Running head: SUBLEXICAL AND LEXICAL SPEECH PERCEPTION	1
1	
2	
3	
4	1
5	2
6	3
7	4
8	5
9	6
10	7
11	8
12	9
13	10
14	11
15	Evaluating the relationship between sublexical and lexical processing in speech
16	
17	perception: Evidence from aphasia
18	
19	Heather Dial and Randi Martin
20	
21	Rice University
22	
23	15
24	16
25	
26	17
27	
28	18
29	
30	19
31	Author Note
32	
33	20
34	Department of Psychology, Rice University, 6100 Main Street, Houston, TX 77005
35	
36	21
37	Corresponding author: Randi Martin, Department of Psychology, Rice University, 6100
38	
39	22
40	Main Street MS-25, Houston, TX 77005, Telephone: 713-348-3417, Email:
41	
42	23
43	rmartin@rice.edu
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	

Abstract

Several studies have reported that aphasic patients may perform substantially better on lexical than sublexical perception tasks (e.g., Miceli, Gainotti, Caltagirone, & Masullo, 1980). These findings challenge claims made by models of speech perception which assume obligatory sublexical processing (e.g., McClelland & Elman, 1986; Norris, 1994). However, prior studies have not closely matched the phonological similarity of targets and distractors or task demands of the sublexical and lexical perception tasks. The current study addressed shortcomings of these prior studies, testing 13 aphasic patients on sublexical and lexical tasks matched in phonological similarity of stimuli and task demands. When the lexical and sublexical tasks were not matched (Experiment 1a), as in prior studies (e.g., Miceli et al., 1980), several patients with impaired sublexical perception were within the control range on tasks tapping lexical perception. In contrast, when the lexical and sublexical tasks (sublexical: syllable discrimination, auditory-written syllable matching (AWSM); lexical: word discrimination, lexical decision, and picture-word matching (PWM)) were matched on these factors (Experiments 1b and 2), in most instances, patients were impaired on both sublexical and lexical tasks relative to controls and performance on the lexical tasks was not significantly greater than that on the sublexical tasks. For two patients, performance on one lexical task was statistically better than that on one sublexical task, but the advantage was not replicated across other task comparisons. The current study is consistent with models of speech perception which assume obligatory sublexical processing and fails to support models that do not require successful sublexical perception in order to access lexical levels (e.g., Goldinger, 1998; Hickok & Poeppel, 2000).

Keywords: sublexical, lexical, speech perception, discrimination, dual route, aphasia

1 Evaluating the relationship between sublexical and lexical processing in speech

2 perception: Evidence from aphasia

3 **1. Introduction** Does the perception of lexical information depend on the ability to
4 map acoustic input to abstract, speech-specific sublexical representations, such as
5 phonemes or syllables, or can acoustic information map directly to a lexical or semantic
6 level? Many well known models of speech perception have assumed the former - that is,
7 that speech sounds must pass through a language specific sublexical level in order to
8 access a lexical level (e.g., Cole & Scott, 1974; McClelland & Elman, 1986; Luce,
9 Goldinger, Auer & Vitevitch, 2000; Luce & Pisoni, 1998; Norris, 1994; Norris, McQueen
10 & Cutler, 2000; Oden & Massaro, 1978). However, other models of speech perception
11 advocate a direct mapping from acoustic to lexical or semantic levels (e.g., Goldinger,
12 1998; Hickok & Poeppel, 2000; Johnson, 1997; Pierrehumbert, 2001). Below we briefly
13 review models exemplifying the two positions.

14 TRACE (McClelland & Elman, 1986), Shortlist (and its successor MERGE;
15 Norris, 1994; Norris, McQueen & Cutler, 2000) and the Neighborhood Activation Model
16 (NAM; Luce & Pisoni, 1998) are among the most influential of models assuming
17 obligatory access to a sublexical level. These models all assume that spectrotemporal
18 analysis parses the auditory signal into abstract, speech-specific, sublexical units of
19 representation, such as phonemes or syllables. These sublexical level representations
20 are subsequently mapped onto a lexical level (Luce & Pisoni, 1998; McClelland &
21 Elman, 1986; Norris, 1994; Norris, McQueen & Cutler, 2000).

22 Although these models assume an obligatory sublexical processing stage, other
23 models assume a direct mapping from acoustic to lexical levels. One class of speech
24 perception models, known as exemplar models, allows for holistic processing of acoustic
25 representations of entire words (e.g., Goldinger, 1998; Johnson, 1997; Pierrehumbert,
26 2001), emphasizing data indicating the influence of specific acoustic features (e.g., pitch

of a speaker's voice) on perception. Recently, researchers have proposed dual route frameworks for speech perception which claim that a) there are separate sublexical and lexical routes¹ (e.g., Hickok & Poeppel, 2000; Hickok & Poeppel, 2004; Hickok & Poeppel, 2007; Hickok, 2014; Poeppel & Hickok, 2004; Majerus, 2013; Scott & Wise, 2004; Wise, Scott, Blank, Mummery, Murphy & Warburton, 2001) and b) the lexical route does not require access to sublexical representations for lexical perception (e.g., Hickok & Poeppel, 2000). For example, Vaden, Piquado and Hickok (2011) state that:

[sublexical] information is only represented on the motor side of speech processing and...[is] not explicitly extracted or represented as a part of spoken word recognition (p. 2672).

According to this argument, sublexical information, and accordingly the sublexical processing route, is only involved in speech perception when it is necessary for subsequent articulation, as in repeating a word. Further, although these two routes share initial processing of the acoustic features of the speech sounds, they subsequently diverge into two streams that operate largely independently up to the generation of articulatory codes in the dorsal route and accessing meaning in the ventral route. Subsequently, the routes converge and interact. Importantly, access to the lexical level via the ventral route does not necessarily require sublexical processing.²

One compelling piece of evidence supporting models with a direct mapping from acoustic to lexical levels is the finding that some brain damaged patients show much better performance on tasks tapping lexical than sublexical speech recognition (Basso,

¹ Further, some researchers have even argued that there are two lexical routes for perception with sublexical information only being processed post-lexically as needed (Gow, 2012). According to this model, patients with sublexical processing deficits would have damage to a post-lexical process, but lexical processing would proceed as normal.

² It is important to note that the exact claims regarding sublexical processing vary across papers. Hickok and Poeppel (2004) state that the dorsal and ventral routes share processing up through the sublexical processing stage (pg. 79-80), whereas Hickok and Poeppel (2000) argue for "sound-based representations of speech" (e.g., pg. 131) that are shared between the routes, leaving open the question of whether these sound-based representations correspond to linguistic units.

Casati & Vignolo, 1977; Blumstein, Baker & Goodglass, 1977; Blumstein, Cooper, Zurif & Caramazza, 1977; Miceli, Gainotti, Caltagirone, & Masullo, 1980). However, there are limitations to these studies which prevent one from drawing strong claims from the findings. For instance, consider the study carried out by Miceli et al. (1980), which is often cited as demonstrating this dissociation between sublexical and lexical perception. Miceli and colleagues had individuals with aphasia perform a sublexical processing task involving CCVC syllable discrimination (e.g., prin-brin; prin-trin) and two lexical processing tasks, involving picture-word matching (with distractor pictures that were semantically related, phonologically related or unrelated), and a sentence comprehension task (where patients acted out sentences such as "Put your hands on the table"). The researchers found that some patients were below the range of controls on CCVC discrimination but not on picture-word matching and sentence comprehension, and vice versa. Miceli et al. (1980) thus concluded that the relationship between the sublexical and lexical levels was more complicated than a simple mapping of phonemes to words. They argued that comprehension of auditorily presented words requires processes different from those involved in perceiving the constituent phonemes.

One means of accommodating such findings would be to argue that while sublexical processing is a necessary step in lexical access, contextual or top-down processes contribute to performance at the lexical but not the sublexical level. This notion is clearly instantiated in interactive speech perception models (e.g., TRACE, McClelland & Elman, 1986) in which information at one level can influence processing at other levels through feedback from higher to lower levels. Such contextual effects are also accommodated by models that do not incorporate feedback, under the assumption that information from different levels is combined in making a decision about perceptual identity (Massaro & Cohen, 1991; Norris, McQueen & Cutler, 2000). Thus, for example, a patient might perform better on lexical than sublexical tasks because of the activation

of semantic information for words (but not sublexical units) that feeds back to the lexical level and stabilizes lexical representations. Further, in more naturalistic tasks like sentence comprehension, meaning coherence or syntactic factors can restrict possible word recognition targets (Tyler & Marslen-Wilson, 1977). In contrast, such top-down contextual factors would play a minimal role in sublexical tasks like syllable discrimination. It seems unlikely, however, that such top-down effects could account for the findings from studies like that of Miceli et al. (1980), as the size of the discrepancy in performance on lexical vs. sublexical tasks was quite large. For example, Basso, Casati and Vignolo (1977) note that some patients have “severe” or “very severe” (as defined on p. 91) phoneme identification deficits “in spite of good comprehension” (p. 93). Findings like this (i.e., good comprehension with severe sublexical perception deficits) would be difficult to accommodate purely on the basis of top-down processing. That is, a severe deficit at the phoneme identification level would lead to weak activation of lexical representations with, in turn, weak activation of semantic information. Such weak semantic activation and weak feedback from a semantic level would not be expected to raise lexical perception to a high level. However, if such top-down processes can account for the dissociations observed in prior studies, then we would expect to see substantially better lexical than sublexical processing, even with stimuli that are closely matched in discriminability.

Upon closer inspection, however, it is clear that the tasks from the patient studies used to assess sublexical and lexical processing were not closely matched in terms of task demands or discriminability of the targets and distractors. The sublexical task required the ability to make fine-grained perceptual distinctions, such as discriminating “prin-brin”, where the two syllables differed in a single distinctive feature of one phoneme. However, such distinctions were not necessary to understand the commands from the sentence comprehension task (e.g., “Put your hands on the table”). Although

1 their picture-word matching task included phonologically related distractors, the
2 phonological lures differed from the target by one or more phonemes (picture-word
3 matching task described in detail in Gainotti, Caltagirone & Ibba, 1975) and when the
4 difference was only one phoneme, the phoneme might differ by more than one
5 distinctive feature from the target. Thus, poorer performance on the sublexical task may
6 have derived solely from the finer phonetic distinctions that were required.

7 Other studies that have observed similar dissociations between performance on
8 tasks tapping sublexical and lexical processing suffer from similar limitations. Basso,
9 Casati and Vignolo (1977) found that several patients with severe impairments on a
10 sublexical processing task requiring discrimination of consonants differing in voice onset
11 time (VOT) (e.g., /ba/ vs. /pa/, which are both bilabial stop consonants but differ in
12 voicing) were in the normal range on a task of auditory comprehension which required
13 patients to act out commands such as “Point to the green rectangle”. Along the same
14 lines, Blumstein, Cooper, Zurif and Caramazza (1977) found no relation between
15 patients’ performance on phoneme discrimination and identification tasks using stimuli
16 varied on VOT and performance on the auditory comprehension subtest of the Boston
17 Diagnostic Aphasia Battery (BDAE; Goodglass, Kaplan, & Barresi, 1972). Finally,
18 Blumstein, Baker and Goodglass (1977) found a similar effect comparing performance
19 on the discrimination of words differing in a single distinctive feature of one phoneme
20 (e.g., rate, raid) and degree of comprehension impairment on a subset of measures from
21 the BDAE (Goodglass, Kaplan, & Barresi, 1972) after excluding Broca’s aphasics who
22 performed at ceiling on both. In fact, the patients who had the most severe deficits on
23 the discrimination tasks had only moderate difficulty on the comprehension measure.
24 These research groups concluded that word comprehension can occur independently of
25 phonemic processing.

As with the Miceli et al. (1980) study, the studies by Basso et al. (1977), Blumstein et al. (1977a) and Blumstein et al. (1977b) failed to match perceptual discriminability of targets and distractors. In all three studies, the sublexical processing tasks required patients to discriminate or identify stimuli that differed by a single feature of a single phoneme. The auditory comprehension tasks, in contrast, did not require such fine-grained distinctions. Basso et al. (1977) and Blumstein et al. (1977a) required participants to act out commands such as “Point to the green rectangle” where the names of distractors items, such as a red rectangle and a green circle, were perceptually distinct from the target. The Blumstein et al. (1977b) study had a similar mismatch in the degree of perceptual discrimination demands of the phoneme and lexical perception tasks.

It should be noted that the lack of correlation between performance on tasks tapping phoneme perception and comprehension from these studies is due in part to some patients performing substantially worse on the comprehension than the phoneme perception measure. Impaired lexical processing in the face of preserved sublexical processing is not difficult to accommodate via models assuming obligatory sublexical processing, as one might assume patients showing this dissociation have preserved sublexical processing, but a disruption at subsequent levels involving lexical representations, semantic representations, or the connections between the two (e.g., Baker, Blumstein & Goodglass, 1981).

Task demands unrelated to speech perception abilities per se might also be the source of the observed dissociations. Some researchers have noted that the tasks used to tap sublexical processing require additional processes, such as short-term or working memory (STM/WM), and therefore may not be appropriate for assessing the role of sublexical processing in comprehension. For instance, syllable discrimination tasks typically involve perceiving and comparing two stimuli and thus may require STM/WM

processes that are not required in deciding if a word matches a picture (e.g., see Hickok & Poeppel, 2004, for discussion). Thus, patients performing poorly on discrimination tasks may do so due to STM/WM impairments. In line with this claim, Hickok and Poeppel (2007) point out that:

speech perception tasks involve some degree of executive control and working memory (p. 394)

In part because of the additional processes required by tasks used to tap sublexical processing (i.e., WM), Hickok and Poeppel (2007) claim that:

speech perception does not necessarily correlate with, or predict, speech recognition (p. 394).

Note that Hickok and Poeppel (2007) define speech perception as tasks tapping sublexical processing (e.g., phoneme discrimination) and speech recognition as tasks tapping lexical processing (e.g., picture-word matching).

2. Current Study

The current study sought to examine whether a dissociation would be observed between impaired sublexical but preserved lexical speech perception for aphasic patients with varying degrees of speech perception deficits, once perceptual discriminability and task factors were controlled for between the sublexical and lexical tasks. Critically, we wanted to determine whether any patient would exhibit a dissociation between performance on sublexical and lexical processing tasks such that they were significantly worse at tasks tapping sublexical compared to lexical processing. As indicated by the foregoing discussion, it is important to verify the existence of this dissociation, given the implications for speech perception models. If lexical perception depends on sublexical perception, as claimed by models assuming obligatory sublexical processing, then one would expect to see a close relation between performance on sublexical and lexical tasks once discriminability and task demands, such as STM/WM

demands and the presence of contextual cues, have been carefully controlled. Further, and most importantly, sublexical processing should influence lexical processing such that patients with sublexical processing deficits would necessarily have lexical processing deficits as well, as evidenced by performance on tasks tapping sublexical and lexical processing. On the other hand, if lexical perception does not depend on the identification of sublexical speech units, as is argued in models assuming a direct association between acoustic and lexical representations, then dissociations may still persist with some patients showing substantially better performance on tasks tapping lexical than sublexical speech perception. In other words, if acoustic information maps directly onto lexical representations without an intermediate sublexical processing stage, then words could be accurately perceived (even when they must be discriminated from phonologically similar words) despite impaired sublexical processing.

The present study took a case series approach in which 13 aphasic patients with varying degrees of speech perception deficits were tested on several tasks designed to assess sublexical or lexical processing. Multiple experiments were used in order to assess the influence of various factors like STM/WM demands or top-down influences and to determine whether evidence would converge on showing that impaired sublexical perception necessarily leads to impaired lexical perception, once these factors were controlled, or instead substantiate the dissociations assumed previously in the literature.

In Experiment 1a, we present data on patients' performance on sublexical and lexical tasks where the stimuli were not closely matched on discriminability and tasks demands in order to demonstrate that we could replicate prior results in the literature under these conditions. In Experiment 1b, we used closely matched stimuli to compare sublexical and lexical processing. In Experiment 1b, participants performed syllable (sublexical) and word (lexical) discrimination tasks where the pairs of items differed by single distinctive feature of a single phoneme and an auditory lexical decision task

(lexical) where nonwords differed from word stimuli by a single distinctive feature of a single phoneme.

To provide converging evidence, Experiment 2 utilized different sublexical and lexical processing tasks, where the lexical task required semantic access. We created a single picture-word matching (PWM) task using pairs of auditory stimuli that differed by a single distinctive feature of a single phoneme (e.g., bear-pear) to assess lexical perception. In order to provide a sublexical task involving similar processes, we created an auditory-written syllable matching (AWSM) task using pairs of stop-consonant syllables that differed by a single distinctive feature of the initial phoneme (e.g., spoken “BA”- written “PA”).

We predicted that if models assuming obligatory sublexical processing were correct, then impairments on sublexical tasks should lead to impairments on lexical tasks. However, a patient might be worse at the lexical tasks due to a lexical or semantic level deficit. Alternatively, if sublexical processing is not a prerequisite for lexical processing, then patients could be better at tasks tapping lexical processing than sublexical processing due to a direct mapping from acoustics to the lexical level (e.g., Goldinger, 1998; Hickok & Poeppel, 2000; Johnson, 1997; Pierrehumbert, 2001).

2.1 General Methods

2.1.1 Participants

Both patients and controls were recruited from existing participant databases in the Brain and Language lab at Rice University and informed consent was obtained in accord with the Rice University IRB regulations. The controls and all but two patients were screened for hearing loss using an audiometer. The two exceptions were not screened during the session of experimental testing and were unavailable for further testing, though it should be noted that one wore hearing aids. Two of the patients who were screened wore hearing aids, but the hearing test was administered without their

hearing aids, providing a conservative measure of hearing thresholds. To provide a measure of hearing loss, we compared patients and controls to international standards for median hearing thresholds as a function of age (International Organization for Standardization (ISO) 7029, 2000). We computed for each patient and control the standard score of their hearing threshold relative to the median and standard deviation from these age-adjusted norms at 500, 1000, 2000, 3000 and 4,000 Hz (see Dalton, Cruickshanks, Klein, Klein, Wiley & Nondahl, 2003) for the left and right ears. We then averaged the resulting numbers across the two ears to obtain a single estimate of hearing loss, with positive numbers indicating hearing thresholds greater than the normative value and negative numbers hearing thresholds less than the normative value. For our 13 patients and 11 controls, the means of these hearing loss values were very close, with both groups having slightly positive means (.25 for the controls and .23 for the patients). In terms of distribution, the standardized hearing threshold values ranged from -.66 to 1.15 for controls and from -1.54 to 2.11 for patients. The one patient who had a value greater than 2.0 standard deviations above the normative value was one of the patients who wore hearing aids during experimental testing but completed the audiometric testing without the hearing aids. Aside from this patient, the greatest hearing loss value for any patient was .73, clearly within the range of our controls. Thus, hearing loss should not be an important contributing factor in any observed relations between sublexical and lexical perception nor in any differences between the results for patients vs. controls. Subjects were compensated \$10 per hour for participation.

Control Subjects. Control data was collected from 11 self-reported neurally healthy age- and education-matched older adults with normal or corrected to normal hearing. The mean age of the participants was 75 years (range 63-79 years) and the mean number of years of education was 16 (range 12-19 years).

Patients. Thirteen individuals with aphasia diagnosis following left-hemisphere stroke participated in the current study. Their mean age was 67 years (range 46 - 89) and the mean number of years of education was 15.4 (range 11-22). Patients were recruited from the Brain and Language Lab database. There were no selection criteria other than left hemisphere stroke with aphasia diagnosis, as we wanted a range of performance on the tasks. Table 1 contains information regarding age, years post-stroke and education for the patients. Appendix A contains more detailed information on the patients.

Table 1

Patients' Age, Years Post-Stroke and Education.

<i>Patient</i>	<i>Age</i>	<i>Years Post-Stroke</i>	<i>Years Of Education</i>
BQ	74	15	16
DZ	46	10	22
DG	54	7	17
EV	58	15	16
HA	89	9	16
KA	58	2	12
MB	67	9	13
NMA	57	6	14
QO	69	4	20
SH	87	10	11
SJ	67	9	13
SQ	74	6	14
SSW	69	2	16

2.1.2 Apparatus

All of the stimuli for the following tasks were administered on an iMac desktop computer running PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993) or via a paper/pencil test book. Auditory stimuli were presented via speakers with volume adjusted to a comfortable level for each participant. Errors were coded using raw data output from Psyscope.

2.1.3 Procedure

All of the tasks were presented across at least four testing sessions over the course of several months for each of the patients. A detailed description of the materials, procedure and results for each of the tasks follows.

3. Experiment 1

The purpose of Experiment 1 was to determine whether the observed double dissociations between performance on tasks tapping sublexical and lexical processing were true double dissociations or artifacts of unmatched stimuli and task demands. First, we wanted to ensure that we would observe this double dissociation in our patient sample (or at least observe some patients with worse performance on sublexical than lexical tasks) when using tasks like those in prior studies. This was important because if we did not observe this pattern in unmatched tasks, no conclusions could be drawn from matched tasks. In other words, a lack of the dissociation of worse sublexical than lexical performance on unmatched tasks, and a similar pattern in matched tasks, could simply indicate that our patient sample does not present the same pattern as in prior research. As such, Experiment 1a sought to replicate the findings of dissociations between sublexical and lexical processing using unmatched stimuli and tasks. After observing this dissociation, we wanted to determine whether it would hold with matched stimuli and task demands. If sublexical processing is not required for lexical processing, then even with matched stimuli and task demands, a patient might exhibit significantly better performance on tasks tapping lexical than sublexical processing. This was the purpose of Experiment 1b.

3.1 Experiment 1a: Sublexical and Lexical Perception with Unmatched Stimuli

3.1.1 Methods. Participants completed three tasks: consonant discrimination, the WAB Auditory Verbal Comprehension test (Kertesz, 1982), and a picture-word matching

task. All three tasks were presented in separate sessions as part of a standard battery given to all patients in the Brain and Language Lab.

Materials/Procedure.

Consonant discrimination. Sublexical processing was assessed using a consonant discrimination task where participants were presented with pairs of consonant-vowel (CV) and vowel-consonant (VC) syllables and were required to determine whether the two items were the same or different. For the nonmatching pairs, stimuli differed by a single distinctive feature, (e.g., ba-pa, ap-ab), half in the initial consonant (i.e., CV syllables) and half in the final consonant (i.e., VC syllables). The task was blocked into a CV and a VC block, each with 54 trials (22 matching and 32 nonmatching). For the nonmatching trials, one-third differed in manner, one-third differed in place and one-third differed in voicing. Each auditory stimulus was spoken by a female voice and pairs of stimuli were presented with a 750 ms delay between the two. This task was presented via PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993), with subject responding with a button press.

WAB Auditory Verbal Comprehension. Patients completed the Auditory Verbal Comprehension section of the Western Aphasia Battery (WAB; Kertesz, 1982) using standard administration procedures. This test contains yes/no questions (e.g., "Does it snow in July?"), an auditory word recognition subtest, and sequential commands. For the auditory word recognition subtest, participants hear words from a given category (i.e., objects, shapes, letters, numbers and colors) and point to the correct item from a set of 6 items in the same category. They also point to objects in the room, again blocked by category (i.e., furniture, body parts, fingers, right-left side of the body). Note that for each of these categories distractors were phonologically unrelated (e.g., for objects: flower, cup, matches, pencil, comb, screwdriver). Finally, in the Sequential Commands subtest they act out commands such as "Point to the chair" and "With the

book point to the comb.” While the Sequential Commands subtest requires abilities beyond lexical access, comprehension of the words in the command is required for good performance.

Single picture-word matching. To provide an additional measure of lexical processing, participants were tested on a PWM task (Martin, Lesch, & Bartha, 1999) where they judged if a spoken word matched a single picture. This task consisted of four sessions where the individual was presented with 54 pictures and asked “Is this a _____?” Subjects responded yes or no for each trial. Each session contained an equal number of correct, unrelated, semantically related and phonologically related trials. Importantly, on the phonologically related distractors trials, the word typically differed by more than one feature of a single phoneme (e.g., book-hook).³

3.1.2 Results and discussion. The data for the three tasks are shown in Table 2. For the PWM task, results are shown both for overall percent correct and for just the trials with phonologically related distractors. As can be seen in Table 2, using tasks and stimuli similar to those in the studies reported in the introduction, some patients performed substantially better at the lexical tasks than the sublexical task (e.g., KA, NMA, SQ). These dissociations were confirmed using Crawford and Howell’s (1998) modified *t*-test for the sublexical and lexical tasks. For the sublexical consonant discrimination task, six of our patients (DZ, HA, KA, NMA, SH, SQ) were significantly impaired relative to age-matched controls (p ’s < .05). For the overall WAB Auditory Verbal Comprehension Test, using norms from 53 individuals with stroke in the non-dominant hemisphere (Kertesz, 2006), three of these patients scored within the control range on the WAB test, with one (SQ) scoring above the control mean. Norms are not available separately for the Auditory Word Recognition subtest. However, we would note that KA, who scored below the control range on the overall WAB test, scored 97%

³ Note that this single PWM task is different than the single PWM created for Experiment 1b.

correct for the word recognition subtest, despite chance performance on consonant discrimination. For the picture-word matching task from our lab, neurally healthy controls score at ceiling. Although none of our patients with impaired consonant discrimination obtained a perfect score, SQ scored 100% for the trials with phonologically related distractors and was close to ceiling for the test overall (99%). Taken together, these data clearly show that some of our patients performed substantially worse on the sublexical than lexical tasks (i.e., NMA, SH and SQ). SQ in particular showed a striking dissociation, with severely impaired performance on consonant discrimination (more than eight standard deviations below the control mean) but normal performance on WAB comprehension and 100% correct on PWM matching for phonologically related foils. Thus, when using unmatched stimuli, we were able replicate dissociations reported previously in the literature (Blumstein et al., 1977a; Blumstein et al., 1977b; Basso et al., 1977; Miceli et al., 1980).

Table 2
Patients' Performance (Percent Correct) on Consonant Discrimination, WAB (Auditory Verbal Comprehension and Auditory Word Recognition) and Single PWM (Overall and Phonological Foils). Asterisks represent significantly impaired performance relative to controls ($p < .05$). Control means and standard deviations are presented where this information is available. Note that the WAB control data reflect patient controls whereas the other control data come from neurally healthy age-matched controls.

Patient	Consonant Discrimination	WAB Auditory Verbal Comprehension	WAB Auditory Word Recognition	Single PWM Overall	Single PWM Phonological Foils
BQ	90	17.6	85	94	98
DZ	85*	13.6*	72	86	96
DG	97	16.9	75	88	98
EV	96	17	95	95	96
HA	79*	13.2*	67	85	89
KA	52*	11.7*	97	86	69
MB	93	20	100	98	98
NMA	85*	16.8	98	97	98
QO	94	19.6	100	100	100
SH	82*	18.4	97	95	94
SJ	95	18.8	98	97	98
SQ	76*	19.1	98	99	100

SSW	87	19.2	98	97	100
Control					
Means	93 (.02)	18.7 (2.1)	N/A	100 (0)	0 (0)

3.2 Experiment 1b: Syllable and Word Perception with Closely Matched Stimuli

In Experiment 1a, which used unmatched sublexical and lexical tasks, there were several patients who showed better performance on tasks tapping lexical than sublexical processing. Having observed these dissociations, we sought to determine whether they would be obtained using closely matched stimuli and tasks. Specifically, we aimed to determine if patients who were impaired in discriminating syllables differing in a single distinctive feature might yet show a preserved ability to discriminate words differing in a single distinctive feature. Such a dissociation should be possible if words are accessed directly from acoustic features without going through an intermediate sublexical level of representation. In Experiment 1b, participants performed a sublexical (syllable discrimination) and two lexical tasks (word discrimination, auditory lexical decision) using stimuli that were closely matched in terms of discriminability. For the two discrimination tasks, the members of a stimulus pair were spoken by different voices (one male and one female) in order to prevent the use of a raw acoustic code in making the discrimination. Due to the STM/WM requirements of the discrimination task, we employed the auditory lexical decision task as an additional lexical processing task to determine if the relation between sublexical and lexical tasks would be minimized when the lexical task had few STM/WM demands. We predicted that if lexical processing depends on sublexical processing, then performance on lexical tasks should be restricted by sublexical task impairments. Alternatively, if lexical processing does not depend on an intermediate sublexical processing stage, then discrimination of lexical items differing by a single feature could be intact despite impaired sublexical processing.

3.2.1 Materials.

Word discrimination. Word stimuli consisted of 90 pairs of single syllable words two to four phonemes in length ($M=3.06$) that differed from each other by a single distinctive feature (i.e., manner, place or voicing) in either the initial or final consonant (e.g., pat-bat, bat-bad). Half of the items differed in the initial consonant and half differed in the final consonant. Appendix C presents the 90 “different” stimulus pairs. For the “different” stimulus pairs, roughly a third of the stimuli differed in manner ($n=32$), a third in place ($n=29$) and a third in voicing ($n=29$), and these were equally distributed between the initial and final consonant. The task consisted of 180 trials, half of which were the same and half of which were different. The 90 “same” trials were created by randomly selecting one of the items from the 90 “different” stimulus pairs.

The stimuli were matched in terms of phonotactic probability (i.e., the frequency of segments or sequences of segments in syllables and words) such that the mean phonotactic probability for items differing in the initial consonant was .06 ($SD=.02$) and for items differing in the final consonant was .05 ($SD=.01$). Further, the difference in the mean phoneme probability between pairs of items was effectively zero (overall: $M=.0038$, $SD=.013$; initial consonant difference: $M=.0068$, $SD=.014$; final consonant difference: $M=.0008$, $SD=.011$). Note that it is important to match phonotactic probability as high phonotactic probability has been shown to facilitate perception of speech, at least at the sublexical level (e.g., Vitevitch & Luce, 1999). All of the items are presented in Appendix B along with the following information: number of phonemes, frequency (Kucera & Francis, 1967), phonological neighborhood density, phonotactic probability and biphone frequency. The information presented in Appendix B was acquired using the item search tool from the Washington University Speech and Hearing Lab Neighborhood Database (Sommers, neighborhoodsearch.wustl.edu).

The stimuli were presented using natural speech tokens with one item spoken by a male and one item spoken by a female to avoid use of raw acoustic code in

discrimination judgments. This manipulation was critical to ensure that participants were comparing abstract representations rather than raw acoustic information to make the discrimination judgments. The two stimuli in each pair were separated by either a 50ms or 1500ms inter-stimulus interval (ISI). The long ISI was included because a previous study (Wolmetz & Rapp, 2011) suggested that some patients show refractory effects in speech perception, where processing of a stimulus leads to degraded or blocked access to a similar representation. Inclusion of the 1500ms condition would reduce the likelihood of refractory effects but place a greater demand on STM while the 50ms condition would reduce STM demands while increasing the likelihood of refractory effects. The same items were used for each ISI condition but counterbalanced across sessions so that the same items were not presented more than once during a single session. Further, none of the 90 “same” or 90 “different” trials repeated within a participant.

Syllable discrimination. CV and VC syllables were created using the stimuli from the single feature difference word discrimination task by removing either the initial or final phoneme (e.g., /pæ/-/bæ/), thus creating 90 stimulus pairs. Half of the items differed in the initial consonant and half differed in the final consonant. The task consisted of 180 trials, half of which were the same and half of which were different. The 90 “same” trials were created by randomly selecting an item from the 90 “different” stimulus pairs.

The stimuli were presented using natural speech tokens with one item spoken by a male and one item spoken by a female to avoid use of a raw acoustic code in discrimination judgments. As with the word discrimination task, stimulus presentation was separated by either a 50ms or 1500ms ISI.

Auditory lexical decision. Stimuli for the auditory lexical decision task were created using 149 randomly selected stimuli from the word discrimination task. These stimuli served as words in the lexical decision task. To create the non-word stimuli, we

changed a single distinctive feature of a single phoneme of either the initial or final consonant of the word stimuli (e.g., bat-bap), thus totaling 298 trials, with half created by changing the initial consonant of a word stimulus and half created by changing the final consonant of a word stimulus. Of the 149 word stimuli, there were three words that required the phoneme to be changed by two distinctive features in order to create a non-word, but the mean number of features changed within a word to create a non-word was essentially one ($M=1.02$). The word and non-word stimuli were presented using natural speech tokens spoken by a female.

3.2.2 Procedure. Because there were two ISI conditions (i.e., 50ms and 1500ms) for the word and syllable discrimination tasks, the tasks were administered across two testing sessions to avoid repetition of stimuli within a single session. In the first testing session, participants performed the word and syllable discrimination tasks, beginning with the word discrimination task with 1500ms ISI. They were told that they would hear a word followed by a pause and a second word, and that their task was to determine whether the two words were the same or different. They were further told that one of the words would be spoken by a male and one by a female but that this should not influence their decision. Finally, they were told that half of the trials were the same and half were different. In all tasks, they were instructed to press one of two buttons that corresponded to responses of “same” or “different”. For all discrimination tasks, they began with five practice trials to ensure that they understood the instructions and subsequently completed 90 experimental trials. Next they performed the syllable discrimination task with 1500ms ISI. They were told that they would be performing a task that was similar to the task they had just completed except that this time they would hear syllables or pieces of words rather than complete words. After completing this task, participants performed the syllable discrimination task with 50ms ISI, followed by the

word discrimination task with 1500 ms ISI. They were told that they would be performing the same type of tasks but with a shorter pause between the stimuli.

In the second testing session, participants performed the same series of tasks in reverse order using the stimuli not presented in session 1. Instructions and practice trials for each task were the same as session 1. After completing the discrimination tasks, participants performed the auditory lexical decision task. In this task, participants were told that they would hear a series of words and nonsense words and that their task was to press a button indicating whether each stimulus was a word or nonsense word. They were told that half of the items were real words and half of the items were nonsense words. They completed six practice trials to ensure they understood the task before completing 298 experimental trials. For the experimental trials, 151 were presented followed by a pause where they could take a break. When they felt ready to continue, they pressed the space bar to complete the additional 147 trials of the task.

3.2.3 Results and discussion. For the discrimination and lexical decision tasks, analyses used d' to remove response bias. For the discrimination tasks, we defined hit rate as the proportion correct on “same” trials and the false alarm rate as the proportion incorrect on “different” trials. For the lexical decision task, we defined hit rate as the proportion of words correctly accepted as words and false alarm rate as the proportion of non-words incorrectly accepted as words. Given that d' values are distorted when the hit rate or false alarm rate is 1 or 0, we adjusted these values using the method described by Stanislaw and Todorov (1999). For rates of 1, the value was replaced by $(n-.5)/n$, and for rates of 0, the value was replaced by $.5/n$, where n is the number of trials going into the computation of hits and false alarms.

Table 3 presents the data from age-matched control participants ($n=11$)⁴, collapsed across the 50ms and 1500ms ISI's for word and syllable discrimination. Paired samples t -tests revealed no difference between control performance in the 50ms compared to 1500ms ISI for word discrimination, $t(10)=-.51$, $p=.62$, or syllable discrimination, $t(10)=-1.01$, $p=.34$, or when we computed the mean for each ISI collapsed across syllable and word discrimination, $t(10)=.82$, $p=.43$.

Table 3.

Mean performance and range of age-matched controls on single feature difference discrimination tasks and auditory lexical decision task.

Task	Mean	d' scores	
		Minimum	Maximum
Word Discrimination	2.44	1.59	2.78
Syllable Discrimination	2.68	1.92	3.52
Lexical Decision	2.44	1.46	3.17

The patients showed a similar pattern, with no significant difference between 50 and 1500 ms ISI's. Using paired-samples t -tests, this was true for syllable discrimination, $t(12)= -.50$, $p=.63$, and word discrimination, $t(12)=.21$, $p=.84$, and when we computed the mean for each ISI collapsed across syllable and word discrimination, $t(12)=.20$, $p=.85$ ⁵. Accordingly, subsequent analyses used the mean across the two ISI's. Table 4 presents the patient data collapsed across the ISI's as well as the auditory lexical decision data.

Table 4.

Patient performance on the single feature discrimination tasks collapsed across 50ms and 1500ms ISI's and the auditory lexical decision task. For patients with better lexical than sublexical task performance, bootstrapping values for the mean difference between lexical and sublexical task is presented, with 95% CI's of the difference included in brackets.

⁴ Note that there were 21 trials overall for which at least four of our initial five control subjects made errors. As such, those trials were excluded from further analyses, thus leaving 172 trials for word discrimination with 50ms ISI, 174 trials for word discrimination with 1500ms ISI, 175 trials for syllable discrimination with 50ms ISI and 178 trials for syllable discrimination with 1500ms ISI.

⁵ Looking at the individual patient data, it appeared that DZ exhibited worse performance on 50 vs. 1500 ms ISI (refractory access deficit) whereas NMA exhibited the opposite pattern (STM deficit). As such, we investigated whether the results of the correlations would differ if we looked at 50 and 1500 ms ISI's separately. However, correlations across the entire group remained similar, so all analyses use data that were averaged across the two ISI's.

<i>Patient</i>	<i>Syllable Discrimination</i>	<i>Word Discrimination</i>	<i>Difference (Word-Syllable Discrimination)</i>	<i>Auditory Lexical Decision</i>	<i>Difference (Lexical Decision-Syllable Discrimination)</i>
BQ	1.45	1.54	.09 [-.53, .35]	.75	
DZ	1.86	1.74		.66	
DG	2.09	2.15	.07 [-.41, .53]	2.38	.33 [-.17, .83]
EV	2.45	2.76	.29 [-.29, .91]	1.84	
HA	.93	.87		.77	
KA	.56	.74	.20 [-.22, .60]	.23	
MB	1.11	1.38	.25 [-.17, .69]	1.00	
NMA	1.98	2.04	.06 [-.50, .60]	1.87	
QO	2.46	2.59	.11 [-.43, .65]	1.96	
SH	1.72	1.56		.21	
SJ	2.10	1.82		1.62	
SQ	2.21	2.35	.12 [-.38, .63]	1.50	
SSW	2.07	1.94		1.84	
Means	1.77	1.81		1.28	

Mean performance on word discrimination and syllable discrimination did not differ significantly for controls ($t(10) = 1.98$, $p = .08$) or patients ($t(12) = -.79$, $p = .45$). Mean performance on auditory lexical decision and syllable discrimination did not differ significantly for controls ($t(10) = 1.77$, $p = .11$), whereas patients performed significantly better on syllable discrimination than auditory lexical decision ($t(12) = 3.67$, $p = .003$). As is evident in Figures 1-3, for patients there were high correlations between syllable and word discrimination.

Although our focus was on individual patients' performance on the sublexical vs. lexical tasks, we assessed the significance of the correlations between the sublexical and lexical tasks across patients, as some prior studies had reported non-significant correlations between tasks tapping sublexical and lexical levels (e.g., Blumstein et al., 1977a). Thus, our aim was simply to assess whether correlations would be obtained when using tasks matched on stimulus discriminability and task demands. Of course, these correlational results should be treated with great caution, given the small sample sizes in our experiments. To assess the overall relation between performance on the

three tasks, Pearson correlation coefficients were calculated; however, given the small sample sizes, we used a bootstrapping procedure to obtain an estimate of the 95% confidence interval (CI) of the correlation coefficient ($n=1000$ bootstrapped samples) rather than the standard parametric method of computing significance, as the former makes no assumptions about the form of the distributions. To calculate the bootstrapped 95% CI of the correlation coefficient, pairs of data (i.e., a single participant's score on tasks X and Y) were randomly sampled with replacement. Each bootstrapped sample had the same number of data points as the original sample. The results for this experiment and subsequent ones are reported in terms of the Pearson r together with a 95% CI determined from the bootstrapping procedure. For Experiment 1b, the results were: 1) syllable discrimination and word discrimination: $r = .96$, 95% CI [.90, .99], 2) syllable discrimination and auditory lexical decision: $r = .74$, 95% CI [.56, .93], and 3) word discrimination and auditory lexical decision: $r = .77$, 95% CI [.63, .90]⁶.

Although the group results indicated that mean performance was not better for the lexical tasks than the sublexical tasks (and, in fact, was significantly worse on auditory lexical decision than syllable discrimination), it is critical to show that this pattern held for the individual patient data (this logic holds for all subsequent experiments as well). Accordingly, for patients with d' values that were larger on word discrimination or auditory lexical decision than on syllable discrimination, we utilized a bootstrapping technique ($n=10,000$) to determine the 95% CI of the difference in d' values between the sublexical and lexical task. Bootstrapping was necessary as there is no prescribed

⁶ Even though all but one patient's hearing thresholds were within normal range, in order to ensure that correlations could not be attributed to hearing loss we computed partial correlations between sublexical and lexical perception, partialling out patient's standard scores for hearing thresholds, with the results reported in Appendix C. As shown there, all correlations that were significant without the covariate (i.e., with the 95% CI exceeding 0) remained high and significant using the bootstrapping procedure, except for correlations with auditory-written word matching (Exp. 2). Note, however, that the sample size for these comparisons was quite small. The same results were obtained using a covariate that used raw thresholds rather than age-adjusted scores.

parametric statistical test for determining at an individual subject level, the 95% CI for differences in d' between two tests (e.g., d' for syllable discrimination vs. d' for word discrimination or auditory lexical decision). In this case, bootstrapping involved using random sampling with replacement of the sample data on a trial-by-trial basis, using probabilities of hits ("yes" trials) and false alarms ("no" trials) to create datasets with similar probabilities. This procedure was carried out 10,000 times to create a bootstrapped distribution of d' values for each task. Finally, we calculated the difference between the d' values for each task, giving us $n=10,000$ difference scores, which were then used to determine the 95% CI of the observed difference in d' values.

There were eight patients whose word discrimination d' values were slightly better than their syllable discrimination d' values. For all eight patients, the 95% CI for the difference between their d' value for syllable vs. word discrimination included zero (see Table 4), indicating that there was no significant difference between performance on the lexical and sublexical processing tasks (p 's from .08 to .77). In this table and in others showing the bootstrapping results, bootstrapped values are not presented for patients with better sublexical than lexical performance as this pattern of performance is compatible with both obligatory sublexical processing and dual route models and is thus not useful in differentiating between the two. There was only one patient (DG) whose performance was better on auditory lexical decision than syllable discrimination, but bootstrapping revealed that the 95% CI for the difference between the two tasks included zero, again indicating that the two d' values were not significantly different from each other, $p = .20$. In other words, critically, no patient performed substantially better on the word discrimination and auditory lexical decision tasks than on the syllable discrimination task (see Figures 1 and 2).

Figure 1. Scatterplot comparing patient performance (d') on the syllable and word discrimination tasks. The red dotted line indicates $y=x$. Note that all data points hovered around $y=x$, indicating roughly equal performance on the two tasks..

Figure 2. Scatterplot comparing patient performance (d') on the syllable discrimination task and auditory lexical decision task. The red dotted line indicates $y=x$. Note that all data points either hovered around $y=x$ or were below the line, indicating better performance on the syllable discrimination task or roughly equal performance on the two tasks.

Figure 3. Scatterplot comparing patient performance (d') on the word discrimination task and auditory lexical decision task. The red dotted line indicates $y=x$.

In addition, patient data was compared to control data using the Crawford and Howell (1998) modified t -test for comparing a single participant's data to a small group (see Table 5). This was done as the control mean d' for the word discrimination task was lower than that for the syllable discrimination task, although not significantly different ($p=.08$), as reported above. By comparing patients to controls, then, we can determine whether a patient is impaired on these tasks while taking into account task difficulty. Importantly, there was no patient who was significantly impaired relative to controls on the sublexical processing task who had normal performance on both lexical processing tasks.

Table 5.

Results of Crawford and Howell's (1998) modified t -tests for evaluating differences between a single participant and a small group of participants. Each patient was compared to the control group. Negative t -values indicate worse performance relative to controls.

	Syllable	Discrimination	Word	Discrimination	Auditory	Lexical Decision
Patient	t	p	t	p	t	p
BQ	-2.46	.03	-2.48	.03	-2.93	.02
DZ	-1.62	.14	-1.92	.08	-3.08	.01
DG	-1.17	.27	-.81	.44	-.10	.92

EV	-.45	.66	.86	.41	-1.04	.32
HA	-3.48	.01	-4.30	<.001	-2.89	.02
KA	-4.23	<.001	-4.66	<.001	-3.82	<.001
MB	-3.13	.01	-2.92	.02	-2.49	.03
NMA	-1.40	.19	-1.10	.30	-.98	.35
QO	-.44	.67	.40	.70	-.83	.43
SH	-1.92	.08	-2.43	.04	-3.86	<.001
SJ	-1.16	.27	-1.70	.12	-1.42	.19
SQ	-.93	.37	-.25	.81	-1.62	.14
SSW	-1.21	.25	-1.38	.20	-1.03	.33

Finally, we determined whether any patient scored substantially better on the lexical tasks than sublexical tasks when patient scores were converted to z-scores based on control performance, which would take into account any differences in difficulty between the two tasks. Table 6 presents the z-scores for the patients on the syllable and word discrimination tasks and the auditory lexical decision task.

Table 6.
Z-scores for each patient compared to the control group for syllable discrimination, word discrimination and auditory lexical decision. More negative values indicate worse performance relative to controls.

Patient	Syllable Discrimination	Word Discrimination	Difference (Word-Syllable Discrimination)	Auditory Lexical Decision	Difference (Lexical Decision-Syllable Discrimination)
BQ	-2.57	-2.59		-3.06	
DZ	-1.70	-2.01		-3.22	
DG	-1.23	-.84	-.41 [-1.59, .76]	-.10	-1.19 [-2.15, -.22]
EV	-.48	.90	-1.30 [-2.94, .13]	-1.09	
HA	-3.64	-4.49		-3.02	-.62 [-1.40, .18]
KA	-4.41	-4.87		-3.99	-.48 [-1.28, .28]
MB	-3.27	-3.05	-.21 [-1.32, .86]	-2.60	-.69 [-1.48, .14]
NMA	-1.46	-1.15	-.34 [-1.73, 1.02]	-1.03	-.39 [-1.39, .69]
QO	-.46	.42	-.83 [-2.23, .50]	-.86	
SH	-2.00	-2.54		-4.03	
SJ	-1.22	-1.78		-1.48	
SQ	-.98	-.26	-.69 [-1.99, .54]	-1.69	
SSW	-1.27	-1.44		-1.07	-.17 [-1.10, .74]

The majority of patients were more impaired on the lexical tasks, as indexed by a more negative z-score. For those patients who performed better on the lexical tasks (for whom difference scores are reported in Table 6), we utilized the same bootstrapping technique as before, converted each score in the distribution to a z-score, and determined the 95% CI for the difference in z-scores for the sublexical and lexical tasks. For syllable and word discrimination, no patient was significantly more impaired on the sublexical than the lexical task (p 's from .08 to .71, see Figure 4). For syllable discrimination and auditory lexical decision, five of the six patients with more negative z-scores on auditory lexical decision were not significantly more impaired on syllable discrimination than on auditory lexical decision (p 's from .09 to .71). One patient (DG) was significantly more impaired on the syllable discrimination task ($p = .02$).

Figure 4. Scatterplot comparing patient performance (z) on the syllable and word discrimination tasks. The red dotted line indicates $y=x$.

Figure 5. Scatterplot comparing patient performance (z) on the syllable discrimination and auditory lexical decision tasks. The red dotted line indicates $y=x$.

Overall, the results from the discrimination and lexical decision tasks are consistent with theories assuming that lexical perception depends on sublexical perception. For word discrimination, no patient performed significantly better on the lexical perception task than the sublexical perception task. For lexical decision, many patients performed significantly worse relative to the syllable discrimination task, but this pattern could be attributed to a disruption of lexical or semantic representations per se. There was a single patient (DG) who was significantly more impaired on the syllable discrimination task relative to controls for auditory lexical decision; however, he was not significantly more impaired on the word discrimination task. Both the syllable and word discrimination tasks require the participant to maintain two items in memory and compare them to each other, which taxes STM/WM. In contrast, the auditory lexical decision task only requires the participant to maintain a single item, minimizing STM/WM demands. Thus, it seems likely that DG's relatively worse performance on the syllable discrimination task than the lexical decision task may be due to the STM/WM demands of the syllable discrimination task.

A concern that might be raised regarding the strength of the correlations between the sublexical task and the two lexical tasks is that they might arise from some other factor such as overall aphasia severity or a deficit in general cognitive abilities (such as working memory or discrimination ability) that affects performance on both sublexical and lexical tasks. To address the first concern regarding overall severity of language deficits, we partialled out the WAB aphasia quotient (Kertesz, 2006; see Appendix A) from the correlations. Both partial correlations were strong and significant using the bootstrapping assessment of significance (syllable discrimination and word discrimination: $r = .95$, 95% CI [.88, .99]; syllable discrimination and auditory lexical decision: $r = .69$, 95% CI [.23, .93]). The second concern regarding general cognitive deficits is harder to address, as the study was not designed to address cognitive abilities on non-language tasks. However, we had performance on a demanding visual arrays task (Cowan et al., 2005) for 10 of the 13 patients tested here, collected as part of a separate study, that can go some way to address this concern. In this task, participants were presented with a visual array composed of from two to five colored rectangles in different locations for 500 ms. This array was followed by a 500 ms visual mask and then a display with a single rectangle. Participants decided whether this single rectangle had the same color as the rectangle in that location in the preceding array. Thus, participants had to rapidly encode and remember the color and spatial location of all of the rectangles to perform the task and then compare the color and location of the probe to memory of the array. Overall, the patients did well on this task, with 9 out of 10 scoring within the range of controls. Because of the manipulation of array size, we could compare performance on this task to the speech perception tasks for conditions of similar difficulty by determining the array size for which the mean d' was similar to that for the speech perception tasks. For syllable and word discrimination (with mean d' s of 1.85 and 1.95, respectively, for these 10 patients), the array size was 3 (mean d' of

1.84). For auditory lexical decision (mean d' of 1.42), the array size was 5 (mean $d'=1.39$). All correlations with visual arrays performance were small and non-significant, with two being negative (syllable discrimination: $r = -.18$, 95% CI $[-.72, .43]$; word discrimination; $r = -.23$, 95% CI $[-.74, .35]$; auditory lexical decision, $r = .23$, 95% CI $[-.29, .71]$). Thus, there was no support from these data that the patients had general language, perceptual or cognitive impairments that were the source of our findings in the speech perception domain.

4. Experiment 2

4.1 Experiment 2a: Mapping Sound to Meaning

Given that a critical claim of the dual route models is that the lexical route is involved in mapping sound to meaning, one might be concerned that the word discrimination and auditory lexical decision tasks used in the previous sections were not valid measures of lexical processing as there was no necessary requirement to map the sound to meaning in either of these tasks. That is, word discrimination might be done on the basis of sublexical codes and auditory lexical decision on the basis of a lexical phonological representation. To address this concern, we created a PWM task in which participants decided whether a spoken word matched the name of a pictured object where on the non-matching trials the word differed from the appropriate name by one distinctive feature of one phoneme. In order to successfully determine whether a picture matches a spoken word, one must either access the meaning of the spoken word and compare it to the meaning of the picture or use the meaning of the picture to access its name to compare to the spoken word.

4.1.1 Materials. Stimuli for the single PWM task were constructed from 28 pairs of spoken words and corresponding pictures where the words differ from each other by a single distinctive feature in the initial or final phoneme. Half of the items differed in the initial phoneme and half in the final phoneme (e.g., pear-bear; log-lock). Each picture

was presented individually two times across the experiment, once with an auditorily presented word that matched the pictured item (e.g., a picture of a log presented with the word log) and once with an auditorily presented word that corresponded to the pair of the picture (e.g., a picture of a log presented with the word lock), thus totaling 112 experimental trials. We were concerned that participants would detect the manipulation, namely that each picture was repeated once with the matching spoken word and once with the non-matching spoken word. Accordingly, we included 40 filler trials consisting of 10 pairs of spoken words that varied in the number of times they were presented with the corresponding matching and non-matching pictures. For example, a filler item may have been presented with the matching picture three times and the non-matching picture twice. Picture and word stimuli were concurrently presented. The word stimuli were recorded by a female and the picture stimuli were acquired using a Google Image search.

4.1.2 Procedure. In this task, participants were presented concurrently with an auditory word stimulus and a single picture and told to determine whether the name of the picture matched the word that they heard. They were instructed to press a button corresponding to “yes” if the word and picture matched and a button corresponding to “no” if the word and picture did not match. They were also informed that half of the trials were matching items and half of the trials were non-matching items.

4.1.3 Results and discussion. Two patients were not available for testing, leaving 11 patients who completed this task. As with the discrimination tasks and lexical decision task presented above, performance was measured using d' scores, with hit rate defined as correctly accepting a word and picture match and false alarm rate defined as incorrectly accepting a word and picture match. Extreme scores (i.e., hit rate and false alarm rates of 1 or 0) were adjusted using the same criteria as previously detailed. Data are presented in Table 7. Note that five patients who were below the control range on

auditory lexical decision were also below the control range on the PWM task (DZ, HA, KA, MB and SH), supporting the claim that some patients have lexical or semantic processing deficits.

Table 7.

Patient and control performance (d') on the single PWM task. For patients with better PWM than syllable discrimination task performance, bootstrapping values for the mean difference is presented, with 95% CI's of the difference included in brackets.

Patient	Syllable Discrimination	PWM	Difference (PWM-Syllable Discrimination)
BQ	1.45	N/A	
DZ	1.86	2.01	.18 [-.47, .89]
DG	2.09	2.72	.73 [-.02, 1.51]
EV	2.45	2.71	.41 [-.37, 1.31]
HA	.93	.43	
KA	.56	.32	
MB	1.11	1.70	.64 [-.04, 1.42]
NMA	1.98	3.34	1.37 [.56, 2.15]
QO	2.46	3.41	1.07 [.21, 2.01]
SH	1.72	1.47	
SJ	2.10	2.14	.11 [-.58, .85]
SQ	2.21	1.99	
SSW	2.07	N/A	
Patient Means	1.77	2.35	
Controls (n=10)		3.42 (2.59-4.17)	

Overall performance on the PWM task was better than on the syllable discrimination task. While this difference did not reach significance for patients ($t(10) = -1.46$, $p = .18$), it was highly significant for controls ($t(9) = -3.75$, $p = .005$). Performance on the single PWM task was significantly correlated with: the syllable discrimination task, $r = .86$, 95% CI [.56, .96], word discrimination task, $r = .88$, 95% CI [.58, .98], and auditory lexical decision task, $r = .82$, 95% CI [.05, .94].

While group mean performance on the PWM task did not significantly exceed that on the syllable discrimination task for the patients, it is critical to determine if any

individual patient showed significantly better performance on PWM. As can be seen in Figure 6, seven patients performed somewhat better on the PWM task than the syllable discrimination task. To assess the significance of these differences, the same bootstrapping procedure described earlier (Experiment 1b) was used to compare performance at an individual level. Table 7 presents the 95% CI's of the difference between d' scores for syllable discrimination and PWM determined via bootstrapping for patients whose PWM d' was larger than their syllable discrimination d' . The bootstrapping analyses indicated that for five of the seven patients there was no significant difference (p 's from .06 to .79) whereas two patients (NMA and QO) did show significantly better PWM performance (p 's = .008 and .015, respectively). Note, though, that the majority of patients were still impaired relative to controls on the PWM task (results from Crawford and Howell's (1998) modified t -test are presented in Table 9) and that the controls also performed better on this task ($M = 3.42$) than on the word discrimination ($M = 2.44$) and auditory lexical decision ($M = 2.44$) tasks. Controls' better performance on this task compared to the other lexical tasks suggests that participants may have benefitted from having a picture target which allowed them to internally generate the name of the picture which they could use as a template for evaluating the auditory input. In fact, a recent study by McMurray and Jongman (2015) demonstrated that visual input can aid in perception of auditory stimuli. In that study, participants were better able to predict which of four vowels would be produced with an isolated fricative (fricative-vowel combinations recorded with vowel subsequently edited out of the speech stream) when they were simultaneously presented with the written fricative and a face revealing the gender of the speaker than when they were presented with the auditory stimulus alone. The possibility of a boost in performance with visual input is addressed further in the next experiment.

Figure 6. Scatterplot comparing patient performance (d') on the syllable discrimination task and single PWM task. The red dotted line indicates $y=x$ and the orange line indicates the minimum control performance on PWM. Note that several patients are substantially above the red-dashed line, indicating better performance on the PWM task than the syllable discrimination task.

4.2 Experiment 2b: Auditory Written Syllable Matching (AWSM)

Given the relatively better performance by patients and controls on the single PWM task as compared to the single feature difference syllable discrimination task, we wanted to create a sublexical task that was more comparable to the PWM task in terms of task demands. In the single PWM task, presentation of the picture could have allowed patients to generate an internal phonological code (i.e., the name of the picture) that they could then compare to the auditorily presented word. Accordingly, to tap sublexical processing, we created a task where patients were presented with a written syllable, which allowed them to generate an internal phonological code that they could compare to the auditory syllable.

4.2.1 Materials. The six stop consonants paired with /b/ (i.e., /pɒ/, /bɒ/, /dɒ/, /tɒ/, /gɒ/, /kɒ/) were used as the auditory stimuli for the AWSM task. The written stimuli were BA, PA, DA, TA, GA and KA, respectively. Each spoken syllable was presented individually twelve times across the experiment, six times with a written syllable that

matched the spoken syllable and six times with a written syllable that did not match (twice for voicing, twice for manner, twice for place), thus totaling 72 experimental trials.

4.2.2 Procedure. In this task, participants were told that they would see a written syllable and hear a spoken syllable, which were presented simultaneously, and had to determine whether the two stimuli were the same or different. Before beginning the task, participants were presented with each written syllable and told to read it aloud. If they made mistakes, the experimenter corrected them and then they went through the materials a second time and their responses were recorded. If they were still unable to read all the stimuli correctly the second time around, they were not tested on the task. The patients who passed the screening then completed six practice trials with experimenter feedback followed by the 72 experimental trials without feedback. They were told to press one button to indicate the written and spoken syllable were the same and a different button to indicate they were different. They were also told that half of the trials were the same and half were different.

4.2.3 Results and discussion. Five patients were unable to complete this task. Of these, two were no longer available for testing (BQ and SSW) and two had no speech output (DZ and HA), preventing us from testing their ability to read these nonsense syllables aloud. The fifth patient, DG, was unable to correctly read the nonsense syllables aloud, even after correction, and as such was excluded from testing. As with the tasks presented above, performance was measured using d' scores, with hit rate defined as correctly accepting a written and spoken syllable match and false alarm rate defined as incorrectly accepting a written and spoken syllable mismatch. Extreme scores (i.e., hit rate and false alarm rates of 1 or 0) were adjusted using the same criteria as previously detailed. Data are presented in Table 8.

Table 8.

Patient and control performance on the single AWSM task.

<i>Patient</i>	<i>d'</i>
BQ	N/A
DZ	N/A
DG	N/A
EV	2.71
HA	N/A
KA	.58
MB	2.41
NMA	3.29
QO	2.60
SH	2.06
SJ	1.43
SQ	2.98
SSW	N/A
Mean	2.30
Controls (n=10)	3.77 (2.6-4.4)

As with the comparison of the PWM task to the single feature difference word discrimination and auditory lexical decision tasks, the majority of patients and controls performed better on AWSM than the single feature difference syllable discrimination task, with 6 of 8 patients performing better on this AWSM task (see Figure 7). However, this difference failed to reach statistical significance, $t(7) = -1.84$, $p = .11$.

We wished to determine if PWM performance would still exceed performance on syllable perception when the syllable processing task also involved internal generation of a target. In contrast to the significant difference between PWM and syllable discrimination where participants were significantly worse on the sublexical task, controls were significantly better at AWSM than PWM, $t(9) = 2.43$, $p = .04$, and there was no significant difference for patients, $t(7) = .53$, $p = .61$. As shown in Figure 8, there was a strong relation between performance on the AWSM and PWM task, $r = .78$, 95% CI [.16, .96].

As in the previous experiments, however, it was more critical to determine

whether any individual demonstrated significantly better lexical than sublexical processing. Two patients (i.e., SJ, QO) performed slightly better on the PWM than the AWSM task (PWM d' was 2.14 for SJ, 3.41 for QO). Using the bootstrapping (procedure described in Experiment 1b), we computed 95% CI's for the difference in d' values between AWSM and PWM, which failed to reach significance for SJ (CI: [-.29, 1.66], $p = .79$) and QO (CI: [-.40, 2.01], $p = .19$).

Figure 7. Scatterplot comparing patient performance (d') on the single feature difference syllable discrimination task and the AWSM task. The red dotted line indicates $y=x$. Note that several patients are substantially above the line, indicating better performance on AWSM than syllable discrimination.

Figure 8. Scatterplot comparing patient performance (d') on the AWSM task and the PWM task. The red dotted line indicates $y=x$. For the two patients who fall above the line, bootstrapping of the d' values indicated no significant difference.

As in Experiment 1b, patient performance was compared to controls using the Crawford and Howell (1998) modified t -test to determine if their performance was significantly worse than controls. The results of these analyses are presented in Table 9.

There was no patient who was significantly impaired relative to controls on the sublexical processing task who had normal performance on the lexical processing task.

Table 9.

Results of Crawford and Howell's (1998) modified t-tests for evaluating differences between a single participant and a small group of participants. Each patient was compared to the control group. Negative t-values indicate worse performance relative to controls.

Patient	AWSM		PWM	
	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
BQ	N/A	N/A	N/A	N/A
DZ	N/A	N/A	-2.81	.02
DG	N/A	N/A	-1.38	.20
EV	-2.00	.08	-1.42	.19
HA	N/A	N/A	-5.95	<.001
KA	-6.04	< .001	-6.17	<.001
MB	-2.57	.03	-3.42	.01
NMA	-.90	.39	-.16	.88
QO	-2.21	.06	-.01	.99
SH	-3.23	.01	-3.88	<.001
SJ	-4.43	<.001	-2.54	.03
SQ	-1.49	.17	-2.85	.02
SSW	N/A	N/A	N/A	N/A

Finally, as in Experiment 1b, we converted all patient scores to z-scores in order to directly compare relative levels of impairment across the tasks. Table 10 presents the z-scores for the patients on the AWSM and PWM tasks (plotted in Figure 9). As shown, five patients were more impaired on AWSM than on PWM. In order to determine if they were significantly more impaired on AWSM, we utilized the same bootstrapping technique as before (converted to z-score, determined 95% CI for difference in z-scores). Four of these patients were not significantly more impaired on the AWSM task (*p*'s from .06 to .68), but one patient (SJ) was (*p* = .03). For this patient, as with patient DG in Experiment 1b, there was not a consistent pattern of worse sublexical task performance across all the tasks. SJ was not significantly more impaired on syllable discrimination relative to word discrimination or auditory lexical decision. More importantly, SJ was not significantly more impaired on syllable discrimination compared

to PWM. Again, this suggests that the difference in performance on AWSM compared to PWM reflects task demands rather than a true dissociation. For example, SJ may have a weaker phonological code for written syllables than for pictures. In fact, SJ is below the control range on tasks involving non-word reading (e.g., PALPA 36, Kay, Lesser & Coltheart, 1992), but performs normally on picture naming (e.g., Philadelphia Naming Test, Roach, Schwartz, Martin, Grewal & Brecher, 1996).

Overall, performance on the task tapping lexical processing was roughly equal to or below that for the task tapping sublexical processing for nearly all patients even for a lexical task that required access to meaning. Thus, the results provide converging evidence with those from Experiment 1b in showing that lexical processing does not consistently exceed sublexical processing for tasks matched in discrimination difficulty.

Table 10.

Z-scores for each patient compared to the control group for AWSM and PWM. More negative values indicate worse performance relative to controls.

<i>Patient</i>	<i>AWSM(z)</i>	<i>PWM(z)</i>	<i>Bootstrapping: Mean [95% CI]</i>
BQ	N/A	N/A	
DZ	N/A	-2.95	
DG	N/A	-1.45	
EV	-2.10	-1.49	-1.15 [-3.14, .60]
HA	N/A	-6.24	
KA	-6.33	-6.47	-.32 [-1.80, 1.27]
MB	-2.69	-3.58	
NMA	-.94	-.17	-1.06 [-3.02, .94]
QO	-2.32	-.01	-2.34 [-4.69, .11]
SH	-3.39	-4.07	
SJ	-4.64	-2.67	-2.15 [-4.05, -.18]
SQ	-1.56	-2.99	
SSW	N/A	N/A

Figure 9. Scatterplot comparing patient performance (z) on the AWSM and PWM tasks. The red dotted line indicates $y=x$. Of the patients falling above the line, only one had a z-score that was significantly larger on PWM than AWSM.

5. General Discussion

Many models of speech perception have assumed obligatory sublexical processing as a prerequisite for lexical processing (e.g., Cole & Scott, 1974; McClelland & Elman, 1986; Luce, Goldinger, Auer & Vitevitch, 2000; Luce & Pisoni, 1998; Norris, 1994; Norris, McQueen & Cutler, 2000; Oden & Massaro, 1978). It was thus surprising that patients with impaired sublexical processing but relatively intact lexical processing were reported in the literature (Basso et al., 1977; Blumstein et al., 1977a; Blumstein et al., 1977b; Miceli et al., 1980). These results led researchers to claim that either a) sublexical and lexical processing operate via largely independent, parallel processing routes (e.g., Hickok & Poeppel, 2000; Hickok & Poeppel, 2004; Hickok & Poeppel, 2007; Hickok, 2014; Poeppel & Hickok, 2004; Majerus, 2013; Scott & Wise, 2004; Wise et al., 2001) or b) inappropriate tasks have been used to tap sublexical processing (e.g., Hickok & Poeppel, 2007). Given the theoretical importance that has been attached to the findings from those patient studies, we closely examined the studies and found that all of them failed to match the degree of phonological similarity of the target and distractor items between the sublexical and lexical processing tasks and many failed to match task demands.

1 In the current study, we were able to replicate the dissociation between impaired
2 sublexical and preserved lexical processing using stimuli and tasks like those from prior
3 studies where close matching of stimuli and task demands was not carried out (e.g.,
4 Miceli et al., 1980)⁷. However, when we tested patients on sublexical and lexical tasks
5 that were closely matched in terms of discriminability of items and task demands, we
6 failed to replicate the striking dissociations showing impaired performance on sublexical
7 tasks but spared performance on lexical tasks reported in the prior literature. For the two
8 sublexical tasks and three lexical tasks employed here, 11 out of 13 patients demonstrated
9 better sublexical performance or roughly equal sublexical and lexical performance. Only
10 two patients demonstrated significantly better lexical than sublexical performance, but
11 this was observed on only one pair of sublexical vs. lexical task comparisons, but not on
12 others. For DG, this difference was only evident when comparing tasks that had different
13 STM/WM requirements, suggesting that his worse performance on syllable
14 discrimination was likely due to the increased STM/WM demands relative to auditory
15 lexical decision. For SJ, this difference was manifested in the comparison of AWSM and
16 PWM, but not in the comparison of syllable discrimination to word discrimination,
17 auditory lexical decision or PWM. This makes it likely that she is less able to generate a
18 stable phonological code for the written syllables in the AWSM task than for the picture
19 names in the PWM task, leading to her worse performance.

20 Overall, the data from the current study support our claim that the findings of
21 studies like Miceli et al. (1980) reflect task artifacts (i.e., task demands and/or stimulus
22 discriminability) rather than true double dissociations between sublexical and lexical

⁷ In a separate experiment not reported in this paper, we had participants perform a word discrimination task where items differed by three features. When we compared this to the single feature difference syllable discrimination task from Experiment 1b, we found that some patients were better at the lexical than sublexical task. This suggests that if researchers fail to match the phonological similarity of targets and distractors between sublexical and lexical processing tasks, even if the two tasks are matched on task demands, patients can present better lexical than sublexical processing.

processing or syllable and word perception. When sublexical and lexical tasks were closely matched in terms of task demands and phonological similarity of targets and distractors, as in the current study, performance on sublexical tasks (i.e., syllable discrimination) was equal to or better than performance on lexical tasks (i.e., word discrimination, auditory lexical decision) for the most part. Additionally, our data are consistent with claims made by and McMurray and Jongman (2015), namely that individuals' perception of auditory stimuli can be aided by simultaneous presentation of related visual stimuli (see also Huettig, Rommers, & Meyer, 2011, for similar suggestions regarding findings from the visual world paradigm).

Rather than finding striking dissociations between lexical and sublexical tasks, we found strong and significant correlations between performance on these tasks for the patients tested here. Although the sample sizes were small, the fact that these correlations were replicated across different task comparisons and that the bootstrapped 95% CI's were far from zero in the majority of comparisons makes it highly unlikely that there is no true relation between sublexical and lexical processing. Both the failure to find consistent and substantially better lexical than sublexical processing and the presence of these correlations are consistent with models assuming obligatory sublexical processing. Of course, it would be valuable to replicate the findings reported here with larger sample sizes to document that both features of the data persist - that is, that patients show better sublexical or roughly equal sublexical and lexical performance, and performance at the two levels is correlated to some extent in that poor sublexical performance necessarily leads to poor lexical performance.

One question that might be raised is whether the failure to find significantly impaired sublexical processing co-existing with good lexical processing owes to the interactivity between sublexical and lexical levels in speech perception (e.g., McClelland & Elman, 1986) and/or, potentially, interactions between dorsal and ventral routes (e.g.,

1 Cloutman, 2013; Ueno et al., 2011). According to this line of reasoning, preserved lexical
2 processing boosted performance at a damaged sublexical level to bring it in line with
3 performance at the lexical level. In other words, resonance between the lexical and
4 sublexical levels served to clean up noisy representations at the sublexical level. Effects
5 such as the well-known phonemic restoration effect (Warren, 1970; Samuel, 1981) and
6 the effect of lexical (i.e., word vs. nonword) context on phonetic categorization (Ganong,
7 1980) would appear to support such a top-down hypothesis. However, as documented
8 by Samuel (1996), the phonemic restoration effect is quite fragile and strongly influenced
9 by the match between the acoustic features of the missing phoneme and the noise that
10 replaces it. The influence of lexical status on phonetic categorization is subtle, involving
11 a small shift in a phoneme category boundary (McClelland & Elman, 1986), rather than
12 any change in accuracy. Moreover, these effects involve presentation of word stimuli
13 whereas the conditions in our experiment supposedly improved by feedback involve
14 presentation of sublexical stimuli (i.e., syllables). As noted by Norris, McQueen, and
15 Cutler (2000), there is no evidence that lexical factors improve sensitivity in phoneme
16 discrimination. As these authors discuss, feedback from the lexical to the sublexical level
17 could impair performance rather than improve it. For example, suppose that a patient is
18 asked to discriminate /pa/ and /ba/ but has a sublexical processing deficit which results
19 in the /ba/ stimulus activating a /p/ representation more than a /b/ representation (though
20 not as much as does the /pa/ stimulus). In that case, lexical representations for words
21 beginning with /p/ will be more activated than those starting with /b/ for the /ba/ stimulus
22 and feedback from the lexical level will only serve to reinforce the erroneous activation of
23 the /p/ representation, making it more likely that the patient would make a discrimination
24 error than if no feedback had occurred.⁸ Thus, top down effects from the lexical to the

⁸ Reasoning along these lines is presented by Norris, McQueen and Cutler (2000) as supporting their contention that speech perception does not involve feedback as it provides no benefit to perception.

sublexical level are quite weak and it is unclear that any such feedback, if it exists, would serve to improve sublexical perception performance.

While the data from the current study provide support for a model of speech perception with sublexical processing as a prerequisite for lexical processing, it would only take a single patient showing substantially better performance on lexical level than sublexical level tasks to negate the conclusions from the current study. One recent study by Wolmetz, Poeppel and Rapp (2011) did report a patient that might match this description. Patient DMN was better at identifying word stimuli compared to nonword stimuli and discriminating pairs of words as opposed to nonwords. However, it is important to note that DMN showed a pattern of identification for words and nonwords that was markedly different from controls and was severely impaired at between-category discrimination for both words ($d'=.41$) and nonwords ($d'=-.15$) relative to a control population (word discrimination mean $d' \sim 2.6$, nonword discrimination mean $d' \sim 2.3$). Although this patient might seem to have substantially better lexical than sublexical perception, we do not view this pattern of performance as problematic for at least two reasons.

First, performance on these two tasks was still related in the sense that DMN was essentially performing at floor – not below the normal range on one and within the normal range on the other. We would view a pattern of performance where a patient is substantially better at lexical level than sublexical level tasks (e.g., within the range of controls on lexical but outside the range on sublexical) as more problematic for our arguments. Second, as discussed in the introduction, top-down influences can affect processing at lower levels; this is an important issue to consider in assessing whether patients perform better on lexical tasks than would be expected based on their sublexical performance. Some models, like TRACE (McClelland & Elman, 1986), assume feedback from semantic to lexical levels (and from lexical to sublexical levels) that could boost

performance on lexical tasks (see also Gow, 2012). Other models, like MERGE, do not assume feedback from higher to lower levels, but instead assume that after processing occurs at the sublexical level, access to lexical and semantic information can be recruited in task performance (Norris, McQueen & Cutler, 2000). For DMN, the potential availability of lexical or semantic information for some of the word stimuli may have allowed him to base some of his comparisons on these codes⁹, leading to slightly better performance in the lexical task, though still far below level of controls. Note, also, that in the current study two of our patients better performance on one of the lexical than sublexical processing tasks that was statistically significant, suggesting that potential top-down factors may have been enough to boost their lexical performance slightly above their sublexical performance.

Although the current data do not support the observed double dissociation between sublexical and lexical processing, evidence from other sources has also been used to support a dissociation between sublexical and lexical phonological processing routes. One line of evidence comes from functional neuroimaging in which it is claimed that sublexical effects appear in the dorsal stream in parietal and frontal areas, with lexical effects appearing in temporal regions. For example, Vaden et al. (2011) had participants passively listen to words which varied in phonotactic frequency, a variable believed to affect sublexical levels of processing. They found that activity in the left inferior frontal gyrus was most closely related to this manipulation, which they argued supported claims of downstream processing of sublexical information. However other researchers have reported activity related to sublexical factors in regions earlier in the processing stream, such as superior temporal gyrus or superior temporal sulcus (e.g., Binder et al., 2005; Dewitt & Rauschecker, 2012; Rimol, Specht, Weis, Savoy & Hugdahl, 2005; see Deschamps, Baum, & Gracco, 2015, for a review). For instance,

⁹ This would have little effect on nonwords, as there is no corresponding semantic representation.

Deschamps et al. (2015) found greater activation in the left superior temporal sulcus for pseudowords of high vs. low phonological complexity (based on sonority differences) during passive listening. Activation in frontal areas for this complexity contrast was found only for visually presented pseudowords – which they attributed to automatic articulation of the visual stimuli. Moreover, they found that the pseudowords in this study activated many regions in common with regions activated in prior studies of words. They suggest that differences between words and pseudowords across some prior studies may be due to the different types of tasks involved. Thus, clearly there is a complex set of results in the neuroimaging literature on speech perception that undermines claims for a clean dissociation between areas involved in lexical and sublexical processing (see also, Price, 2012).

Another line of evidence taken to support a direct mapping of acoustic information to lexical levels has come from behavioral studies of young adults. For example, Goldinger (1998) reported that individuals tend to imitate surface features (e.g., pitch) of spoken word and nonword tokens in an immediate, but not a delayed, shadowing task. Goldinger (1998) argued that these data support claims that episodic aspects of speech stimuli are represented in the speech perception stream but that these aspects decay with time, hence the difference in immediate versus delayed shadowing. In line with this claim, a computational model of speech perception based on episodic acoustic representations (rather than abstract phonemic representations) was able to accommodate these data. Note, however, that researchers have recently argued that speech representations can be discrete and categorical, such as for sublexical phonemic representations, while still maintaining continuous information present in the speech signal (e.g., McQueen, Cutler & Norris, 2006; Smolensky, Goldrick & Mathis, 2014).

5.2 Limitations and Future Directions.

The present findings provide strong evidence that prior reports of dissociations between good lexical and poor sublexical processing in aphasia were most likely due to a failure to match perceptual discriminability and task demands for the tasks at the two levels. Consequently, these prior findings cannot be taken as evidence for separable routes for sublexical and lexical processing. The present findings showed no patient with consistently better lexical than sublexical processing and substantial correlations between performance on the tasks at the two levels, with both of these findings instead being consistent with obligatory sublexical processing models. However, the case for obligatory sublexical processing could be strengthened considerably by documenting that variables affecting sublexical performance have predictable influences on lexical performance. For example, if lexical processing depends on sublexical processing, then one would predict that the extent to which sublexical variables such as biphone probability or sonority gradients affect individual patients' sublexical perception would also affect their lexical processing. That is, for example, a patient who showed an exaggerated effect of biphone frequency on sublexical perception would be predicted to show an exaggerated effect of this variable on lexical perception. Such findings revealing more detailed correspondence between sublexical and lexical levels would provide stronger evidence for a dependence of lexical processing on sublexical processing. To our knowledge, no such study has been reported that has related the size of effects of sublexical variables to their effects on lexical perception in aphasic individuals. Thus, this would be an important avenue to pursue in future research.

5.3 Concluding comments on implications for dual route models

Though dual route models with a specific neuroanatomical basis like that of Hickok and Poeppel have been proposed relatively recently (Hickok & Poeppel, 2000), cognitive models of language processes with a dual route framework (though typically without a specified neural basis) are common in the neuropsychological literature,

particularly for reading and repetition (e.g., Coltheart, Rastle, Perry, Langdon & Ziegler, 2001; Dell, Martin & Schwartz, 2007; Hanley, Dell, Kay & Baron, 2004; Hanley & Kay, 1997; Hillis & Caramazza, 1991; McCarthy & Warrington, 1984; Nozari, Kittredge, Dell & Schwartz, 2010). Critically, many of these models assume that sublexical processing is shared between the two routes and the routes do not become activated until after sublexical processing occurs. A similar approach could be applied in the speech perception domain. That is, one might assume that there are separable routes for translation to speech output and for accessing meaning, but assume that sublexical processing is shared by the two routes and must be accomplished before processing branches into the separate routes. Such an approach would preserve most aspects of the dual route model proposed by Hickok and Poeppel (2000), but provide a better account of the findings reported here that would be consistent with an influence of sublexical speech processing on word recognition. Of course, the localization of the shared sublexical processing would need to be established.

In summary, the current study provides support for models of speech perception where processing of sublexical information is a prerequisite for processing of lexical information, as is the case in TRACE (McClelland & Elman, 1986), NAM (Luce & Pisoni, 1998) and Shortlist/MERGE (Norris, 1994; Norris, McQueen & Cutler, 2000). On the other hand, we failed to find support for models that do not require passage through sublexical levels to reach lexical levels, such as the episodic theory of speech perception (e.g., Goldinger, 1998) or dual route models of speech perception (Hickok & Poeppel, 2000; Hickok & Poeppel, 2004; Hickok & Poeppel, 2007; Hickok, 2014; Poeppel & Hickok, 2004; Majerus, 2013; Scott & Wise, 2004; Wise et al., 2001). That being said, dual route models could easily accommodate these findings by making explicit the claim that sublexical processing is shared between the two processing streams and must be successfully carried out in order to proceed down either path.

1 Acknowledgements: We are grateful to Simon Fischer-Baum and David Lane of the
2 Department of Psychology at Rice University for their assistance and statistical guidance
3 on this project. We would also like to thank our research assistants Blaine Tomkins,
4 Tristanne Mock and Maddie Flores. Finally, we would like to thank the patients for their
5 participation in this study.

6 Funding: This research was supported in part by funding for the Neuroplasticity
7 Research Laboratory from the T.L.L. Temple Foundation.

8
9

References

- Baker, E., Blumstein, S., & Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychologia*, 19, 1-15. doi:10.1016/0028-3932(81)90039-7
- Basso, A., Casati, G., & Vignolo, L. (1977). Phonemic identification defect in aphasia. *Cortex*, 13, 85-95. doi:10.1016/s0010-9452(77)80057-9
- Blumstein, S. E., Baker, E., & Goodglass, H. (1977a). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, 15(1), 19-30. doi:10.1016/0028-3932(77)90111-7
- Blumstein, S., Cooper, W., Zurif, E. & Caramazza, A. (1977b). The perception and production of voice-onset time in aphasia. *Neuropsychologia*, 15, 371-383. doi:10.1016/0028-3932(77)90089-6
- Cloutman, L. (2013). Interaction between ventral and dorsal processing streams: Where, when and how? *Brain and Language*, 127, 251-263. doi:10.1016/j.bandl.2012.08.003
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interface environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25(2), 257-271. doi:10.3758/bf03204507
- Cole, R. A., & Scott, B. (1974). Toward a theory of speech perception. *Psychological Review*, 81(4), 348. doi:10.1037/h0036656
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204. doi: 10.1037/0033-295X.108.1.204
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51, 42-100. doi:10.1016/j.cogpsych.2004.12.001.
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12(4), 482-486.
- Dalton, D. Cruickshanks, K., Klein, B., Klein, R., Wiley, T., & Nondahl, D. (2003). The impact of hearing loss on quality of life in older adults. *The Gerontologist*, 43, 661-668. doi:10.1093/geront/43.5.661
- Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language*, 56(4), 490-520. doi:10.1016/j.jml.2006.05.007
- Deschamps, I., Baum, S., & Gracco, V. (2015). Phonological processing in speech perception: What do sonority differences tell us? *Brain and Language*, 149, 77-83. <http://dx.doi.org/10.1016/j.bandl.2015.06.008>
- DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, 109, 505-514. doi:10.1073/pnas.1113427109
- Gainotti, G., Caltagirone, C., & Ibba, A. (1975). Semantic and phonemic aspects of auditory language comprehension in aphasia. *Linguistics*, 13, 15-30. doi:10.1515/ling.1975.13.154-155.15
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110-125. <http://dx.doi.org/10.1037/0096-1523.6.1.110>
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251-279. doi:10.1037/0033-295x.105.2.251

- Goodglass, H., Kaplan, E., & Barresi, B. (1972). Boston Diagnostic Aphasia Test. *Philadelphia: Lea and Febiger*.
- Gow, David. (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language*, 121, 273-288. doi:10.1016/j.bandl.2012.03.005
- Hanley, J., Dell, G. S., Kay, J., & Baron, R. (2004). Evidence for the involvement of a nonlexical route in the repetition of familiar words: A comparison of single and dual route models of auditory repetition. *Cognitive Neuropsychology*, 21(2-4), 147-158. doi:10.1080/02643290342000339
- Hanley, J. & Kay, J. (1997). An effect of imageability on the production of phonological errors in auditory repetition. *Cognitive Neuropsychology*, 14(8), 1065-1084. doi:10.1080/026432997381277
- Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience*, 29, 2-20. doi:10.1080/01690965.2013.834370
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4(4), 131-138. doi:10.1016/s1364-6613(00)01463-7
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67-99. doi:10.1016/j.cognition.2003.11.001
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 292-402. doi:10.1038/nrn2113
- Hillis, A. E., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language*, 40(1), 106-144. doi:10.1016/0093-934x(91)90119-I
- Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Thames Valley Test Company.
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151-171. doi:10.1016/j.actpsy.2010.11.003
- International Organization for Standardization. (2000). *Acoustics-Statistical Distribution of Hearing Thresholds as a Function of Age*. ISO 7029. Geneva: ISO.
- Johnson, K. (1997). *Speech perception without speaker normalization*. In Johnson & Mullennix (Eds.), *Talker variability* (145-165). San Diego: Academic Press.
- in speech processing, ed. by Keith Johnson and John W. Mullennix, 145-65.
- Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language processing in aphasia*. Psychology Press. doi: 10.1080/02687039608248403
- Kertesz, A. (1982). *Western Aphasia Battery Test Manual*. Psychological Corp.
- Kertesz, A. (2006). *Western Aphasia Battery-Revised Test Manual*. Psychological Corp.
- Kucera, H., & Francis, N. (1967). *Computational analysis of present-day American English*. Brown University Press.
- Luce, P., Goldinger, S., Auer, E., & Vitevitch, M. (2000). Phonetic priming, neighborhood activation and PARSYN. *Perception and Psychophysics*, 62, 615-625. doi:10.3758/bf03212113
- Luce, P., & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36. doi:10.1097/00003446-199802000-00001
- Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, 7, 1-16. doi:10.3389/fnhum.2013.00357

- Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63. doi:10.1016/0010-0285(78)90018-X
- Martin, R.C., Lesch, M., & Bartha, M. (1999). Independent of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, 41, 3-29. doi:10.1006/jmla.1999.2637
- Massaro, D.W., & Cohen, M.M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology*, 23, 558-614. doi:10.1016/0010-0285(91)90006-a
- McCarthy, R., & Warrington, E. K. (1984). A two-route model of speech production. *Brain*, 107(2), 463-485. doi:10.1093/brain/107.2.463
- McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86. doi:10.1016/0010-0285(86)90015-0
- McQueen, J., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, 30, 1113-1126. doi:10.1207/s15516709cog0000_79
- McMurray, B., & Jongman, A. (2015). What comes after /f/? Prediction in speech derives from data-explanatory processes. *Psychological Science*, 1-10. doi:10.1177/0956797615609578.
- Miceli, G., Gainotti, G., Caltagirone, C., & Masullo, C. (1980). Some aspects of phonological impairment in aphasia. *Brain and Language*, 11, 159-169. doi:10.1016/0093-934x(80)90117-0
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234. doi:10.1016/0010-0277(94)90043-4
- Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-370. doi:10.1017/s0140525x00003241
- Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language*, 63(4), 541-559. doi:10.1016/j.jml.2010.08.001
- Oden, G.C., & Massaro, D.W. (1978). Integration of featural information in speech perception. *Psychological Review*, 85, 172-191. doi:10.1037/0033-295x.85.3.172
- Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Bybee & Hopper (Eds.), *Frequency effects and emergent grammar* (45-64). Amsterdam: John Benjamins.
- Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition*, 92, 1-12. doi:10.1016/j.cognition.2003.11.001
- Price, C. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62, 816-847. doi:10.1016/j.neuroimage.2012.04.062
- Roach, A., Schwartz, M., Martin, N., Grewal, R., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, 24, 121-133.
- Rimol, L., Specht, K., Weis, S., Savoy, R., & Hugdahl, K. Processing of sub-syllabic speech units in the posterior temporal lobe: An fMRI study. *NeuroImage*, 26, 1059-1067. doi:10.1016/j.neuroimage.2005.03.028
- Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110, 474-494.
- Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, 125, 28-51. <http://dx.doi.org/10.1037/0096-3445.125.1.28>
- Scott, S., & Wise, R. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition*, 92, 13-45. doi:10.1016/j.cognition.2002.12.002

- Smolensky, P., Goldrick, M., & Mathis, D. (2014). Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. *Cognitive Science*, 38(6), 1102-1138.
- Sommers, M. (n.d.). WU Speech & Hearing Lab Neighborhood Database Site. Retrieved March 30, 2015, from <http://neighborhoodsearch.wustl.edu/Neighborhood/Home.asp>.
- Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 16(6), 683-692. doi:10.1016/s0022-5371(77)80027-3
- Vaden J., K. I., Piquado, T., & Hickok, G. (2011). Sublexical properties of spoken words modulate activity in Broca's area but not superior temporal cortex: implications for models of speech recognition. *Journal of cognitive neuroscience*, 23(10), 2665-2674. doi:10.1162/jocn.2011.21620
- Vitevitch, M., & Luce, P. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408. doi:10.1006/jmla.1998.2618
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392-393.
- Wise, R., Scott, S., Blank, S., Mummery, C., Murphy, K., & Warburton, E. (2001). Separate neural subsystems within 'Wernicke's area'. *Brain*, 124(1), 83-95. doi:10.1093/brain/124.1.83
- Wolmetz, M., Poeppel, D., & Rapp, B. (2011). What does the right hemisphere know about phoneme categories?. *Journal of Cognitive Neuroscience*, 23(3), 552-569. doi:10.1162/jocn.2010.21495

Appendix A

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
BQ	Anomic	71.8	94	95	93	94	100	Inferior frontal; insula; inferior, middle and superior temporal extending into Heschl's gyrus; inferior and superior parietal extending into supramarginal and angular gyri; middle occipital; pre- and post-central gyri

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
DZ	Broca's	29.4	87	0	98	81	75	Inferior and middle frontal; insula; middle and superior temporal extending into Heschl's gyrus and temporal pole; inferior and superior parietal extending into supramarginal and angular gyri; middle and superior occipital; pre- and post-central gyri

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
DG	Conduction	71.1	88	50	98	90	88	Inferior, middle and superior temporal extending into Heschl's gyrus and temporal pole; inferior and superior parietal extending into supramarginal and angular gyri; middle and superior occipital; fusiform gyrus; precentral gyrus
EV	Anomic	93.6	95	86	100	92	98	Inferior and middle frontal; insula
HA	Broca's	19.4	85	0	98	94	98	Inferior frontal; insula; superior temporal extending into Heschl's gyrus and temporal pole; inferior parietal and supramarginal gyrus; pre- and post-central gyri

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
KA	Wernicke's	68.5	86	83	97	85	98	Inferior frontal; left and right insula; middle and superior temporal; right inferior, middle and superior temporal; left and right temporal poles; left and right Heschl's gyri; right supramarginal and angular gyri; right middle and superior occipital
MB	Conduction	94.4	98	89	98	100	100	Inferior and superior parietal extending into supramarginal and angular gyri; middle and superior occipital; postcentral gyrus

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
NMA	Conduction	83.6	97	94	100	81	100	Middle and inferior frontal; insula; inferior, middle and superior temporal extending into Heschl's gyrus and temporal pole; inferior parietal and angular gyrus; inferior and middle occipital
QO	Anomic	95	100	98	100	98	100	Middle and superior temporal; supramarginal gyrus
SH	Anomic	84.8	95	93	97	98	100	Middle and superior temporal extending into temporal pole; inferior parietal; supramarginal and angular gyri; middle and superior occipital
SJ	Anomic	88.8	97	97	100	98	100	Middle and superior temporal extending into Heschl's gyrus; inferior parietal; supramarginal and angular gyri

<i>Patient</i>	<i>Aphasia Type</i>	<i>WAB Aphasia Quotient</i>	<i>Single PWM*</i>	<i>Philadelphia Naming Test</i>	<i>PALPA 24: Visual Lexical Decision</i>	<i>Pyramids and Palm Trees: 3 Pictures</i>	<i>PALPA 48: Written Word to Picture Matching</i>	<i>Lesion Location**</i>
SQ	Anomic	75.9	99	90	98	90	90	Inferior, middle and superior frontal; insula; temporal pole; inferior parietal; supramarginal gyrus; pre- and post-central gyrus and supplementary motor area
SSW	Anomic	91***	97	88	-	94	100	No information available

*the single PWM task data presented here is the same as Experiment 1a

** All lesions are in the left hemisphere unless otherwise stated, with lesion defined as regions with at least 10% damage using Automated Anatomical Labelling

***participant was not tested on the object naming portion of WAB and was not available for further testing. score reflects total out of 88 possible points

Pyramids and Palm Trees from Howard and Patterson (1992); PALPA from Kay, et al. (1992), Philadelphia Naming Test from Roach et al., 1996

1

Appendix B

<i>Word Stimuli</i>	<i>Phonetic Transcription</i>	<i># of Phonemes</i>	<i>Frequency</i>	<i>Phonological Neighborhood Density</i>	<i>Average Phonotactic Probability</i>	<i><u>BiphoneFrequency</u></i>		
						<i>Biphone 1</i>	<i>Biphone 2</i>	<i>Biphone 3</i>
bad	b@d	3	143	30	0.056	0.0059	0.0024	
bag	b@g	3	42	28	0.050	0.0059	0.0028	
bait	Bet	3	2	33	0.049	0.0017	0.0030	
bake	Bek	3	12	29	0.045	0.0017	0.0020	
ban	b@n	3	7	30	0.076	0.0059	0.0144	
base	Bes	3	91	27	0.053	0.0017	0.0030	
bat	b@t	3	18	40	0.066	0.0059	0.0059	
bent	bEnt	4	34	17	0.083	0.0041	0.0095	0.0158
bid	bId	3	22	25	0.062	0.0041	0.0032	
big	bIg	3	360	20	0.055	0.0041	0.0035	
bill	bIl	3	143	35	0.074	0.0041	0.0090	
bin	bIn	3	2481	27	0.081	0.0041	0.0095	
bit	bIt	3	101	35	0.071	0.0041	0.0059	
boat	Bot	3	72	32	0.056	0.0023	0.0041	
bud	b^d	3	9	23	0.043	0.0034	0.0010	
bun	b^n	3	1	30	0.062	0.0034	0.0057	
bus	b^s	3	35	20	0.056	0.0034	0.0039	
but	b^t	3	4393	35	0.052	0.0034	0.0024	
buzz	b^z	3	13	15	0.037	0.0034	0.0010	
cab	k@b	3	12	22	0.066	0.0122	0.0025	
cage	keJ	3	9	17	0.044	0.0020	0.0007	
cake	Kek	3	13	26	0.058	0.0020	0.0020	
calf	k@f	3	11	19	0.064	0.0122	0.0013	
call	Kcl	3	188	26	0.076	0.0166	0.0059	
calve	k@v	3	1	17	0.065	0.0122	0.0019	
cap	k@p	3	27	30	0.070	0.0122	0.0048	
cat	k@t	3	23	35	0.079	0.0122	0.0059	
cave	Kev	3	9	22	0.049	0.0020	0.0015	
crane	Kren	4	5	15	0.061	0.0094	0.0053	0.0018
craze	Krez	4	2	11	0.053	0.0094	0.0053	0.0009
cup	k^p	3	45	18	0.056	0.0043	0.0012	
cut	k^t	3	192	25	0.066	0.0043	0.0024	
dad	d@d	3	15	21	0.056	0.0019	0.0024	
dam	d@m	3	39	22	0.060	0.0019	0.0049	
date	Det	3	103	24	0.049	0.0015	0.0030	
deed	Did	3	8	18	0.041	0.0030	0.0022	

dip	dIp	3	6	25	0.062	0.0152	0.0049	
dock	Dak	3	28	22	0.055	0.0023	0.0034	
dog	Dcg	3	75	8	0.043	0.0023	0.0007	
dole	Dol	3	1	33	0.058	0.0014	0.0055	
done	d^n	3	321	30	0.062	0.0024	0.0057	
dose	Dos	3	11	12	0.060	0.0014	0.0024	
dote	Dot	3	1	25	0.056	0.0014	0.0041	
dove	Dov	3	4	16	0.038	0.0024	0.0012	
doze	Doz	3	1	20	0.040	0.0014	0.0013	
dub	d^b	3	1	22	0.039	0.0024	0.0036	
duck	d^k	3	9	25	0.048	0.0024	0.0019	
dust	d^st	4	70	8	0.065	0.0024	0.0039	0.0232
fair	fEr	3	84	28	0.068	0.0034	0.0071	
fan	f@n	3	18	21	0.074	0.0034	0.0144	
file	fYl	3	81	27	0.028	0.0017	0.0003	
fin	fIn	3	3	30	0.080	0.0047	0.0095	
fit	flt	3	75	28	0.070	0.0047	0.0059	
found	fWnd	4	536	14	0.051	0.0006	0.0040	0.0131
gall	Gcl	3	7	24	0.053	0.0020	0.0059	
gap	g@p	3	17	20	0.048	0.0033	0.0048	
gill	gIl	3	2	28	0.065	0.0012	0.0090	
goat	Got	3	6	26	0.047	0.0015	0.0041	
gun	g^n	3	118	20	0.054	0.0016	0.0057	
gut	g^t	3	1	24	0.044	0.0016	0.0024	
had	h@d	3	5133	31	0.052	0.0052	0.0024	
half	h@f	3	275	18	0.046	0.0052	0.0013	
halve	h@v	3	3941	17	0.047	0.0052	0.0019	
ham	h@m	3	19	26	0.056	0.0052	0.0049	
has	h@z	3	2439	20	0.066	0.0052	0.0071	
have	h@v	3	3941	17	0.047	0.0052	0.0019	
jam	J@m	3	6	16	0.048	0.0011	0.0049	
jest	Jest	4	1	18	0.064	0.0030	0.0069	0.0232
just	J^st	4	872	12	0.055	0.0014	0.0039	0.0232
kill	kIl	3	63	36	0.088	0.0020	0.0090	
late	Let	3	179	32	0.043	0.0019	0.0030	
leaf	Lif	3	12	25	0.029	0.0024	0.0009	
leash	liS	3	3	16	0.025	0.0024	0.0002	
let	lEt	3	384	30	0.058	0.0040	0.0049	
lick	llk	3	3	32	0.061	0.0062	0.0086	
life	lYf	3	715	20	0.029	0.0036	0.0012	
lip	lIp	3	18	29	0.056	0.0062	0.0049	
live	llv	3	177	15	0.051	0.0062	0.0031	
loaf	Lof	3	4	16	0.034	0.0022	0.0002	

lobe	Lob	3	3	15	0.036	0.0022	0.0013	
lone	Lon	3	54	33	0.060	0.0022	0.0022	
loose	Lus	3	53	23	0.045	0.0020	0.0014	
lose	Luz	3	58	17	0.045	0.0020	0.0014	
man	m@n	3	1207	26	0.078	0.0101	0.0144	
mess	mEs	3	22	18	0.070	0.0059	0.0069	
met	mEt	3	132	26	0.065	0.0059	0.0049	
moon	Mun	3	60	21	0.055	0.0001	0.0005	
need	Nid	3	361	21	0.031	0.0017	0.0022	
nip	nIp	3	3	25	0.052	0.0019	0.0049	
noon	Nun	3	25	19	0.043	0.0005	0.0005	
pad	p@d	3	8	26	0.067	0.0087	0.0024	
pair	pEr	3	58	26	0.081	0.0087	0.0071	
pass	p@s	3	89	24	0.081	0.0087	0.0071	
pat	p@t	3	35	39	0.077	0.0087	0.0059	
peck	pEk	3	5	20	0.070	0.0045	0.0071	
peg	pEg	3	4	10	0.058	0.0045	0.0016	
pill	pIl	3	15	36	0.085	0.0048	0.0090	
pin	pIn	3	16	31	0.092	0.0048	0.0095	
pit	pIt	3	14	33	0.082	0.0048	0.0059	
rage	reJ	3	16	19	0.030	0.0031	0.0007	
raid	red	3	10	29	0.039	0.0031	0.0023	
rain	ren	3	80	38	0.058	0.0031	0.0030	
rate	ret	3	209	39	0.048	0.0031	0.0030	
rave	rev	3	1	26	0.034	0.0031	0.0015	
rib	rIb	3	1	19	0.057	0.0171	0.0023	
rid	rId	3	19	29	0.061	0.0171	0.0032	
ring	rIG	3	49	23	0.053	0.0171	0.0032	
rip	rIp	3	6	32	0.061	0.0171	0.0049	
sack	s@k	3	8	28	0.078	0.0046	0.0054	
sad	s@d	3	35	25	0.073	0.0046	0.0024	
sag	s@g	3	4	22	0.067	0.0046	0.0028	
same	sem	3	686	21	0.060	0.0018	0.0016	
sane	sen	3	8	33	0.076	0.0018	0.0030	
sap	s@p	3	1	27	0.073	0.0046	0.0048	
sat	s@t	3	150	32	0.083	0.0046	0.0059	
send	sEnd	4	74	18	0.084	0.0096	0.0095	0.0131
shake	Sek	3	17	24	0.031	0.0007	0.0020	
sip	sIp	3	2	28	0.079	0.0096	0.0049	
sit	sIt	3	67	36	0.088	0.0096	0.0059	
sound	sWnd	4	204	10	0.062	0.0007	0.0040	0.0131
sub	s^b	3	5	17	0.056	0.0059	0.0036	
sue	su	2	18	33	0.062	0.0023		

sum	s^m	3	1662	23	0.064	0.0059	0.0051	
sun	s^n	3	278	26	0.079	0.0059	0.0057	
tack	t@k	3	4	37	0.059	0.0039	0.0054	
tad	t@d	3			0.054	0.0039	0.0024	
tag	t@g	3	5	21	0.047	0.0039	0.0028	
take	tek	3	611	25	0.042	0.0012	0.0020	
tap	t@p	3	18	26	0.054	0.0039	0.0048	
tend	tEnd	4	43	16	0.069	0.0025	0.0095	0.0131
tin	tIn	3	12	30	0.079	0.0025	0.0095	
tip	tIp	3	22	25	0.059	0.0025	0.0049	
tub	t^b	3	13	17	0.037	0.0014	0.0036	
tuck	t^k	3	2	28	0.046	0.0014	0.0019	
tug	t^g	3	3	18	0.034	0.0014	0.0013	
van	v@n	3	32	12	0.066	0.0025	0.0144	
vase	ves	3	4	16	0.043	0.0012	0.0030	
vent	vEnt	4	10	16	0.076	0.0039	0.0095	0.0158
vest	vEst	4	4	19	0.058	0.0035	0.0069	0.0232
vile	vYl	3	5	22	0.020	0.0025	0.0003	
vine	vYn	3	4	21	0.051	0.0025	0.0048	
wade	wed	3	2	24	0.029	0.0014	0.0023	
wage	weJ	3	56	17	0.020	0.0014	0.0007	
wet	wEt	3	53	29	0.053	0.0024	0.0049	
win	wIn	3	55	31	0.071	0.0046	0.0095	
wine	wYn	3	72	30	0.050	0.0016	0.0048	
wing	wIG	3	18	18	0.043	0.0046	0.0032	
wish	wIS	3	110	13	0.041	0.0046	0.0012	
wit	wIt	3	20	31	0.061	0.0046	0.0059	
zap	z@p	3			0.040	0.0000	0.0048	
zip	zIp	3	1	14	0.045	0.0005	0.0049	
zit	zIt	3			0.055	0.0005	0.0059	
zone	zon	3			0.049	0.0005	0.0022	
zoo	zU	2			0.012	0.0001		

1

2

Appendix C

<i>Item 1</i>	<i>Item 2</i>
ban	man
base	vase
bent	vent
bid	bill
bin	win
cab	calve

cap	calf
craze	crane
dam	jam
date	late
deed	need
doze	dole
dub	dove
dust	just
fan	van
had	has
have	ham
lobe	loaf
met	mess
pair	fair
pat	pass
pin	fin
raid	rain
sub	sum
take	shake
tend	send
tip	sip
vine	wine
wade	wage
wit	wish
zip	nip
zone	lone
bag	bad
bake	bait
big	bid
cage	cave
cake	take
cat	pat
cut	cup
done	bun
goat	dote
goat	boat
gun	done
jest	vest
kill	pill
late	rate
leash	leaf
let	wet
lick	lip

moon	noon
rage	rave
rid	rib
same	sane
sit	fit
sit	sip
sound	found
sum	sun
tack	tap
tin	pin
tug	tub
wing	ring
bus	buzz
but	bud
call	gall
cap	gap
cap	cab
cut	gut
dock	dog
dose	doze
file	vile
half	halve
kill	gill
life	live
loose	lose
pat	bat
pat	pad
peck	peg
pit	bit
rip	rib
sack	sag
sap	zap
sat	sad
sit	zit
sue	zoo
tack	tag
tad	dad
tip	dip
tuck	duck
tuck	tug
tuck	tug

1

2

Appendix C

Partial correlations controlling for hearing thresholds for tasks from Experiment 1b and Experiment 2. Partial correlations using Pure Tone Average (Dalton et al., 2003) as a covariate were nearly identical to those presented here. 95% CI's were obtained using a bootstrapping procedure with n=1000 samples. Note that SD=syllable discrimination, WD=word discrimination, ALD=auditory lexical decision, PWM=picture word matching, AWSM=auditory written syllable matching.

Tasks	<i>r</i>	95% CI	n
SD and WD	0.95	.89 - .99	12
SD and ALD	0.73	.52 - .91	12
WD and ALD	0.75	.60 - .90	12
SD and PWM	0.83	.62 - .95	11
WD and PWM	0.85	.60 - .98	11
ALD and PWM	0.82	.70 - .95	11
SD and AWSM	0.17	-.35 - .71	8
WD and AWSM	0.37	-.59 - .81	8
ALD and AWSM	0.35	-.46 - .96	8
PWM and AWSM	0.43	-.64 - .94	8

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Author Note

Department of Psychology, Rice University, 6100 Main Street, Houston, TX 77005

This research was supported in part by funding for the Neuroplasticity Research

Laboratory from the T.L.L. Temple Foundation.

Corresponding author: Randi Martin, Department of Psychology, Rice University, 6100

Main Street MS-25, Houston, TX 77005, Telephone: 713-348-3417, Email:

rmartin@rice.edu

Acknowledgements: We are grateful to Simon Fischer-Baum and David Lane of the

Department of Psychology at Rice University for their assistance and statistical guidance

on this project. We would also like to thank our research assistants Blaine Tomkins,

Tristanne Mock and Maddie Flores. Finally, we would like to thank the patients for their

participation in this study.