002) findings supporting a central or pre-central locus for lexical frequency effects. It was predicted that when picture naming was the primary task, the slowing of naming reaction times caused by lower-frequency naming targets would be reflected in secondary tone identification reaction times. If secondary tone reaction times were not affected by the lexical frequency of primary naming targets, it would suggest that frequency-sensitive processing either does not compete with any component of this particular secondary task or is not resource-limited. In the context of the bottleneck and resource models reviewed above, this would suggest a primarily post-central or response execution locus for the effects of word frequency.

3.2 RESEARCH DESIGN AND METHODS

3.2.1 Participants

Twenty-six subjects (17 females, 9 males) participated in Experiment 1. They were recruited through the University of Pittsburgh Department of Psychology Research Participation Program. All participants met the following inclusion criteria: American English as their native language; aged 18 to 49 years old; pure-tone thresholds ≤ 35 dB HL at 500, 1000, 2000, and 3000 Hz in at least one ear; 20/40 vision or better, aided or unaided, screened using the reduced Snellen chart; word finding skills above the 5th percentile for their age cohort based on a modified version of the Brief Picture Naming: Nouns subtest of the *Test of Adolescent/Adult Word Finding* (TAWF) (German, 1990); and negative self-reported history of communication disorder, learning disability, neurological illness, head injury, and psychiatric illness. Two potential subjects were excluded from participation for failure to meet the naming criterion. Also, data from two participants, both female, were excluded because of examiner error in administering the task instructions. The remaining 24 participants whose data were submitted to analysis ranged in age from 18 to 23 years old (m = 18.3, sd = 1.05).

3.2.2 Apparatus and Stimuli

Data were collected in a sound-attenuated booth on a Dell Latitude D620 laptop PC using E-Prime (Schneider, Eschman, & Zuccolotto, 2002). Picture stimuli were presented on a CRT monitor with a 60 Hz refresh rate. Tone stimuli were presented binaurally via headphones. Manual reaction times were collected via a PI Engineering X-Keys keypad. Vocal reaction times were collected by the integrated voice key included in the Psychology Software Tools serial response box, using a lapel microphone. Vocal responses were recorded on a separate laptop PC using Adobe Audition. The tone stimuli were recorded to the same files in order to make naming reaction times recoverable from the audio files for comparison with values recorded by E-Prime via voice key when necessary.

Voice keys measure vocal response latencies with a degree of constant error because they detect a response only when the acoustic energy associated with the response has risen beyond some threshold value. This means that the time stamp for the naming response is always logged slightly after the articulatory gestures for the response have in fact begun and after the response is visible in the acoustic waveform. Because of this constant error, the IRI values logged by E-Prime misrepresented response order for a small number of low-IRI trials. In the current picture-first experiment, this would have resulted in the invalid exclusion of some trials. In order avoid this, response order for all trials with logged IRIs > -150 ms and < 0 ms was determined by visual inspection of the acoustic waveform. Whenever the acoustic waveform indicated that response order in fact matched presentation order (< 0.2% of trials), the trial was included in the analysis as described below. In order to maintain consistency of the voice key measurements the naming RTs themselves were not adjusted. In Experiment 2, in which the tone was presented first, an analogous procedure was used to avoid invalid inclusion of mis-ordered trials, affecting the response order coding of 0.1% of trials. Both procedures were used in Experiment 3, in which the stimuli were presented in both orders, and the response order coding of 2.2% of trials was affected.

Picture-naming stimuli were taken primarily from the University of California San Diego Center for Research in Language International Picture Naming Project (CRL-IPNP) online database of object pictures and associated normative data (Szekely et al., 2003; Szekely et al., 2004). Ten additional pictures collected from various sources were included to permit balancing of stimulus lists on length and onset characteristics. Two sets of 144 object pictures were constructed, one with high frequency naming targets, and one with low frequency naming targets. Set construction was based on the written word frequency counts published by Zeno, Ivens, Millard, and Duvvuri (1995). The high frequency targets all had raw frequencies of $\geq$ 329 and the low frequency items had raw frequencies of $\leq$ 274 in the Zeno et al. corpus. These cutoffs corresponded to the 53rd and 46th percentiles, respectively, among the 486 non-redundant names in the CRL database also found as single entries in the Zeno et al. corpus. The naming targets and their associated log frequency values are listed in Appendix B. The Zeno et al. frequency counts have been shown to account for a greater proportion of the variance in both written naming and written lexical decision times by younger and older normal subjects (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004) than the frequency counts included in the CRL-IPNP database, which were taken from previously published norms (Baayen et al., 1995; Kucera & Francis, 1967). In the set of 288 words chosen for this study, the Zeno et al. frequency norms correlated highly (0.78) with the norms published in the CRL-IPNP database. Both frequency measures also correlated moderately with the naming reaction times published in the CRL-IPNP database (-0.47 and -0.37 for the Zeno et al. and CRL norms, respectively).

The high and low frequency lists did not significantly differ with regard to length in phonemes or syllables. Because the phonetic features of word onsets have been shown to account for a significant proportion of the variance in spoken word reaction times collected by voice key (Balota et al., 2004; Kessler, Treiman, & Mullennix, 2002), the lists were also constructed to have similar onset characteristics. The two lists were equated for consonant onset and syllable-initial stress, and had equal numbers of vowel-initial words (7 each). The lexical frequency characteristics, naming reaction times (from the CRL database), and length in phonemes and syllables for the high and low frequency lists are summarized in Appendix B. Also included in Appendix B are the results of statistical tests demonstrating significant differences between the high and low frequency lists in lexical frequency and reaction time, and nonsignificant differences in length.

The stimulus lists were not equated for age of acquisition. Although both subjective age of acquisition ratings and objective measures drawn from psycholinguistic studies have been shown to account for substantial variance in naming times and are highly correlated with lexical frequency counts, there are a number of reasons to be cautious in applying them to the current experiments. First, subjective age of acquisition ratings are highly correlated with factors that affect the age at which a word is learned, such as frequency, imageability, length, and familiarity. Because the age at which words are learned is itself what the subjective ratings seek to measure, this creates a "circularity problem" that makes their interpretation problematic (Zevin & Seidenberg, 2002; Zevin & Seidenberg, 2004). The objective measures taken from studies of children represent the outcome of psycholinguistic processes, are subject to the same confounds and difficulty in interpretation as subjective ratings (Zevin & Seidenberg, 2004). Finally, although age of acquisition measures have occasionally shown larger effects on naming and lexical decision reaction than were predicted based on frequency alone, and have shown significant effects between word lists equated for frequency, both current theory and evidence suggest that age of acquisition and lexical frequency may affect similar components of word processing (Bonin, Barry, Méot, & Chalard, 2004; Ghyselinck, Lewis, & Brysbaert, 2004).

Thus, it was concluded that the statistical benefits of maximizing the number and variety of stimuli outweigh the potential theoretical advantages of controlling for age of acquisition.

Tone identification with manual response was chosen as the competing task for three reasons. First, this choice minimizes the potential for interference with the naming task due to input or output modality. Also, the tone task was chosen to minimize potential interference related to language processing, because the dual-task models under study are concerned with domain-general mechanisms of attention. While tone identification is certainly susceptible to verbal mediation, it is reasonable to assume that any language-processing load imposed by the task would be small relative to other components, and relative to the competing naming task. Finally, tone identification tasks have been productively used in with the PRP method in a large number of published studies. This was deemed an important consideration because predictions of the current investigations depend in large part upon the replication of previous findings. Tone stimuli were be 285 ms in duration and low (400 Hz), medium (1000 Hz), or high (2500 Hz) in pitch.

3.2.3 Procedure

All data were collected in a single session for each participant requiring approximately 60 minutes. First, participants completed the informed consent process and screening tasks described above. Next they performed 36 picture naming and 36 tone identification practice trials in isolation, followed by two blocks of 36 dual-task practice trials identical in structure to the experimental trials, but using different stimuli. Instructions for both single and dual-task trials were presented verbally and in writing on the computer screen. The text of the instructions is provided in 0. Instructions highlighted the importance of fast, accurate responses, and directed participants to respond to the naming task first on all trials, emphasizing the speed of this first response. Following practice blocks, subjects were given general feedback on accuracy and reaction time, and were reinstructed as necessary to insure compliance with instructions. After the practice blocks, participants performed four blocks of 72 experimental dual-task trials, with approximately 2 minutes of rest provided between blocks.

Each trial began with a fixation cross on the computer screen, and subjects pressed a green key with their right index finger to initiate the trial. The screen went blank when the green button was pressed and the picture stimulus appeared 500 ms later. The tone stimulus was presented 50, 150, or 900 ms following the picture stimulus. The picture remained on the screen until the voice key detected a response. The keypad used to collect the tone responses had three buttons labeled low, medium, and high. Participants responded to the tone stimulus using their left hand by pressing the low button with their ring finger, the medium button with their middle finger, or the high button with their index finger. The trial ended when responses to both tasks were detected. The next trial began 1000 ms later. The author observed all experimental trials and coded naming responses online. Any naming response that was not a fluent production of the target name without hesitation or false start was coded as incorrect. Voice key failures were also coded online by the author. Inter-rater reliability for the coding of naming responses was checked by having a second listener, a licensed speech-language pathologist with psycholinguistic research experience, listen to the audio recordings of three randomly chosen subjects and re-code their naming responses. Reliability was calculated as the number of agreements divided by the total number of trials examined. Coding errors found during reliability checking were verified by the author and corrected.

3.2.4 Design

Each block of 72 trials contained four trials representing the 18 possible combinations of the following variables: lexical frequency (high, low), stimulus onset asynchrony (50, 150, 900 ms) and tone pitch (low, medium, high). The presentation order of conditions was pseudo-randomized such that the entire design was repeated every 18 trials. Across subjects, each picture stimulus occurred approximately equally often in each condition.

3.2.5 Analysis and Hypotheses

Naming and tone identification reaction times were the dependent variables of primary interest. Excluded from the reaction time analyses were trials on which either response was incorrect, responses were mis-ordered, a voice key failure occurred, or either reaction time was determined to be an outlier according to the non-recursive procedure developed by Van Selst and Jolicoeur (1994). A single ANOVA for naming RT and tone identification RT was performed (alpha = 0.05), with three repeated factors: task (primary naming, secondary tone ID), lexical frequency (high, low), and stimulus onset asynchrony (50 ms, 150 ms, 900 ms). Error rates for each task were also examined separately, using two-way ANOVAs with lexical frequency and SOA as repeated factors, to insure that no speed-accuracy tradeoff had occurred. The following hypotheses were of primary interest for addressing Specific Aim A: To investigate the temporal locus of cognitive performance limitations resulting from lexical frequency effects.

Hypothesis A.1: *Picture naming reaction times will be significantly longer for low frequency words than for high frequency words.*

Hypothesis A.2: *Tone identification reaction times will be progressively and significantly longer when the tone stimuli are presented at progressively shorter stimulus onset asynchronies following a picture-naming stimulus.*

Hypothesis A.3: *Secondary task tone identification reaction times will be significantly longer when tone stimuli follow pictures with low frequency names than when they follow pictures with high frequency names. Across the lower values of stimulus onset asynchrony, lexical frequency will have an approximately equal effect on naming and tone ID reaction times.*

In addition to the above hypotheses related directly to Specific Aim A, the following prediction is made based on previous findings using the PRP method:

Hypothesis A.4: *Stimulus onset asynchrony will have a nonsignificant effect on naming reaction time, but a large and significant effect on tone identification reaction time.*

Power analyses were conducted using the MANOVA procedure in SPSS as described by D'Amico, Neilands, and Zambarano (2001), which provides estimates of effect size and power for repeated measures designs given sample size, estimated cell means and standard deviations, and estimated correlations between repeated measures. A pilot study, described in Appendix D, provided estimates for the cell means, standard deviations, and correlations between conditions. The power analysis revealed that a sample size of 10 or less would be sufficient to detect the expected extremely large main effects (partial eta-squared >.50) of task, SOA, and lexical frequency, as well as the expected task by SOA interaction. However, in order to have sufficient power to rule out an unexpected task by lexical frequency interaction reflecting a smaller effect of lexical frequency on tone ID than on naming (see Hypothesis A.3 above), a larger sample size was necessary. Twenty-four subjects were estimated to provide power of .80 for detecting a standardized large effect size (partial eta-squared = .20), equivalent to an approximately 30%

reduction (or 90% increase) in lexical frequency effects on secondary tone ID reaction time, as opposed to primary naming reaction time, in the shorter SOA conditions.

3.3 RESULTS

Inter-rater reliability for coding of naming responses was 98.6%. Trials containing voice key failures (2.4%), response order errors (0.9%), naming errors (15.7%), tone ID errors (11.3%) or RT outliers (2.3% of naming responses, 1.8% of tone responses) were excluded from the reaction time analyses.

3.3.1 Error Rates

Examination of error rates by condition, averaged across subjects, did not suggest the presence of a speed-accuracy trade-off for either task. For the naming task, the main effect of lexical frequency was significant, F(1, 23) = 66.386, MSe = 0.005, $p < 0.001$, effect size (ES) = 0.743[1], with more errors on low frequency items. Neither the main effect of SOA, F(2, 46) = 0.693, MSe = 0.003, p = 0.505, ES = 0.029, nor the interaction, F(1.679, 38.627) = 0.119, MSe = 0.004, p =0.854, ES = 0.005, was significant[2].

For the tone ID task, only the main effect of SOA was significant, F(2, 46) = 3.314, MSe = 0.001, p = 0.045, ES = 0.13, with error rates decreasing as SOA increased. Neither the main

---

[1] Effect sizes are reported throughout as partial eta squared.

[2] Where the sphericity assumption was violated, the Huynh-Feldt correction to the degrees of freedom was used throughout.

effect of lexical frequency, F(1, 23) = 1.766, MSe = 0.002, p = 0.197, ES = 0.071, nor the interaction, F(2, 46) = 2.093, MSe = 0.002, p = 0.135, ES = 0.083, was significant. Mean error rates by task and condition are presented in Table 1.

Table 1. Experiment 1 error rates by task and condition.

Task	Lexical Frequency	Stimulus Onset Asynchrony (SOA)		
		50	150	900
Picture Naming	Low	0.203	0.200	0.207
	High	0.109	0.105	0.122
Tone ID	Low	0.117	0.114	0.117
	High	0.126	0.109	0.087

3.3.2 Reaction Times

The mean reaction times for each task and condition are displayed in Figure 4. The results for each hypothesis are presented below:

Hypothesis A.1: *Picture naming reaction times will be significantly longer for low frequency words than for high frequency words.* The main effect of lexical frequency was significant, F(1, 23) = 101.087, MSe = 11786, p < 0.001, ES = 0.815. Picture naming RTs were 986 ms for low frequency words and 865 ms for high frequency words, averaged across SOA conditions.

Hypothesis A.2: *Tone identification reaction times will be progressively and significantly longer when the tone stimuli are presented at progressively shorter stimulus onset asynchronies following a picture-naming stimulus.* The main effect of SOA was significant, F(2,

46) = 272.674, MSe = 12226, p < 0.001, ES = 0.922, as was the Task x SOA interaction, F(1.541, 35.432) = 540.128, MSe = 6008, p < 0.001, ES = 0.959. As SOA decreased, tone RTs increased. Tone RTs averaged 1535 ms, 1472 ms, and 904 ms in the 50 ms, 150 ms, and 900 ms SOA conditions, respectively.

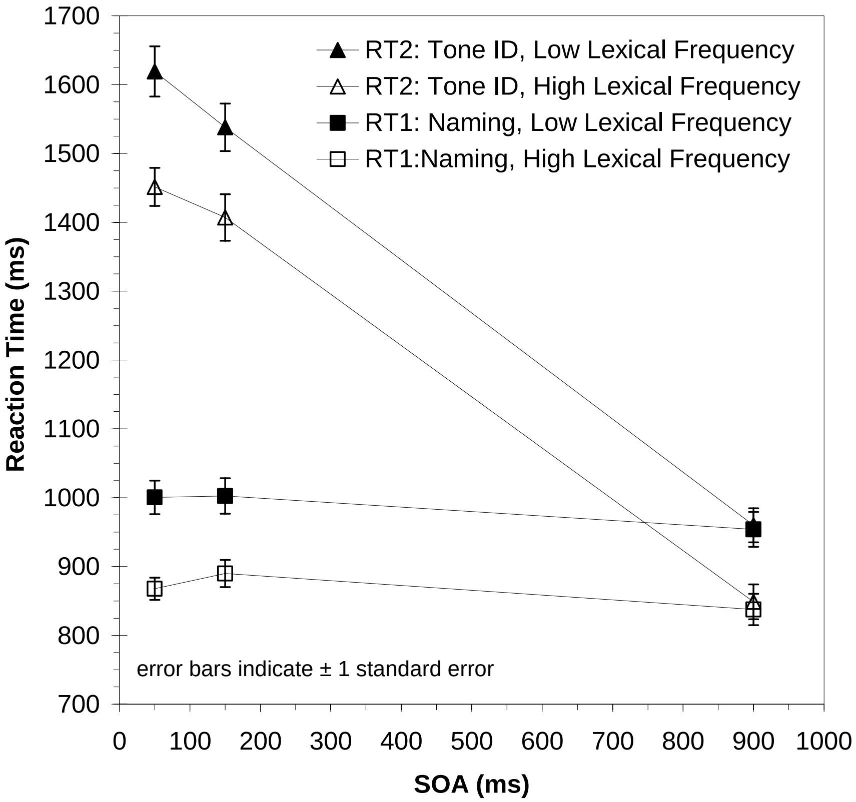

Figure 4. Mean reaction times for Experiment 1 by task and condition.

Hypothesis A.3: *Secondary task tone identification reaction times will be significantly longer when tone stimuli follow pictures with low frequency names than when they follow pictures with high frequency names. Across the lower values of stimulus onset asynchrony, lexical frequency will have an approximately equal effect on naming and tone ID reaction times.* As noted above under Hypothesis A.1, the main effect of lexical frequency was significant. The

Task x Lexical Frequency interaction was not significant, F(1, 23) = 2.444, MSe = 1892, p = 0.132, ES = 0.096, suggesting that lexical frequency affected naming and tone RTs similarly. Tone responses following low and high frequency names averaged 1372 ms and 1236 ms, respectively. The word frequency manipulation affected tone RTs by 136 ms and naming RTs by 121 ms.

Hypothesis A.4: *Stimulus onset asynchrony will have a nonsignificant effect on naming reaction time, but a large and significant effect on tone identification reaction time.* As noted above, the Task x SOA interaction was significant, consistent with the impression from Figure 4 that SOA had a larger effect on secondary tone RTs than primary naming RTs. In order assess whether the effect of SOA on naming RTs was reliable, the difference between the mean naming RTs at 150 and 900 ms was examined in a post-hoc test using the Scheffe correction for multiple comparisons. These means were selected for post-hoc testing because they represented the largest difference between SOA conditions for the naming task. The test was not significant, contrast estimate = 50.5, SE = 17.5, 95%CI = ± 86.9.

In addition to the effects described above, the main effect of task was also significant, F(1, 23) = 501.917, MSe = 20570, p < 0.001, ES = 0.956, reflecting the fact that secondary task tone RTs were on average longer than primary task naming RTs. The Lexical Frequency x SOA was not significant, F(2, 46) = 1.234, MSe = 7186, p = 0.301, ES = 0.051. The three-way Task x Lexical Frequency x SOA interaction was significant, F(2, 46) = 3.544, MSe = 672, p = 0.037, ES = 0.134. Given the significant three-way interaction, the Task x Lexical Frequency interaction was examined separately at each level of SOA, again using the Scheffe method to correct for multiple comparisons. None of the contrasts was significant, suggesting that the null task-by-frequency interaction was consistent across the SOA conditions. Thus, the interpretation of the three-way interaction is unclear, and does not have any direct bearing on the experimental predictions.

3.4 DISCUSSION OF EXPERIMENT 1

The predictions of Experiment 1 were confirmed. As predicted by both dual-task models under discussion, secondary task tone RTs slowed significantly as SOA decreased. The observation that lexical frequency affected naming and tone responses to approximately the same degree (121ms effect for naming vs. 135ms effect for tone ID) is consistent with the conclusion that frequency effects operate at the central, response selection stage of the dual-task models under consideration. As noted above, however, this result is also consistent with a pre-central, perceptual analysis locus for frequency-sensitive lexical processing. Experiment 2, in which the presentation order of the tasks was reversed, was conducted in order to investigate this possibility.

4.0 EXPERIMENT 2

4.1 RATIONALE

The purpose of Experiment 2 was to further constrain the interpretation of the results of Experiment 1 by testing whether lexical frequency effects on naming reaction time interact with stimulus onset asynchrony when picture naming is the secondary task. According to both the central bottleneck and central resource models, a null interaction in the presence of the expected main effects of lexical frequency and stimulus onset asynchrony, taken together with the results of Experiment 1, would support a central locus for frequency-sensitive lexical processing. If, however, the frequency effect is smaller at shorter stimulus onset asynchronies, this would suggest that frequency-sensitive processing shares neither structures nor resources with any component of the competing task, or that frequency-sensitive processing is not resource-limited. In the context of the central bottleneck and central resource models, this pattern of results would suggest a primarily pre-central or perceptual locus for the effects of lexical frequency.

4.2 RESEARCH DESIGN AND METHODS

4.2.1 Participants

Twenty-seven subjects (21 females, 6 males) participated in Experiment 2. None of the participants in this experiment participated in Experiment 1. They were recruited through the University of Pittsburgh Department of Psychology Research Participation Program. Inclusion criteria were identical to Experiment 1. Two potential subjects were excluded from participation for failure to meet the naming criterion. Data from three participants, all females, were excluded from analysis because of examiner error resulting in incorrect assignment of stimuli to conditions (n = 2) or failure to log a complete data set (n = 1). The remaining 24 participants whose data were submitted to analysis were all 18 or 19 years old (m = 18.2, sd = 0.41).

4.2.2 Apparatus and Stimuli

Apparatus and stimuli were the same as in Experiment 1.

4.2.3 Procedure

The procedure was identical to Experiment 1, except, that the order of tasks was reversed. The tone identification task was presented first during single-task practice blocks, and the tone stimulus occurred first on all experimental trials.

4.2.4 Design

The design was identical to Experiment 1.

4.2.5 Analysis and Hypotheses

The primary objective of Experiment 2 was to investigate whether the lexical frequency effect on secondary naming task reaction times is smaller in shorter SOA conditions. Thus, secondary naming task reaction times were analyzed in a single ANOVA (alpha = .05) with lexical frequency (high, low) and SOA (50 ms, 150 ms, 900 ms) as repeated factors. Predictions regarding primary task tone RTs, which are not directly related to Specific Aim A, will be examined in a separate, identical ANOVA. In all other respects, the analyses for Experiment 2 were identical to Experiment 1. The following hypotheses with regard to secondary naming reaction time were proposed to further address Specific Aim A: To investigate the locus of cognitive performance limitations resulting from lexical frequency effects:

Hypothesis A.1: *Picture-naming reaction times will be significantly longer for low frequency words than for high frequency words.*

Hypothesis A.5: *Picture-naming reaction times will become progressively and significantly longer when the picture stimuli are presented at progressively shorter stimulus onset asynchronies following a tone identification stimulus.*

Hypothesis A.6: *The effect of lexical frequency on secondary task picture-naming reaction times will be similar at all levels of stimulus onset asynchrony.*

The following hypotheses concerning primary task RTs, and not directly related to Specific Aim A, were also made:

Hypothesis A.7: *Stimulus onset asynchrony will have a nonsignificant main effect on tone identification reaction time.*

Hypothesis A.8: *Lexical frequency will have a nonsignificant effect on tone identification reaction time*.

The power analysis for Experiment 2 employed the same procedure and strategy as Experiment 1, and was focused solely on providing adequate power for addressing the primary question of interest, the presence vs. absence of a Frequency x SOA interaction on naming reaction time. It was estimated that a sample size of 24 participants would provide power of approximately 0.80 for detecting an unexpected large (ES = 0.19) Lexical Frequency x SOA interaction effect on naming reaction time, consistent with near-total washout of lexical frequency effects in the shorter SOA conditions.

Performing both the ANOVA of main interest and the ANOVA on tone ID reaction times at p = .05 resulted in a family-wise type I error rate of 0.0975. This moderate increase over the traditional 0.05 type 1 error rate was deemed acceptable because it was conservative in the sense that it was slightly biased to disconfirm the experimental prediction of a null Lexical Frequency x SOA interaction.

4.3 RESULTS

Inter-rater reliability for coding of naming responses was 99.1%. Trials containing voice key failures (3.7%), response order errors (0.7%), naming errors (15.1%), tone ID errors (6.2%) or RT outliers (2.3% of naming responses, 2.0% of tone responses) were excluded from the reaction time analyses.

4.3.1 Error Rates

Examination of error rates by condition, averaged across subjects, did not suggest the presence of a reliable speed-accuracy trade-off for either task. For the tone ID task, as in Experiment 1, only the main effect of SOA was significant, F(2, 46) = 4.031, MSe = 0.001, p = 0.024, ES = 0.149, with error rates increasing in the shortest SOA condition. Neither the main effect of lexical frequency, F(1, 23) = 0.211, MSe = 0.002, p = 0.650, ES = 0.009, nor the interaction, F(2, 46) = 1.732, MSe = 0.001, p = 0.188, ES = 0.07, was significant.

For the naming task, the main effect of lexical frequency was significant, F(1, 23) = 56.739, MSe = 0.008, $p < 0.001$, effect size (ES) = 0.712, with more errors on low frequency items. Although the mean error rate increased slightly with SOA, neither the main effect of SOA, F(2, 46) = 1.955, MSe = 0.003, p = 0.153, ES = 0.078, nor the interaction, F(2, 46) = 0.268, MSe = 0.003, p =0.766, ES = 0.012, was significant. Error rates by task and condition are presented in Table 2.

Table 2. Experiment 2 error rates by task and condition.

Task	Lexical Frequency	Stimulus Onset Asynchrony (SOA)		
		50	150	900
Tone ID	Low	0.066	0.058	0.055
	High	0.078	0.051	0.060
Picture Naming	Low	0.210	0.211	0.221
	High	0.091	0.098	0.118

4.3.2 Reaction Times

The mean reaction times for each task and condition are displayed in Figure 5.

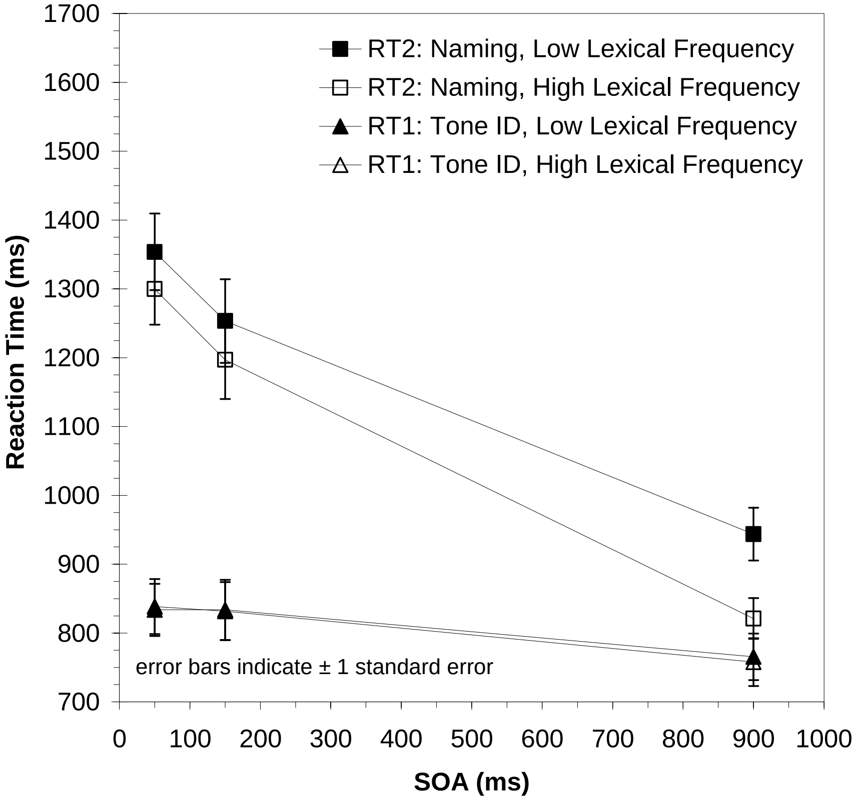

Figure 5. Mean reaction times for Experiment 2 by task and condition.

The results for each hypothesis are presented below, beginning with those concerning primary-task tone identification RTs.

4.3.2.1 RT1: Tone ID

Hypothesis A.7: *Stimulus onset asynchrony will have a nonsignificant main effect on tone identification reaction time.* Contrary to prediction, the main effect of SOA on primary-task tone RTs was significant, F(1.665, 38.290) = 12.560, MSe = 8099, p < 0.001, ES = 0.353. Tone RTs increased at shorter SOAs, averaging 836 ms, 833 ms, and 761 ms in the 50, 150, and 900 ms SOA conditions.

Hypothesis A.8: *Lexical frequency will have a nonsignificant effect on tone identification reaction time*. There was no main effect of lexical frequency on primary-task tone RTs, F(1, 23) = 0.074, MSe = 1098, p = 0.788, ES = 0.003. Tone RTs averaged 811 ms when they preceded low frequency naming responses and 809 ms preceding high frequency naming responses. The Lexical Frequency x SOA interaction was also not significant, F(1.65, 37.942) = 0.282, MSe = 2031, p = 0.713, ES = 0.012.

4.3.2.2 RT2: Naming

Hypothesis A.1: *Picture-naming reaction times will be significantly longer for low frequency words than for high frequency words.* The main effect of lexical frequency was significant, F(1, 23) = 46.235, MSe = 4685, p < 0.001, ES = 0.668. Secondary-task naming RTs were on average 77 ms slower in the low frequency conditions than in the high frequency conditions.

Hypothesis A.5: *Picture-naming reaction times will become significantly and progressively longer when the picture stimuli are presented at progressively shorter stimulus onset asynchronies following a tone identification stimulus.* The main effect of SOA was significant, F(1.32, 30.35) = 169.123, MSe = 23307, p < 0.001, ES = 0.88. Secondary-task naming RTs slowed as SOA decreased, averaging 1327 ms, 1225 ms, and 882 ms in the 50, 150, and 900 ms SOA conditions, respectively.

Hypothesis A.6: *The effect of lexical frequency on secondary task picture-naming reaction times will be similar at all levels of stimulus onset asynchrony.* Contrary to prediction, the Lexical Frequency x SOA interaction was significant, F(2, 46) = 8.306, MSe = 2197, p = 0.001, ES = 0.265. The effect of word frequency on naming RTs averaged 54 ms (t = 3.38, one-tailed p = 0.001), 56 ms (t = 3.78, p < 0.001), and 123 ms (t = 7.32, p < 0.001) in the 50, 150, and 900 ms SOA conditions, respectively. Using the Bonferroni correction to adjust the family-wise type I error rate for multiple comparisons (0.05 ÷ 15 possible pairwise comparisons = 0.003), the observed differences were reliable at each level of SOA.

4.4 DISCUSSION OF EXPERIMENT 2

Contrary to prediction, a significant effect of SOA on primary-task tone identification RTs was found. The increase in RT1 at short SOAs is suggestive of resource-sharing, and was not predicted because the instructions for Experiment 2 emphasized the speed of the first response. This should have caused participants to allocate all or virtually all processing capacity to the primary task, resulting in a flat RT1-SOA curve. One interpretation of the negative RT1-SOA curve is that, despite the instruction to emphasize task 1 (tone ID), subjects nevertheless allocated some resources to task 2 (picture naming) central processing during task 1 central processing. Alternatively, this result can be accounted for by the central bottleneck model augmented with the assumption that participants grouped a significant proportion of their responses at short SOAs only. As discussed above in Section 2.1.4, grouped trials are typically expected to have inter-response intervals (IRIs) close to zero. To evaluate whether response grouping could have accounted for the effect of SOA on RT1, the distribution of IRIs was plotted according to SOA condition. As shown in Figure 6, the mode for both short-SOA distributions was between 400 and 400 ms, with no indication of a local mode close to zero.

Also, the RT1 data from Experiment 2 were re-analyzed with trials having IRIs < 200 ms excluded. The main effect of SOA on RT1 remained significant, F(2, 46) = 10.922, MSe = 6602, p < 0.001, ES = 0.322, suggesting that it was not due to response grouping. However, the central resource model predicts a more dramatic negative slope to the RT1-SOA curve across the lower values of SOA (50 to 150ms) than was actually observed. In order to account for the observed data, the central resource model would require relaxation of the assumptions that available resource capacity and/or the allocation ratio do not vary systematically with SOA.

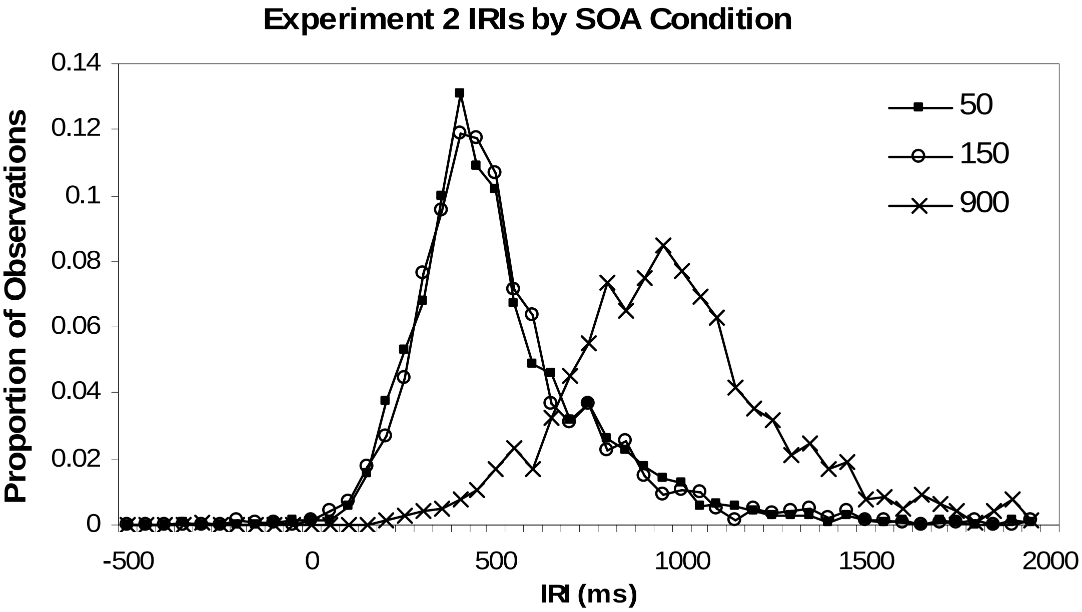

Figure 6. Distribution of IRIs for Experiment 2. Proportion of occurrences is plotted on the y-axis.

The predictions regarding the main effects of lexical frequency and SOA on picture naming RTs were upheld, but the prediction of a nonsignificant Lexical Frequency x SOA interaction was disconfirmed. Post-hoc testing revealed that there was a significant effect of lexical frequency at each level of SOA, but the effect was smaller in the shorter SOA conditions.

This partial underadditivity of lexical frequency with SOA suggests that *some, but not all* of the additional processing occurring on low frequency trials proceeded in parallel with the primary tone ID task. In the context of the dual-task models discussed above, these results are consistent with the view that lexical frequency effects, as operationalized in the current study, participate in *both the pre-central and central stages.*

Experiment 1 demonstrated a 121 ms lexical frequency effect that carried over approximately additively onto secondary-task tone ID RTs. This effectively ruled out the possibility of substantial post-central, response execution locus for frequency-sensitive lexical processing in picture naming. Experiment 2 demonstrated a similar 123 ms effect of lexical frequency on secondary-task naming RTs, but only in the longest SOA condition. At shorter SOAs, the frequency effect was smaller (54-56 ms), although still reliable. This result, taken together with the findings of Experiment 1, suggests that the frequency effect obtained with the current stimuli operate at both pre-central and central stages of the dual-task model. Of the ~120 ms frequency effect observed in the 900ms SOA condition, approximately 65 ms "washed out" in the shorter SOA conditions, suggesting that this processing time was absorbed in the delay caused by the central bottleneck or resource limitation. The maintenance of a 55 ms frequency effect in the short SOA conditions further suggests that there was some additional processing time due to the frequency manipulation that was not absorbed into the delay. In the context of the dual-task models, this latter effect that remained in the short SOA conditions may be localized to the central, response selection stage of processing. The former (65 ms) portion of the frequency effect that was absorbed into the delay imposed by the bottleneck may be localized to the pre-central, perceptual analysis stage.

A time diagram demonstrating the logic of this conclusion is presented in Figure 7. The upper panel shows that, at a short SOA the pre-central portion of the frequency effect is absorbed in the delay imposed by the central processing of the primary task. Thus, only the central portion of the frequency effect contributes to lengthening of secondary-task naming RT. In the lower panel, at a long SOA, both portions of the manipulation contribute to a larger difference in naming RT between high and low frequency conditions. The results of Experiments 1 and 2 are consistent with one another and with the conclusion that lexical frequency effects as realized in the current stimulus set participate in both the pre-central and central stages of the CB and resource models.

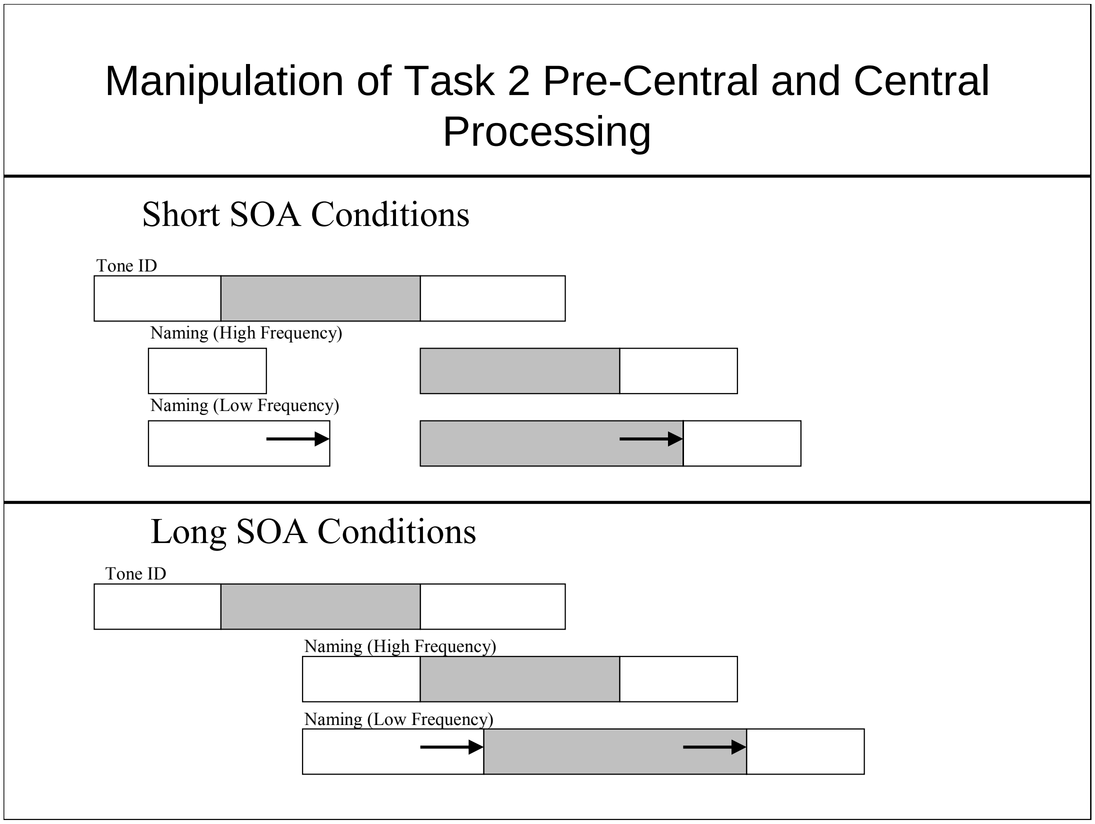

Figure 7. A schematic time diagram demonstrating combined effects of task 2 pre-central and central manipulation.

As noted above, the high and low frequency stimulus lists were balanced for length in syllables and phonemes, and for initial phoneme. They were not, however, balanced for age of acquisition, name agreement, image agreement, or object recognition time, all variables that tend to correlate with both lexical frequency and naming RT[3]. Each of these variables is discussed briefly in turn below.

Unlike the other three variables considered here, age of acquisition is solely a property of the target word, and as discussed in Chapter 3 above, is highly correlated with factors that affect the age at which a word is learned, such as frequency, length, and familiarity. Also, although age of acquisition measures have occasionally shown larger effects on naming and lexical decision reaction than were predicted based on frequency alone, and have shown significant effects between word lists equated for frequency, both current theory and evidence suggest that age of acquisition and lexical frequency affect similar components of word processing (Bonin et al., 2004; Ghyselinck et al., 2004). Thus, of the variables considered here, age of acquisition appears to be the least likely to have contributed to the pre-central, as opposed to central, effects observed in Experiment 2.

---

[3] At the time the proposal for the present work was advanced, the author was unaware of the findings of Sullivan and Macchi (2002) and Dell'Acqua et al. (in press) suggesting a pre-central locus for lexical-semantic processing in picture-naming. Based on prior PRP dual-task studies, which only found clear evidence for pre-central effects in the case of relatively peripheral manipulations of spatial attention, visual intensity, or visual distortion (Johnston et al., 1995; McCann & Johnston, 1992; Pashler & Johnston, 1989), it was assumed that age of acquisition, name agreement, object agreement, and object recognition time would all contribute substantially to response selection processing and have negligible effects at the pre-central stage. In light of the finding of partial underadditivity of lexical frequency with SOA in Experiment 2, this assumption required re-examination.

Name agreement is the extent to which participants assign the same name to a given picture, and takes into account both the number of alternative names and the proportion of participants using each alternative. Name agreement has been shown to correlate negatively with naming RT (Barry et al., 1997; Bonin, Chalard, Meot, & Fayol, 2002; Szekely et al., 2003). In the current stimulus set, name agreement correlated 0.31 with lexical frequency and -0.67 with naming RT (normative data obtained from the CRL-IPNP database). Name agreement may affect the picture naming process in at least two distinct ways (Barry et al., 1997). First, if a picture is difficult to interpret visually, or is difficult to distinguish from visually similar, *incorrect* alternatives, it may have low name agreement. In this case, the operative difficulty would seem to occur at the stage of recognizing the pictured object. Second, a picture may have low name agreement because there are multiple correct alternative names available, e.g., couch/sofa, stairs/staircase, present/gift/box. In this case, poor name agreement seems to have its source subsequent to the conceptual stage, perhaps at lemma selection. In support of this view, Vitkovich and Tyrell (1995) found that low name agreement pictures of the first type had slower object/non-object decision RTs than low name agreement pictures of the second type.

Image agreement refers to the rated extent to which a picture matches a respondent's mental image of the target name. Like name agreement, it has consistently demonstrated relatively strong negative correlations with naming RT (Barry et al., 1997; Bonin et al., 2002; Szekely et al., 2003), and also with naming difficulty of the pictures from the Boston Naming Test (Himmanen, Gentles, & Sailor, 2003). In the current stimulus sets, rated image agreement correlated 0.22 with lexical frequency and -0.56 with naming RT. Barry and colleagues (1997) proposed that that image agreement effects can be localized to the object recognition stage of the picture naming process.

Object recognition time has most often been measured by presenting subjects with a word followed by a picture, and requiring a speeded yes/no response indicating whether the picture matches the word. Using this method, both Wingfield (1968) and Jescheniak and Levelt (1994) found small, non-significant differences in object recognition time for their high and low frequency stimulus lists. Elsewhere, Levelt (2002) has argued that object recognition time constitutes an important experimental control that must be exercised when attempting to isolate the effects of lexical frequency on naming RT. Object recognition times for the current stimuli were obtained by conducting an experiment modeled on Jescheniak and Levelt (1994, Experiment 2). Object recognition time correlated weakly ($r = 0.11$, 1-tailed $p = 0.03$) with lexical frequency and moderately (0.39) with naming RT in the current stimulus set. The high and low frequency picture sets used in Experiments 1 and 2 differed in object recognition RT by 18 ms on average, t (1, 286) = 2.302, 1-tailed $p = .011$. The method and results of this experiment are described in detail in Appendix E.

In an attempt to isolate the effect of lexical frequency on naming RT in the context of the dual-task method employed here, re-analysis of the Experiment 1 and 2 data was undertaken. First, a subset of the naming stimuli employed in Experiments 1 and 2 balanced on name agreement, image agreement, and object recognition time was identified. These variables were chosen for additional balancing because each is at least partially a property of the stimulus picture, and each could be hypothesized to operate at perceptual stages of a picture naming task. Also, given the findings discussed above regarding the underadditivity of picture-word interference effects with SOA (Dell'Acqua et al., in press; Sullivan & Macchi, 2002), it seems possible or even likely that some degree of lexical semantic processing occurs at pre-central stages of picture naming, providing additional motivation for the inclusion of name agreement as a balancing variable. The picture subset thus identified included 99 high frequency and 99 low frequency words, and maintained the previous balance on length and initial phoneme, as well as a significant split on the lexical frequency variable. Subjects' RT means for each condition in experiments 1 and 2 were recalculated, including the outlier screening procedure. These means obtained from the reduced, more balanced stimulus set were submitted to ANOVAs identical to those described above.

The RT means for the Experiment 1 re-analysis are displayed in Figure 8. The results of the statistical analysis were essentially identical to the original. Most importantly, the main effect of lexical frequency was again significant, $F(1, 23) = 29.293$, $MSe = 9121$, $p < 0.001$, $ES = 0.56$, while the Task x Lexical Frequency interaction was not, $F(1, 23) = 0.039$, $MSe = 1590$, $p = 0.845$, $ES = 0.002$. The three-way Task x Lexical Frequency x SOA interaction also failed to reach significance in the re-analysis, $F(2, 46) = 1.132$, $MSe = 1060$, $p = 0.331$, $ES = 0.047$.

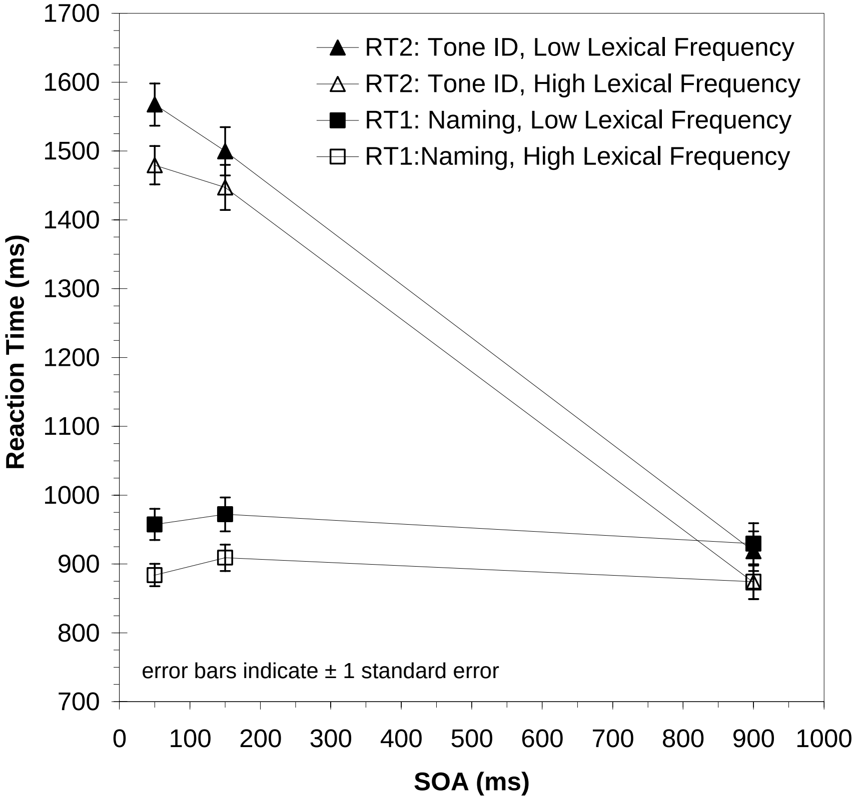

Figure 8. Mean reaction times from Experiment 1 re-analysis with naming stimuli balanced on name agreement, image agreement, and object recognition time, by task and condition.

The RT means for the Experiment 2 re-analysis are displayed in Figure 9. For naming RT2, the main effects of lexical frequency, F(1, 23) = 9.089, MSe = 5133, p = 0.006, ES = 0.283, and SOA, F(1.304, 30.002) = 177.775, MSe = 22229, p < 0.001, ES = 0.885, were again significant. The lexical frequency manipulation slowed naming RTs by 25 ms at 50ms SOA, by 27 ms at 150ms SOA and by 56 ms at 900ms SOA. The Lexical Frequency x SOA interaction, however, was not significant in the re-analysis F(2, 46) = 1.621, MSe = 2171, p = 0.209, ES = 0.066.

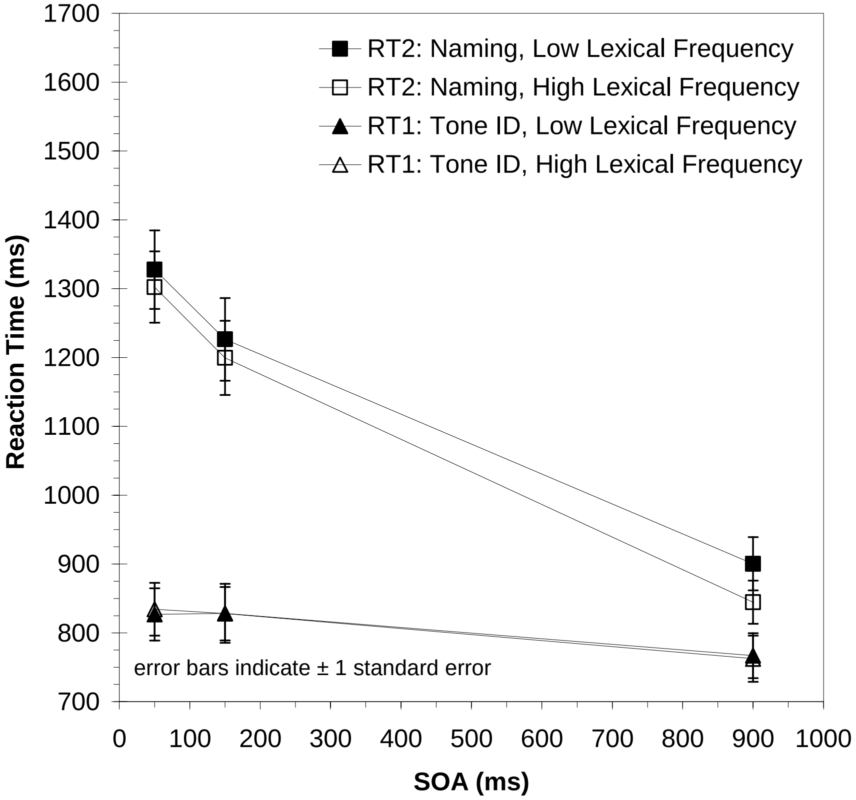

Figure 9. Mean reaction times from Experiment 2 re-analysis with naming stimuli balanced on name agreement, image agreement, and object recognition time, by task and condition.

The results of the re-analysis of Experiments 1 and 2, using a reduced set of stimuli that was balanced for a wider range of potentially confounding variables, were consistent with the predictions of those experiments. Specifically, in Experiment 1, lexical frequency affected primary naming and secondary tone RTs equally, consistent with a central, response selection effect in the CB model. In Experiment 2, lexical frequency slowed secondary picture naming RTs significantly, and although the absolute differences between the high and low frequency conditions were smaller in the shorter SOA conditions, the interaction was not statistically significant. The hypothesis that lexical frequency effects operate in the central, response selection stage of the dual-task models under study remains tenable.

5.0 EXPERIMENT 3

5.1 RATIONALE

Because of the emphasis on the primary task in the traditional fixed-order PRP method, Experiments 1 and 2 did not optimally address the question of whether the observed dual-task interference is better described as a structural bottleneck or a shared central resource that is limited in quantity and flexibly deployed. To accomplish this goal, Experiment 3 used the same tasks as Experiments 1 and 2, presented in both orders and with equal emphasis in task instructions, in order to maximize the possibility that capacity sharing would occur (Tombu & Jolicoeur, 2002a). The central resource model used to derive the predictions for Hypotheses B.1-B.3 was the same as that published by Navon and Miller (2002) and Tombu and Jolicoeur (2003). The simplifying assumptions were the same as those discussed above in section 2.1.4 above, with the additional assumption that the total available resource capacity does not vary systematically with the experimental factors of lexical frequency or SOA. Also, as proposed by Tombu and Jolicoeur (2002a), it was hypothesized that an increase in task 2 difficulty on a given trial would result in an increase in the proportion of resources allocated to task 2 on that trial, with a corresponding decrease in the proportion allocated to task 1.

5.2 RESEARCH DESIGN AND METHODS

5.2.1 Participants

Participants were 108 healthy individuals (86 females, 22 males) aged 18 to 44 (mean = 23.2, sd = 4.3), none of whom had participated in either of the prior experiments. They were recruited from the University of Pittsburgh community and the VA Pittsburgh Healthcare System via fliers, e-mail solicitation, and oral presentations to classes in the University of Pittsburgh School of Health and Rehabilitation Sciences. Inclusion criteria were identical to Experiments 1 and 2. Two potential subjects were excluded from participation for failure to meet the naming criterion and two were excluded for reporting a native language other than English. Data from five participants, all females, were excluded from analysis because of software errors resulting in the mis-timing of stimulus presentation. One additional female subject was excluded because of examiner error resulting in incorrect assignment of stimuli to conditions. Finally, 39 participants who completed the protocol were excluded from analysis for failure to obtain a sufficient number of valid trials per condition for the tone-primary conditions, as described below. Thus, data from 63 participants were submitted to analysis. These participants included 51 females and 12 males, and ranged in age from 18 to 44 (m = 23.2, sd = 4.5). All Experiment 3 participants who met the selection criteria and completed the protocol were paid for their participation.

5.2.2 Apparatus and Stimuli

The apparatus was the same as in Experiments 1 and 2. The tone stimuli were also identical to the first two experiments. The naming stimuli for the tone-first conditions were the 99 high frequency and 99 low frequency stimuli identified in the re-analysis of Experiment 1 and 2 data. As described above, these two lists were balanced on initial phoneme, length in phonemes and syllables, name agreement, image agreement, and object recognition time. An additional, non-overlapping set of 198 pictures was identified (99 high frequency, 99 low frequency) for the picture-first conditions. Because of the limited number of stimuli remaining to choose from, this latter set was not balanced for any confounding variables. The balanced high and low frequency lists were used for the tone-first conditions because these were the conditions most critical for evaluating the predictions of the central resource model. Also, the central resource model predictions for the tone-primary conditions were most clear when it could be assumed that the task 2 difficulty manipulation was localized to the central stage. The resource model predictions for the naming-primary conditions did not require such an assumption. The naming targets for Experiment 3 and their characteristics are summarized in Appendix F.

5.2.3 Procedure

Informed consent, screening, and data collection required approximately 75 minutes for each participant. Following the informed consent process and screening, participants performed 36 picture naming and 36 tone identification practice trials in isolation, followed by two blocks of 36 dual-task practice trials identical in structure to the experimental trials, but using different stimuli. Instructions for both single and dual-task trials were presented verbally and in writing. The text of the instructions is provided in 0. As before, instructions highlighted the importance of fast, accurate responses. Participants were instructed that they could respond to the tasks in either order, and that they should give equal attention or effort to both. Following dual-task practice and experimental blocks in which subjects responded in presentation order on ≤ 33% of trials in any condition, they were told that they were giving more attention to the favored task, and re-instructed to give equal attention or effort to both tasks. After the practice blocks, participants performed three blocks of 108 experimental dual-task trials and one block of 72 experimental dual-task trials, with approximately 3 minutes of rest provided between blocks.

Each trial began with a fixation cross on the computer screen, and subjects pressed a green key with their right index finger to initiate the trial. The screen went blank when the green button was pressed and the first stimulus was presented 500 ms later. The second stimulus was presented 50, 150, or 900 ms after the first. The picture remained on the screen until the voice key detected a response. The trial ended when responses to both tasks were detected. The next trial began 1000 ms later. The picture stimulus was presented first on half of the trials, and the tone stimulus was presented first on the remaining half.

Procedures for coding the naming responses and reliability checking were as in Experiments 1 and 2, except that data from six randomly selected subjects were re-coded for reliability checking.

5.2.4 Design

Each block of 108 trials contained three trials representing the 36 possible combinations of the following variables: task order (picture-first, tone-first), lexical frequency (high, low), stimulus onset asynchrony (50, 150, 900 ms) and tone pitch (low, medium, high). The final block of 72 trials contained two trials representing each of the 36 possible trial types. The presentation order of conditions was pseudo-randomized such that the entire design was repeated every 36 trials. Across the 63 subjects whose data were included in the analyses, each picture stimulus was presented equally often in each condition.

5.2.5 Analysis and Hypotheses

The primary objectives of Experiment 3 were to investigate whether the RT1 increased with decreasing SOA and whether the lexical frequency of secondary-task naming targets affected primary-task tone RT1. Because of the necessary assumption that task 1 central processing begins before task 2 central processing begins and finishes before the end of task 2 central processing, only trials on which response order matched presentation order and on which both responses were correct were included in the analyses (Tombu & Jolicoeur, 2002a)[4]. Also, because the tone-primary conditions were the most critical for testing the predictions of the resource model, an a priori criterion of ≥ 10 correct, valid responses in presentation order in each tone-primary condition was set for each subject to be included in the analyses. Any subjects failing to obtain at least 10 correct, valid trials on which response order matched presentation order were excluded from analysis.

Four separate ANOVAs were performed, each with reaction time to one task as the dependent variable: 1) primary naming (naming RT on trials where the picture stimulus was presented and responded to first), 2) secondary tone identification (tone RT on trials where the tone was presented and responded to second), 3) primary tone identification (tone RT on trials where the tone was presented and responded to first), and 4) secondary naming (naming RT on

---

[4] As described above in Section 3.2.2, because of the constant error inherent in RTs measured by voice key, the recorded acoustic waveform of the tone stimuli and vocal responses was used to insure that all picture-first trials with logged IRIs between 0 and -150 ms had been validly excluded and all tone-first trials with logged IRIs between 0 and +150 had been validly included. This procedure changed the response order coding of 2.1% of picture-first trials and 2.4% of tone-first trials.

trials where the picture was presented and responded to second). Each ANOVA had lexical frequency (high, low) and stimulus onset asynchrony (50 ms, 150 ms, 900 ms) as within-subject factors. Also, for the tone-primary conditions, planned orthogonal comparisons were used to test the simple main effect of lexical frequency at each level of SOA. Error rates were examined in similar ANOVAs to inspect for speed-accuracy tradeoffs. The following hypotheses, which concern the effects of SOA and secondary task difficulty manipulation on primary task reaction time, all address Specific Aim B: To investigate whether dual-task performance limitations in single word production are more consistent with a central resource or central bottleneck**.**

Hypothesis B.1: *Primary task reaction times will increase significantly as stimulus onset asynchrony is decreased. This hypothesis is identical for both tone identification and picture naming when they are in primary position.*

Hypothesis B.2: *Primary task tone identification reaction times will be significantly longer when secondary naming task targets are low frequency words as opposed to high frequency words.*

Hypothesis B.3: *Stimulus onset asynchrony and secondary task difficulty will interact significantly in their effects on primary task tone identification reaction times such that lexical frequency will have a significantly larger effect at shorter stimulus onset asynchronies than at longer stimulus onset asynchronies.*

The alpha level was set at 0.05 for the two primary-task ANOVAs, and 0.001 for the two secondary-task ANOVAs, which provided for a family-wise type 1 error rate of 0.1. This increase over the traditional 0.05 value was justified because it represents an appropriate balance of type I error rate, type II error rate, and practical demands related to sample size. Also, the specific combination of methods and predictions of Experiment 3 were relatively novel, and the predicted main and interaction effects of secondary task difficulty on primary task reaction time were modest in size but theoretically important. The rationale for performing primary and secondary-task ANOVAs at different alpha levels was that it helped to minimize the family-wise type I error rate while still providing adequate statistical power to test the experimental predictions.

Power analysis for Experiment 3 was targeted at providing adequate power to detect main and interaction effects of SOA and lexical frequency on primary tone ID reaction times, because a) these represent the most distinctive predictions of the central resource model vis-à-vis the central bottleneck model, and b) they are likely to be the smallest predicted effects observed in the analysis. Based on the pilot study described in Appendix D, projected effect size for the Lexical Frequency x SOA interaction was small-to-moderate (partial eta-squared = 0.066). Tombu and Jolicoeur (2002a), found a similarly-sized (non-significant) interaction of stimulus onset asynchrony and secondary task difficulty on task 1 reaction time. In the current design, the sample size of 63 subjects provided power of 0.75 for detecting this effect. Although the pilot data did not suggest a substantial main effect of lexical frequency on tone RT1, Tombu and Jolicoeur (2002a) found an analogous effect of task 2 difficulty on RT1 with an effect size of 0.22. The current study had power >0.80 for detecting the main effect of lexical frequency on tone RT1 of approximately half that size, 0.12. Based on Experiment 2, the main effect of stimulus onset asynchrony on tone RT1 was projected to be larger (partial eta-squared = 0.33), with the sample size providing power >.99 for detecting this effect.

With regard to the secondary-task ANOVAs, the predicted effects of SOA and lexical frequency on task 2 reaction time were essentially identical to the analogous effects predicted in

Experiments 1 and 2. With an alpha level of 0.001, a sample size of 63 subjects provided power >.99 for detecting these main effects on secondary task reaction time.

5.3 RESULTS

Inter-rater reliability for coding of naming responses was 99.0%. A total of 39 participants who completed the protocol failed to obtain ≥ 10 valid trials in all tone-primary conditions, and were excluded from further analysis. For the tone-primary conditions, data from 63 participants were submitted to analysis. For the naming-primary conditions, the analyses included the 31 participants who had ≥ 8 trials per condition in those conditions, in addition to meeting the ≥ 10 trials per condition for the tone-primary conditions. Trials containing operator or software errors (< 0.2% of trials), voice key failures (2.6%), response order not matching presentation order (21.9%), naming errors (11.9%), tone ID errors (7.6%) or RT outliers (1.4% of naming responses, 1.2% of tone responses) were excluded from the reaction time analyses. The mean number of trials per condition included in the reaction time analyses is summarized by task order and SOA condition in Table 3.

Table 3. Mean number of trials per condition included in Experiment 3 RT analyses by condition. Standard deviations are given in parentheses.

Task Presentation Order	Lexical Frequency	Stimulus Onset Asynchrony (SOA)		
		50	150	900
Picture-First	Low	11.5 (3.3)	13.5 (3.4)	20.5 (2.9)
	High	11.7 (2.6)	14.7 (3.5)	21.9 (2.8)
Tone-First	Low	15.9 (3.9)	18.8 (4.0)	23.5 (3.3)
	High	16.3 (4.0)	19.3 (3.9)	23.9 (3.7)

5.3.1 Error Rates: Naming and Tone ID on Naming-Primary Trials

Error rates for the naming-primary trials, i.e., trials on which the naming task was presented and responded to first, are presented by task and condition in Table 4. For the primary naming task, the main effect of lexical frequency was significant, F(1, 30) = 24.319, MSe = 0.006, $p < 0.001$, ES = 0.448, with more errors on low frequency names. The main effect of SOA was not significant, F(2, 60) = 2.739, MSe = 0.007, p = 0.073, ES = 0.084, nor was the interaction, F(1.534, 46.013) = 0.435, MSe = 0.007, p = 0.597, ES = 0.014.

For the secondary tone ID task on the naming-primary trials, only the main effect of SOA was significant, F(2, 60) = 10.555, MSe = 0.004, $p < 0.001$, ES = 0.26, with error rates increasing as SOA decreased. Neither the main effect of lexical frequency, F(1, 30) < 0.001, MSe = 0.004, p = 1.00, ES = 0.00, nor the interaction, F(2, 60) = 0.444, MSe = 0.006, p = 0.644, ES = 0.015, were significant. No speed-accuracy trade-offs were evident in either the primary naming responses or the secondary tone responses.

Table 4. Experiment 3 error rates by task and condition for the naming-primary trials.

Task	Lexical Frequency	Stimulus Onset Asynchrony (SOA)		
		50	150	900
Picture Naming	Low	0.210	0.244	0.251
	High	0.168	0.177	0.196
Tone ID	Low	0.129	0.127	0.075
	High	0.121	0.120	0.089

5.3.2 Error Rates: Tone ID and Naming on Tone-Primary Trials

Error rates for the tone-primary trials, i.e., trials on which the tone ID task was presented and responded to first, are presented by task and condition in Table 5. For the primary tone ID task, the main effect of lexical frequency was not significant, F(1, 62) = 1.092, MSe = 0.003, p = 0.300, ES = 0.017, nor was the main effect of SOA, F(2, 124) = 2.983, MSe = 0.003, p = 0.054, ES = 0.046. The interaction was also not significant, F(2, 124) = 0.771, MSe = 0.002, p = 0.771, ES = 0.004.

For the naming task on the tone-primary trials, the main effect of lexical frequency was not significant, F(1, 62) = 0.336, MSe = 0.005, p = 0.564, ES = 0.005, nor was the main effect of SOA, F(1.773, 109.947) = 1.527, MSe = 0.009, p = 0.223, ES = 0.024. The interaction was also not significant, F(2, 124) = 0.250, MSe = 0.005, p = 0.779, ES = 0.004. There was no indication of a reliable speed-accuracy trade-off in either the primary tone or the secondary naming responses.

Table 5. Experiment 3 error rates by task and condition for the tone-primary trials.

Task	Lexical Frequency	Stimulus Onset Asynchrony (SOA)		
		50	150	900
Tone ID	Low	0.076	0.094	0.076
	High	0.081	0.096	0.087
Picture Naming	Low	0.141	0.124	0.141
	High	0.139	0.125	0.129

5.3.3 Naming RT1: Naming RTs When the Naming Task was Presented and Responded to First

The mean reaction times for the naming-primary, tone-secondary trials are displayed by task and condition in Figure 10.

Hypothesis B.1: *Primary task reaction times will increase significantly as stimulus onset asynchrony is decreased.* This prediction was not upheld for the naming-primary trials. The main effect of SOA was not significant, F(1.681, 50.433) = 1.483, MSe = 10248, p = 0.235, ES = 0.047. Observed power was 0.278. Average naming RT1s were 986 ms, 999 ms, and 968 ms in the 50, 150, and 900 ms SOA conditions, respectively.

As expected, the main effect of lexical frequency was significant, F(1, 30) = 46.877, MSe = 5470, p < 0.001, ES = 0.610, with naming RT1 averaging 1022 ms for low frequency items and 947 ms for high frequency items. The Lexical Frequency x SOA interaction was not significant, F(2, 60) = 0.214, MSe = 6565, p = 0.808, ES = 0.007.

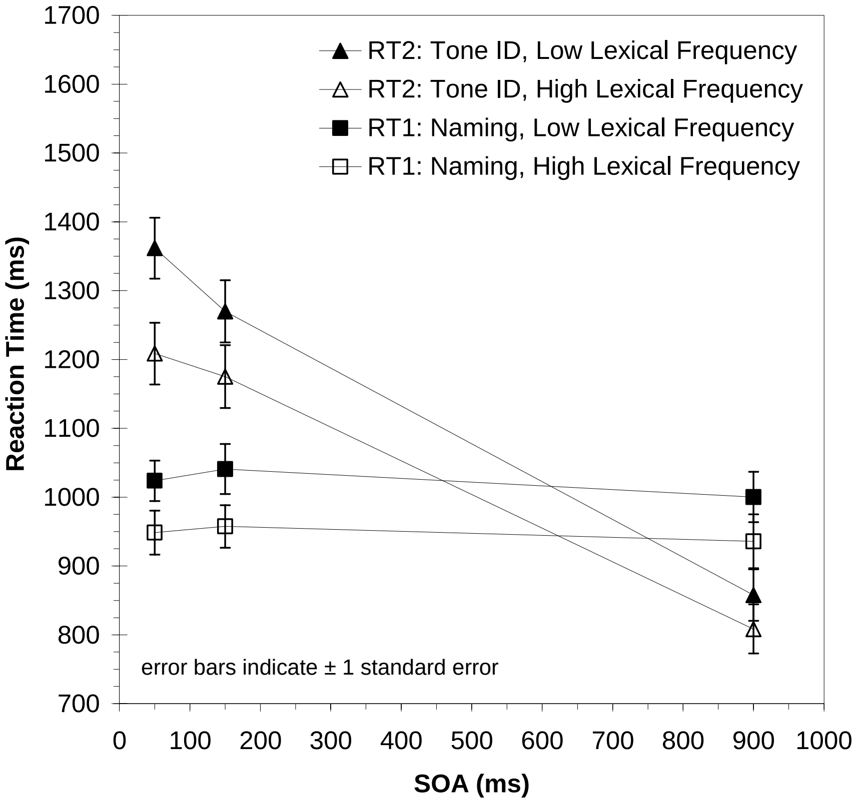

Figure 10. Mean reaction times for Experiment 3 naming-primary trials by task and condition.

5.3.4 Tone RT2: Tone RTs When the Naming Task was presented and Responded to First

The expected main effect of lexical frequency was significant, F(1, 30) = 56.461, MSe = 8100, $p < 0.001$, ES = 0.653. Secondary tone RTs averaged 1163 ms and 1064 ms in the low and high lexical frequency conditions, respectively. The main effect of SOA was also significant in the expected direction, F(2, 60) = 296.908, MSe = 12522, $p < 0.001$, ES = 0.908. Tone ID RTs increased with decreasing SOA, averaging 1285 ms, 1223 ms, and 833 ms in the 50, 150, and 900 ms SOA conditions, respectively. The Lexical Frequency x SOA interaction was not significant at the specified alpha level of 0.001, F(2, 60) = 3.968, MSe = 10542, p = 0.024, ES =

0.117, although there was a trend for the lexical frequency effect to increase with decreasing SOA. The lexical frequency effect on secondary-task tone RTs averaged 153 ms, 95 ms, and 49 ms in the 50, 150, and 900 ms SOA conditions, respectively.

5.3.5 Tone RT1: Tone RTs When the Tone ID Task was presented and Responded to First

The mean reaction times for the tone-primary, naming-secondary trials are displayed by task and condition in Figure 11.

Hypothesis B.1: *Primary task reaction times will increase significantly as stimulus onset asynchrony is decreased*. This prediction was upheld for the tone-primary trials. The main effect of SOA on primary-task tone RTs was significant, F(1.449, 89.857) = 5.836, MSe = 8410, p = 0.009, ES = 0.086. Tone RTs increased with decreasing SOA, averaging 736 ms, 722 ms, and 702 ms in the 50, 150, and 900 ms SOA conditions.

Hypothesis B.2: *Primary task tone identification reaction times will be significantly longer when secondary naming task targets are low frequency words as opposed to high frequency words.* This prediction was not upheld. The main effect of lexical frequency on primary-task tone RTs was not significant, F(1, 62) = 0.358, MSe = 2481, p = 0.552, ES = 0.006. Observed power was 0.091. Tone RTs averaged 722 ms and 718 ms in the low and lexical high frequency conditions, respectively.

Hypothesis B.3: *Stimulus onset asynchrony and secondary task difficulty will interact significantly in their effects on primary task tone identification reaction times such that lexical frequency will have a larger effect at shorter stimulus onset asynchronies.* This prediction was not upheld. Although the pattern of cell means was in the expected direction, the Lexical Frequency x SOA interaction was not significant, F(1.77, 109.724) = 0.837, MSe = 3497, p = 0.423, ES = 0.013. Observed power was 0.181. The effect of lexical frequency on primary task tone RTs averaged 11 ms, 4 ms, and -7 ms in the 50, 150, and 900 ms SOA conditions. The planned comparisons also failed to reach significance for both the 50 ms SOA condition (t = 1.01, 95%CI = ± 23) and the 150 ms condition (t = 0.427, 95%CI = ± 20).

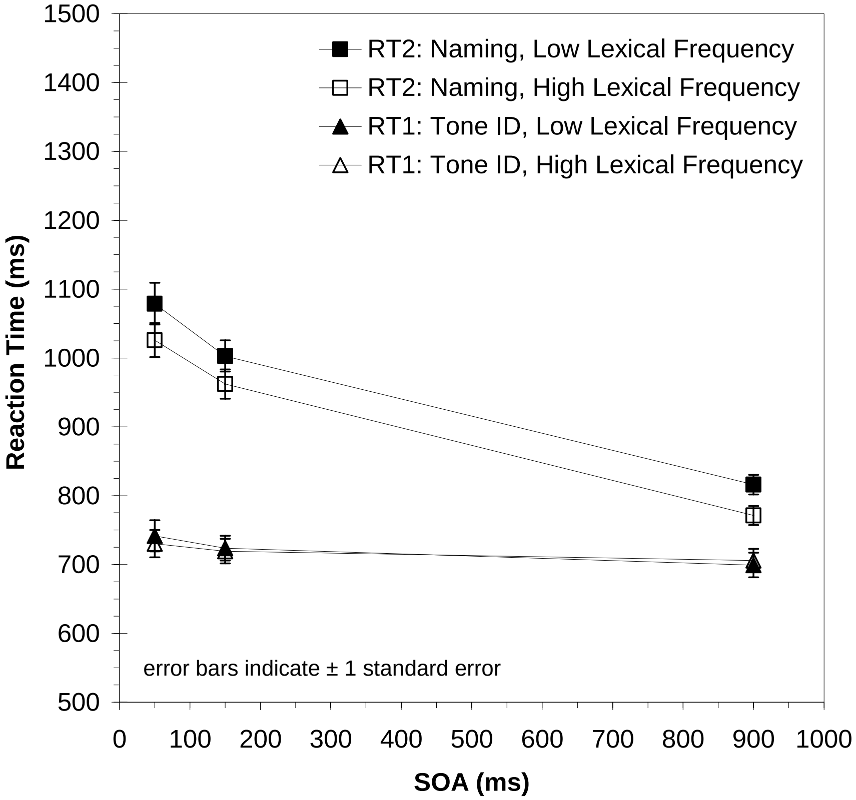

Figure 11. Mean reaction times for Experiment 3 tone-primary trials by task and condition.

5.3.6 Naming RT2: Naming RTs When the Tone ID Task was Presented and Responded to First

The expected main effect of lexical frequency on secondary-task naming RTs was significant, F(1, 62) = 36.897, MSe = 5457, $p < 0.001$, ES = 0.373. Secondary-task naming RTs averaged 966 ms for low frequency items and 920 ms for high frequency items. The main effect of SOA was also significant as expected, F(1.535, 95.172) = 172.133, MSe = 17096, $p < 0.001$, ES = 0.735. Secondary naming RTs increased as SOA decreased, averaging 1052 ms, 982 ms, and 794 ms in the 50, 150, and 900 ms SOA conditions, respectively. The interaction was not significant, F(1.840, 114.058) = 0.278, MSe = 4876, p = 0.740, ES = 0.004, nor was there any trend of underadditivity of lexical frequency with SOA. The average lexical frequency effect was 53 ms, 41 ms, and 45 ms in the 50, 150, and 900 ms SOA conditions, respectively.

5.3.7 Additional Analyses

5.3.7.1 Further Examination of Lexical Frequency Effects on Tone RT1

The predicted interaction of lexical frequency and SOA on tone RT1 was not statistically significant, but the pattern of cell means was in the expected direction. This prediction was based on the hypothesis that a more difficult secondary naming stimulus would cause online re-allocation of resources toward the naming task, causing slowing of tone RT1 in low lexical frequency, short SOA conditions. Thus, this prediction regarding tone RT1 depends on the presence of a lexical frequency effect on naming RT2. Given the inter-subject variability inherent in reaction time data, it was reasonable to ask what proportion of the 63 participants were responsible for the observed lexical frequency effect on naming RT2. Fifty-one participants demonstrated secondary-task naming RTs that were on average slower in the low frequency conditions than in the high frequency conditions. The average frequency effect on naming RT in this subgroup was 63 ms (sd = 53) and it ranged from 6 ms to 275 ms for individual participants. The frequency effect in the 12 participants excluded from this subgroup averaged -26 ms (sd = 20), ranging from -1 ms to -67 ms, i.e. in the opposite direction.

When the tone-primary RT1 data from these 51 participants were analyzed separately, the predicted interaction was significant in the expected direction, F(1.839, 91.938) = 4.133, MSe = 2778, p = 0.022, ES = 0.076. The data from this re-analysis are displayed in Figure 12.

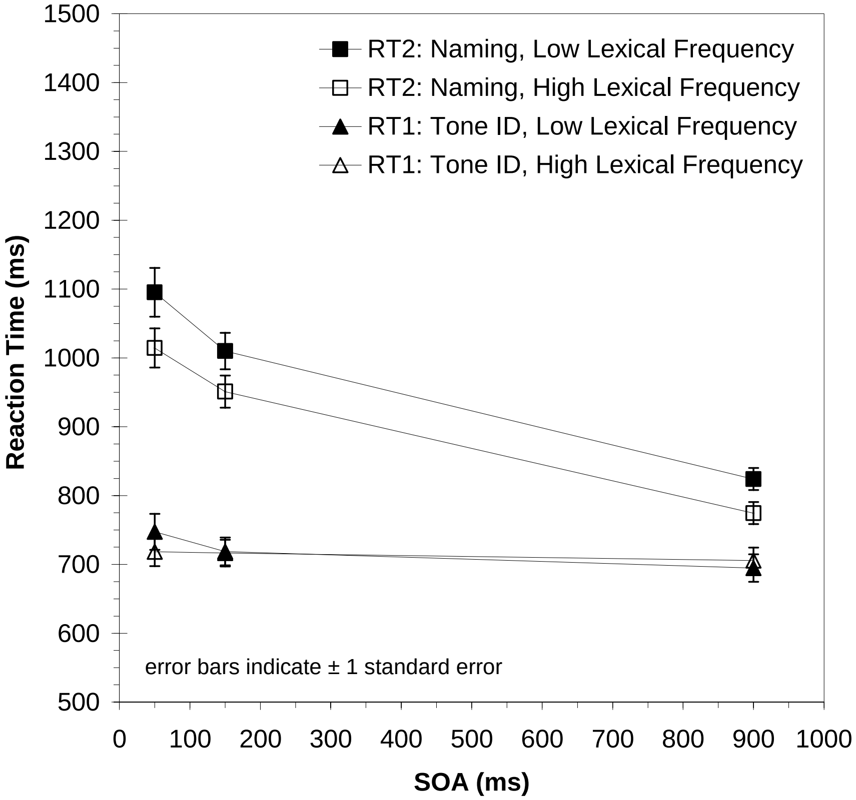

Figure 12. Mean reaction times for Experiment 3 tone-primary trials by task and condition, including only participants demonstrating a lexical frequency effect on naming RT2 (n = 51).

For these participants, there was a 29 ms lexical frequency effect on tone RT1 at 50 ms SOA, a 2 ms effect at 150 ms SOA, and a -11 ms effect at 900 ms SOA. The simple effect of lexical frequency was significant at 50 ms SOA (t = 2.45, 95%CI = ± 24), but not in the 150 ms (t = 0.21, 95%CI = ± 21) or 900 ms (t = 1.51, 95%CI = ± 15) SOA conditions. The main effect of SOA on tone RT1 remained significant, F(1.422, 71.098) = 3.984, MSe = 9777, p = 0.036, ES = 0.074, with means of 733 ms, 718 ms, and 700 ms in the 50, 150, and 900 ms SOA conditions, respectively. The findings with regard to naming RT2 were not appreciably different in this re-analysis, except for a non-significant trend for the size of the lexical frequency effect to increase with decreasing SOA, F(2, 100) = 1.604, MSe = 4124, p = 0.206, ES = 0.031. The frequency effect on naming RT2 was 81 ms, 59 ms, and 50 ms in the 50, 150, and 900 ms SOA conditions, respectively.

5.3.7.2 Evaluation of the Response Grouping Hypothesis

As discussed above in Sections 2.1.4 and 4.4, proponents of the bottleneck model have often attributed increases in RT1 at short SOAs to response grouping. They propose that on some proportion of trials, participants select their responses to the two tasks serially, but do not produce the response to the first stimulus until the second response is ready, and then emit them together as a single unit. Furthermore, the account goes, participants employ this response pattern more frequently in shorter SOA conditions. One way to evaluate the response grouping hypothesis is to examine the distribution of inter-response intervals (IRIs), with the assumption that trials with grouped responses should have IRIs close to zero. In the present data, the RT1 increased at short SOAs only for the tone-primary trials. If the IRI distributions for the short-SOA, tone-first trials have modes close to zero, this suggests that response grouping could account for the observed slowing of RT1 in those conditions.

The IRI distributions for the trials on which the tone was presented first are displayed in Figure 13. The analogous data for the picture-first trials are displayed in Figure 14.

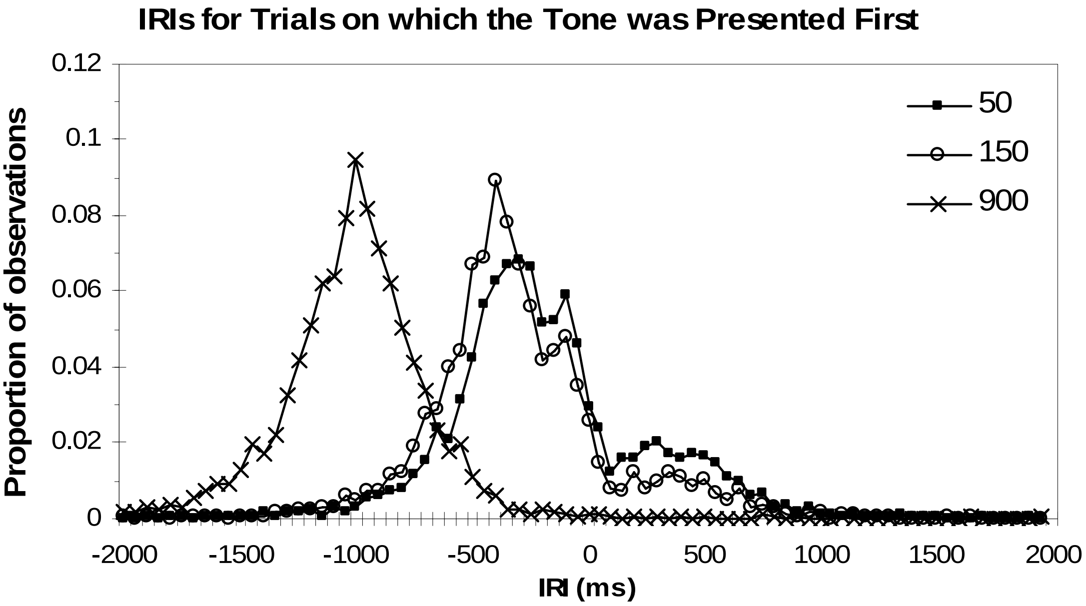

Figure 13. Distribution of IRIs for trials on which the tone was presented first, by SOA condition. Negative IRI values represent trials on which the tone response was produced first.

In both figures, positive IRI values indicate trials on which the naming response was produced first, and negative IRI values represent trials where the tone response was produced first. Also, trials on which either response was incorrect were excluded from both figures. Finally, when interpreting these figures, it is important to bear in mind that the naming RTs were collected by voice key. As discussed above in Section 3.2.2, the voice key always logged the time stamp for the naming response slightly after the articulatory gestures for the response had in fact begun. For the purposes of the current analysis, the practical effect is that the actual IRI distributions are all likely shifted slightly in the positive direction (i.e., to the right) along the x-axis compared to their representations in Figures 12 and 13.

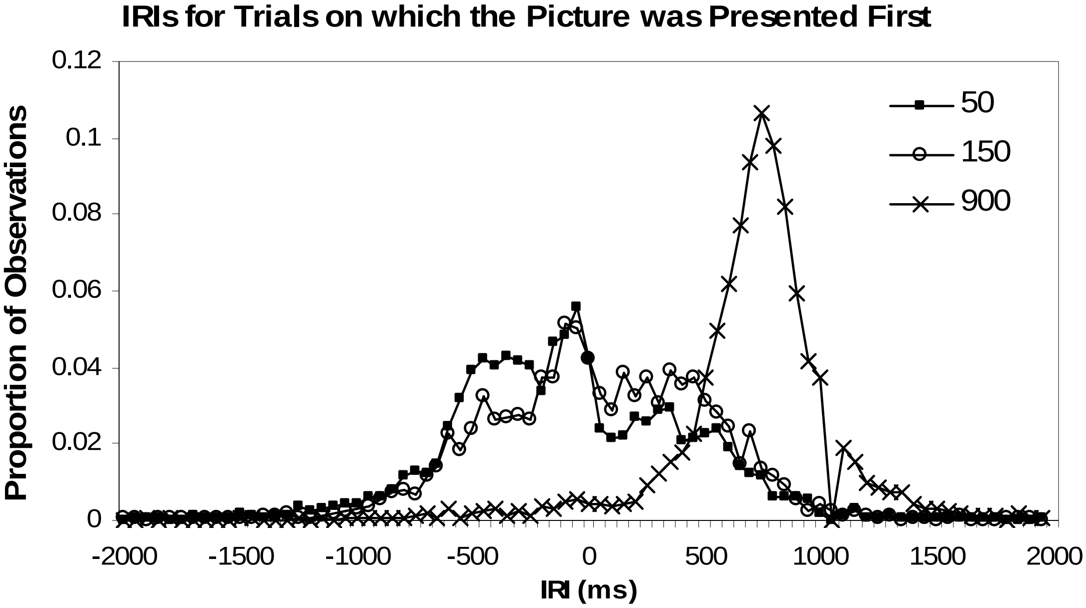

Figure 14. Distribution of IRIs for trials on which the picture was presented first. Positive IRI values represent trials on which the naming response was produced first.

It is apparent from Figure 13 that on trials where the tone was presented first, participants produced the tone response first on the majority of trials in all three SOA conditions. In contrast, Figure 14 shows that on the picture-first trials, participants responded out of presentation order much more often in the two short SOA conditions. This is not surprising, because response order was specifically not constrained, and these participants were included in the analyses on the basis that they responded in presentation order on a minimum number of tone-first trials.

Figure 13 also reveals that the mode for the tone-first IRI distribution in the 50 ms SOA condition was between -350 and -300 ms, and the mode for the 150 ms SOA condition was between -450 and -400 ms. However, there was a second, local mode in both short SOA conditions between -100 and -51 ms, and a non-negligible number of IRIs between -50 and -1 ms. This suggests that response grouping may have been responsible for the observed slowing of tone RT1 at short SOAs. Inspection of the IRI data for the picture-first trials, displayed in Figure 14, reveals modes in the 50 and 150 ms SOA distributions between -1 and -100 ms. Among the positive IRIs representing trials included in the RT analyses, the mode for both short SOA conditions was between 0 and 50 ms. Thus, it seems that response grouping may have been occurring approximately equally as often for the short SOA picture-first trials as for the short SOA tone-first trials. The number of trials with absolute IRI values < 200 ms, expressed as a proportion of the total number of trials included in the RT analyses (i.e., correct trials where response order matched presentation order) is presented for each task order and SOA condition in Table 6.

Table 6. Proportion of Experiment 3 trials with absolute IRIs < 200 ms by task order and SOA condition.

Task Presentation Order	Stimulus Onset Asynchrony (SOA)		
	50	150	900
Tone-First	0.288	0.205	0.005
Picture-First	0.284	0.282	0.021

Given that there was a substantial number of tone-primary trials with IRIs close to zero, the data from those conditions were re-analyzed to determine whether the finding that RT1 increased with decreasing SOA was robust to the exclusion of these low-IRI, potentially grouped trials. It was not. When the data from trials on which the tone was presented and responded to first were re-analyzed with all trials having absolute IRIs < 200 ms excluded, the main effect of SOA was no longer significant, F(1.427, 88.454) = 1.258, MSe = 88.454, p = 0.28, ES = 0.02. Tone RT1 averaged 707 ms, 691 ms, and 700 ms in the 50, 150, and 900 ms SOA conditions respectively. No other aspects of the analysis for either tone RT1 or naming RT2 were substantially different. The data from this re-analysis are presented in Figure 15.

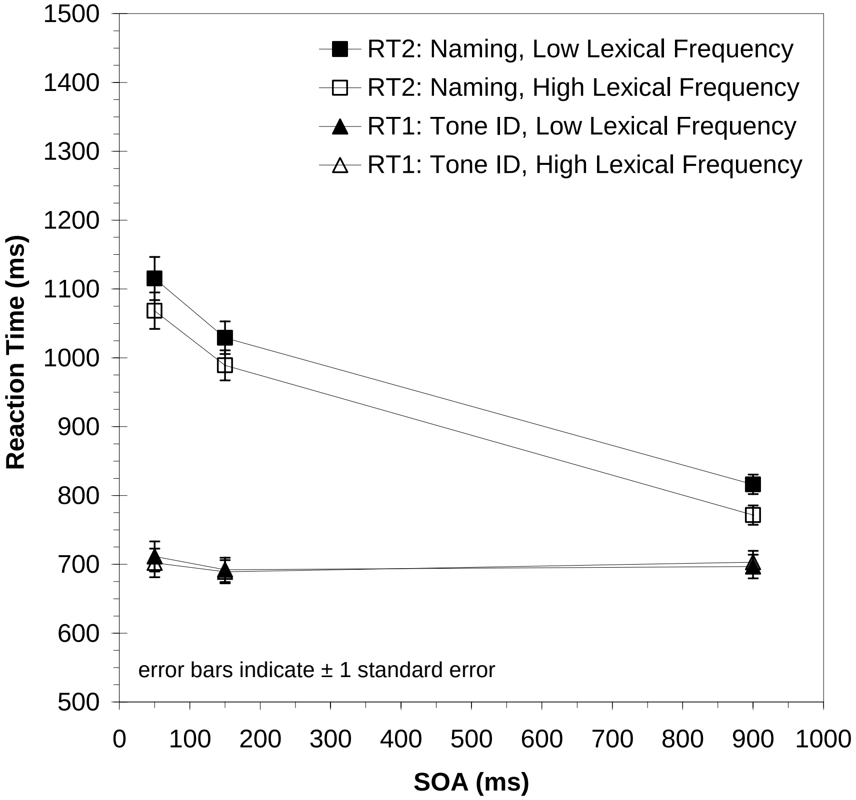

Figure 15. Mean reaction times for Experiment 3 tone-primary trials by task and condition, excluding trials with IRIs < 200 ms.

A similar re-analysis including only the 51 participants who demonstrated a lexical frequency effect on naming RT2 was also undertaken. The data from this re-analysis are presented in Figure 16.

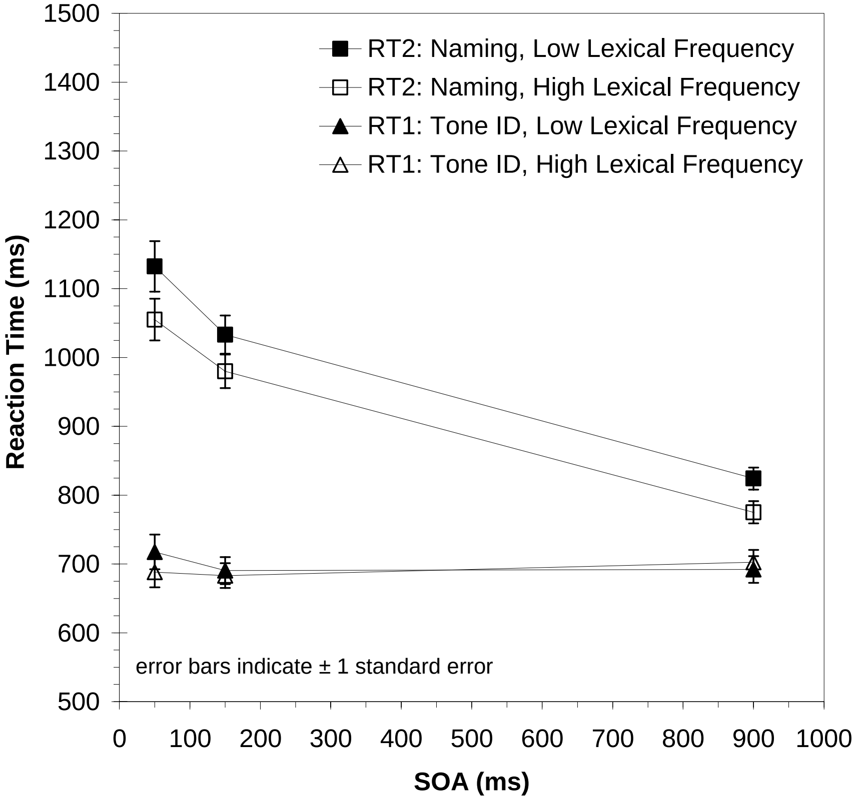

Figure 16. Mean reaction times for Experiment 3 tone-primary trials by task and condition, including only participants demonstrating a frequency effect on naming RT2 (n = 51), and excluding trials with IRIs < 200 ms.

As reported in the preceding section, when the data from these participants were analyzed separately from the minority who failed to demonstrate a lexical frequency effect on naming RT2, the predicted interaction of lexical frequency and SOA on tone RT1 was obtained in addition to the predicted main effect of SOA. With the low IRI, potentially grouped trials excluded, the main effect of SOA was no longer significant, F(1.455, 72.77) = 0.927, MSe = 9338, p = 0.373, ES = 0.018. Tone RT1 averaged 703 ms, 687 ms, and 697 ms in the 50, 150, and 900 ms SOA conditions, respectively. The interaction effect, however, remained significant, F(2, 100) = 3.973, MSe = 10118, p = 0.022, ES = 0.74. In this analysis, the lexical frequency effect on tone RT1 was 29 ms, 7 ms, and -10 ms in the 50, 150, and 900 ms SOA conditions, respectively. The simple effect of lexical frequency was significant only at 50 ms SOA (t = 2.35, 95%CI = ± 25). No other aspects of the analysis were substantially different.

5.4 DISCUSSION OF EXPERIMENT 3

The aim of Experiment 3 was to investigate whether dual-task performance limitations associated with single word production are more consistent with the central resource or central bottleneck model. As discussed by Navon and Miller (2002) and Tombu and Jolicoeur (2002a), the most distinctive predictions of the two models concern RT1, and in particular the effects of SOA and secondary task factors on RT1. Consistent with the central resource model, it was predicted that RT1 would increase with decreasing SOA for both orders of task presentation. This prediction was upheld only in the tone-primary conditions. On naming-primary trials, RT1 increased by 31 ms as SOA decreased from 900 ms to 150 ms, but the trend was not significant and there was no further increase in naming RT1 as SOA decreased to 50 ms. This suggests that participants engaged in limited-capacity parallel processing on tone-primary trials, but processed the tasks serially when the naming stimulus was presented and responded to first. However, it should be noted that the analysis of the naming-primary conditions had less statistical power (observed power = 0.28) than the tone-primary analysis (observed power = 0.77), because fewer participants were included. When the criterion for inclusion in the naming-primary analysis was reduced from eight trials per condition to four trials per condition, providing a sample size of 51 participants, the main effect of SOA approached significance (p = 0.056, ES = 0.057, observed power = 0.55). Further consideration of this issue will be deferred to the general discussion in the following chapter.

While the finding of a negative RT1-SOA curve is consistent with the resource model, it is also consistent with the central bottleneck model augmented by response grouping. Indeed, further examination of the tone-primary RT data demonstrated a substantial number of low-IRI trials consistent with this latter hypothesis. Furthermore, when the low-IRI trials were removed from the analysis, the main effect of SOA on tone RT1 was no longer significant. However, there are a number of reasons to be cautious about accepting serial processing with response grouping as the explanation for the slowing of tone RT1 at short SOAs, and this issue will be considered further in the general discussion below.

It was also predicted in Experiment 3 that the lexical frequency of secondary-task naming targets would affect tone RT1, but only in the short SOA conditions. This prediction was based on the hypothesis that the arrival of a more difficult secondary task during primary-task central processing would cause a shift in the allocation of processing resources away from the primary task. The prediction required the assumption that the lexical frequency manipulation affected central processing of the secondary naming task. Based on the results of Experiment 2, this assumption was tenable, and it was confirmed by Experiment 3, which showed no hint of an underadditive Lexical Frequency x SOA interaction. Regarding the predicted interaction effect on tone RT1, the cell means demonstrated the expected pattern across SOA and lexical frequency conditions, but the differences were small and failed to reach significance in the full participant sample. However, when the analysis excluded the minority of participants who failed to demonstrate a lexical frequency effect on the secondary naming task itself, the predicted interaction was significant, with tone RT1 demonstrating a 29 ms lexical frequency effect in the 50 ms SOA condition. Furthermore, this particular result was robust to the exclusion of low-IRI, potentially grouped responses.

In summary, the major predictions of Experiment 3 were partially confirmed. The data were largely consistent with the central resource model, but aspects of the results are also compatible with the central bottleneck model with response grouping. These issues will be discussed further in the context of the results of Experiments 1 and 2 in the following chapter.

6.0 GENERAL DISCUSSION

In addition to the question of whether the central bottleneck or central resource model accounts better for dual-past performance, the present results also have implications for models of word production. Consistent with Ferreira and Pashler (2002), Experiment 1 found that lexical frequency affected primary-task naming and secondary-task tone ID approximately equally. Experiment 2 attempted to replicate McCann and colleagues' (2000) finding of a null Lexical Frequency x SOA interaction on secondary task naming RTs. The initial analysis showed a partially underadditive interaction of lexical frequency and SOA on secondary-task naming times, but the frequency manipulation was confounded by additional correlated factors including name agreement, image agreement, and object recognition time. In a follow-up analysis employing high and low frequency sub-lists balanced on these variables, the interaction was no longer statistically significant, but a trend of underadditivity remained. Experiment 3, using the same more balanced stimulus lists, provided a stronger replication of McCann et al. (2000), with no hint of an underadditive trend. Taken together, these results are most consistent with the conclusion that frequency-sensitive processing in word production occurs in the central stage of the dual-task models under discussion.

It is also useful to consider the present findings in the context of recent PRP studies using naming tasks presented in secondary position with unrelated and semantically related distractor words, (Dell'Acqua et al., in press; Sullivan & Macchi, 2002). The source of relative slowing of naming RT in the presence of semantically-related vs. unrelated distractors has been plausibly argued to reside at the stage of lemma selection within current models of word production (Schriefers et al., 1990). In these PRP studies, slowing of naming RT2 associated with semantically-related distractors was clearly underadditive with SOA. This result is most consistent with a pre-central locus for at least some aspects of lexical-semantic processing. On the other hand, the present results (and those of McCann et al., 2000) demonstrate additivity of word frequency effects with SOA, suggesting that frequency impacts some later stage of processing. The differential behavior of secondary-task picture-word interference and lexical frequency effects in the PRP method provides converging evidence with prior work placing the locus of word frequency effects at the stage of phonological word form access (Ferreira & Pashler, 2002; Griffin & Bock, 1998; Jescheniak & Levelt, 1994).

The clear evidence that lexical frequency affects the central stage(s) of dual-task processing makes the hypothesis regarding the effects of lexical frequency on primary-task tone RT relatively straightforward. In the central resource model, if an increase in secondary-task central processing demand causes a shift in allocation ratio toward the secondary task, RT1 should be affected at shorter SOAs, where there is overlap in the central processing of the two tasks. The interaction consistent with this hypothesis was not significant in the analysis of the tone RT1 data from Experiment 3, but the pattern of cell means was in the expected direction. Also, the effect of lexical frequency on tone RT1 depended on the presence of a frequency effect on the secondary naming task itself. Twelve of the 63 participants included in the Experiment 3 tone RT1 analysis failed to demonstrate slower naming of low frequency words when performance was averaged across SOA conditions. This could be due to stimulus factors, as the naming stimuli were pseudorandomly assigned to SOA conditions, or it could be due to subject factors, especially given the relatively poor split-half and test-retest reliability of psycholinguistic manipulations such as semantic priming on RT (Stolz, Besner, & Carr, 2005). When these 12 participants were excluded from the analysis, the Lexical Frequency x SOA interaction on tone RT1 was significant, as was the 29 ms simple effect of frequency in the 50 ms SOA condition. Confidence in this finding is certainly mitigated by the selective exclusion of participants and the increase in the possibility of type I error associated with multiple analyses of related data sets, but the result is nevertheless consistent with the resource model, and with the notion that allocation ratio may be influenced by task demand (Kahneman, 1973). The result is inconsistent with the central bottleneck model, which predicts that secondary task factors should have no influence on RT1.

The central bottleneck and resource models also make differential predictions regarding the main effect of SOA on RT1, with the bottleneck model again predicting a null effect. The central resource model, on the other hand, predicts that as SOA is decreased and there is greater overlap of central processing between the two tasks, RT1 should increase, provided that that some proportion of available processing resources are re-allocated from task 1 central processing to task 2 central processing as soon as it begins. Both Experiment 2 and the tone-primary trials in Experiment 3 produced data (at least partially) consistent with this prediction of the resource model, and inconsistent with the central bottleneck model. In the case of Experiment 2, this was unexpected because the emphasis given to the tone ID task in the instructions should have encouraged serial processing, even if resource sharing were possible. At the same time, at least two fixed-order PRP experiments have demonstrated SOA effects on RT1 consistent with resource sharing (Cleland et al, 2006; Tombu & Jolicoeur, 2005). However, it should be noted that the particular shape of the RT1-SOA curve obtained in Experiment 2 is inconsistent with the strict version of the central resource model, which assumes that neither allocation ratio nor available resource supply vary systematically with SOA. Given the observed 72 ms difference in RT1 between the 900 and 150 ms SOA conditions, the central resource model predicts a definite negative slope to the portion of the RT1-SOA curve between 50 and 150 ms, amounting to a difference on the order of 25 ms. As illustrated in Figure 5, this portion of the curve was essentially flat, corresponding to a negligible 3 ms difference when averaged across lexical frequency conditions. To account for this pattern in the RT1 data, the resource model would require relaxation of the assumptions of non-systematic variation in capacity and/or allocation ratio, with the ad hoc assumption that the total amount of resource capacity directed to task 1 was *greater* in the 50 ms SOA conditions than the 150 ms SOA conditions.

Tombu and Jolicoeur (2003), whose initial presentation of the resource model assumed that available capacity was fixed, discussed ways in which this assumption might be modified. One possibility they considered is that some processing resources might be allocated neither to task 1 nor task 2, but instead to "overhead" costs. These overhead costs, which they likened to the costs of concurrence discussed by Navon and Miller (1979), might be used for such activities as keeping the response mappings for both tasks activated and remaining prepared for the arrival of new stimuli. They presented a modified version of their central capacity sharing model that included a parameter for overhead costs, with the assumption that overhead costs should decrease over time. They based this assumption on the idea that, as a trial progresses, task 1 is more likely to have been completed, thereby freeing up capacity previously used to keep its response mappings activated and in order. This version of the central resource model actually predicts that *less* resource capacity should be available in the 50 ms SOA condition, and so does not assist the model in accounting for the flat portion of the RT1-SOA curve observed in Experiment 2.

However, the assumption that overhead costs decrease over time is itself a potentially empirical question that requires examination. If a significant proportion of resources are devoted to remaining prepared for the arrival of the second stimulus (S2), overhead costs might decrease as a step function at the point S2 arrives. This would have the effect of making more processing capacity available to both tasks in shorter SOA conditions, so long as task 1 processing is still ongoing. Also, subjects might devote progressively more resources to preparation for the arrival of S2 as time elapses and the probability of its occurrence increases. In Experiment 2, if some proportion of resources were required to remain prepared for the appearance of the naming S2, the earlier appearance of the naming stimulus (in the 50 ms SOA condition) could have conceivably made additional resources available to the tone ID task earlier in the course of its processing. If the majority of the additional capacity made available for task processing upon the appearance of the naming stimulus was allocated to the primary tone ID task, it would benefit from that increased capacity for 100 ms longer in the 50 ms SOA condition than in the 150 ms SOA condition. This explanation is ad hoc and highly speculative, but it could nevertheless enable the central resource model, so modified, to account for the observed results.

As noted above, the tone-primary conditions of Experiment 3 also demonstrated a significant main effect of SOA on RT1, and in this case the result was more consistent with the strict version of the central resource model, demonstrating a more definite negative slope of the RT1-SOA curve between 50 and 150 ms. However, unlike Experiment 2, additional analyses of the Experiment 3 tone RT1 data indicated that the SOA effect could be attributed to participants' tendency to group their responses to the two stimuli in the shorter SOA conditions. This explanation is attractive because its is established that subjects in PRP experiments do engage in response grouping under certain circumstances (Pashler & Johnston, 1989), and the IRI distributions illustrated in Figure 14 strongly suggest that participants in Experiment 3 were coordinating a substantial minority of their responses in a manner consistent with the grouping hypothesis.

The assertion that participants in Experiment 3 were grouping their responses does not, however, selectively favor the central bottleneck model over the central resource model. While response grouping does enable the bottleneck model to account for the effect of SOA on RT1, it is a strategy that is equally available regardless of whether tasks are processed serially or in parallel (Navon & Miller, 2002; Tombu & Jolicoeur, 2003). Far from being theoretically informative, response grouping is perhaps best regarded as a nuisance factor to be minimized through task design and instructions on the one hand, and trimmed from the data after collection where grouping has occurred (Navon & Miller, 2002). The strategy of trimming low-IRI trials to differentiate between resource-like and grouping effects on RT1 would seem to be appropriate when the effect is maintained and the resource explanation is favored, as in the present Experiment 2 and in Tombu and Jolicoeur (2002a; 2005). However, there is reason to be cautious in interpreting a disappearance of the RT1-SOA effect with the exclusion of low-IRI trials, as in the present Experiment 3. This is because resource models predict that IRI should decrease with decreasing SOA, even across the shortest range of SOAs (Kahneman, 1973; Navon & Miller, 2002)[5]. The central bottleneck model (un-augmented by grouping), by

---

[5] As pointed out by Navon and Miller (2002), the prediction of a continuously positive slope for the IRI-SOA function is mathematically redundant with the prediction of a negative RT1-SOA curve, given that, like the

contrast, predicts that IRI should be constant across lower values of SOA. By trimming low-IRI trials, one may in some cases be excluding exactly the data points that would support the resource model.

Moreover, the 200 ms IRI criterion for determining which responses were grouped was necessarily arbitrary, although it did receive some support from the existence of a local mode centered at -100 ms in the IRI distribution for tone-primary responses (see Figure 14). The criterion was taken from a fixed-presentation-order PRP study that required two manual responses (Tombu & Jolicoeur, 2005). Interestingly, Tombu and Jolicoeur's (2002b) prior variable-order PRP study produced IRI distributions with very few observations between -200 and +200 ms for either order of task presentation. In this study, the response to the tone ID task was manual, while the response to a polygon matching task was given by means of pedals. It is possible that the present study's pairing of a vocal and a manual response in the context of variable task order in was more susceptible to grouping, perhaps because of humans' tendency to coordinate short, fast movements, especially of the arms, hands, and head, with stressed syllables in speech (Bull & Connely, 1985). Perhaps vocal-pedal output combinations would be less likely to lead to coordinated response patterns suggestive of grouping. Further research targeted at describing the vocal-manual or vocal-pedal response patterns obtained under conditions encouraging grouping would also be helpful in defining an a priori criterion for identifying grouped trials.

In any case, the effect of SOA on RT1 in Experiment 2 was robust to the exclusion of low-IRI trials, as was the Lexical Frequency x SOA interaction in the follow-up analysis of

---

central bottleneck model, the central resource model predicts a -1 slope for RT2-SOA curve across lower values of SOA.

Experiment 3 tone RT1 data. Where these particular results are concerned, response grouping does not seem to have played a role.

It is worth considering why, across the three experiments reported here, primary-task tone RTs demonstrated significant increases with decreasing SOA, while primary-task naming RTs did not. In this regard, the results of the three experiments were quite consistent: tone RT1 was significantly longer in the short SOA conditions than in the 900 ms SOA condition, while there was a nonsignificant trend for naming RT1 to be slowest at 150 ms, followed by the 50 and 900 ms conditions, in that order. In both cases, the lack of significance for the effect of SOA on naming RT1 could be due to low statistical power. In Experiment 3, because of the a priori decision to base the criterion for inclusion in the analysis on the number of valid tone-first trials, the naming-first analysis included fewer subjects ($n = 31$ vs. $n = 63$ in the tone-primary conditions), and fewer trials per condition (an average difference of 4 trials per condition). The observed power for the naming-primary trials was 0.28. In Experiment 1, the data were analyzed using a single ANOVA including both RT1 and RT2, and the effect of SOA on RT1 was evaluated post-hoc, using the Scheffé method to correct for multiple comparisons. When the naming RT1 data from Experiment 1 were analyzed separately in the same design as the tone RT1 data from Experiment 2, the main effect of SOA was significant, $F(1.479, 34.013) = 5.695$, $MSe = 45060$, $p = 0.013$, $ES = .198$, with average RTs of 934 ms, 946 ms, and 896 ms in the 50, 150, and 900 ms SOA conditions, respectively. Also, when the naming RT1 data from Experiment 1 were directly compared with the Experiment 2 tone RT1 data in an ANOVA with SOA and lexical frequency as repeated factors, and task as a between-subjects factor, the task by SOA interaction was not significant. A similar analysis using a completely within-subjects ANOVA to compare the tone RT1 and naming RT1 data from Experiment 3 produced a similar null result. Although the apparent finding of an asymmetrical performance trade between the two tasks appears to support the notion that language processing may be particularly demanding (in the sense of "more mandatory"), the present data actually do not offer it any particular support.

Comparison of the SOA effect on RT2 across the three experiments reported here is potentially more interesting. The central bottleneck and central resource models make identical predictions: RT2 should increase dramatically as SOA decreases, and the slope of the RT2-SOA curve should approach -1 across lower values of SOA. Thus, both models predict that the size of the PRP effect on naming RT2 should be similar in Experiment 2 and Experiment 3. Comparison of Figures 9 and 11 reveals that this was not the case. The PRP effect averaged 443 ms in Experiment 2[6] and 259 ms in Experiment 3. The 184 ms difference was significant, $t(1, 85) = 5.414$, $p < 0.001$, 95%CI = ± 67. Comparison of the PRP effect on tone RT2 in Experiments 1 and 3 was confounded by differences in the naming stimulus lists, but produced similar results: a 631 ms PRP effect in Experiment 1 and a 451 in Experiment 3. Again, the difference was significant, $t (1, 85) = 5.123$, $p < 0.001$, 95%CI = ± 69.

The central bottleneck model has no internally motivated way to account for this result, but here again, relaxation of the assumption of fixed capacity permits the resource model accommodate the finding, if one assumes that more capacity was available under variable task-order conditions and instructions to give equal attention to both tasks. On first examination, this seems unlikely, because variable task order conditions should impose more overhead costs (in

---

[6] This comparison employed the data from the follow-up analysis of Experiment 2 including only the more balanced naming stimulus lists, because these were the same stimulus lists used for the tone-primary conditions in Experiment 3.

line with the discussion above) than conditions in which task order is predictable. On the other hand, it could be argued that the more demanding variable order conditions raised participants' effort level and correspondingly increased the available resource supply (Kahneman, 1973; Tombu and Jolicoeur, 2003).

However, the central resource model also predicts that a decrease in the PRP effect associated with an increase in available resource supply should be accompanied by a similar decrease in RT1. In contrast to the 184 ms difference between the PRP effects observed in Experiments 2 and 3, there was a non-significant 62 ms difference between the experiments in tone RT1 in the 900 ms SOA condition, $t(1, 85) = 1.812$, $p = 0.073$, 95%CI = ± 68. The fact that the reduction in dual-task interference associated with variable task order as opposed to fixed task order exceeded the reduction in RT1 by a factor of three is consistent with the notion that participants in the variable order conditions were processing some component(s) of one or both tasks in parallel in ways not accounted for by the central resource model. Put differently, some processing component of one or both tasks that required capacity-limited resources in the fixed order conditions may have been performed without capacity limitations in the variable-order conditions.

The differences between performance in Experiment 3 and the preceding experiments could also be due to differential subject selection. Participants in Experiment 3 were on average slightly older, but this would be expected to produce slower RTs and more dual-task interference, rather than less. The more important concern is likely the fact that Experiment 3 participants were included in the analysis only if they achieved a minimum number of valid trials in the tone-primary conditions. Although the major determinant of whether or not subjects met this criterion was their relative preference in response order, error rates also played a role, with included subjects generally committing fewer errors than those who were excluded. Thus subject selection may have contributed to generally superior performance in Experiment 3 compared to the preceding experiments. Additionally, Experiment 3 participants were paid for their participation, although the payment was not contingent on performance, while participants in Experiments 1 and 2 received course credit. However, regardless of whether the observed differences in primary-task performance and dual-task interference are due to subject selection, variability in task presentation order, instructions, or some combination, the theoretical implications are the same. Some component or components of one or both tasks were apparently resource-limited in Experiment 2 but not in Experiment 3.

7.0 CONCLUSIONS

The first specific aim of the present experiments was to investigate the locus of cognitive processing limitations resulting from lexical frequency effects in word production. The results are relatively clear and consistent with the small number of previous PRP studies that have manipulated word frequency in the context of naming tasks. Frequency effects in picture naming appear to be additive with SOA, suggesting that they participate in some resource-limited stage of processing, regardless of whether the operative resource is construed as a divisible and flexibly allocatable quantity, or as a unitary processor accessible to only one task at a time. This finding provides an important constraint on models of word production, given that other manipulations of naming RT, namely semantic interference (Dell'Acqua et al., in press; Sullivan & Macchi, 2002) and phonological facilitation (Ferreira and Pashler, 2002) have been shown to interact with SOA in ways consistent with the central bottleneck and central resource models.

The findings relevant to the second specific aim, to investigate whether dual-task performance limitations in word production are more consistent with the central bottleneck or central resource model, are less straightforward, but still informative. Many of the results, including the effects of SOA on RT1 observed in both Experiments 2 and 3 and the Lexical Frequency x SOA interaction on tone RT1 in Experiment 3, are clearly inconsistent with the central bottleneck model. In the case of the Experiment 3 RT1-SOA effect, the central bottleneck model augmented with response grouping offers a potential alternative (though theoretically unsatisfying) explanation. However, in the other two cases response grouping did not appear to play a role. On balance, the data are more consistent with the central resource model than with the central bottleneck model.

At the same time, however, there are aspects of the data that present significant difficulties for the central resource model as well. In order to explain the particular shape of the RT1-SOA curve observed in Experiment 2, the resource model requires relaxation of the assumption that capacity and/or allocation ratio do not vary systematically with SOA, along with fairly specific ad hoc assumptions about how concurrence costs or allocation ratios manifest over time. The resource model is only partially able to account for substantial differences in dual-task cost between Experiment 3 and the preceding experiments. Manipulation of factors potentially including task order, task instruction, and subject selection, caused a greater reduction in dual-task slowing than the central resource model predicts.

The direct relevance of the present work to the study of aphasia is limited, but it does provide some support for the theoretical notions underlying resource-allocation and capacity-limitation views of normal and disordered language processing. The dual-task models studied in this investigation are potentially useful in understanding how language processing in persons with aphasia differs from language processing in healthy individuals. PRP experiments designed to encourage resource-limited parallel processing may help to illuminate the extent to which persons with aphasia can be considered to have language-specific capacity limitations, impairments of resource allocation, or both. At the same time, PRP studies that encourage serial processing may be useful in isolating aspects of aphasic language impairment within staged models of lexical access. Both avenues of investigation would benefit from further studies of normal individuals designed to better specify what components of language and non-language processing participate in central, as opposed to pre-central or post-central processing in these dual-task models.

APPENDIX A

MATHEMATICAL PRESENTATION OF THE CENTRAL RESOURCE MODEL

The central resource model as presented by Navon and Miller (2002) and Tombu and Jolicoeur (2003) has the following assumptions:

1. The central stage of processing is the only one that is resource-limited.
2. Allocation ratio (p) remains constant during a trial.
3. The available resource capacity (q) does not vary systematically with the experimental factors, including SOA.
4. There is no cost of online re-allocation of resources from one task to the other
5. The central resource capacity is undifferentiated, i.e., it may be applied to any task domain with equal efficiency and productivity.
6. Task 1 central processing begins before Task 2 central processing begins.
7. Task 1 central processing ends before Task 2 central processing ends.

$$RT1_{short} = A1/p + B1/(p*q) + (1-1/p)*(SOA + A2) + C1 \qquad (1)$$

$$RT1_{long} = A1 + B1/q + C1 \qquad (2)$$

$$RT2_{short} = (B2 + B1)/q + A1 - SOA + C1 \qquad (3)$$

$$RT2_{long} = A2 + B2/q + C2 \qquad (4)$$

RT_{short} refers to short SOA conditions where there is likely to be overlap in central processing of tasks 1 and 2. For the present experiments, these were assumed to be the 50 ms and 150 ms SOA conditions. RT_{long} refers to long SOA conditions where overlap of central processing is not expected. For the present experiments, this was assumed to be the case for the 900 ms condition. The letter p represents the proportion of resources allocated to task 1 central processing, and takes values between 0.5 and 1. The letter q represents available resource capacity and takes values between 0 and 1. A1 refers to the time needed to complete perceptual encoding or other pre-central processing for task 1, and A2 is the corresponding quantity for task 2. C1 and C2 refer to the time needed to complete response execution or other post-central processing for tasks 1 and 2, respectively. B1 and B2 refer to the demand of tasks 1 and 2, respectively, for central resources. The demand divided by the processing rate yields the time to complete central processing.

APPENDIX B

LEXICAL AND PICTURE CHARACTERISTICS FOR EXPERIMENT 1 AND 2 NAMING STIMULI

Tables 7 and 8 present the high and low frequency picture naming targets for the tone-primary trials of Experiment 3, and their associated log frequency values. Table 9 summarizes the lexical and picture characteristics for the high and low frequency lists, and presents the results of independent samples t-tests for the following variables: Lexical frequency, rated image agreement, name agreement, naming RT, object recognition RT, rated age of acquisition, number of phonemes, and number of syllables. Table 10 presents the correlations among those variables within the combined high and low frequency lists.

Table 7. Low Frequency picture naming targets for Experiments 1 and 2

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
dustpan	1.79	pliers	3.78	rake	4.29	tractor	4.71	crab	5.09	anchor	5.40
highchair	1.79	banjo	3.83	vest	4.30	lemon	4.75	cactus	5.09	puzzle	5.42
seahorse	1.79	peacock	3.83	seesaw	4.32	onion	4.76	rooster	5.10	guitar	5.42
teepee	2.20	sailboat	3.87	hinge	4.34	flute	4.77	cherry	5.11	spoon	5.43
bra	2.40	penguin	3.87	funnel	4.34	banana	4.77	lizard	5.12	cannon	5.43
waffle	2.48	llama	3.91	sock	4.38	skunk	4.78	nut	5.14	robot	5.44
stroller	2.48	pear	3.91	maze	4.41	parrot	4.80	shovel	5.15	towel	5.45
blimp	2.71	badge	4.01	mailbox	4.42	pizza	4.80	stool	5.18	shower	5.45
tweezers	2.77	clamp	4.03	wallet	4.44	glove	4.84	dime	5.21	tire	5.46
diaper	3.18	acorn	4.04	necklace	4.44	scarf	4.84	couch	5.21	whip	5.48
ladle	3.18	dice	4.08	waiter	4.44	torch	4.91	dolphin	5.21	comb	5.48
mixer	3.40	quotes	4.09	vase	4.44	peach	4.93	hose	5.22	donkey	5.48
tripod	3.40	crib	4.11	toaster	4.45	cork	4.97	peanut	5.22	glue	5.50
backpack	3.40	panda	4.13	lobster	4.48	boot	4.98	broom	5.23	grapes	5.51
igloo	3.43	faucet	4.14	wig	4.50	snail	5.00	raccoon	5.23	sweater	5.52
thimble	3.53	mop	4.14	carrot	4.50	medal	5.00	ax	5.26	chimney	5.52
thermos	3.56	razor	4.19	swan	4.55	pirate	5.02	bug	5.27	skirt	5.53
umpire	3.58	celery	4.22	cookie	4.56	lettuce	5.04	moose	5.27	flashlight	5.54
hanger	3.61	hoof	4.22	walnut	4.58	bride	5.04	toe	5.31	knight	5.54
hammock	3.61	walrus	4.23	skis	4.58	toilet	5.04	pitcher	5.32	pillow	5.55
slipper	3.64	genie	4.23	bathtub	4.61	clown	5.05	rainbow	5.33	bucket	5.55
fireman	3.64	crackers	4.28	giraffe	4.62	scissors	5.05	fountain	5.34	ant	5.57
snowman	3.64	wrench	4.28	windmill	4.64	heel	5.06	fan	5.35	bomb	5.58
radish	3.74	leopard	4.29	hoe	4.65	plug	5.06	screw	5.35	fork	5.62

Table 8. High Frequency picture naming targets for Experiments 1 and 2

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
volcano	5.80	hammer	6.08	whale	6.38	apple	6.64	truck	7.12	mountain	7.83
drum	5.81	sink	6.08	globe	6.39	lips	6.67	roof	7.14	indian	7.86
pants	5.83	turkey	6.09	belt	6.40	flower	6.70	cross	7.15	watch	7.93
vacuum	5.84	tiger	6.10	pencil	6.43	chest	6.72	newspaper	7.15	letter	7.97
squirrel	5.85	arrow	6.13	microscope	6.45	fence	6.73	hospital	7.17	present	8.01
whistle	5.87	piano	6.15	ruler	6.45	lion	6.76	wheel	7.17	horse	8.03
thumb	5.88	flag	6.16	cake	6.48	clock	6.76	bread	7.22	map	8.06
cigarette	5.89	telescope	6.16	quarter	6.48	mirror	6.77	hat	7.24	window	8.10
skeleton	5.89	thread	6.21	magnet	6.50	wolf	6.78	salt	7.26	hair	8.23
statue	5.90	soldier	6.21	chicken	6.51	bottle	6.78	telephone	7.27	heart	8.26
lock	5.92	lightning	6.23	jar	6.53	finger	6.81	corn	7.27	woman	8.27
spider	5.93	lamp	6.23	orange	6.54	stairs	6.81	desert	7.32	table	8.37
button	5.95	jacket	6.25	balloon	6.54	fox	6.82	television	7.35	fish	8.41
saddle	5.96	tent	6.25	bicycle	6.54	cloud	6.83	smoke	7.38	picture	8.43
candle	5.97	slide	6.25	football	6.54	dress	7.00	train	7.42	car	8.45
shoe	5.99	butter	6.25	pipe	6.56	frog	7.00	nose	7.44	king	8.45
sword	6.01	elephant	6.29	camera	6.58	rabbit	7.01	radio	7.47	paper	8.53
crown	6.01	dragon	6.30	pig	6.62	rope	7.01	music	7.59	book	8.56
swing	6.02	rocket	6.30	pool	6.62	wagon	7.01	arm	7.65	sun	8.86
package	6.04	stove	6.30	nest	6.63	plate	7.03	baby	7.76	hand	8.95
priest	6.05	pen	6.33	basket	6.63	ring	7.05	doctor	7.77	city	8.99
turtle	6.06	ghost	6.33	castle	6.64	shoulder	7.06	iron	7.79	house	9.41
ladder	6.06	pan	6.34	snake	6.64	bridge	7.07	glass	7.80	well	9.60
seal	6.07	glasses	6.36	brush	6.64	desk	7.11	box	7.81	can	10.78

Table 9. Summary of the characteristics[7] of the Experiment 1 and 2 picture stimuli.

Variable	High Frequency		Low Frequency		t	Error df	1-tailed p-value
	mean	sd	mean	sd			
Log Frequency	6.91	0.91	4.54	0.87	22.756	286	<0.001
Image Agreement	6.03	0.55	5.69	0.65	4.733	276	<0.001
Name Agreement	0.91	0.12	0.85	0.15	4.02	276	<0.001
Naming RT	910	153	1078	212	-7.658	276	<0.001
Object Recognition RT	488	49	506	75	-2.302	286	0.011
Age of Acquisition Rating	4.64	0.96	5.62	1.22	-7.512	276	<0.001
Number of Phonemes	4.40	1.3	4.44	1.1	-0.299	286	0.383
Number of Syllables	1.58	0.69	1.60	.51	-.392	286	0.348

---

[7] Lexical Frequency counts were obtained from Zeno et al. (1995). Object Recognition RTs were obtained from the experiment described in Appendix E. Data for all other variables were obtained from the CRL-IPNP database (Szekely et al., 2004).

Table 10. Correlations between picture and lexical characteristics for Experiment 1 and 2 naming stimuli. Correlations denoted with an asterisk (*) are significant at $p < 0.05$.

	Log Frequency	Image Agreement	Name Agreement	Naming RT	Object Recognition RT	Age of Acquisition (AoA)	#Phonemes
Image Agr.	0.22*						
Name Agr.	0.24*	0.38*					
Name RT	-0.48*	-0.56*	-0.56*				
Object RT	-0.11*	-0.36*	-0.29*	0.39*			
AoA	-0.52*	-0.25*	-0.25*	0.61*	0.19*		
#Phonemes	-0.16	0.06	-0.11*	0.13*	-0.08	0.18*	
#Syllables	-0.14	0.04	-0.11*	0.07	-0.05	0.10*	0.72*

APPENDIX C

TASK INSTRUCTIONS

Experiment 1: Naming Primary

Initial Instructions

In this experiment, you're going to be performing two tasks: picture naming and tone identification. In the naming task, you will see a picture on the computer screen and say the name of the picture. In the tone ID task, you will hear a tone and press a button to indicate whether it is high, medium, or low in pitch. You will have the chance to practice each task separately, and then together, before doing the experimental trials in which both tasks will be presented. For both tasks, it is important that you respond as quickly as you can without sacrificing accuracy.

Single-Task Naming Practice

On the naming task, it is important to do the following things: First, try to speak at a relatively loud and constant volume across trials. Second, give only one word in each response and avoid making any extra sounds like saying "uh" or smacking your lips before you begin speaking your response. Third, respond to each picture as quickly as you can without sacrificing accuracy. You will see a "+" on the screen at the beginning of each trial. Press the green button when you are ready to begin the trial. The picture will disappear when you say your response. If you have any questions, ask now. Press the green button when you are ready to begin.

Single-Task Tone ID Practice

On each trial of the tone ID task, you will hear one of three tones and press a button to indicate whether it was low, medium, or high in pitch. Always use your left hand, and your left ring finger on the "Low" button, your left middle finger on the "Medium" button, and your left index finger on the "High" button. Before practicing the tone ID task, you will familiarize yourself with the three tones. [The subject heard each of the three tones once, in response to pressing the low, medium, and high buttons in that order.]

When you see a "+" on the screen, rest your left ring finger on the "Low" button, your left milddle finger on the "Middle" button, and your left index finger on the "High" button. Always respond with these fingers on these buttons. When you are ready, press the green button with your right index finger to begin the next trial. After the "+" disappears, you will hear a high, medium, or low tone. Press the button indicating which tone you heard. Respond as quickly as you can without sacrificing accuracy. You will now do nine practice trials with feedback provided after each trial. If you have any questions, ask now. [Participants performed nine trials, three with each pitch in random order, and received accuracy feedback after each trial.].

Good Job! Now you will do thirty-six practice trials without feedback. Remember to always respond as quickly as you can without sacrificing accuracy, using your left ring finger for low tones, your middle finger for medium tones, and your index finger for high tones. Press the green button when you are ready to begin.

Dual-Task Practice

Now you will practice the two tasks together. There will be two blocks of 36 dual-task practice trials each. On each trial, after you press the green button, you will see a picture and then hear a tone. Say the name of the picture first, and then press the button indicating which tone you heard. Always say the name of the picture first, as quickly as you can without sacrificing accuracy and then press the button for the tone as quickly as you can after that, also without sacrificing accuracy. Don't wait to say the name of the picture until after you've made a decision about the tone. It's important that you respond first to the picture on each trial, as quickly as you can. Press the green button when you are ready to begin.

Dual-Task Experimental Trials

Now you will do four blocks of experimental trials. Each block will be twice as long as the practice blocks you just did. When you see the "+" on the screen, press the green button with your right index finger to signal that you are ready to begin the trial. After you press the green button, you will see a picture and then hear a tone. Say the name of the picture, and then press the button indicating which tone you heard. Always say the name of the picture first, as quickly as you can without sacrificing accuracy and then press the button for the tone as quickly as you can after that, also without sacrificing accuracy. Press the green button when you are ready to begin.

Experiment 2: Tone Primary

Initial Instructions

In this experiment, you're going to be performing two tasks: tone identification and picture naming. In the tone ID task, you will hear a tone and press a button to indicate whether it is high, medium, or low in pitch. In the naming task, you will see a picture on the computer screen and say the name of the picture. You will have the chance to practice each task separately, and then together, before doing the experimental trials in which both tasks will be presented. For both tasks, it is important that you respond as quickly as you can without sacrificing accuracy.

Single-Task Practice

The single-task practice instructions were identical to Experiment 1.

Dual-Task Practice

Now you will practice the two tasks together. There will be two blocks of 36 dual-task practice trials each. On each trial, after you press the green button, you will hear a tone and then see a picture. Press the button indicating which tone you heard, and then say the name of the picture. Always press the button for the tone first, as quickly as you can without sacrificing accuracy, and then say the name of the picture as quickly as you can after that, also without sacrificing accuracy. Don't wait to respond to the tone until after you've made a decision about the picture. It's important that you respond first to the tone on each trial, as quickly as you can. Press the green button when you are ready to begin.

Dual-Task Experimental Trials

Now you will do four blocks of experimental trials. Each block will be twice as long as the practice blocks you just did. On each trial, after you press the green button, you will hear a tone and then see a picture. Press the button indicating which tone you heard, and then say the name of the picture. Always press the button for the tone first, as quickly as you can without sacrificing accuracy, and then say the name of the picture as quickly as you can after that, also without sacrificing accuracy. Press the green button when you are ready to begin.

Experiment 3: Variable Task Order

The initial and single-task practice instructions were identical to Experiment 1.

Dual-Task Practice

Now you will practice the two tasks together. When you see a "+" on the screen, press the green button with your right index finger to signal that you are ready for the next trial. After you press the green button, you will see a picture and hear a tone. Sometimes the picture will come first, and sometimes the tone will come first. You should respond as before, making each individual response as quickly as you can without sacrificing accuracy. You may respond in either order, but it is important that you give equal attention or effort to both tasks. Press the green button when you are ready to begin.

Contingency Instructions

The following instructions were given between blocks if, on the previous block, the participant responded in task presentation order on ≤ 33% of trials in any condition: You seem to be giving more attention to the [the name of the task being responded to first on more trials] than to the other task. Remember, you should give equal attention or effort to both tasks.

Dual-Task Experimental Trials

These instructions were identical to the dual-task practice trials, except that the first sentence was omitted and replaced with: Now you will do the experimental trials. There will be four blocks of experimental trials, and each one will be two or three times as long as the practice blocks you just did.

APPENDIX D

PILOT STUDY

INTRODUCTION

Experiment 3 represented a substantial modification of traditional PRP methods, and included picture-naming, a task that has been few studies of PRP dual-task performance. Also, the analysis plan for Experiment 3 permitted the use of data only from trials in which both tasks were responded to correctly and in the order of presentation. For these reasons, a pilot study was conducted to determine whether Experiment 3 as proposed was likely to provide enough valid data to address the experimental questions.

METHOD

Participants

Eight subjects (3 male, 5 female) ages 18-48 (mean = 29, SD = 8.9) participated in the pilot experiment. They all provided informed consent and met the same inclusion criteria as participants in Experiments 1-3. Two potential subjects who provided informed consent to participate were excluded because they failed to meet criterion on the picture-naming screening test.

Apparatus and Stimuli

Stimulus presentation and data collection were performed using a locally-designed software program, Stimulate (Necessity Consulting, 2006), run on a Dell Latitude X1 notebook computer. Visual stimuli were displayed to participants on an ELO TouchSystems 1525L monitor. Tones were presented binaurally through headphones. Vocal reaction times were collected by voice key, and manual reaction times were collected using an X-Keys SE keypad.

Procedure

Procedures were similar to those for Experiment 3. Briefly, on each trial, after pressing a green "ready" button with the right index finger, subjects saw a picture and heard one of three tones. Subjects were required to name the picture as quickly as possible and to press a button with the left hand as quickly as possible indicating whether they heard a high, medium, or low tone. On half of the trials the picture appeared first, and on the other half, the tone was presented first. Stimulus onset asynchrony values were 50, 150, and 900 ms in both presentation orders. Subjects were instructed to give equal attention or effort to both tasks, that they were free respond in either order, and that each response should be as fast and accurate as possible. Prior to the experimental trials, subjects performed 36 practice trials of each task by itself, and then 2 blocks of 36 dual-task practice trials each. If, during the first dual-task practice block, a participant responded first to either task on a substantial majority of trials (>70%), they were informed that they appeared to be giving more attention to that task than to the other, and were re-instructed to give equal attention to both tasks.

Design

Four blocks of 72 trials each were run for each subject. Trial blocks contained two trials representing each of the 36 possible combinations of the following variables: Task order (naming primary, tone primary), lexical frequency (high, low), stimulus onset asynchrony (50, 150, 900 ms), and tone pitch (high, medium, low). The order of presentation of conditions was pseudo-randomized such that the entire design was repeated every 36 trials.

RESULTS

Trials were excluded from analysis based on the following criteria, in the following order: (1) the voice key failed to accurately register the vocal response; (2) the subject gave a naming response that was other than the target response or which included an audible false start or self-correction; (3) the tone ID response was incorrect; and (4) the response order did not match the stimulus presentation order. Voice key failures occurred on 1.2% of trials, naming errors on an additional 12.1%, and tone ID errors on additional 6.5%. A further 22.3% of trials on which the response order failed to match stimulus presentation order were excluded. Finally, each subject's data were screened for reaction time outliers using the recursive procedure described by Van Selst and Jolicoeur (1994), resulting in the exclusion of an additional 2.5% of trials.

Following exclusion of invalid trials as described above, five subjects met the criterion of ≥ 4 valid trials in each cell of the design for analysis of the naming-primary trials, and seven subjects met the criterion for the tone-primary trials. The number of valid trials per condition are summarized in Table 11 for subjects meeting the criterion for inclusion in reaction time analyses and in Table 2 for the entire sample.

Table 11. Summary of valid trials per condition for pilot subjects with ≥ 4 valid trials in all naming-primary (n = 5) or tone-primary (n = 7) conditions.

	Naming Primary						Tone ID Primary					
	High Frequency			Low Frequency			High Frequency			Low Frequency		
SOA	50ms	150ms	900ms	50ms	150ms	900ms	50ms	150ms	900ms	50ms	150ms	900ms
Mean	11.0	12.0	17.6	8.8	12.2	15.2	15.0	17.7	18.8	12.5	13.0	17.5
(SD)	(2.6)	(4.6)	(2.7)	(3.3)	(5.0)	(3.6)	(4.3)	(3.4)	(3.1)	(3.3)	(2.8)	(2.2)
Range	8-15	9-20	15-22	5-13	6-18	12-20	7-20	13-23	14-22	10-19	9-17	14-20

Table 12. Summary of valid trials per condition for all pilot subjects (n = 8).

	Naming Primary						Tone ID Primary					
	High Frequency			Low Frequency			High Frequency			Low Frequency		
SOA	50ms	150ms	900ms	50ms	150ms	900ms	50ms	150ms	900ms	50ms	150ms	900ms
Mean	7.8	9.4	17.5	6.9	10.0	15.5	14.8	16.5	18.3	11.6	13.6	18.0
(SD)	(4.9)	(5.1)	(2.2)	(3.7)	(4.9)	(3.0)	(4.9)	(4.2)	(3.2)	(4.5)	(4.0)	(2.4)
Range	2-15	3-20	15-22	3-13	5-18	12-20	7-20	9-23	14-22	3-19	9-21	14-22

Given the large amount of excluded data and the need to balance the lexical characteristics of the naming stimuli across conditions, a post-hoc analysis of the characteristics of the naming items that were actually included in the RT analyses was conducted. For each subject, the mean lexical frequency, length in phonemes, length in syllables, and single-task naming reaction time (from the CRL database) for the items actually entered into the RT analyses described below were examined for each cell of the design. In all cases, the desired large difference in lexical frequency was obtained, and in no case did imbalance on the length or the single-task RT variable appear to account for the minority of individual instances where the expected effect of lexical frequency on dual-task naming RT failed to emerge. Also, the lexical frequency, length, and single-task RT data were entered as dependent variables into ANOVAs analogous to those used to analyze the experimental reaction time data. Large, reliable differences in lexical frequency and single-task RT were obtained where expected in all cases. There was a trend for low lexical frequency items entered into the RT analyses to be slightly longer than high-frequency items, by approximately 0.1 to 0.15 phonemes and .05 syllables on average. However, this trend was not reliable across SOA conditions. Also, the pattern of differences in stimulus length across conditions did not suggest that these variables contributed to observed trends or effects of lexical frequency and SOA on dual-task RT reported below.

As in Experiment 3, separate ANOVAs with lexical frequency and SOA as repeated factors were conducted for each of the dependent variables (primary naming RT, secondary tone ID RT, primary tone ID RT, and secondary naming RT). The alpha level was set at 0.05 for the primary-task ANOVAs and 0.001 for the secondary-task analyses. The reaction time data for the naming-primary trials are presented in Figure 4, and the data for the tone-primary trials are presented in Figure 5.

<u>Naming performance on naming-primary trials:</u> The expected trend of longer naming RTs in the low lexical frequency conditions was observed. This trend did not reach significance with the current small sample size, $F(1,4) = 2.822$, $p = 0.17$, $ES = 0.41$. There was also a trend for longer RTs in the 900ms SOA condition, which failed to reach significance, $F(2,8) = 1.140$, $p = 0.36$, $ES = 0.22$. Finally, the interaction trend suggesting a reduced frequency effect in the 50ms SOA condition also failed to reach significance, $F(2,8) = 2.84$, $p = 0.12$, $ES = 0.42$.

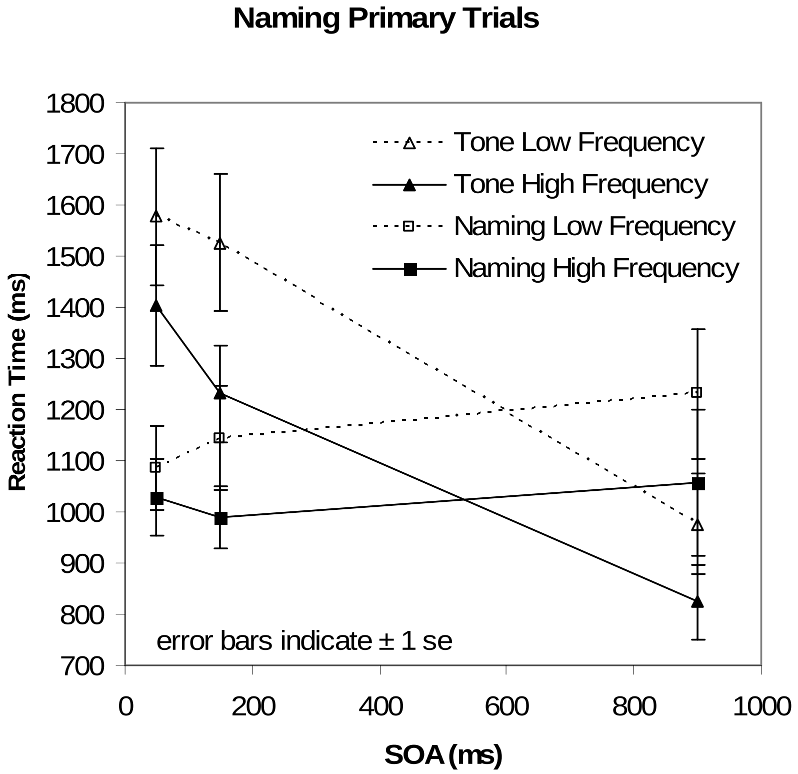

Figure 17. Mean reaction times for naming-primary trials.

Tone ID performance on naming-primary trials: The expected trend of longer RTs for tone ID responses following low-frequency picture names was observed at all SOAs, but did not reach significance with the current sample size, F(1,4) = 4.804, p = 0.09, ES = 0.55. The predicted effect of SOA was statistically significant, F(2,8) = 73.646, p < 0.001, ES = 0.95. There was no prominent trend of an interaction effect, F(2,8) = 1.629, p = 0.26, ES = 0.29.

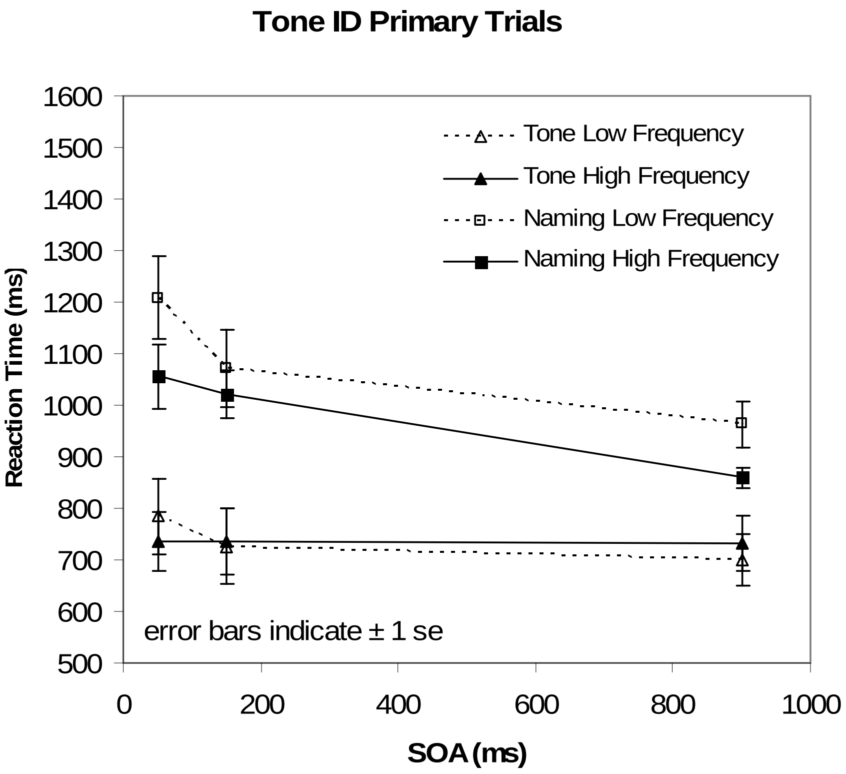

Figure 18. Mean reaction times for tone-primary trials.

Tone ID performance on tone-primary trials: There was no substantial trend for either lexical frequency, F(1,6) = 0.007, p = 0.936, ES = 0.001, or SOA, F(2,12) = 0.674, p = 0.528, ES = .10, on primary-task Tone ID RTs. There was a slight interaction trend, F(2,12) = 2.389, p = 0.134, ES = 0.29, characterized by increasing RTs at shorter SOAs for low frequency trials only.

Naming performance on tone-primary trials: The expected effects of lexical frequency, F(1,6) = 19.102, p = 0.005, ES = 0.76, and SOA, F(2,12) = 14.889, p = 0.001, ES = 0.71, were observed, although only the latter reached statistical significance in the current pilot sample, given the strict alpha level. Also, a non-significant interaction trend characterized by a smaller difference between the high and low lexical frequency items at the 150ms SOA was observed, F(1,6) = 2.630, p = 0.11, ES = 0.31.

DISCUSSION

Although a minority of subjects failed to respond correctly to both tasks in the presentation order on a sufficient number of trials to support the RT analyses, particularly in the naming-primary conditions, the results of the pilot study suggested that the proposed methods would obtain enough valid data to support analysis and interpretation of the results. One possible reason for the finding that fewer correct responses in task presentation order were obtained in the naming-primary conditions is that the average reaction times to the Tone ID task were shorter, by 200-300ms. Thus, even if subjects were processing the two tasks in the correct order as proposed by the central resource model, the proportion of short-SOA naming-primary trials on which the tone ID response is given first might be large. Alternately, subjects might simply be choosing to complete the easier task first on a majority of trials.

One could argue that this imbalance in responses to the naming-primary trials presented a threat to the validity and potential informativeness of Experiment 3, on the grounds that it suggested that the relatively even resource allocation ratios necessary for differentiating the resource from the bottleneck model were unlikely to be obtained. While this imbalance was a concern, it was not, however, a fatal flaw. First, when the data for the four subjects who produced sufficient data in all naming- and tone-primary conditions were analyzed separately, the trends for the effects of lexical frequency and SOA were identical to those observed in the larger samples. Second, fixed-order PRP experiments have in some cases shown RT1 effects consistent with resource-sharing (e.g., Carrier & Pashler, 1995; Cleland et al., 2006; Tombu & Jolicoeur, 2005). Third, the tone-primary trials were in any case more important for testing the predictions of the central capacity model, because they permitted examination of the main and interaction effects of both SOA and second-task difficulty on primary task RTs. The naming-primary trials, while potentially helpful in inducing subjects to give equal effort to both tasks and informative for distinguishing between the two models under consideration, only permitted examination of the main effect of SOA on primary task RT.

The extremely small sample size precluded strong conclusions from the reaction time results of this pilot experiment, but certain observed trends did suggest that Experiment 3 was likely to produce data that would informatively address the specific aims of this proposal. First, as predicted, there was a trend of longer average primary and secondary task naming RTs in the low lexical frequency conditions across all levels of SOA. This trend was also observed as expected in the secondary task tone RTs. Second, there was prominent and statistically reliable slowing of second-task RTs at short SOAs for both naming and tone ID. This is consistent with the robust PRP effect predicted by both the bottleneck and capacity models.

APPENDIX E

OBJECT RECOGNITION EXPERIMENT

METHOD

Participants

Participants were 27 University of Pittsburgh Students (18 men, 10 women) ranging in age from 18 to 24 (mean = 18.9, sd = 1.4) recruited through the Department of Psychology Research Participation Program. They met the same selection criteria as participants in Experiments 1-3. One potential subject was excluded from participation because he reported a native language other than English. Three participants were excluded from the analyses due to examiner errors in administering the protocol. The 24 participants whose data were submitted to analysis included 15 men and 9 women ranging in age from 18 to 24 (mean = 19.0, sd = 1.4).

Apparatus and Stimuli

The apparatus was the same as that used for Experiments 1-3, except that there was no audio recording, and manual reaction time responses were collected via the Psychology Software Tools Serial Response Box. The picture stimuli were the 288 object pictures used in Experiments 1 and 2, and 44 additional pictures (22 high frequency, 22 low frequency). Word stimuli were the

332 names of the picture stimuli, and an additional non-overlapping set of 166 object picture names also taken from the CRL-IPNP database.

Procedure

All data were collected in a single session for each participant requiring approximately 30 minutes. Following informed consent and screening, participants completed 12 practice trials, and then four blocks of 83 experimental trials each. Instructions were presented verbally and in writing on the computer screen. Participants were instructed to respond to each picture as quickly as possible without sacrificing too much accuracy or becoming careless, and to make all responses using only their dominant index finger.

Each trial began with a fixation cross. When subjects pressed a button labeled "ready", the cross disappeared, and after a 200 ms delay a word was displayed on the screen for 1000 ms. Two hundred milliseconds following the offset of the word, a picture stimulus was presented. Participants were instructed to press a button labeled "yes" if the picture stimulus matched the word, or a button labeled "no" if it did not. The "yes" and "no" buttons were located immediately to the left and right, respectively, of the "ready" button. The picture remained on the screen until a response was detected, or until 2000 ms had elapsed. Following the response, a feedback screen displayed the RT for that trial and the mean percent correct for that trial block for 1000 ms. The next trial began following a 500 ms inter-trial interval.

Design

For each odd-numbered participant, half of the high and low frequency pictures were randomly assigned to the "yes" condition, and the other half were assigned to the "no" condition. For each even-numbered participant, the mapping of high and low frequency pictures was exactly the opposite of that given to the preceding odd-numbered participant. For the "no" trials, pictures were paired with word stimuli drawn from the same half of the lexical frequency distribution as their target names. Also, on the "no" trials, the words were pseudorandomly assigned to pictures for each participant such that they bore no semantic or associative relationship to the picture, and did not share an initial phoneme or rhyme with the picture's target name. Each participant received 166 "yes" trials and 166 "no" trials, with both trial types evenly split between high and low frequency pictures. Across participants, each picture was presented equally often in the "yes" and "no" conditions. Each block of 83 experimental trials was as evenly split as possible between the four trial types.

ANALYSIS AND RESULTS

Only the 288 pictures used in Experiments 1 and 2 were included in the analyses. Error rates were examined by participants averaged across items, to inspect for any speed-accuracy trade-off. A two-way ANOVA with lexical frequency (high, low) and trial type (yes, no) as repeated factors was conducted. Error rates were low across conditions, 1.4% for high frequency pictures, 1.5% for low frequency pictures, 1.7% for "yes" trials and 1.3% for "no" trials. Neither the main effect of frequency, $F(1, 23) = 0.055$, $p = 0.816$, nor trial type, $F(1, 23) = 1.396$, $p = 0.249$, was significant. The interaction was significant, $F(1, 23) = 6.338$, $p = 0.019$, with "yes" trials showing a difference in the expected direction (1.2% high frequency, 2.1% low frequency), and "no" trials demonstrating the opposite trend (1.7% high frequency, 0.9% low frequency). Examination of the simple effects suggested that the effect of frequency was more reliable in the "yes" conditions than in the "no" conditions. Thus, there was no evidence for any speed-accuracy trade-off.

Reaction times were analyzed by item, averaged across participants[8]. Erred trials and RT outliers, determined by the same procedure as Experiments 1-3, were excluded. The RT means are presented by condition in Table 13.

Table 13. Object recognition reaction time means in milliseconds. Standard errors are given in parentheses.

Lexical Frequency	Trial Type	
	Yes	No
High	488 (5.3)	510 (3.3)
Low	506 (5.3)	512 (3.3)

A two-way ANOVA with trial type as a repeated factor and lexical frequency as a between-items factor was conducted. The main effect of trial type was significant, F(1, 286) = 10.308, MSe = 2671, p = 0.001, ES = 0.035, as was the effect of lexical frequency, F(1, 286) = 4.396, MSe = 2970, p = 0.037, ES = 0.015. Although there was a trend for lexical frequency to have a greater effect on "yes" trials, the interaction was not significant, F(1, 286) = 3.215, MSe = 2671, p = 0.074, ES = 0.011.

DISCUSSION

The results demonstrate a small, but reliable effect of lexical frequency on object recognition time for the picture stimuli included in Experiments 1 and 2. When evaluating differences in object recognition time between the balanced high and low frequency stimulus lists used in the

---

[8] Reaction time analyses by participants, averaged across items produced exactly the same pattern of results as those reported here.

follow-up analyses for Experiments 1 and 2, and the tone-primary conditions in Experiment 3, the mean object recognition time for each object was taken from the "yes" conditions, because, despite the lack of a significant interaction effect, this was the condition where lexical frequency effects were most evident. Also, this was the same strategy employed by Jescheniak and Levelt (1994).

APPENDIX F

LEXICAL AND PICTURE CHARACTERISTICS FOR EXPERIMENT 3 NAMING STIMULI

Tables 14 and 15 present the low and high frequency picture naming targets for the tone-primary trials of Experiment 3, and their associated log frequency values. Table 16 summarizes the lexical and picture characteristics for the high and low frequency lists, and presents the results of independent samples t-tests for the following variables: Lexical frequency, rated image agreement, name agreement, naming RT, object recognition RT, rated age of acquisition, number of phonemes, and number of syllables. Table 17 presents the correlations among those variables within the combined high and low frequency lists. Tables 18-21 present the analogous data for the stimuli presented on Experiment 3 naming-primary trials. It should be noted that object recognition RTs were unavailable for a majority of the naming-primary stimuli.

Table 14. Low Frequency picture naming targets for Experiment 3 tone-primary trials.

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
highchair	1.79	funnel	4.34	cherry	5.11
seahorse	1.79	sock	4.38	lizard	5.12
bra	2.40	mailbox	4.42	shovel	5.15
stroller	2.49	vase	4.44	stool	5.18
tweezers	2.77	waiter	4.44	dolphin	5.21
ladle	3.18	toaster	4.45	couch	5.21
backpack	3.40	lobster	4.48	hose	5.22
tripod	3.40	carrot	4.50	peanut	5.22
igloo	3.43	wig	4.50	broom	5.23
thimble	3.53	swan	4.55	raccoon	5.23
hammock	3.61	skis	4.59	moose	5.27
hanger	3.61	bathtub	4.61	fountain	5.34
fireman	3.64	giraffe	4.63	fan	5.35
snowman	3.64	windmill	4.64	screw	5.35
banjo	3.83	lemon	4.75	anchor	5.40
peacock	3.83	onion	4.76	guitar	5.43
penguin	3.87	banana	4.77	puzzle	5.43
sailboat	3.87	skunk	4.78	spoon	5.43
llama	3.91	parrot	4.80	cannon	5.43
pear	3.91	pizza	4.80	robot	5.44
acorn	4.04	glove	4.84	towel	5.45
dice	4.08	scarf	4.84	shower	5.45
crib	4.11	boot	4.98	tire	5.46
mop	4.14	snail	4.98	comb	5.48
razor	4.19	medal	5.00	donkey	5.48
hoof	4.22	pirate	5.02	grapes	5.51
genie	4.23	bride	5.04	chimney	5.52
walrus	4.23	toilet	5.04	flashlight	5.54
crackers	4.28	clown	5.05	knight	5.54
wrench	4.28	scissors	5.05	pillow	5.55
rake	4.29	plug	5.06	bucket	5.55
vest	4.30	cactus	5.09	bomb	5.58
seesaw	4.32	crab	5.09	fork	5.62

Table 15. High Frequency picture naming targets for Experiment 3 tone-primary trials.

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
volcano	5.80	pan	6.34	plate	7.03
pants	5.83	glasses	6.37	shoulder	7.06
vacuum	5.84	whale	6.38	bridge	7.07
squirrel	5.85	globe	6.39	truck	7.12
thumb	5.88	microscope	6.45	roof	7.14
cigarette	5.89	ruler	6.45	bread	7.22
skeleton	5.89	magnet	6.50	hat	7.24
statue	5.90	chicken	6.51	salt	7.26
lock	5.92	jar	6.53	telephone	7.27
spider	5.94	balloon	6.54	corn	7.27
button	5.95	bicycle	6.54	desert	7.32
saddle	5.96	football	6.54	smoke	7.38
candle	5.97	pipe	6.56	radio	7.47
shoe	5.99	camera	6.58	arm	7.65
sword	6.01	pool	6.62	baby	7.76
crown	6.01	basket	6.63	doctor	7.77
package	6.04	nest	6.63	glass	7.80
seal	6.07	castle	6.64	mountain	7.83
hammer	6.08	brush	6.64	indian	7.86
sink	6.08	apple	6.64	letter	7.97
turkey	6.09	lips	6.67	present	8.01
piano	6.15	flower	6.70	hair	8.23
telescope	6.16	chest	6.72	woman	8.27
soldier	6.21	clock	6.76	table	8.37
lamp	6.23	mirror	6.77	fish	8.41
lightning	6.23	bottle	6.78	picture	8.43
jacket	6.25	finger	6.81	car	8.45
tent	6.25	stairs	6.81	king	8.45
butter	6.25	fox	6.82	paper	8.53
elephant	6.29	cloud	6.83	hand	8.95
dragon	6.30	frog	7.00	city	8.99
rocket	6.30	rope	7.01	house	9.41
stove	6.30	wagon	7.01	well	9.60

Table 16. Summary of picture and lexical characteristics of the Experiment 3 tone-primary picture stimuli.

Variable	High Frequency		Low Frequency		t	Error df	1-tailed p-value
	mean	sd	mean	sd			
Log Frequency	6.87	0..88	4.59	0.83	18.629	196	<0.001
Image Agreement	5.97	0.56	5.90	0.52	0.941	194	0.174
Name Agreement	0.91	0.10	0.91	0.09	-0.347	194	0.364
Naming RT	928	149	1010	179	-0.348	194	<0.001
Object Recognition RT	491	52	488	45	0.382	196	0.351
Age of Acquisition Rating	4.65	0.93	5.54	1.20	-5.817	194	<0.001
Number of Phonemes	4.49	1.3	4.51	1.08	-0.119	196	0.453
Number of Syllables	1.64	0.68	1.64	.50	0	196	0.500

Table 17. Correlations among picture and lexical characteristics of the Experiment 3 naming stimuli for tone-primary trials. Correlations denoted with an asterisk (*) are significant at $p < 0.05$.

	Log Frequency	Image Agreement	Name Agreement	Naming RT	Object Recognition RT	Age of Acquisition (AoA)	#Phonemes
Image Agr.	0.02						
Name Agr.	-0.03	0.21*					
Name RT	-0.32*	-0.45*	-0.40*				
Object RT	0.06	-0.28*	-0.15*	0.31*			
AoA	-0.52*	-0.16*	-0.10	0.58*	0.20*		
#Phonemes	-0.16*	0.13*	-0.10	0.14*	0.001	0.20*	
#Syllables	-0.12*	0.05	-0.08	0.09	0.03	0.17*	0.70*

Table 18. Low Frequency picture naming targets for Experiment 3 naming-primary trials.

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
pinecone	0.69	anvil	4.28	ax	5.26
lightbulb	0.69	leopard	4.29	bug	5.27
trashcan	1.10	hippo	4.30	suitcase	5.30
carousel	1.39	hinge	4.34	toe	5.31
stoplight	1.79	maze	4.41	pitcher	5.32
dustpan	1.79	wallet	4.44	pyramid	5.32
artichoke	1.79	necklace	4.44	rainbow	5.33
popsicle	1.95	spaghetti	4.45	helmet	5.33
mousetrap	2.08	gorilla	4.51	butterfly	5.34
lawnmower	2.08	balcony	4.55	lighthouse	5.35
yoyo	2.08	cookie	4.56	purse	5.37
teepee	2.20	harp	4.58	peas	5.39
corkscrew	2.64	walnut	4.59	umbrella	5.44
clothespin	2.71	octopus	4.63	whip	5.48
blimp	2.71	parachute	4.64	glue	5.50
slingshot	2.71	hoe	4.65	sweater	5.52
diaper	3.18	wheelchair	4.65	skirt	5.53
spatula	3.18	tractor	4.71	typewriter	5.54
pitchfork	3.33	hamburger	4.85	sailor	5.56
paintbrush	3.40	torch	4.91	ant	5.57
zipper	3.43	trumpet	4.92	curtains	5.60
ladybug	3.53	violin	4.92	paw	5.62
pelican	3.56	peach	4.93	mask	5.62
thermos	3.56	kangaroo	4.94	bricks	5.63
slipper	3.64	cork	4.97	dentist	5.63
pliers	3.78	lettuce	5.04	sled	5.63
antlers	3.97	knot	5.04	worm	5.64
badge	4.01	heel	5.06	cowboy	5.65
zebra	4.01	eskimo	5.08	shark	5.65
clamp	4.03	rooster	5.10	envelope	5.66
panda	4.13	nut	5.14	sandwich	5.67
faucet	4.14	tomato	5.17	rug	5.68
celery	4.22	dime	5.21	hook	5.69

Table 19. High Frequency picture naming targets for Experiment 3 naming-primary trials.

Word	Log Freq.	Word	Log Freq.	Word	Log Freq.
cane	5.70	slide	6.25	leg	7.00
potato	5.71	cage	6.27	rabbit	7.01
drawer	5.72	bow	6.28	paint	7.01
doll	5.72	duck	6.28	coat	7.04
monkey	5.73	needle	6.35	jack	7.11
grave	5.73	tank	6.35	block	7.12
closet	5.74	airplane	6.39	newspaper	7.15
camel	5.76	log	6.39	sheep	7.18
policeman	5.76	belt	6.40	bag	7.20
drill	5.76	wine	6.41	neck	7.22
barrel	5.77	pencil	6.43	tail	7.38
tear	5.79	quarter	6.48	train	7.42
beard	5.79	bowl	6.49	safe	7.49
feather	5.79	orange	6.54	fly	7.50
trash	5.80	knife	6.54	bird	7.58
drum	5.81	shirt	6.58	rose	7.58
bee	5.83	pot	6.61	music	7.59
beaver	5.83	nurse	6.63	boat	7.61
bat	5.87	tape	6.64	cat	7.63
whistle	5.88	snake	6.64	iron	7.79
goat	5.96	net	6.65	teeth	7.87
swing	6.02	cup	6.65	bear	7.88
eagle	6.03	deer	6.67	gas	7.89
priest	6.05	shell	6.68	wood	7.97
turtle	6.06	fence	6.73	rain	8.04
ladder	6.06	lion	6.76	window	8.10
canoe	6.06	wheat	6.76	road	8.13
wing	6.09	wolf	6.78	floor	8.14
tiger	6.11	gun	6.80	rock	8.17
arrow	6.13	chain	6.84	girl	8.23
flag	6.16	branch	6.95	top	8.49
thread	6.21	scale	6.95	man	9.69
bench	6.21	dress	7.00	can	10.78

Table 20. Summary of picture and lexical characteristics of the Experiment 3 naming-primary picture stimuli.

Variable	High Frequency		Low Frequency		t	Error df	1-tailed p-value
	mean	sd	mean	sd			
Log Frequency	6.74	0.90	4.36	1.29	15.094	196	<0.001
Image Agreement	5.70	0.69	5.56	0.64	1.394	191	0.082
Name Agreement	0.82	0.18	0.77	0.17	1.901	191	0.029
Naming RT	1029	208	1131	216	-3.340	191	<0.001
Object Recognition RT	480	40	537	104	-2.698	64	0.004
Age of Acquisition Rating	5.00	1.39	5.75	1.19	-4.036	191	<0.001
Number of Phonemes	3.73	1.03	5.17	1.60	-7.544	196	<0.001
Number of Syllables	1.30	0.52	1.91	0.73	-6.712	196	<0.001

Table 21. Correlations among picture and lexical characteristics of the Experiment 3 naming stimuli for naming-primary trials. Correlations denoted with an asterisk (*) are significant at $p < 0.05$.

	Log Frequency	Image Agreement	Name Agreement	Naming RT	Object Recognition RT	Age of Acquisition (AoA)	#Phonemes
Image Agr.	0.004						
Name Agr.	0.15*	0.44*					
Name RT	-0.31*	-0.52*	-0.57*				
Object RT	-0.24*	-0.38*	-0.30*	0.33*			
AoA	-0.37*	-0.11	-0.25*	0.47*	0.03		
#Phonemes	-0.57*	0.14*	-0.03	0.07	-0.23*	0.29*	
#Syllables	-0.49*	0.15*	-0.09	0.01	-0.16	0.16*	0.82*

BIBLIOGRAPHY

Allen, P. A., Lien, M.-C., Murphy, M. D., Sanders, R. E., Judge, K. S., & McCann, R. S. (2002). Age differences in overlapping-task performance: Evidence for efficient parallel processing in older adults. *Psychology and Aging, 17,* 505-519.

Allen, P. A., Wallace, B., & Weber, T. A. (1995). Influence of case type, word frequency, and exposure duration on visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 914-934.

Arvedson, J. C. (1986). *Effect of lexical decisions on auditory semantic judgments using divided attention in adults with left and right hemisphere damage.* Unpublished Doctoral Dissertation, University of Wisconsin-Madison.

Arvedson, J. C. & McNeil, M. R. (1987). Accuracy and response times for semantic judgments and lexical decisions with left and right hemisphere lesions. *Clinical Aphasiology, 17,* 188-200.

Baayen, R. H., Piepenbrick, R., & Gulikers, L. (1995). *The CELEX Lexical Database (Release 2) [CD-ROM]*. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283-316.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology, 50A,* 560-585.

Bates, E., Frederici, A., & Wulfeck, B. (1987). Comprehension in aphasia: A cross-linguistic study. *Brain and Language, 32,* 19-67.

Blackwell, A. & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience, 7,* 228-257.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50,* 456-476.

Bonin, P., Chalard, M., Meot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology, 93,* 89-114.

Bonin, P. & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology, 14,* 289-313.

Brookshire, R. H. & Nicholas, L. (1984). Comprehension of directly and nondirectly stated main ideas and details in discourse by brain-damaged and non-brain-damaged listeners. *Brain and Language, 21,* 36.

Brookshire, R. H. & Nicholas, L. E. (1980). Verification of active and passive sentences by aphasic and nonaphasic subjects. *Journal of Speech and Hearing Research, 23,* 878-893.

Bull, P. & Connely, G. (1985). Body movement and emphasis in speech. *Journal of Nonverbal Behavior, 9,* 169-187.

Campbell, T. F. & McNeil, M. R. (1985). Effects of presentation rate and divided attention on auditory comprehension in children with acquired language disorder. *Journal of Speech and Hearing Research, 28,* 513-520.

Caplan, D. (1987). *Neurolinguistics and linguistic aphasiology*. Cambridge, MA: Cambridge University Press.

Caplan, D. & Waters, G. S. (1996). Syntactic processing in sentence comprehension under dual-task conditions in aphasic patients. *Language and Cognitive Processes, 11,* 525-551.

Caramazza, A., Bi, Y., Costa, A., & Miozzo, M. (2004). What determines the speed of lexical access: homophone or specific-word frequency? A reply to Jescheniak et al. (2003). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 278-282.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1430-1450.

Carrier, L. M. & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1339-1348.

Clark, H. M. & Robin, D. A. (1995). Sense of effort during a lexical decision task: Resource allocation deficits following brain damage. *American Journal of Speech-Language Pathology, 4,* 143-147.

Cleland, A. A., Gaskell, M. G., Quinlan, P. T., & Tamminen, J. (2006). Frequency effects in spoken and visual word recognition: Evidence from dual-task methodologies. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 104-119.

Craik, K. J. W. (1947). Theory of the human operator in control systems I. The operator as an engineering system. *British Journal of Psychology, 38,* 142-148.

Craik, K. J. W. (1948). Theory of the human operator in control systems II. Man as an element in a control system. *British Journal of Psychology, 38,* 56-61.

Crisman, L. G. (1971). *Response variability in naming behavior of aphasic patients.* Unpublished master's thesis, University of Pittsburgh.

Cutting, J. C. & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 318-344.

D'Amico, E. J., Neilands, T. B., & Zambarano, R. (2001). Power analysis for multivariate and repeated measures designs: a flexible approach using the SPSS MANOVA procedure. *Behavior Research Methods, Instruments, & Computers, 33,* 479-484.

Darley, F. L. (1976). Maximizing input to the aphasic patient: a review of research. In R.H.Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp. 1-21). Minneapolis, MN: BRK Publishers.

Darley, F. L. (1982). *Aphasia*. Philadelphia, PA: W.B. Saunders.

De Bleser, R. (1988). Localisation of aphasia: Science or fiction. In G.Denes, C. Semenza, & P. Bisiacchi (Eds.), *Perspectives on cognitive neuropsychology* (pp. 161-185). Hillsdale, NJ: Lawrence Erlbaum.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 965-980.

Dell'Acqua, R., Job, R., Peressotti, F., & Pascali, A. (in press). The picture-word interference effect is not a Stroop effect. *Psychonomic Bulletin & Review.*

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283-321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313-349.

Dell, G. S. & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychological Review, 98,* 604-614.

Dell, G. S. & Reich, P. A. (1981). Stages in verbal sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20,* 629.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801-838.

Der, G. & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging, 21,* 62-73.

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review, 108,* 759-788.

Duffy, J. R. & Coelho, C. A. (2001). Schuell's stimulation approach to rehabilitation. In R.Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (4th ed., pp. 341-382). Baltimore, MD: Lippincott Williams & Wilkins.

Eggert, G. H. (1977). *Wernicke's works on aphasia: A sourcebook and review*. The Hague: Mouton.

Erickson, R. J., Goldinger, S. D., & LaPointe, L. L. (1996). Auditory vigilance in aphasic individuals: Detecting nonlinguistic stimuli with full or divided attention. *Brain and Cognition, 30,* 244-253.

Ernest-Baron, C. R., Brookshire, R. H., & Nicholas, L. (1987). Story structure and retelling of narratives by aphasic and non-brain-damaged adults. *Journal of Speech and Hearing Research, 30,* 44-49.

Fagot, C. & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1058-1079.

Fay, D. & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8,* 505-520.

Ferraro, F. R. & Hansen, C. L. (2002). Orthographic neighborhood size, number of word meanings, and number of higher frequency neighbors. *Brain and Language, 82,* 200-205.

Ferreira, V. S. & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1187-1199.

Fisher, S. (1984). Central capacity limits in consistent mapping visual search tasks: Four channels or more? *Cognitive Psychology, 16,* 449-484.

Freed, D. B., Marshall, R. C., & Chulantseff, E. A. (1996). Picture naming variability: A methodological consideration of inconsistent naming responses in fluent and nonfluent aphasia. In R.H.Brookshire (Ed.), (pp. 193-205). Austin, TX: Pro-Ed.

Freud, S. (1953). *On aphasia*. New York: International Universities Press, Inc. (originally published in German, 1891).

Garrett, M. F. (1975). The anaylsis of sentence production. In G.Bower (Ed.), *Psychology of learning and motivation* (pp. 133-177). New York: Academic Press.

Garrett, M. F. (1980). The anaylsis of sentence production. In B.Butterwoth (Ed.), *Language production* (pp. 177-120). New York: Academic Press.

German, D. J. (1990). *Test of Adolescent/Adult Word Finding*. Austin, TX: Pro-Ed.

Geschwind, N. (1965a). Disconnexion syndromes in animals and man. Part I. *Brain, 88,* 237-294.

Geschwind, N. (1965b). Disconnexion syndromes in animals and man. Part II. *Brain, 88,* 585-644.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: a review of the literature and a new multi-task investigation. *Acta Psychologica, 115,* 43-67.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders*. (3rd ed.) Baltimore: Lippincott, Williams, & Wilkins.

Gopher, D., Brickner, M., & Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: Evidence for multiple resources. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 146-157.

Gopher, D. & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors, 26,* 519-532.

Griffin, Z. M. & Bock, K. (1998). Constraint word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313-338.

Guenther, F. H. & Perkell, J. S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B.Maasen, R. Kent, H. Peters, P. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 29-49). Oxford: Oxford University Press.

Hageman, C. F. (1980). *Attentional mechanisms underlying patterns of auditory comprehension brain-damaged aphasic, non-aphasic, and normal listeners.* Unpublished doctoral dissertation, University of Colorado.

Hageman, C. F., McNeil, M. R., Rucci-Zimmer, S., & Cariski, D. M. (1982). The reliability of patterns of auditory processing deficits: Evidence from the Revised Token Test. In R.H.Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp. 230-234). Minneapolis, MN: BRK.

Head, H. (1926). *Aphasia and kindred disorders*. (vols. I & II) London: Cambridge University Press.

Himmanen, S. A., Gentles, K., & Sailor, K. (2003). Rated familiarity, visual complexity, and image agreement and their relation to naming difficulty for items from the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology, 25,* 1178-1185.

Howard, D., Patterson, K., Franklin, S., Morton, J., & Orchard-Lisle, V. (1984). Variability and consistency in naming by aphasic patients. *Advances in Neurology, 42,* 263-276.

Jescheniak, J. D. & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824-843.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. (2003). Specific-word frequency is not all that counts in speech production: comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 432-438.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6,* 365-369.

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

Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R.Parasuraman & D. R. Davies (Eds.), *Varieties of Attention* (pp. 29-62). Orlando, FL: Academic Press, Inc.

Kempen, G. & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185-209.

Kertesz, A. (1979). *Aphasia and associated disorders*. New York: Grune & Stratton.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. Journal of Memory and Language 47[1], 145-171.

Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language, 41,* 275-288.

Kleiss, J. A. & Lane, D. M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 200-210.

Knowles, W. B. (1963). Operator loading tasks. *Human Factors, 5,* 151-161.

Kreindler, A. & Fradis, A. (1968). *Performances in aphasia: A neurodynamical diagnostic and psychological study*. Paris: Gauthier-Villars.

Kucera, H. & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

La Heij, W., Puerta-Melguizo, C., van Oostrumm, M., & Starreveld, P. A. (1999). Picture naming: Identical priming and word frequency interact. *Acta Psychologica, 102,* 77-95.

LaPointe, L. L. & Erickson, R. J. (1991). Auditory vigilance during divided task attention in aphasic individuals. *Aphasiology, 5,* 511-520.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-38.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (2002). Picture naming and word frequency: Comments on Alario, Costa, and Caramazza, *Langauge and Cognitive Processes,* 17(3), 299-319. *Language and Cognitive Processes, 17,* 663-671.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: a study of picture naming. *Psychological Review, 98,* 122-142.

Lien, M.-C., Allen, P. A., Ruthruff, E., Grabbe, J., McCann, R. S., & Remington, R. W. (2006). Visual word recognition without central attention: evidence for greater automaticity with advancing age. *Psychology and Aging, 21,* 431-447.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behaviors. *Psychological Review, 89,* 483-506.

Marshall, J. C. (1986). The description and interpretation of aphasic language disorder. *Neuropsychologia, 24,* 5-24.

Matthews, G. & Margetts, I. (1991). Self-report arousal and divided attention: A study of performance operating characteristics. *Human Performance, 4,* 107-125.

McCann, R. S. & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14-24.

McCann, R. S. & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 471-484.

McCann, R. S., Remington, R. W., & Van Selst, M. (2000). A dual-task investigation of automaticity in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1352-1370.

McNeil, M. R. (1982). The nature of aphasia in adults. In N.J.Lass, L. V. McReynolds, J. L. Northern, & D. E. Yoder (Eds.), *Speech, Language, and Hearing (Vol. II): Pathologies of Speech and Language* (pp. 692-740). Philadelphia: Saunders.

McNeil, M. R. (1988). Aphasia in the adult. In N.J.Lass, L. V. McReynolds, J. Northern, & D. E. Yoder (Eds.), *Handbook of speech-language pathology and audiology* (pp. 738-786). Toronto: D.C. Becker, Inc.

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

McNeil, M. R. & Hageman, C. F. (1979). Auditory processing deficits in aphasia evidenced on the Revised Token Test: Incidence and prediction of across subtest and across item within subtest patterns. In R.H.Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp. 47-69). Minneapolis, MN: BRK.

McNeil, M. R. & Kimelman, M. D. Z. (1986). Toward an integrative information-processing structure of auditory comprehension and processing in adult aphasia. *Seminars in Speech and Language, 7,* 123-146.

McNeil, M. R., Odell, K., & Campbell, T. F. (1982). The frequency and amplitude of fluctuating auditory processing in aphasic and nonaphasic brain-damaged persons. In R.H.Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp. 220-229). Minneapolis, MN: BRK.

McNeil, M. R., Odell, K., & Tseng, C. H. (1991). Toward the integration of resource allocation into a general theory of aphasia. In T.E.Prescott (Ed.), *Clinical Aphasiology* (pp. 21-39). Austin, TX: Pro-Ed.

Meschyan, G. & Hernandez, A. (2002). Age of acquisition and word frequency: determinants of object-naming speed and accuracy. *Memory and Cognition, 30,* 262-269.

Miozzo, M. & Caramazza, A. (2005). The representation of homophones: evidence from the distractor-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1360-1371.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11,* 671-717.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica, 27,* 84-92.

Morrison, C. M. & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91,* 167-180.

Murray, L. L. (2000). The effects of varying attentional demands on the word retrieval skills of adults with aphasia, right hemisphere brain damage, or no brain damage. *Brain and Language, 72,* 40-72.

Murray, L. L., Holland, A. L., & Beeson, P. M. (1997a). Accuracy monitoring and task demand evaluation in aphasia. *Aphasiology, 11,* 401-414.

Murray, L. L., Holland, A. L., & Beeson, P. M. (1997b). Auditory processing in individuals with mild aphasia: a study of resource allocation. *Journal of Speech, Language, and Hearing Research, 40,* 792-808.

Navon, D. (1984). Resources- A theoretical soup stone? *Psychological Review, 91,* 216-234.

Navon, D. (1990). Exploring two methods for estimating performance tradeoff. *Bulletin of the Psychonomic Society, 28,* 155-157.

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

Navon, D. & Gopher, D. (1980). Task difficulty, resources, and dual-task performance. In R.S.Nickerson (Ed.), *Attention and Performance* (pp. 297-315). Hillsdale, NJ: Erlbaum.

Navon, D. & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology, 44,* 193-251.

Nicholas, L. & Brookshire, R. H. (1986). Consistency of the effects of rate of speech on brain-damaged adults' comprehension of narrative discourse. *Journal of Speech and Hearing Research, 29,* 462-470.

Ninio, A. & Kahneman, D. (1974). Reaction time in focused and in divided attention. *Journal of Experimental Psychology, 103,* 394-399.

Oldfield, R. C. & Wingfield, A. (1965). Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology, 17,* 273-281.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 358-377.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence of a two component theory of divided attention in simple tasks. *Cognitive Psychology, 21,* 469-514.

Pashler, H. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance, 16,* 826-842.

Pashler, H. (1991). Shifting visual attention and selecting motor responses: Distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 1023-1040.

Pashler, H. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220-244.

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance, 20,* 330-342.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Pashler, H. & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *The Quarterly Journal of Experimental Psychology, 41A,* 19-45.

Puleo, J. S. & Pastore, R. E. (1978). Critical-band effects in two-channel auditory signal detection. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 153-163.

Reynolds, M. & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1303-1323.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107-142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249-284.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starrveld and La Heij. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 246-251.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86-102.

Schuell, H., Jenkins, J. J., & Jimenez-Pabon, E. (1964). *Aphasia in adults: Diagnosis, prognosis, and treatment*. New York: Harper and Row.

Schwartz, M. F. (1984). What the classical aphasia categories can't do for us, and why. *Brain and Language, 21,* 3-8.

Schwartz, M. F., Saffran, E. M., Bloch, D. E., & Dell, G. S. (1994). Disordered speech production in aphasic and normal speakers. *Brain and Language, 47,* 52-88.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 876-900.

Shewan, C. M. (1976). Error patterns in auditory comprehension of adult aphasics. *Cortex, 12,* 336.

Shewan, C. M. & Canter, G. J. (1971). Effects of vocabulary, syntax, and sentence length on auditory comprehension in aphasic patients. *Cortex, 7,* 209-226.

Shuster, L. I. (2004). Forum: Resource theory and aphasia reconsidered: Why alternative theories can better guide our research. *Aphasiology, 18,* 811-854.

Siakaluk, P. D., Sears, C. R., & Lupker, S. J. (2002). Orthographic neighborhood effects in lexical decision:The effects of nonword orthographic neighborhood size. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 661-681.

Silkes, J., McNeil, M. R., & Drton, M. (2004). Simulation of aphasic naming performance in non-brain-damaged adults. *Journal of Speech, Language, and Hearing Research, 47,* 610-623.

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

Smith, M. C. (1969). The effect of varying information on the psychological refractory period. *Acta Psychologica, 30,* 220-231.

Stolz, J. A., Besner, D., & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Journal of Visual Cognition, 12,* 284-336.

Sullivan, M. P. & Macchi, C. (2002). A dual-task study of lemma selection in the picture-word interference task: Is competition strategically or automatically minimized? Paper presented to the annual meeting of the Cognitive Science Association for Interdisciplinary Learning, Hood River, OR.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G. et al. (2003). Timed picture naming: extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers, 35,* 621-633.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D. et al. (2004). A new online resource for psycholinguistic studies. *Journal of Memory and Language, 51,* 247-250.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14,* 1-35.

Tombu, M. & Jolicoeur, P. (2002a). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological Research, 66,* 274-286.

Tombu, M. & Jolicoeur, P. (2002b). Does size rescaling require central attention? *Canadian Journal of Experimental Psychology, 56,* 10-17.

Tombu, M. & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 3-18.

Tombu, M. & Jolicoeur, P. (2005). Testing the Predictions of the Central Capacity Sharing Model. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 790-802.

Treisman, A. & Fearnley, J. S. (1971). Can simultaneous speech be classified in parallel? *Perception and Psychophysics, 10,* 1-7.

Tseng, C. H., McNeil, M. R., & Milenkovic, P. (1993). An investigation of attention allocation deficits in aphasia. *Brain and Language, 45,* 276-296.

Valachovic, A. M., Smith, B., Elisevich, K., Jacobson, G., & Fisk, J. (1998). Language and its management in the surgical epilepsy patient. In A.F.Johnson & B. H. Jacobson (Eds.), *Medical speech-language pathology: A practitioner's guide* (pp. 425-464). New York: Thieme.

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M.R.McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 1-25). New York: Thieme.

Van Selst, M. & Jolicoeur, P. (1994). A solution to the effect of sample size and skew on outlier elimination. *Quarterly Journal of Experimental Psychology, 47A,* 631-650.

Vidulich, M. & Wickens, C. D. (1986). Causes of dissociation between subjective workload measures and performance: Caveats for the use of subjective assessments. *Applied Ergonomics, 17,* 291-296.

Vitkovich, M. & Tyrell, L. (1995). Sources of disgreement in object naming. *Quarterly Journal of Experimental Psychology, 48A,* 822-848.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance: A review and theory. *British Journal of Psychology, 43,* 2-19.

Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quarterly Journal of Experimental Psychology, 11,* 193-210.

Welford, A. T. (1967). Single-channel operation in the brain. *Acta Psychologica, 27,* 5-22.

Wheeldon, L. R. & Monsell, S. (1992). The locus of repetition priming of spoken word production. *The Quarterly Journal of Experimental Psychology, 44A,* 723-761.

Wickens, C. D. (1984). Processing resources in attention. In R.Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63-102). New York: Academic Press.

Wilson, S. M., Saygin, A. P., Schleicher, E., Dick, F., & Bates, E. (2003). Grammaticality judgment under non-optimal processing conditions: Deficits observed in normal participants resemble those observed in aphasic patients. Abstracts from the Academy of Aphasia [Special Issue]. Brain and Language 87, 67-68.

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *The American Journal of Psychology, 81,* 226-234.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science.

Zevin, J. D. & Seidenberg, M. S. (2002). Age of Acquisition Effects in Word Reading and Other Tasks. *Journal of Memory and Language, 47,* 1-29.

Zevin, J. D. & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: tests of cumulative frequency and frequency trajectory. *Memory and Cognition, 32,* 31-38.

Ziegler, J. C. & Perry, C. (1998). No more problems in Coltheart's neighborhood: Resolving neighborhood conflicts in the lexical decision task. *Cognition, 68,* B62.