

Cerebellum and Speech Perception: A Functional Magnetic Resonance Imaging Study

Klaus Mathiak, Ingo Hertrich, Wolfgang Grodd,
and Hermann Ackermann

Abstract

■ A variety of data indicate that the cerebellum participates in perceptual tasks requiring the precise representation of temporal information. Access to the word form of a lexical item requires, among other functions, the processing of durational parameters of verbal utterances. Therefore, cerebellar dysfunctions must be expected to impair word recognition. In order to specify the topography of the assumed cerebellar speech perception mechanism, a functional magnetic resonance imaging study was performed using the German lexical items “Boden” ([bodn], Engl. “floor”) and “Boten” ([botn], “messengers”) as test materials. The contrast in sound structure of these two lexical items can be signaled either by the length of the wordmedial pause (closure time, CLT; an exclusively temporal measure) or by the aspiration noise of wordmedial “d” or “t” (voice onset time, VOT; an intrasegmental cue). A previous study found bilateral cerebellar disorders to compromise word recognition based on CLT whereas the encoding of VOT remained unimpaired. In the present study, two series of “Boden–Boten” utterances were

resynthesized, systematically varying either in CLT or VOT. Subjects had to identify both words “Boden” and “Boten” by analysis of either the durational parameter CLT or the VOT aspiration segment. In a subtraction design, CLT categorization as compared to VOT identification (CLT – VOT) yielded a significant hemodynamic response of the right cerebellar hemisphere (neocerebellum Crus I) and the frontal lobe (anterior to Broca’s area). The reversed contrast (VOT – CLT) resulted in a single activation cluster located at the level of the supratemporal plane of the dominant hemisphere. These findings provide first evidence for a distinct contribution of the right cerebellar hemisphere to speech perception in terms of encoding of durational parameters of verbal utterances. Verbal working memory tasks, lexical response selection, and auditory imagery of word strings have been reported to elicit activation clusters of a similar location. Conceivably, representation of the temporal structure of speech sound sequences represents the common denominator of cerebellar participation in cognitive tasks acting on a phonetic code. ■

INTRODUCTION

A longstanding premise of clinical neurology, tracing back to the early nineteenth century, holds that cerebellar lesions and diseases compromise motor capabilities, such as the coordination of voluntary movements and the control of balance, while sparing sensory and cognitive functions (Gilman, Bloedel, & Lechtenberg, 1981; Dow & Moruzzi, 1958). Contrary to this notion, Ivry and Keele (1989) observed disrupted time perception in patients with cerebellar pathology: Subjects performed worse when asked to compare a pair of time intervals (each bound by clicks), the second item being either shorter or longer than the first one. Subsequent studies found, in addition, impaired representation of temporal information within the somatosensory and visual domains (see Ivry & Fiez, 2000). Since at least some cerebellar motor deficits, for example, increased variability of finger tapping in response to external

stimuli, might be due to disordered central timing mechanisms, Keele and Ivry (1990) considered the cerebellum an “internal clock” (Braitenberg, 1967) supporting precise temporal representation across motor and sensory functions (Kawashima et al., 2000; Jueptner et al., 1995; for a “generalized timing hypothesis,” see Ivry & Fiez, 2000).

Within some limits, distinct acoustic parameters characterize the various classes of speech sounds (phonemes) of any language system (Ladefoged & Maddieson, 1997). Among others, durational measures contribute to phoneme specification. The generalized timing hypothesis predicts impaired processing of those features of the acoustic speech signal. Accordingly, a previous clinical study found cerebellar atrophy to compromise temporal aspects of speech perception. For example, the English word “rabbit” involves a short period of silence (closure time, CLT), signaling the intraword stop consonant “b.” Prolongation of CLT from 20 to 130 msec gives rise to a phoneme boundary effect: Stimuli with short intraword pause (<40 msec) are recognized as the lexical item “rabbit,” whereas closure intervals

above approximately 80 msec yield the percept “rapid” (Liberman, 1996). During fluent speech, as a rule, no aspiration noise is produced in association with the wordmedial stop consonants. It must be expected, therefore, that recognition of the words “rabbit” and “rapid” under these conditions depends solely upon the encoding of a durational acoustic cue, that is, CLT signaling the intraword stop consonants “b” or “p.” Using a German analogue of the “rabbit–rapid” paradigm (“Boden” [bodn] = short CLT, Engl. “floor” vs. “Boten” [botn] = long CLT, “messengers,” see Figure 1a), patients with diffuse cerebellar atrophy, indeed, did not exhibit any significant phoneme boundary effect (Ackermann, Gräber, Hertrich, & Daum, 1997). Besides CLT, voice onset time (VOT) represents a further durational parameter contributing to the specification of speech sounds. For example, the syllables [da] and [t^ba] differ in VOT of the initial stop consonant, namely, the interval between initial burst and vowel onset (compare Figure 1b). English-speaking adults asked to label monosyllabic stimuli extending in VOT from less than 10 msec, yielding the percept [da], to 80 msec, recognized as [t^ba], show a rather abrupt shift between these two response categories at a value of about 30 msec (Lisker & Abramson, 1964, 1970). Similar identification curves across a continuum of verbal utterances varying in VOT of the initial stop consonant have been documented in cerebellar subjects

and their controls ([ba] vs. [p^ba], Ivry & Gopal, 1992; “dick” [dik], Engl. “thick” vs. “Tick” [t^bik], “tic,” Ackermann et al., 1997). It is well established that listeners may utilize during phonological encoding all available acoustic cues of the speech signal. In the case of VOT processing, the sound energy of the aspiration phase, that is, the loudness of this noise segment, can contribute to the perceived voicing contrast of stop consonants, such as “d” and “t” (see Ackermann et al., 1997 for further references). Consequently, categorical VOT distinction may merely reflect intrasegmental backward masking, that is, aspiration noise must exceed a given intensity threshold in order to be detected (Miller, Wier, Pastore, Kelly, & Dooling, 1976). In contrast to CLT, categorical VOT representations, therefore, do not necessarily require explicit time measurements. In accordance with the general timing hypothesis, cerebellar disorders, thus, seem to disrupt exclusively the temporal aspects of speech sound encoding.

Functional compartmentalization is a widely recognized principle of cerebellar organization (Dichgans, 1984). So far, disordered speech perception has only been found in patients with diffuse cerebellar atrophy. On the basis of these data, functional topography of speech perception in the cerebellum cannot be further specified. Functional magnetic resonance imaging (fMRI) enables detection and localization of cerebellar hemodynamic responses, like those in supratentorial structures (Grodd, Hülsmann, Lotze, Wildgruber, & Erb, 2001; Ellerman et al., 1994). Motor, perceptual, and cognitive tasks studied include odor perception (Yousem et al., 1997), speech production (Wildgruber, Ackermann, & Grodd, 2001; Hirano et al., 1997), and verbal working memory (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997; for a review, see Cabeza & Nyberg, 2000). The current study used fMRI to identify the cerebellar subsystem engaged in the encoding of temporal aspects of verbal utterances. In analogy to the clinical study on speech perception in cerebellar disorders, a series of 15 stimuli systematically varying in the durational parameter CLT was generated. It extended from a prototypical item of “Boden” (CLT = 27 msec) at one extreme to “Boten” (CLT = 167 msec) at the other (see Figure 1a for illustration). Word recognition of these “Boden–Boten” stimuli was required during the active condition.

The design of the present study sought to minimize any interference from other cognitive processes. Clinical and functional imaging data indicate that the cerebellum contributes to distinct lexical operations or verbal response selection (Fiez & Raichle, 1997). Furthermore, this part of the brain has been assumed to participate in the regulation of attentional resources, as well as executive functions (Allen, Buxton, Wong, & Courchesne, 1998; Courchesne & Allen, 1997; Akshoomoff & Courchesne, 1992). Hierarchical subtraction designs comparing phonetic discrimination either with passive

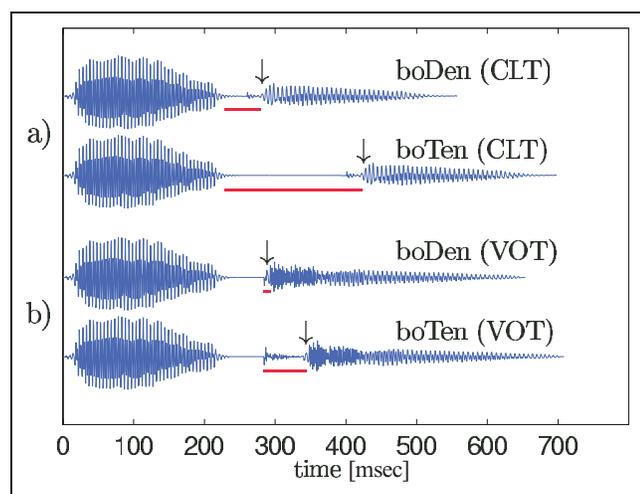


Figure 1. Prototypical examples of the two series of “Boden–Boten” utterances. (a) The first two waveforms differ in CLT, that is, the underlined pause signaling wordmedial “d” or “t,” respectively, all the other components being identical (arrows = voice onset of the second syllable). The utterance with short CLT ([bodn]) is recognized as the word “Boden” and with long CLT ([botn]) as “Boten.” (b) The lower two acoustic speech signals differ only in VOT, that is, the underlined noise segment extending from the burst to the beginning of the vocalic segment (= onset of the schwa vowel; see arrows). A series of equally spaced stimuli was derived from these two pairs of acoustic signals (CLT: 27 . . . 167 msec, VOT: 9 . . . 69 msec). The arrows mark the onset of the voiced portion of the second syllable.

listening to acoustic stimuli or encoding of nonspeech auditory materials as a control condition may therefore be confounded by interactions with other cognitive processes (see Ivry & Fiez, 2000). In order to exclude such interference, subtraction was conducted between two behaviorally identical tasks. Besides intersegmental CLT, the difference in sound structure between the words “Boden” and “Boten” may be signaled by the VOT of the wordmedial stop consonants “d” and “t” ([bod↔n] vs. [bot^b↔n]), the “t” sound being characterized by an aspiration noise preceding the schwa vowel; see Figure 1b). Thus, a second series of verbal utterances with constant CLT but systematically varied wordmedial VOT was generated. Subjects had to distinguish the lexical items “Boden” versus “Boten” thus either by analysis of the intersegmental durational parameter (CLT, active condition) or by the characteristics of the noise segment (VOT, reference condition). Furthermore, a monosyllabic control condition ([d↔n] vs. [t^b↔n], MON) was added in order to allow assessment of eventual segmentation effects (see Burton, Small, & Blumstein, 2000). Randomized and balanced blocks of these three stimulus series (CLT, VOT, and MON conditions) served as the test materials. Hemodynamic responses were measured by means of fMRI across the entire brain covering both supra- and infratentorial structures. Furthermore, reaction times (RTs) were recorded and identification curves were determined across conditions. Based on findings obtained with verb generation and other tasks, impaired error detection and correction has been considered a possible mechanism of cerebellar disorders in the linguistic domain (see Ivry & Fiez, 2000). Furthermore, a correlation between task difficulty and degree of cerebellar activation has been assumed (Ivry, 1997). RT measure-

Figure 2. Group averages of (a) response probabilities for the percept “boTen” (or “Ten”) and (b) RTs plotted against CLT and VOT. Arrows indicate the average transition point of the (s slope) identification curves as a correlate of the phoneme boundary effect. There is an increase in RT in response to ambiguous stimuli, that is, events near the respective categorical boundary. The dashed line in the RT plots represents the onset of the voiced part of the last syllable (= auditory onset; equivalent to arrow inserts in Figure 1). With CLT and VOT lengthening, this time point increases. The trend of RT in the CLT, conceivably, reflects the later onset of this relevant cue.

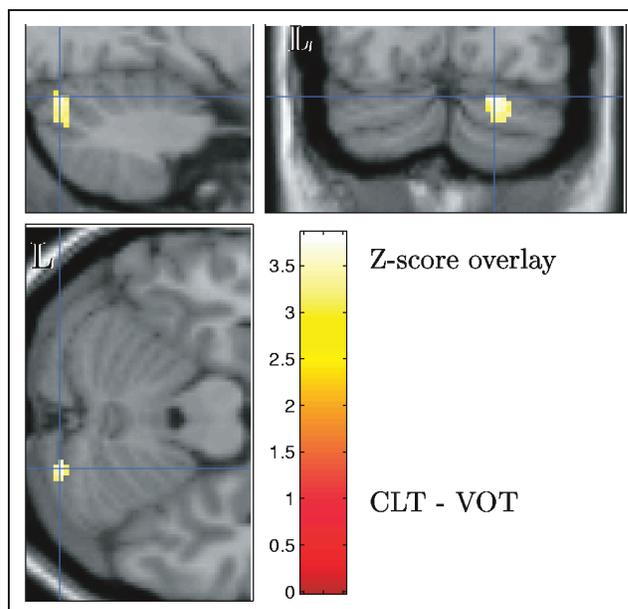
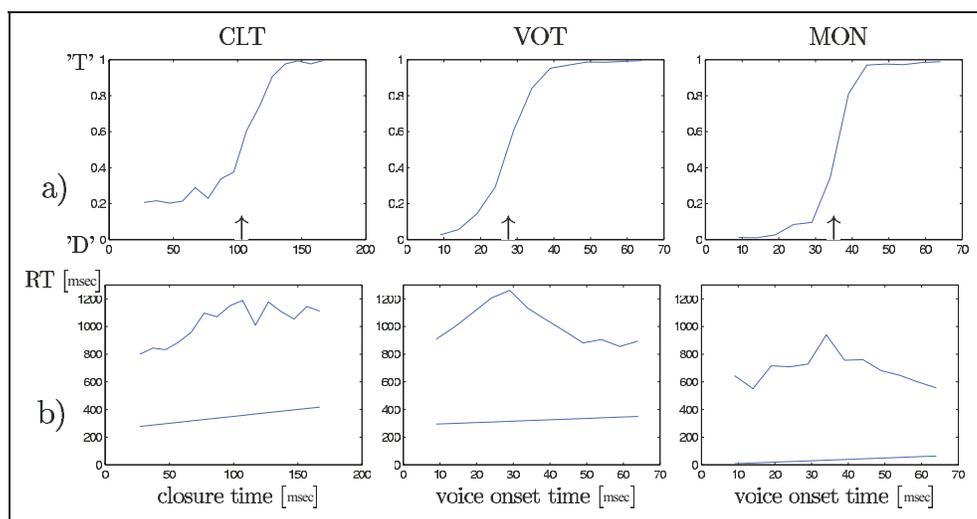


Figure 3. Significant BOLD effects at the level of infratentorial structures (Z score overlay on normalized anatomical images, threshold at $Z > 3.1$, corresponds to $p < .001$, uncorrected). Within the right cerebellar hemisphere, decoding of the intersegmental closure time (CLT) yielded a significantly higher BOLD effect than intrasyllabic voice onset time (VOT). The displayed activation cluster must be assigned to lateral aspects of Crus I of the neocerebellum (see Schmahmann et al., 2000).

ments allowed assessment of behavioral performance and task difficulty, mitigating these concerns.

In summary, the objectives of the present study were (1) to identify cerebellar structures that participate in word recognition and (2) to explore the cerebral networks involved in the decoding of the two distinct phonetic contrasts considered.

RESULTS

Behavioral Data

Voice Onset Time (VOT)

During MR measurements, subjects were asked to press a key either with the left or right index finger in order to indicate—under forced choice conditions—the perceived lexical item. The side of motor response, visually indicated on a screen, was balanced across blocks. As expected, all participants showed a clear-cut phoneme boundary effect during application of the stimulus series varying in VOT: Individual identification curves were characterized by a rather abrupt transition between the two response categories “Boten” and “Boden” (group average of calculated transition points = 27.3 ± 4.5 msec; Figure 2a). Accordingly, short VOT values (<20 msec) elicited a significantly higher number of “Boden” percepts as compared to the remaining items. Their long counterparts (>55 msec) were significantly more often categorized as “Boten” ($p < .001$ each; Wilcoxon test). Analysis of variance revealed significantly delayed responses to VOT values adjacent to the transition point of the identification curves ($29 \text{ msec} \leq \text{VOT}$

≤ 34 msec, $\text{RT} = 121 \pm 435$ msec) as compared to events located towards the extrema of the stimulus series ($\text{VOT} < 29$ msec or $\text{VOT} > 34$ msec, $\text{RT} = 949 \pm 18$ msec; Figure 2b).

Monosyllables (MON)

The MON control condition yielded results analogous to the VOT series. The group average of the calculated perceptual boundary amounted to 35.0 ± 3.1 msec. Ambiguous stimuli around this transition point resulted in prolonged RT (824 ± 28 msec) as compared to the prototypical items (group mean: -652 ± 17 msec).

Closure Time (CLT)

Across groups, prototypical stimuli ($\text{CLT} \leq 37$ msec or $\text{CLT} \geq 157$ msec) were significantly more often categorized according to the respective consonants as compared to the remaining items ($p < .001$ each). Furthermore, the subjects presented a shorter RT in response to the prototypical stimulus events (993 ± 15 msec) as

Table 1. Hemodynamic Activation of Supratentorial Structures during Word Recognition Tasks Based on a Subtraction Approach: Activation Clusters ($Z > 4.5$, based on $p < .05$, Volume Corrected) in Terms of their Highest Z Score, Respective Talairach Coordinates, Cluster Size, and Brodmann’s Area

Contrast Anatomical Substrate	Z Score Maximum	Talairach			Cluster Size (cm^3)	Brodmann’s Area
		x	y	z		
<i>CLT – VOT</i>						
L Inferior frontal gyrus	5.2	–46	40	–19	0.4	47
<i>VOT – CLT</i>						
L Superior temporal gyrus	6.6	–51	–13	4	4.2	48
<i>VOT – MON</i>						
L Occipital pole	9.2	–22	–92	–9	4.6	18
R Occipital pole	7.7	28	–94	–5	1.9	18
L Superior parietal lobule	7.3	–28	–58	51	8.5	39 ^a
R Superior parietal lobule	4.9	40	–50	50	0.2	39 ^a
L Superior temporal gyrus	5.9	–53	–15	4	1.7	48
R Superior temporal gyrus	5.9	59	–12	–1	1.1	48
R Inferior frontal gyrus	6.6	59	17	–4	1.6	47
<i>MON – VOT</i>						
Sinus saggitalis superior	4.9	0	61	17	0.1	

CLT – VOT = closure time against voice onset time (main contrast); VOT – CLT = reversed subtraction; VOT – MON = bisyllables – monosyllables; MON – VOT = monosyllables – bisyllables (control contrasts). For the sake of readability, several smaller ($\leq 0.3 \text{ cm}^3$) VOT – MON clusters adjacent to the right temporal and frontal areas are not listed (compare with Figure 5).

^aAlso parts of Brodmann’s areas 7 and 40.

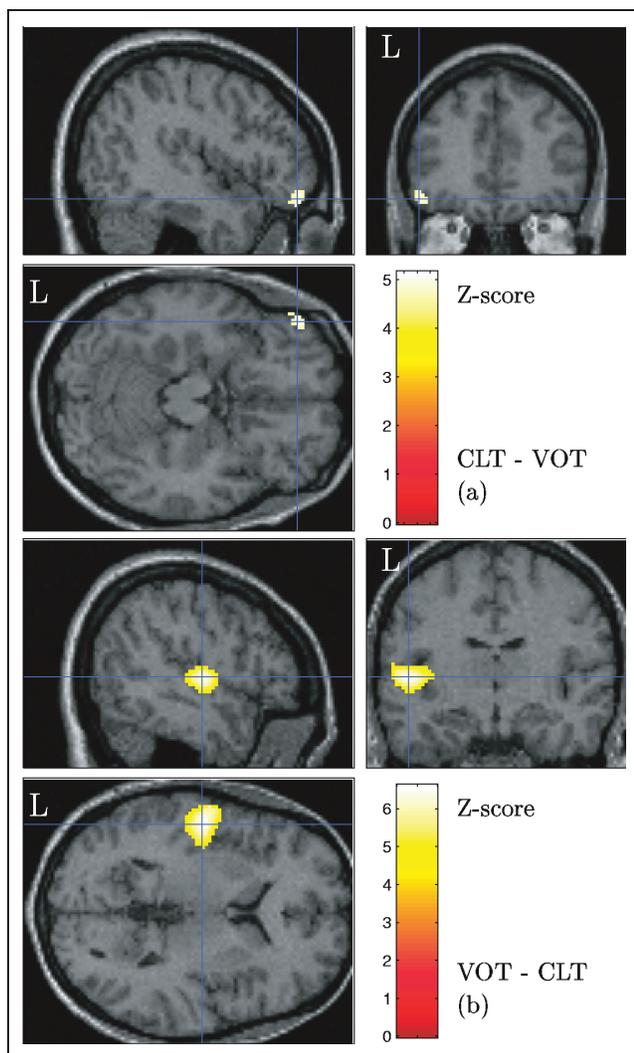


Figure 4. Significant hemodynamic effects at the level of supratentorial structures (Z score overlay on normalized anatomical images, threshold at $Z > 3.1$, corresponds to $p < .001$, uncorrected). (a) Upper panel: The main contrast (BOLD signal during encoding of CLT – BOLD signal obtained during processing of VOT) resulted in an activation cluster adjacent to Broca’s area, that is, contralateral to the cerebellar response observed under the same conditions. (b) Lower panel: The inverse subtraction procedure, that is, VOT – CLT, revealed significantly higher engagement of the left supratemporal plane in processing VOT as compared to CLT.

compared to ambiguous CLT stimuli ($87 \text{ msec} \leq \text{CLT} \leq 107 \text{ msec}$; $\text{RT} = 1137 \pm 31 \text{ msec}$). However, these stimuli yielded more variable results: Only 10 out of a total of 13 participants exhibited a clear-cut phoneme boundary effect (group mean of transition point = $103.6 \pm 9.6 \text{ msec}$; Figure 2a), 1 participant showed a considerably lowered phoneme boundary at about 57 msec, and the remaining 2 subjects recognized all stimuli, regardless of CLT length, as pertaining to the lexical category “Boten.” Remarkably, no significant differences in RT between the two stimulus series (VOT vs. CLT) emerged (estimated effect size $3 \pm 20 \text{ msec}$, $p \geq .5$; Figure 2b). Accordingly, informal inquiry

of the subjects revealed no conscious distinction between the CLT and VOT condition.

Cerebellar Signal Changes

At the level of the cerebellum (manually predefined volume of interest on the anatomical template), the main contrast (CLT – VOT blocks) yielded a single cluster of hemodynamic responses located superior to the horizontal fissure (lateral Crus I; Schmahmann, Doyon, Toga, & Evans, 2000) within the right hemisphere (cluster size $k_E = 2.4 \text{ cm}^3$, $p < .05$, volume corrected according to the Gaussian-field theory in repeated-measures ANOVA; Talairach coordinates of highest Z score: $(20 \ -82 \ -16) \text{ mm}$ with descriptive $p = .004$; Figure 3). The reversed (VOT – CLT), as well as the control subtractions (MON – VOT and VOT – MON), did not show any significant activation clusters within the cerebellum.

Cortical Signal Changes

Table 1 summarizes locations of the observed supratentorial hemodynamic effects in terms of Talairach coordinates and Brodmann’s areas. As a complement to right hemisphere cerebellar activation, left prefrontal responses anterior inferior to Broca’s area emerged in the main contrast CLT – VOT (Figure 4a). The inverted contrast (VOT – CLT) yielded a single significant cluster within the left supratemporal plane (Figure 4b).

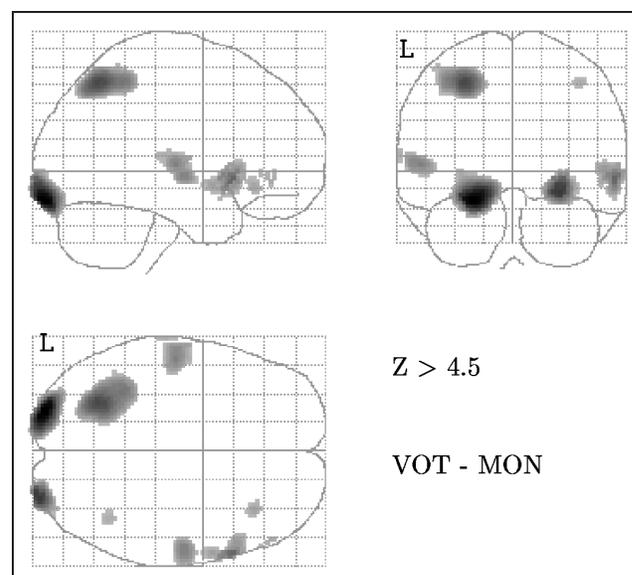


Figure 5. Descriptive display of signal changes in terms of maximum intensity projections on the normalized SPM-glass brain. Sensory areas, as well as parietal and right prefrontal regions, exhibited stronger activation in response to two-syllable utterances than to monosyllables (VOT – MON contrast).

During the MON control condition, a single syllable was applied in the auditory domain and each visual hemifield. In contrast, the bisyllabic VOT stimuli were longer and more complex. Accordingly, the contrast VOT – MON yielded signal changes both in visual and auditory cortices (see Table 1 and Figure 5). The remaining signal changes presumably refer to bilateral activation of the superior parietal lobule and to the right hemisphere Broca analogue. The negative control (MON – VOT, i.e., monosyllabic against bisyllabic items) resulted in a significant fMRI signal increase restricted to the venous sinus. This intravasal effect most likely reflects global perfusion changes.

DISCUSSION

The present study documented cerebellar involvement during a word recognition task: A distinct right hemisphere cerebellar activation cluster superior to the horizontal fissure emerged when identification of a lexical item solely depended upon the encoding of a durational speech parameter. These findings accord with previous clinical data demonstrating deficient perception of temporal speech cues in subjects with cerebellar atrophy. Active (CLT) and reference (VOT) conditions were balanced with respect to response side and lexical demands. Subjects' RTs did not differ significantly. The obtained data point at an inherent capability of the cerebellum to provide precise representation of temporal information even within the speech perception domain.

Speech Perception Networks

In most instances, the syndrome of pure word deafness or verbal auditory agnosia—a clinical model of disrupted speech sound perception—is bound to bilateral lesions of the auditory cortex. Therefore, the extraction of linguistic information in terms of phonetic features from the acoustic speech signal has been attributed to the supratemporal plane of both hemispheres (for a review, see Ackermann & Mathiak, 1999). Nevertheless, compromised phoneme discrimination or syllable identification also has been found in a subgroup of patients with anterior perisylvian lesion including Broca's aphasia (Blumstein, Baker, & Goodglass, 1977; Blumstein, Cooper, Zurif, & Caramazza, 1977) and in cerebellar subjects (Ackermann et al., 1997). More recent functional imaging studies provided further evidence for a contribution of structures outside the superior temporal cortex to speech sound encoding in terms of phoneme discrimination or identification. Based on a hierarchical subtraction design, the seminal positron emission tomography (PET) studies of Zatorre, Evans, Meyer, and Gjedde (1992) found increased blood flow at the level of Broca's area during a phonetic discrimination task—participants were asked to respond to pairs of monosyllables with the

same final consonant—as compared to passive listening to the same items. A series of subsequent PET investigations found a left frontal activity during various tasks of phonetic processing (rhyming of letter names: Sergent, Zuck, Lévesque, & MacDonald, 1992; detection of pre-specified phonemes: Démonet, Price, Wise, & Frackowiak, 1994; for a review, see Poeppel et al., 1996, and Zatorre, Meyer, Gjedde, & Evans, 1996). The magnitude of activity at or near Broca's area correlates with task demands, such as the rate of external stimuli, and was more anterior for perception as compared to speech production (Price et al., 1996). Furthermore, the acoustic characteristics of the test materials have been found relevant in this regard: Fiez et al. (1995) observed, for instance, significant left frontal responses during detection of tone triplets, syllables, and words, whereas, in contrast, relatively long steady-state vowel targets (duration = 250 msec) did not show a similar effect. Recruitment of left frontal areas, therefore, seems to depend upon the presence of rapidly changing spectral cues in the acoustic speech signal.

As concerns cerebellar contributions to speech perception, inconsistent data have been reported so far. For example, Fiez et al. (1995) found activation of bilateral medial and left-sided lateral aspects in three participants serving as a hypothesis-generating subgroup. These findings, however, could not be replicated in a second independent sample of subjects. A recent fMRI study observed a hemodynamic response of the right cerebellum during the discrimination of monosyllabic consonant–vowel–consonant items (Burton et al., 2000). The present investigation extends these previous functional imaging studies, providing firm evidence for a task-dependent contribution of the cerebellum to phonetic processing.

Considering the available clinical and functional imaging data, auditory association areas might represent the core area of speech sound processing that join up with frontal or cerebellar areas when syllable or word recognition requires further computational support. In accordance with this suggestion, the computed negative contrast (VOT – CLT) indicated that the intrasegmental parameter VOT is extracted and presented within the superior temporal lobe. Likewise, previous electrophysiological studies both in monkeys (Steinschneider, Schroeder, Arezzo, & Vaughan, 1994) and in humans (Ackermann, Lutzenberger, & Hertrich, 1999; Mathiak, Hertrich, Lutzenberger, & Ackermann, 1999, 2000, 2001; Steinschneider, Volkov, Noh, Garell, & Howard, 1999) revealed distinct and localized responses to consonant contrasts at the level of the auditory cortex.

Cerebellum and Cognition

A variety of cerebellar contributions to the cognitive domain have been proposed, including support of lexical processes and control of attentional resources,

as well as “coordination” of extracerebellar mental and sensory functions. For example, several PET studies documented right hemispheric cerebellar activation associated with lexical tasks, such as verbal response selection (Fiez & Raichle, 1997). The test materials of the present study referred to the same pair of lexical items. Therefore, the CLT – VOT contrast cannot reflect lexical or semantic operations.

Clinical studies established an influence of cerebellar pathology on practice-related learning. Using similar test materials as the considered PET investigations, a patient suffering from an infarction at the level of the right cerebellar hemisphere exhibited a larger number of incorrect answers and a rather constant RT across blocks (normals showed a significant decrease; Fiez, Petersen, Cheney, & Raichle, 1992). Thus, learning or task demands might be expected to activate the cerebellum during speech perception. However, the current study’s finding that the superior temporal area was more active in the VOT condition provides a strong argument against that assumption: If right cerebellar and left prefrontal activation during the CLT condition were reflecting additional demands, then they would not be accompanied by decreased superior temporal activity. Furthermore, RT measures were comparable in response both to VOT and CLT processing. Admittedly, two subjects failed throughout to discriminate the words “Boden” and “Boten” on the basis of CLT. These findings might reflect a personal trait to omit or shift the turning point of the identification curve as observed in a subgroup of patients with Parkinson’s disease (Gräber, Hertrich, Daum, Spieker, & Ackermann, in press).

The cerebellum has reciprocal connections with a multitude of supratentorial brain structures (Schmahmann & Pandya, 1997; Leiner, Leiner, & Dow, 1993). Based on the analysis of a wide range of tactile afferent projections, the cerebellum has been considered to facilitate the efficiency of sensory neuronal processing but not to be obligatory for these functions (Bower, 1997). Moreover, unselective, as well as relatively delayed, responses in the cerebellum were considered to argue against its pure timing character (Rao et al., 1997; Rao, Mayer, & Harrington, 2001). Accordingly, the cerebellum has been viewed as a “master computational system” distributing attentional resources across different cerebral processors and sensory channels (Akshoomoff, Courchesne, & Townsend, 1997; Courchesne & Allen, 1997). In a similar vein, Molinari, Leggio, and Silveri (1997) suggested that the organ acts as a “controller,” mediating sequential organization of the various subcomponents of complex cognitive tasks. In contrast, due to the general redundancy of articulatory cues, isolating specific physical features might disclose obligatory and autonomous cerebellar tasks as suggested by the previously cited lesion study (Ackermann et al., 1997). The active and control conditions of the present study differed solely in their demands on precise repre-

sentation of time intervals of the acoustic signal and thus acted upon a single subsystem of the speech and language apparatus. Taken together, the present results indicate that the brain makes use of inherent cerebellar capabilities in temporal computations to facilitate phoneme specification and, consecutively, word recognition.

Cerebellum and Speech Processing

Discrimination of monosyllabic lexical items exclusively differing in wordinitial VOT (e.g., “dip–tip”) was found to elicit significant right hemisphere cerebellar activation after subtraction of the responses to paired stimuli varying in several sounds (e.g., “dip–ten”; Burton et al., 2000). This recent fMRI study addressed the contribution of Broca’s area to phonological segmentation. At first glance, the suggestion that cerebellar activity is associated with VOT encoding is at variance with the present finding. However, the generalized timing hypothesis merely predicts that encoding of phonetic features based on purely durational parameters of the acoustic speech signal necessarily depends upon cerebellar structures. In accordance with this suggestion, processing of CLT elicits cerebellar activation, and cerebellar dysfunctions impair perceptual encoding of CLT. The generalized timing hypothesis does not exclude that processing of VOT—a parameter varying both in sound energy and duration—also recruits the cerebellum. In contrast to CLT, however, encoding of VOT does not exclusively depend upon precise representation of temporal information. Thus, the auditory cortex may categorize the syllable by the intrasegmental information.

Cerebellar disorders may give rise to ataxic dysarthria, which is a distinct syndrome of compromised articulation and phonation (for a review, see Ackermann & Hertrich, 2000). Clinical data indicate that the medial superior cerebellum supports these “lower-level” aspects of speech production (see Ackermann, Vogel, Petersen, & Poremba, 1992). In line with these findings, PET studies of single word processing based on a hierarchical subtraction design reported bilateral activation of medial superior portions of the cerebellum in association with articulatory processes (Petersen, Fox, Posner, & Mintun, 1989; Petersen, Fox, Snyder, & Raichle, 1990). Considering the findings of the present study, the cerebellum’s contribution to acoustic communication might be twofold: (a) participation in speech perception and verbal working memory, among other functions, through a lateral system of phonetic representations based on linguistic units and (b) a more medial loop integrating sensory–motor functions in association with the generation of innervation patterns acting on the vocal tract musculature.

The present study revealed a cerebellar contribution to speech perception at the level of word recognition. On a higher processing level, previous PET studies had documented right cerebellar hemodynamic responses

during lexical operations, such as verb generation (for a review, see Fiez & Raichle, 1997). Moreover, covert production of word strings (auditory imagery, “inner speech”; Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Ackermann, Wildgruber, Daum, & Grodd, 1998), as well as verbal memory tasks (e.g., Schumacher et al., 1996; Paulesu, Frith, & Frackowiak, 1993), exhibited right hemispheric neocerebellar activation accompanied by contralateral frontal responses adjacent to Broca’s area. Jueptner et al. (1995) demonstrated the involvement of both cerebellar hemispheres and, in particular, the vermis during the discrimination of non-speech periods. These patterns might recur on the same basic neurobiological mechanisms such as the computation of temporal cues. It remains to be specified whether the hypothesized “internal clock” function of the cerebellum can be characterized by topographically distinct modes.

Outlook

Assuming higher proficiency of the cerebellum than the cerebrum in precise computation of durational parameters, infratentorial disorders might compromise any cognitive function depending upon internal temporal representation. As does sensory processing, auditory imagery and working memory rehearsal mechanisms require encoding of temporal structures of word and sentence forms. Because besides timing, other parameters, for example, spectral features, contribute to the specification of phonetic and phonological information, the engagement of cerebellar structures during cognition and the severity of neuropsychological deficits in cerebellar disorders might depend upon strategy and task demands, such as temporal constraints.

METHODS

Participants

A total of 13 German volunteers (eight women; age 19–33, median 26) participated in the present study. Subjects were selected on the basis of right-handedness according to the Edinburgh handedness scale. None of them presented medical or neurological dysfunctions at clinical examination or an increased hearing threshold.

Stimuli

The contrast between the wordmedial consonants “d” and “t” of the German lexical items “Boden” and “Boten” may be signaled by variation of either CLT or VOT (see Introduction and Figure 1). As a rule, the former variants emerge during fluent sentences spoken in colloquial style whereas the latter cues are used in

clear enunciations, for example, words in isolation. Both types of utterances were recorded from a female actor instructed accordingly. Three series of test materials were created by means of digital cut-and-paste techniques: (a) A “Boten” item was chosen that realized the “t” sound as a lengthened occlusion pause followed by a nasal plosion [botn]. A continuum between “Boden” [bodn] and “Boten” was created by exclusively varying the silent wordmedial CLT between [bo] and [tn] from 27 to 167 msec in 15 steps (CLT condition; Figure 1a). (b) Based on a “Boten” production that exhibited a clear-cut aspiration noise followed by a schwa vowel [bot^b↔n], the second syllable [t^b↔n] was varied with respect to VOT ranging from 9 to 69 msec in 12 steps. Thus, a continuum between the monosyllables [d↔n] and [t^b↔n] was obtained (MON condition). (c) By pasting these syllables after the (invariant) initial syllable [bo] of the utterances from the CLT condition, a complementary series of “Boden–Boten” stimuli was derived, varying in VOT of the wordmedial “d–t” plosive ([bod↔n]—[bot^b↔n]; VOT condition). Figure 1 illustrates the temporal structure of the stimuli.

Experimental Procedure

Magnetic resonance imaging was performed on a 1.5-T scanner (Siemens Vision, Erlangen, Germany) equipped with gradient boosters. For each subject, a T1-weighted isometric data set served as anatomical reference (TE = 5 msec, TR = 120 msec, flip angle = 90°, matrix size = 256 × 256, 196 partitions, voxel size = 1 × 1 × 1 mm). Echo-planar imaging with sinusoidal gradient ramps was used for the detection of blood oxygenation level dependent (BOLD) effect with full brain coverage (28 slices per acquisition cycle, matrix size = 64 × 64, voxel size = 4 × 4 × 3.3 mm, TE = 54 msec, TA = 3.3 sec, TR = 5.4 sec). Slices were tilted occipito-caudally from axial position to obtain maximal volume coverage, including the entire cerebellum. The duration of each session comprising 132 volume acquisitions amounted to 11 min and 40 sec. The total scanning time including anatomical and functional imaging varied between 90 and 115 min.

Functional imaging sessions were designed as a “block design.” A white screen with a red fixation cross was displayed (baseline condition) during six TR cycles at the beginning and at the end of each session. The remaining 120 volume acquisitions were alternating with one auditory stimulus each. During five scanning periods (a “block”), sounds were selected randomly for a fixed task. The series of 24 tasks were pseudorandomized as CLT, VOT, or MON conditions (eight times each). During the CLT and VOT conditions, forced choice responses for “Boden” against “Boten” were requested. (Similarly, during the MON condition, the prototypes were “Den” and “Ten.”) For each task, with the onset of its first scan, the display figured one of the minimal pairs—“boDen” or “boTen”—on the left and the other

one on the right side of the screen (170 cm distance, 25° visual field). Presentation and, thus, response sides were randomized and balanced across conditions to cancel out motor activation. Responses were recorded from optically operating buttons at each side.

Onset of each stimulus was jittered between 3.75 and 3.95 sec after scanner onset leaving 0.45–0.65 sec between the end of the scanner noise and the beginning of the speech stimulus. Thus, no fixed relationship between stimulus duration and time to the following scanner noise emerged. For each session and subject, randomization and jitter were calculated independently. Stimuli were applied via directly driven, noise-canceling headphones (modified dynamic system without proper magnet; see Baumgart et al., 1998). The scanner noise amounted to 82.3 dB SPL, and stimuli were adjusted to 85.0 dB SPL.

Data Analysis

RTs of the behavioral responses were measured with reference to the onset of each acoustic test item. Response probabilities were compared using the Wilcoxon test. RTs were entered into a linear regression model. A repeated-measures ANOVA estimated the effect of ambiguous stimuli (close to perceptual boundary: $29 \text{ msec} \leq \text{VOT} \leq 34 \text{ msec}$ or $87 \text{ msec} \leq \text{VOT} \leq 107 \text{ msec}$). The variables CLT or VOT were entered into a logistic generalized linear model, that is, binomial response distribution with logistic linking, for each subject. The categorical boundary was set at 50% chance level for the responses. Group statistics were calculated across these values.

According to the statistical standard procedure of SPM99 (Friston & Turner, 1997), motion correction and coregistration was conducted. To reduce variability of the fMRI signal due to varying relationships between brainstem and cortex across subjects, separate normalizations and statistical analyses for cerebrum and cerebellum were performed. As concerns the latter structure, a mask was applied covering the cerebellum of the anatomical template image. The Talairach space (Talairach & Tournoux, 1988) served as a reference. Condition indicators folded with a modeled hemodynamic response function were entered into a repeated-measures ANOVA. Correction for multiple testing throughout the image voxels was calculated assuming a correlation structure of a Gaussian field after applying a smoothing kernel (12 mm full width at half maximum). Group statistics were obtained from the estimated contrasts and residual variance of each individual. In order to test the a priori hypothesis derived from the previous lesion study, intensity changes within cerebellar structures were considered if significant on the cluster level ($p < .05$, corrected). For the supplementary a posteriori analysis of the entire brain, dominated by the telencephalon, activations were thresholded voxel-wise accord-

ing to a volume-corrected error probability of $p < .05$. BOLD signal increase in response to the active (CLT) as compared to the reference (VOT) condition served as the main contrast. As a control, the remaining three subtractions of interest were analyzed as well, namely, reference – active (VOT – CLT), reference – control (VOT – MON), and control – reference (MON – VOT).

Acknowledgments

This study was supported by the German Research Foundation (SFB 550/B1 and SPP-ZIZAS Ac 55/5-1). The authors thank Nell Zink for revising the manuscript and Maike Borutta for excellent technical assistance and helpful discussions.

Reprint requests should be sent to Klaus Mathiak, Universität Tübingen, MEG-Zentrum, Otfried-Müller-Str. 47, 72076 Tübingen, Germany, or via e-mail: klaus.mathiak@med.uni-tuebingen.de.

The data reported in this experiment have been deposited in The fMRI Data Center (<http://www.fmridc.org>). The accession number is 2-2001-112EW.

REFERENCES

- Ackermann, H., Gräber, S., Hertrich, I., & Daum, I. (1997). Categorical speech perception in cerebellar disorders. *Brain and Language*, *60*, 323–332.
- Ackermann, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics*, *13*, 95–116.
- Ackermann, H., Lutzenberger, W., & Hertrich, I. (1999). Hemispheric lateralization of the neural encoding of temporal speech features: A whole-head magnetencephalography study. *Cognitive Brain Research*, *7*, 511–518.
- Ackermann, H., & Mathiak, K. (1999). Syndromes, pathologic-anatomical basis, and pathophysiology of central hearing disorders (pure word deafness, auditive agnosia, cortical deafness). *Fortschritte der Neurologie-Psychiatrie*, *67*, 509–523 (in German).
- Ackermann, H., Vogel, M., Petersen, D., & Poremba, M. (1992). Speech deficits in ischaemic cerebellar lesions. *Journal of Neurology*, *239*, 223–227.
- Ackermann, H., Wildgruber, D., Daum, I., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans. *Neuroscience Letters*, *247*, 187–190.
- Akshoomoff, N. A., & Courchesne, E. (1992). A new role for the cerebellum in cognitive operations. *Behavioral Neuroscience*, *106*, 731–738.
- Akshoomoff, N. A., Courchesne, E., & Townsend, J. (1997). Attention coordination and anticipatory control. *International Reviews on Neurobiology*, *41*, 575–598.
- Allen, G., Buxton, R. B., Wong, E. C., & Courchesne, E. (1998). Attentional activation of the cerebellum independent of motor involvement. *Science*, *275*, 1940–1943.
- Baumgart, F., Kaulisch, T., Tempelmann, C., Gaschler-Markefski, B., Tegeler, C., Schindler, F., Stiller, D., & Scheich, H. (1998). Electrodynamic headphones and woofers for application in magnetic resonance imaging scanners. *Medical Physics*, *25*, 2068–2070.
- Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, *15*, 19–30.

- Blumstein, S. E., Cooper, W. E., Zurif, E. B., & Caramazza, A. (1977). The perception and production of voice-onset time in aphasia. *Neuropsychologia*, *15*, 371–383.
- Bower, J. M. (1997). Control of sensory data acquisition. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 1241–1253). San Diego, CA: Academic Press. (*International Review of Neurobiology*, *41*).
- Braitenberg, V. (1967). Is the cerebellar cortex a biological clock in the millisecond range? *Progress in Brain Research*, *25*, 334–346.
- Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, *12*, 679–690.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II. An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning and Memory*, *4*, 1–35.
- Desmond, J. E., Gabrieli, J. D., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *Journal of Neuroscience*, *17*, 9675–9685.
- Démonet, J.-F., Price, C. J., Wise, R. J. S., & Frackowiak, R. J. S. (1994). A PET study of cognitive strategies in normal subjects during language tasks: Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, *117*, 671–682.
- Dichgans, J. (1984). Clinical symptoms of cerebellar dysfunction and their topodiagnosical significance. *Human Neurobiology*, *2*, 269–279.
- Dow, R. S., & Moruzzi, G. (1958). *The physiology and pathology of the cerebellum*. Minneapolis, MN: University of Minnesota Press.
- Ellerman, J. M., Flament, D., Kim, S. G., Fu, Q. G., Merkle, H., Ebner, T. J., & Ugurbil, K. (1994). Spatial patterns of functional activation of the cerebellum investigated using high field (4 T) MRI. *NMR in Biomedicine*, *7*, 63–68.
- Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage: A single case study. *Brain*, *115*, 155–178.
- Fiez, J. A., & Raichle, M. E. (1997). Linguistic processing. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 233–254). San Diego, CA: Academic Press. (*International Review of Neurobiology*, *41*).
- Fiez, J. A., Raichle, M. E., Miezin, F. M., Petersen, S. E., Tallal, P., & Katz, W. F. (1995). PET studies of auditory and phonological processing: Effects of stimulus characteristics and task demands. *Journal of Cognitive Neuroscience*, *7*, 357–375.
- Friston, K. J., & Turner, R. (1997). *SPMcourse: Short course notes*. London: Institute of Neurology, UCL. (<http://www.fil.ion.ucl.ac.uk/spm/course/notes.html>).
- Gilman, S., Bloedel, J. R., & Lechtenberg, R. (1981). *Disorders of the cerebellum*. Philadelphia, PA: Davis.
- Gräber, S., Hertrich, I., Daum, I., Spieker, S., & Ackermann, H. (in press). Speech perception deficits in Parkinson's disease: Underestimation of time intervals comprises identification of durational phonetic contrasts. *Brain and Language*.
- Grodd, W., Hülsmann, E., Lotze, M., Wildgruber, D., & Erb, M. (2001). Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Human Brain Mapping*, *13*, 55–73.
- Hirano, S., Naito, Y., Okazawa, H., Kojima, H., Hanjo, I., Ishizu, K., Yenokura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1997). Cortical activation by monaural speech sound stimulation demonstrated by positron emission tomography. *Experimental Brain Research*, *113*, 75–80.
- Ivry, R. B. (1997). Cerebellar timing systems. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 233–254). San Diego, CA: Academic Press. (*International Review of Neurobiology*, *41*).
- Ivry, R. B., & Fiez, J. A. (2000). Cerebellar contributions to cognition and imagery. In M. Gazzaniga (Ed.), *The new cognitive neuroscience* (pp. 999–1011). New York: Plenum.
- Ivry, R. B., & Gopal, H. S. (1992). Speech production and perception in patients with cerebellar lesions. In D. Meyer & S. Kornblum (Eds.), *Attention and performance* (vol. XIV, pp. 771–802). Hillsdale, NJ: Erlbaum.
- Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*, 136–152.
- Jueptner, M., Rijntjes, M., Weiller, C., Faiss, J. H., Timmann, D., Mueller, S. P., & Diener, H. C. (1995). Localization of a cerebellar timing process using PET. *Neurology*, *45*, 1540–1545.
- Kawashima, R., Okuda, J., Umetsu, A., Sugiura, M., Inoue, K., Suzuki, K., Tabuchi, M., Tsukiura, T., Narayan, S. L., Nagasaka, T., Yanagawa, I., Fujii, T., Takahashi, S., Fukuda, H., & Yamadori, A. (2000). Human cerebellum plays an important role in memory-timed finger movement: An fMRI study. *Journal of Neurophysiology*, *83*, 1079–1087.
- Keele, S. W., & Ivry, R. Y. (1990). Does the cerebellum provide a common computation for diverse tasks? A timing hypothesis. In A. Diamond (Ed.), *The cerebellum and cognition* (pp. 179–211). New York: New York Academy of Science.
- Ladefoged, P., & Maddieson, I. (1997). *The sounds of the world's languages*. Oxford, UK: Blackwell.
- Leiner, H. C., Leiner, A. L., & Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends in Neuroscience*, *16*, 444–447.
- Lieberman, A. M. (1996). *Speech: A special code*. Cambridge: MIT Press.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422.
- Lisker, L., & Abramson, A. S. (1970). The voicing dimension: Some experiments in comparative phonetics. In M. Romportl & P. Janota (Eds.), *Proceedings of the sixth international congress of phonetic sciences* (pp. 563–567). Prague, Czech Republic: Academica.
- Mathiak, K., Hertrich, I., Lutzenberger, W., & Ackermann, H. (1999). Pre-attentive processing of consonant–vowel syllables at the supratemporal plane: A whole-head magnetencephalography study. *Cognitive Brain Research*, *8*, 251–257.
- Mathiak, K., Hertrich, I., Lutzenberger, W., & Ackermann, H. (2000). Encoding of temporal speech features (formant transients) during binaural and dichotic stimulus application: A whole-head magnetencephalography study. *Cognitive Brain Research*, *10*, 125–131.
- Mathiak, K., Hertrich, I., Lutzenberger, W., & Ackermann, H. (2001). Neuronal correlates of duplex perception: A whole-head magnetencephalography study. *NeuroReport*, *12*, 501–506.
- Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J., & Dooling, R. J. (1976). Discrimination and labeling of noise–buzz sequences with varying noise lead time: An example of categorical perception. *Journal of the Acoustical Society of America*, *60*, 410–417.
- Molinari, M., Leggio, M. G., & Silveri, M. C. (1997). Verbal

- fluency and agrammatism. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 325–339). San Diego, CA: Academic Press. (*International Review of Neurobiology*, 41).
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362, 342–345.
- Petersen, S. E., Fox, P. T., Posner, M. I., & Mintun, M. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, 1, 153–170.
- Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, 249, 1041–1044.
- Poeppel, D., Yellin, E., Phillips, C., Roberts, T. P., Rowley, H. A., Wexler, K., & Marantz, A. (1996). Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. *Cognitive Brain Research*, 4, 231–242.
- Price, C. J., Wise, R. J. S., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., Frackowiak, R. S. J., & Friston, K. J. (1996). Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain*, 119, 919–931.
- Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., Cox, R. W., Stein, E. A., & Binder, J. R. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *NeuroReport*, 8, 1987–1993.
- Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, 4, 317–323.
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport*, 11, 1997–2000.
- Schmahmann, J. D., Doyon, J., Toga, A. W., & Evans, A. C. (2000). *MRI atlas of the human cerebellum*. San Diego, CA: Academic Press.
- Schmahmann, J. D., & Pandya, D. N. (1997). The cerebrocerebellar system. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 31–60). San Diego, CA: Academic Press. (*International Review of Neurobiology*, 41).
- Schumacher, E. H., Lauber, E., Awh, E., Jonides, J., & Smith, E. E., & Koeppel, R. A. (1996). PET evidence for an amodal verbal working memory system. *Neuroimage*, 3, 78–88.
- Sergent, J., Zuck, E., Lévesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex*, 2, 68–80.
- Steinschneider, M., Schroeder, C. E., Arezzo, J. C., & Vaughan, H. G., Jr. (1994). Speech-evoked activity in primary auditory cortex: Effects of voice onset time. *Electroencephalography and Clinical Neurophysiology*, 92, 30–43.
- Steinschneider, M., Volkov, I. O., Noh, M. D., Garell, P. C., & Howard, M. A., 3rd. (1999). Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *Journal of Neurophysiology*, 82, 2346–2357.
- Talairach, J., & Tournoux, P. (1988). *Talairach: Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- Wildgruber, D., Ackermann, H., & Grodd, W. (2001). Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: Effects of syllable repetition rate evaluated by fMRI. *Neuroimage*, 13, 101–109.
- Yousem, D. M., Williams, S. C., Howard, R. O., Andrew, C., Simmons, A., Allin, M., Geckle, R. J., Suskind, D., Bullmore, E. T., Brammer, M. J., & Doty, R. L. (1997). Functional MR imaging during odor stimulation: Preliminary data. *Radiology*, 204, 833–838.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256, 846–849.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication and reanalysis. *Cerebral Cortex*, 6, 21–30.