Brain and Language **73**, 442–455 (2000)
doi:10.1006/brln.2000.2319, available online at http://www.idealibrary.com on IDEAL®

Perception of Dynamic Acoustic Patterns by an Individual with Unilateral Verbal Auditory Agnosia

Emily Wang,* Richard K. Peach,*,† Yi Xu,‡ Michael Schneck,†
and Charles Manry II*

**Department of Communication Disorders and Sciences, Rush University and Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; †Department of Neurological Sciences, Rush University and Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; and ‡Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois*

Previous studies have found that subjects diagnosed with verbal auditory agnosia (VAA) from bilateral brain lesions may experience difficulties at the prephonemic level of acoustic processing. In this case study, we administered a series of speech and nonspeech discrimination tests to an individual with unilateral VAA as a result of left-temporal-lobe damage. The results indicated that the subject's ability to perceive steady-state acoustic stimuli was relatively intact but his ability to perceive dynamic stimuli was drastically reduced. We conclude that this particular aspect of acoustic processing may be a major contributing factor that disables speech perception in subjects with unilateral VAA. © 2000 Academic Press

Key Words: verbal auditory agnosia; pure word-deafness; speech acoustics; speech perception; auditory perception.

INTRODUCTION

An individual with verbal auditory agnosia (VAA), or pure word-deafness, is unable to perceive spoken language in the absence of a significant hearing impairment (Goldstein, 1974; Benson, 1996). His reading, spontaneous writing, and speaking abilities are relatively preserved. Since verbal auditory agnosia is a relatively rare disorder, studies reporting findings on this disorder have been mainly case reports. The pathology of VAA, according to Benson (1996, p. 304), "involves the primary auditory cortex (Heschl's gyrus) or connections between the thalamus and this area." However, reported cases indicate that VAA could have resulted from either cortical or subcortical, unilateral or bilateral brain lesions (e.g., Kanshepolsky, Kelley, & Waggener, 1973; Saffran, Marin, & Yeni-Komshian, 1976; Auerbach, Allard, Naeser, Alexander, & Albert, 1982; Buchtel & Stewart, 1989; Kazui, Naritomi, Sawada, & Inoue, 1990; Godefroy, Leys, Furby, Reuck, Daems, Rondepierre, Debachy, Deleume, & Desaulty, 1995). Auerbach et al. (1982) suggested that there might be two different types of VAA: one is prephonemic in nature due to bilateral temporal lobe lesions and the other phonemic in nature due to left unilateral lesions because the former could be explained by deficits in the temporal auditory acuity observed in these patients while the latter could not. A review of the literature on unilateral VAA, however, indicates the following: on the one hand, hearing acuity, as defined by the ability to perceive static pitch, timbre, and loudness and even sound localization, is adequate in these right-handed individuals with left temporal lobe lesions; therefore, the speech-perception deficits observed cannot be readily explained by the normal or near-normal hearing acuity (e.g., Gazzaniga, Glass, Sarno, & Posner, 1973; Albert & Bear, 1974; Denes & Semenza, 1975; Saffran, Marin, & Yeni-Komshian, 1976). On the other hand, the ability to process various other aspects of acoustic signals may be impaired in these individuals, which may account for in part, if not all, the speech-perception deficits observed (Albert and Bear, 1974; Saffran *et al.*, 1976). For example, Albert and Bear (1974) found that their subject had a significant deficit in auditory temporal processing for both linguistic and nonlinguistic acoustic signals. On a nonlinguistic click-fusion task, he consistently fused two clicks at intervals of 15 ms while the normal controls were able to distinguish them as two at 1- to 3-ms separations. On a linguistic task, they found that if they reduced the speaking rate to 1/3 of the normal rate, the subject's comprehension significantly improved. One interesting observation was reported in several studies that investigated VAA involving either left unilateral or bilateral lesions. That is, the subjects were able to perceive simple vowels but not CV or CVC syllables (e.g., Denes & Semenza, 1975; Auerbach *et al.*, 1982; Kazui *et al.*, 1990; Godefroy *et al.*, 1995). These findings indicate that (1) prephonemic processing may also be affected in the VAA associated with left unilateral lesions and (2) the ability for processing steady-state acoustic information (such as vowels) vs the ability for processing rapidly changing acoustic information (such as formant transitions) may be differentially affected in these individuals.

The acoustic signal of speech is characterized by rapid temporal changes in fundamental frequency, intensity, and spectral properties. The ability to understand spoken language or speech thus requires the ability to process the temporal variations in the acoustic signal of speech. If one's ability for processing such variations is impaired, his or her ability to process spoken language or speech is limited. The reverse, however, is not necessarily true. Intact auditory processing ability may not guarantee intact comprehension

---

We thank all our subjects for their participation, Dr. Patricia McCarthy, and the anonymous reviewers for their valuable comments on an earlier version of this article.

Address correspondence and reprint requests to Emily Wang, 1653 West Congress Parkway, 203 SENN, Rush-Presbyterian-St. Luke's Medical Center, Chicago, IL 60612. Fax: (312) 243-8280. E-mail: ewang@rush.edu.

0093-934X/00 $35.00

of spoken language since language processing also requires the ability to analyze other aspects of language, such as syntax, semantics, and pragmatics. The fact that a VAA subject can comprehend written language suggests that his or her ability to process these aspects of the language is largely retained. Further, the fact that a VAA subject can perceive simple vowels but not CV or CVC syllables suggests that it is the ability to process the temporal changes in the speech signal that is probably responsible for the observed impairment in auditory comprehension of spoken language or speech.

There are a variety of temporal changes in speech signals. There are click-like sounds known as transients with durations as short as only a few milliseconds (Fant, 1973); there are rapid spectral changes near the boundaries of consonants and vowels known as formant transitions whose durations are typically around tens of milliseconds (Lehiste & Peterson, 1961); there are also relatively slower spectral changes within a vowel or diphthong with durations as long as hundreds of milliseconds (Lehiste & Peterson, 1961). Past speech perception studies have found that all of the temporal changes discussed above carry important cues for the identification of phonemes—the smallest sound units in speech. Both the transients and formant transitions are important for identifying stops (Liberman, Delattre, & Cooper, 1952); slower formant changes are important for identifying diphthongs and glides (Gay, 1968; Bladon, 1985). Kraus, McGee, and Koch (1998) reported their findings on neurophysiologic responses and behavioral discrimination of speech sound contrasts in normal and learning-disabled children. They presented the children with two pairs of phonetic contrasts (a /da/–/ga/ contrast and a /ba/–/wa/ contrast) and measured both their neurophysiologic and behavioral responses (for a detailed description of the study, please see Kraus *et al.*, 1998). As a group, the learning-disabled children had more difficulty in discriminating the /da/–/ga/ contrast than the /ba/–/wa/ contrast. If we look at the spectral analyses, it is clear that /wa/ in the /ba/ – /wa/ contrast has a longer F2 transition than those of stops. Thus, it is possible that this longer transition is within the temporal processing capacity of these children while the shorter transitions (in syllables with stops) are not. This finding, although observed in a different clinical population, nonetheless indicates that the ability to process relatively short dynamic patterns in speech acoustic signals is a crucial one since the transition differences among these stimuli are in tens of milliseconds. Albert and Bear (1974) observed that their subject could comprehend spoken language significantly better when the examiner reduced his speaking rate considerably to 1/3 of the normal rate. Although these authors did not test the subject's ability to process rapid formant transitions or even moderately slow temporal changes such as those in diphthongs, their finding that VAA subjects had greater perceptual fusion of sequence of clicks than normals suggests that the subject's ability to process release bursts in consonants is probably hampered. The ability of individuals with VAA to process temporal changes lasting for tens or hundreds of milliseconds, however, is still unclear. If this ability is impaired, it is conceivable that failure in speech perception should also occur.

In the present study, we examined the ability to process dynamic versus static variations in acoustic signals in a subject who presented with clinical behaviors typical of unilateral verbal auditory agnosia. We focused in particular on the subject's ability to process rapid or moderately slow temporal acoustic changes lasting either tens or hundreds of milliseconds. In light of existing findings about VAA, and based on what is known about speech perception in general and in learning-disabled children, we hypothesized that our subject's ability to process the rapid dynamic acoustic variations was impaired, but his ability to process moderately rapid dynamic acoustic variations was preserved. Hence, our subject should be able to process steady-state sounds such as vowels and moderately rapid dynamic acoustic variations such as those in diphthongs, but not able to process rapid dynamic acoustic variations such as those in stops or CV syllables. To test this hypothesis, we used various speech and nonspeech signals as stimuli in a series of perception tests administered to the subjects.

CASE DESCRIPTION

The subject is a 31-year-old, right-handed man, with a known history of testicular cancer for which he underwent surgery and chemotherapy 10 years prior to being seen for this study. He also was diagnosed with congenital hypertrophic cardiomyopathy at 5 years of age. A pacemaker was placed after syncopal episodes at age 24. He was admitted to Rush-Presbyterian-St. Luke's Medical Center (RPSLMC) with a diagnosis of cerebrovascular accident (CVA) following sudden onset of confusion, aphasia, and right hemiparesis. Upon admission, the subject underwent a carotid angiogram with urokinase infusion that revealed an occlusion of the left middle cerebral artery, and the intraclot lysis with urokinase was successful. A noncontrast computed tomographic (CT) scan of the head was performed that revealed an infarct in the territory of the left middle cerebral artery. Repeat noncontrast CT scan performed 2 days later revealed continued evolution of focal infarction in the left cerebral hemisphere involving the cortical and subcortical white matter of the left temporal lobe and extending superiorly into the frontoparietal region (see Fig. 1).

Neurological Testing

The neurological examination on admission revealed that the subject was confused, but occasionally uttered a clear spontaneous word. He had dense comprehensive deficits. He also demonstrated a right facial droop and had a right-sided weakness. Two days after admission, the neurological examination revealed the following: the subject's Mini Mental State (Folstein, Folstein, & McHugh, 1975) was appropriate to written input but not to spo-

FIG. 1. A noncontrast CT scan performed 2 days postadmission which revealed continued evolution of focal infarction in the left cerebral hemisphere involving the cortical and subcortical white matter of the left temporal lobe and extending superiorly into the frontoparietal region.

ken conversation. The subject achieved a score of 21/30 on the Mini Mental State Exam when the test items were presented visually rather than auditorily. His spontaneous speech was described as being appropriate with occasional paraphasias. He was unable to comprehend spoken language or follow verbal commands but could follow written commands and read and write. He responded to a tuning fork. His cranial nerve response was 20/30 on the right and 20/20 on the left. Strength was 5/5 bilaterally for the upper extremities. Reflexes were symmetric. His fine motor ability was slightly decreased in the right hand. He had right hand and right foot sensory loss. He had no symptoms of ataxia and his gait was intact.

Initial Language Testing

The Western Aphasia Battery (WAB) was administered 4 days after the admission. At the time of testing, the subject demonstrated appropriate pragmatic skills and fluent grammatical verbal output of normal phrase length with phonemic paraphasias. Auditory comprehension was severely impaired even for simple yes/no questions. Repetition was severely impaired, while naming to confrontation was moderately to markedly impaired. Reading comprehension was relatively preserved. The subject clearly exhibited contrasting abilities in comprehending language materials presented visually rather than auditorily. His performance was consistent with a diagnosis of pure word-deafness, i.e., relatively preserved visuoverbal language with severely impaired auditory-verbal comprehension (see Table 2).

TABLE 1
Pure Tone Threshold (in decibels) Results at 10-Day Postonset

Frequency (Hz)	Left ear	Right ear
250	15	25
500	20	25
750	25	35
1000	30	45
2000	30	40
4000	25	45
8000	10	25

Audiological Testing

An audiological evaluation was conducted 10 days postadmission that included pure-tone audiometry, tympanometry, acoustic reflex testing, speech awareness threshold, and speech reception and recognition testing. The results from pure-tone threshold audiometry (see Table 1) showed that the subject's hearing was within normal limits for the left ear at 250–2 and 4–8 kHz and a mild sensorineural hearing loss was detected at 1–2 kHz. For the right ear, his hearing was within normal limits at 250–500 kHz and 8 kHz but a mild-to-moderate sensorineural hearing loss was detected at 750–4 kHz. The hearing loss observed in this subject might have been related to unknown etiology since he denied family history of hearing loss, history of noise exposure, and history of ototoxic medication. Tympanogram and speech awareness were within normal limits, and acoustic reflexes were present at expected levels for both ears. Speech-reception and -recognition testing could not be completed due to the fact that the subject could not understand any speech information even when he was provided with a list of the words for speech-reception testing.

METHODS

Language Testing

The WAB test was readministered in a face-to-face clinical setting without any effort to conceal the examiner's face. However, even with such visual aids as lip reading to assist his

TABLE 2
Results from WAB Testing

WAB subtest	Initial	1 week postonset	1 month postonset	5 months postonset
Spontaneous Speech (20)	15	16	18	19
Auditory Comprehension (10)	3.85	4.45	9.75	9.75
Repetition (10)	1.4	3.3	7.2	7.8
Naming (10)	3.1	6.8	8.4	9.0
Aphasia Quotient (93.8)	46.7	61.1	86.7	91.1
Reading (100)	68	100	94	100

auditory comprehension, the subject continued to demonstrate severely impaired comprehension for language materials presented auditorily. Treatment was provided by a licensed speech-language pathologist on a daily basis after the initial language evaluation. Emphasis was placed on training the subject to use visual information to aid his auditory comprehension. Consistent improvement was observed in this area.

As can be seen from Table 2, the subject's *inability* to comprehend and repeat information presented auditorily was in clear contrast with his *ability* to comprehend complex information presented visually. While the subject's ability to comprehend written language reached a ceiling 7 days postonset, his ability to comprehend spoken language remained limited. He continued to improve on the auditory comprehension and repetition subtests at 1 month postonset, although these scores largely reflected his improved ability to use visual cues following treatment. Evidence for this conclusion was obtained from the subject's performance on a series of discrimination/identification tasks that were administered concurrently with language testing.

Perception Testing—Subject

Perception testing was performed 1 week postonset of the subject's CVA. The stimuli were delivered via headphones in a quiet room at an intensity level of approximately 70 dB SPL. Two speech perception tests were administered: identification and discrimination. Twenty-four minimal pairs of unique, single-syllable natural speech stimuli were used for both tests, with half of them differing in vowels and the other half differing in initial or final consonants (Appendix A). For the identification test, each stimulus was presented twice in a random order with an ISI of 5000 ms, resulting in a total of 96 responses. The subject's task was to circle on an answer sheet the word he just heard from the pair of written stimuli presented visually. For the discrimination test, the stimuli were presented in the original minimal pairs. To eliminate the order effect, the stimuli were counterbalanced, i.e., each stimulus was presented as the first one as well as the last one within each pair. After counterbalancing, each pair of stimuli was presented twice randomly to the subject, resulting in a total of 192 responses. The subject's task was to circle the word "same" or "different" on the answer sheet after he heard each stimulus pair. Prior to the testing, written instructions for the tasks were provided to the subject. Practice runs were conducted until he felt comfortable with the tasks and was able to carry them out as required. The subject performed at chance level for both tests (45.83% for identification and 46.32% for discrimination), indicating an inability to perceive the phonetic differences in the stimuli.

At 1 month postonset, a series of discrimination tests were administered (see Table 3). For Tests A and B (see below), the stimuli were delivered via headphones in a sound-treated suite. For Tests C–F, the stimuli were delivered via headphones in a quiet room at an intensity level of approximately 70 dB SPL. All natural speech stimuli were produced by a male native speaker of American English from the same region as the subject. They then were digitized and sequenced using the SoundEdit program (Macromedia).

TABLE 3
Result Summary from Perception Testing at 1 Month Postonset

Discrimination test	No. of total responses	% Correct—subject (n = 1)	% Correct—controls (n = 15)
Amplitude	20	95	—
Frequency	20	95	—
Simple Vowels	72	100	—
Diphthongs	72	100	—
Word Stress	64	60.93	99.69
Sine Wave with Dynamic Changes	64	60.93	98.13

(A) *Intensity discrimination.* The subject's ability to detect differences in sound intensity was measured. The discrimination test for intensity consisted of four unique pairs of steady-state pure-tone stimuli. Two intensity levels were used: 60 and 75 dB HL at 2 kHz. These levels were minimally 20 dB HL above the subject's pure-tone thresholds (see Table 1). Each stimulus was presented at a duration of 1000 ms with an ISI of 500 ms. Each pair was presented five times in a random order resulting in a total of 20 responses. The subject's task was to circle the word "same" or "different" on the answer sheet after he heard each pair of stimuli. He performed with 95% accuracy.

(B) *Frequency discrimination.* The subject's ability to detect differences in sound frequency was measured. The Pitch Pattern Sequences (Pinheiro, 1977) were used for this task. The test consists of four pairs of steady-state pure-tone stimuli. Two frequency levels are used: 880 and 1430 Hz. The stimuli were delivered at 85 dB HL, which is 50 dB HL above the subject's pure-tone-average (PTA) threshold. Each pair was presented five times with an ISI of 500 ms in a random order, resulting in a total of 20 responses. Again, the subject performed with 95% accuracy.

(C) *Simple-vowel discrimination.* The discrimination test for simple vowels consisted of three naturally produced tense vowels [a, i, u]. After counterbalancing, there were nine pairs of unique stimuli. Each pair was presented eight times with an ISI of 500 ms in a random order, resulting in a total of 72 responses. The subject performed with 100% accuracy.

(D) *Diphthong discrimination.* The discrimination test for diphthongs consisted of three naturally produced diphthongs ([aɪ, ou, eɪ] at 590, 570, and 530 ms respectively). After counterbalancing, there were nine pairs of unique stimuli. Each pair was presented eight times with an ISI of 500 ms in a random order, resulting in a total of 72 responses. The subject performed with 100% accuracy.

(E) *Word-stress discrimination.* The discrimination test for word stress consisted of two pairs of naturally produced words differing in the primary word stress, e.g., "Whitehouse" vs "white house," "hotdog" vs "hot dog." After counterbalancing, there were eight pairs of unique stimuli. Each pair was presented eight times with an ISI of 500 ms in a random order, resulting in a total of 64 responses. The subject performed with 60.9% accuracy. A one-sample t test with a hypothesized mean of 0.5 was conducted to determine whether the performance was different from chance. The result was not statistically significant ($t = -1.779$, $p = .080$). A one-factor ANOVA was conducted to determine if the subject's performance was any different for the two stimuli. The result was not statistically significant ($F = .064$, $p = .802$).

(F) *Discrimination of sine waves with dynamic changes.* The discrimination test for sine waves with dynamic changes consisted of two pairs of synthesized stimuli. One pair had no steady state (rising–falling), i.e., their frequencies changed continuously (800–1250 and 1250–

800 Hz). The duration of the stimuli was 302 ms. The other pair had a short onset ramp (650–1000 Hz signaling "up" and 1350–1000 Hz signaling "down") followed by a steady-state tone of 1000 Hz. The duration of the onset ramp was 50 ms for both "up" and "down." The total durations of the stimuli (including the duration of steady-state tone) were 292 ms for "up" and 312 ms for "down." After counterbalancing, there were eight pairs of unique stimuli, i.e., four pairs for the "rising–falling" type and four for the "up–down" type. Each pair was presented eight times with an ISI of 500 ms in random order, resulting in a total of 64 responses. The subject performed with 60.9% accuracy. A one-sample t test with a hypothesized mean of .5 was conducted to determine whether the performance was different from chance. The result was not statistically significant ($t = -1.779$, $p = .08$), indicating that the subject's performance was no better than chance. A one-factor ANOVA was conducted to determine if there was any difference in the subject's performance for the two types of stimuli (i.e., continuous changing vs with an onset ramp). The result was not statistically significant ($F = .578$, $p = .450$).

Perception Testing—Normal Controls

In order to determine if the stimuli were too difficult to discriminate even for normal subjects, 15 control subjects participated in discrimination tests E and F (see below). The subjects were 1st- and 2nd-year graduate students in the speech-language pathology program at Rush University with ages ranging from 22.5 to 34.4 years old. All control subjects were free from speech and hearing impairment. Two of the 15 control subjects were older than the subject. Both discrimination tests were conducted under the same experimental conditions used for the subject. The results are presented below.

(E) *Word-stress discrimination.* The mean percentage correct for the discrimination of word stress was 99.69% (ranging from 95.31 to 100% correct). A one-sample t test with a hypothesized mean of 0.5 was performed. The result was statistically significant ($t = 279.874$, $p < .0001$), indicating that the control subjects' performance as a group was well above chance. A one-factor ANOVA was conducted to compare the performance of the subject with the group means of the control subjects. The result was statistically significant ($F = 6.324$, $p < .0001$), indicating that as a group, the control subjects performed better than the subject.

(F) *Discrimination of sine waves with dynamic changes.* The mean percentage correct for discrimination of sine waves with dynamic changes was 99.97% (ranging from 99.90 to 100% correct) A one-sample t test with a hypothesized mean of .5 was performed. The result was statistically significant ($t = 91.790$, $p < .0001$), indicating that the control subjects' performance as a group was well above chance. A one-factor ANOVA was conducted to compare the performance of the subject with the group means of the control subjects. The result was statistically significant ($F = 6.080$, $p < .0001$), indicating that the control subjects performed better than the subject.

In order to determine if the control subjects performed differently for the two types of stimuli, the percentage correct scores were examined. Of the 15 control subjects, 14 scored 100% correct and one 99.9% correct for the rising–falling type, while seven scored 100% correct and eight ranged from 99.8 to 99.9% correct for the up–down type. A one-factor ANOVA was conducted to determine if the observed difference between the two types of stimuli was statistically significant. The result was indeed statistically significant ($F = 7.314$, $p < .0115$). This seems to indicate that for the normal subjects, the up–down type of stimuli was perceptually more challenging than the rising–falling type, although they all performed well above chance.

DISCUSSION

Although verbal auditory agnosia has been described as an inability to interpret the meanings of verbal sounds (a phonemic decoding disorder, see Goldstein, 1974), in the absence of hearing difficulties, recent studies have shown that certain aspects of processing speech acoustic signals may be affected (Kanshepolsky *et al.*, 1973; Albert & Bear, 1974; Saffran *et al.*, 1976; Buchtel & Stewart, 1989). Even though such changes were observed more frequently in cases with bilateral temporal cerebral lesions (e.g., Kanshepolsky *et al.*, 1973; Glass, Sarno, & Posner, 1973; Auerbach *et al.*, 1982), they were observed in cases with left unilateral temporal cerebral lesions as well. Albert & Bear (1974) reported decreased ability to handle temporal resolution in their subject, and Saffran, Marin, & Yeni-Komshian (1976) reported the need for increased intensity level for sound processing in their subject. These findings suggest that it is possible that in certain cases of unilateral VAA, similar to VAA involving bilateral temporal cerebral lesions, acoustic processing of speech signals may be affected at the prephonemic level rather than solely at the phonemic level, as suggested by Auerbach *et al.* (1982).

In this study, we have found that a subject with verbal auditory agnosia from focal infarction of the left cerebral hemisphere, had no difficulty detecting differences in the intensity and frequency of nonspeech stimuli nor among simple vowels or even diphthongs. However, the same subject showed great difficulty discriminating sine waves with dynamic changes, minimal-pair words, and word stress. His difficulty could not be explained by changes in his hearing since all stimuli were presented well above his pure-tone thresholds. Inspection of those stimuli that the subject could discriminate and those he could not revealed that it was his ability to handle dynamic changes in acoustic signals that was really impaired differentially.

The subject's inability to perceive rapid dynamic variations lasting 50 ms is consistent with his inability to discriminate different consonant places of articulation (such as /ba/ vs /da/). It has been well established that formant transitions, whose duration is usually around 50 ms, are critical for identifying consonant place of articulation (Liberman *et al.*, 1952). VAA subjects who are unable to process either click-like sounds (Albert & Bear, 1974), which are similar to the release bursts in stop consonants, or pure-tone glides with durations similar to that of formant transitions, as found in the present study, have lost the ability to detect two of the most important acoustic cues for the perception of stop consonants. It follows therefore that such subjects cannot properly perceive stop consonants in speech. Phillips and Farmer (1990) analyzed existing evidence and concluded that the temporal processing deficit seen in VAA subjects was largely restricted to sound elements that are very brief in duration (i.e., from milliseconds to tens of milliseconds). In this study, we have taken a step further to show that the deficits in VAA subjects are not only in their inability to process sound elements that are short in duration, but also their inability to process dynamic acoustic changes that occur in a short period of time.

What is surprising is the finding that the auditory impairment of this VAA subject was so severe that he could not even discriminate pure-tone glides (rising–falling) as long as 300 ms. For the normal controls in this study, the continuous tone glides (rising–falling type) were slightly easier to discriminate than the transitionlike tone glides (up–down type). However, the subject's performance suggests that the two types of dynamic stimuli, rising–falling and up–down, were equally challenging to him and that he could not perform better than chance for either type of stimuli. Many dynamic patterns in speech last only as long as several hundred milliseconds. The average syllable duration in spontaneous speech is 256–333 ms (summarized by Ohde & Sharf, 1992, p. 271), and the speaking rate ranges from 150 to 250 words/min or four to seven syllables/s (Klatt, 1976). The dynamic spectral changes of vowels, diphthongs, and glides all occur within the duration of a syllable. If VAA subjects cannot detect dynamic changes in acoustic signals lasting as long as 300 ms, it is unlikely that they can process speech utterances in any practical way.

It is interesting that this subject was able to discriminate the diphthong pairs used in this study. There may be two reasons for this result. First, the duration of these stimuli might have been sufficiently long for him. In a study examining the patterns of discrimination between tone glides that were consonant–vowel-(CV) and vowel–consonant- (VC) like formant transitions, Van Wieringen and Pols (1995) found that as the transition duration increased, the discriminability improved even when the total duration of the stimulus remained constant. The improved discriminability was due to the fact that as the transition duration increased, all difference limens decreased, the spectral change decreased, and the processing time increased. The longest transition duration these authors used was 50 ms. In the current study, the subject was unable to discriminate stimuli that were as long as 310 ms in duration, indicating a serious problem in his ability to process dynamic changes. However, the shortest duration of our diphthong stimuli was 530 ms (since they were all produced in isolation). This may have exceeded his difference limen for detecting dynamic changes. The second reason for his successful discrimination of the diphthong stimuli may be the fact that they were not really minimal pairs in the sense that the only difference in each pair is in the dynamics rather than anything else. The three diphthongs used in this study, /aɪ/, /ou/ and /eɪ/, differed not only in their dynamic patterns, but also in their starting and ending formant values. In other words, to the subject, these diphthongs may have sounded just like different steady-state vowels.

Speech perception involves a long chain of events beginning with hearing. The acoustic signal picked up by the ears has to be processed further to reveal patterns critical for the identification of different speech sounds. As argued by Liberman and Mattingly (1985, p. 28), speech perception "uses all the information in the stimulus that is relevant to phonetic structures: every potential cue proves to be an actual cue." If, however, somewhere along the way in the auditory system, the mechanism for detecting dynamic changes in the acoustic signal is damaged, as found in VAA, conversion of dynamic patterns into the phonetic forms needed for higher processing would be hindered, making it impossible for the listener to recognize spoken words or sentences. In humans, the ability to process dynamic information such as formant transitions and transients in speech signals is highly remarkable. In normals, speech signals can be processed at a speed of seven syllables/s (Klatt, 1976). However, when the auditory system is damaged, this ability may be greatly affected, ranging from greatly reduced to nonfunctional, depending on the amount and the location of the damage. Evidence supporting this argument can also be found in various case studies showing that VAA subjects could discriminate simple vowels but not CV or CVC syllables (e.g., Denes & Semenza, 1975; Auerbach *et al.*, 1982; Kazui *et al.*, 1990; Godefroy *et al.*, 1995) and in the phenomenon of rate-dependent comprehension as observed by Albert and Bear (1974).

In conclusion, based on the findings of the present study, it seems that what is critically affected in this type of unilateral verbal auditory agnosia, or at least in this subject, is the ability to process dynamic patterns in acoustic signal, which in turn impairs the subject's ability to perceive speech.

APPENDIX A

Perception Testing Stimuli-Minimal Pairs

Item No.	Phonetic transcription	Gloss	Phonetic transcription	Gloss
1	bæt	bat	pæt	pat
2	fæt	fat	kæt	cat
3	lɛd	led	let	let
4	læb	lab	læd	lad
5	biːt	beat	miːt	meat
6	niːt	neat	siːt	seat
7	hiːl	heal	hiːt	heat
8	sɛθ	Seth	set	set
9	gɛt	get	met	met
10	lɛd	led	ted	Ted
11	diːl	deal	diːd	deed
12	fiːt	feet	fiːd	feed
13	iːt	eat	æt	at
14	hiːt	heat	hæt	hat
15	piːtʃ	peach	pætʃ	patch
16	liːk	leak	læk	lack
17	fiːd	feed	fed	fed
18	miːt	meat	mɛt	met
19	biːd	bead	bed	bed
20	siːt	seat	set	set
21	pet	pet	pæt	pat
22	set	set	sæt	sat
23	lɛd	led	læd	lad
24	ʃed	shed	ʃæd	shad

REFERENCES

Albert, M., & Bear, D. (1974). Time to understand. *Brain,* **97,** 373–384.

Auerbach, A. H., Allard, T., Naeser, M., Alexander, M. P., and Albert, M. L. (1982). Pure word deafness-analysis of a case with bilateral lesions and a defect at the prephonemic level. *Brain,* **105,** 271–300.

Benson, D. F., & Ardila, A. (1996). *Aphasia.* New York: Oxford Univ. Press.

Bladon, A. (1985). Diphthongs: A case study of dynamic auditory processing. *Speech Communication,* **4,** 145–154.

Buchtel, H. A., & Stewart, J. D. (1989). Auditory agnosia: Apperceptive or associative disorder? *Brain and Language,* **37,** 12–25.

Denes, G., & Semenza, C. (1975). Auditory modality specific anomia: Evidence from a case of pure word deafness. *Cortex,* **11,** 401–411.

Fant, G. (1973). *Speech sounds and features.* Cambridge, MA: M.I.T. Press.

Folstein, M. D., Folstein, S. E., & McHugh, P. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatry Research,* **12,** 189–198.

Gay, T. J. (1968). Effect of speaking rate on diphthong formant movements. *Journal of the Acoustical Society of America,* **44,** 1570–1573.

Gazzaniga, M. S., Glass, A. V., Sarno, M. T., & Posner, J. B. (1973). Pure word deafness and hemispheric dynamics: A case history. *Cortex,* **9,** 136–143.

Godefroy, O., Leys, D., Furby, A., Reuck, J. D., Daems, C., Rondepierre, P., Debachy, B., Deleume, J.-F., & Desaulty, A. (1995). Psychoacoustical deficits related to bilateral subcortical hemorrhages—a case with apperceptive auditory agnosia. *Cortex,* **31,** 149–159.

Goldstein, M. N. 1974. Auditory agnosia for speech ("Pure word-deafness"). *Brain and Language,* **1,** 195–204.

Kanshepolsky, J., Kelley, J. J., & Waggener, J. D. (1973). A cortical auditory disorder. Clinical audiologic and pathologic aspects. *Neurology,* **23,** 699–705.

Kazui, S., Naritomi, H., Sawada, T., & Inoue, N. (1990). Subcortical auditory agnosia. *Brain and Language,* **38,** 476–487.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of Acoustic Society of America,* **59**(5), 1208–1221.

Kraus, N., McGee, T. J., & Koch, D. B. (1998). Speech sound representation, perception, and plasticity: A neurophysiologic perspective. *Audiology and Neuro-Otology,* **3,** 168–182.

Lehiste, I., & Peterson, G. E. (1961). Transitions, glides and diphthongs. *Journal of the Acoustical Society of America,* **33,** 268–277.

Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1952). The role of stimulus variables in the perception of stop consonants. *American Journal of Psychology,* **65,** 497–516.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition,* **21,** 1–36.

Ohde, R. N., & Sharf, D. J. (1992). Phonetic analysis of normal and abnormal speech. New York: Merrill.

Phillips, D. P., & Farmer, M. E. (1990). Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. *Behavioral Brain Research,* **40,** 85–94.

Pinheiro, M. L. (1977). Auditory pattern reversal in auditory perception in patients with left and right hemisphere lesions. *Ohio Journal of Speech and Hearing,* **12,** 9–20.

Saffran, E. M., Marin, O. S. M., & Yeni-Komshian, G. H. (1976). An analysis of speech perception in word deafness. *Brain and Language,* **3,** 209–228.

Van Wieringen, A., & Pols, L. C. W. (1995). Discrimination of single and complex consonant-vowel- and vowel-consonant-like formant transitions. *Journal of Acoustical Society,* **98**(3), 1304–1312.