

Neuronal activity in the human lateral temporal lobe

I. Responses to speech

O. Creutzfeldt¹, G. Ojemann², and E. Lettich²

¹ Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, P.O.B. 2841, D-3400 Göttingen-Nikolausberg, Federal Republic of Germany

² Department of Neurological Surgery, RI 20, University of Washington, Seattle, WA 98195, USA

Summary. Single and multiple unit neuronal activity was recorded from the cortex of the lateral temporal lobe in conscious humans during open brain surgery for the treatment of epilepsy. Recordings were obtained from the right and left superior, middle and inferior temporal gyrus of 34 patients (41 recording sites). Recordings were restricted to regions to be resected during subsequent surgery. This excluded recordings from language areas proper. Neuronal responses to words and sentences presented over a loudspeaker and during free conversation were recorded. No significant differences between the right and left hemisphere were obvious. All neurons in the *superior temporal gyrus* responded to various aspects of spoken language with temporally well defined activation/inhibition patterns, but not or only little to non-linguistic noises or tones. Excitatory responses were typically short or prolonged (up to several hundred ms) bursts of discharges at rates above 20/sec, reaching peak rates of 50–100/s. Such responses could be specifically related to certain combinations of consonants suggesting a function in categorization, they could depend on word length, could differentiate between polysyllabic and compound words of the same length or could be unspecifically related to language as such. No formant specific responses were found, but the prolonged excitations across syllables suggest that consonant/vowel combinations may play a role for some activation patterns. Responses of some neurons (or neuronal populations) depended on the attention paid to the words and sentences, or the task connected with them (repeat words, speech addressed to the patient demanding something). Neurons in the *middle and inferior temporal gyrus* were only little affected by listening to single words or sentences, but some

were unspecifically activated by words or while listening to sentences. Excitatory responses varied within a limited range of discharge rates usually below 5–10/s. Phonetic distortion of spoken language could reduce responses in superior temporal gyrus neurons, but also the slight changes in discharge rate of middle temporal neurons could be absent during distorted and uncomprehensible speech sounds. We conclude that superior temporal gyrus neuron responses reflect some general phonetic but not semantic aspects of spoken language. Middle and inferior temporal gyrus neurons do not signal phonetic aspects of language, but may be involved in understanding language under certain conditions.

Key words: Single unit recording – Temporal lobe – Listening to speech – Man

Introduction

The lateral human temporal lobe receives thalamic afferents from the medial geniculate body and the pulvinar, and association fibres from the prestriate cortex and the frontal lobe. The primary auditory cortex is located in its most dorsal parts, while the more ventral aspects were considered as auditory and visual association areas. The caudal perisylvian cortex on the language dominant hemisphere is part of Wernicke's area and necessary for a variety of linguistic functions such as understanding of spoken language, naming and speaking. The function of the rostral portion, the temporal pole, is not known. The intermediate portion, below the motor and somato-sensory cortex, contains auditory and visual association fields (for details and references see Creutzfeldt 1983). The auditory association cortex is thought to be confined to the superi-

or temporal gyrus (Galaburda and Sanides 1980), while the middle and inferior temporal gyrus are considered as visual association areas, homologous to the infero-temporal cortex of subhuman primates.

Details of the involvement of the lateral human temporal association cortex in auditory and visual functions are not known. Experiments on primates suggest that the infero-temporal cortex is involved in visual object recognition and memory (Mishkin et al. 1983; Ungerleider 1985), while increase of regional blood flow (RCBF) during speech (Lassen and Larsen 1980; Ingvar 1983), local cerebral glucose metabolism (Mazziotta et al. 1983; Raichle 1987) as well as evoked potential studies (see Hilliard and Picton 1987) suggest that large parts of the human lateral temporal lobe on both sides maybe involved in linguistic functions. Clinical experience with lesion, symptoms during temporal lobe seizures and deficits during electrical stimulation indicate that the lateral and basal temporal lobe is involved in functions related to emotions, memory and consciousness (Jackson 1898; Luria 1966; Penfield and Jasper 1954; Penfield and Perrot 1963; Ojemann 1983; Ojemann and Creutzfeldt 1987). In view of this, it is difficult to predict to what extent and how temporal lobe neuronal activity might be involved in various functions related to language, memory and emotions.

During temporal lobe surgery under local anesthesia, we had the opportunity to record in consenting conscious humans with microelectrodes the activity of single neurons or groups of neurons in various locations of the right and left temporal lobe convexity during auditory, visual and linguistic tasks. As microelectrode recording is an invasive technique, it was restricted to brain tissue which, after electro-corticographic and functional exploration with electrical stimulation, was subsequently to be resected. This, of course, precluded recordings from Wernicke's language area, *sensu strictu*, of the dominant hemisphere. With these constraints, our analysis included microelectrode recordings predominantly from the lateral surface of anterior, superior, middle and inferior temporal gyri, with a few recordings from the posterior third of the superior temporal gyrus of the non-speech dominant hemisphere.

In this paper we report on single and grouped unit activity during listening to speech, including acoustically deteriorated speech. It will be demonstrated that neuronal activities in the superior and middle temporal gyri were affected differently by speech signals, but without obvious hemispheric differences. In subsequent papers we report on re-

sponses of the same population of neurons to the subjects own voice, to music and to more complex tasks involving memory and cognition (Creutzfeldt et al. 1989a, b). A detailed statistical analysis of dominant temporal lobe activity related to language and memory (Ojemann et al. 1989) and preliminary reports on speech related responses have been published (Creutzfeldt et al. 1987).

Methods

We report here on 41 successful microelectrode recordings from 34 patients (23 recordings in 21 patients from the left, all except one the speech dominant and 18 recordings in 13 patients from the right hemisphere). The recording sites are shown in Fig. 1 and are mostly located on the middle and superior temporal gyrus in the anterior half of the temporal lobe below the face representation of the sensori-motor cortex. A few recording sites were located further posterior, just below the perisylvian language area in the left, in the posterior third of the superior temporal gyrus in the right hemisphere and in the inferior temporal gyrus. In several patients we also did microelectrode recordings in the temporal pole region during electro-corticography (ECoG) without formal testing, but during informal speaking. These recording sites are not included in Fig. 1. The total number of unit activities tested and analyzed is higher (well above 100), as some recordings were done with two microelectrodes, repeated at different depth after a unit got lost, and multiple unit recordings were analyzed for different populations of units as distinguished by amplitude.

All patients suffered from pharmacologically intractable temporal lobe seizures and were referred to the Department of Neurological Surgery, University of Washington, Seattle/WA for temporal lobe surgery under local anesthesia, guided by ECoG and functional mapping, following evaluation at the University of Washington, Epilepsy Center. Preoperative evaluation included seizure monitoring with EEG- and video recording as well as amyltal-test (Wada and Rasmussen 1960) for determination of the speech dominant hemisphere. All patients and their relatives were informed before the operation about the intraoperative microelectrode recording and its purely scientific nature, and they gave their informed consent. During the operation they were also informed about the beginning of the microelectrode recording and could refuse it. This was done by two patients. The investigation had the approval of the Human Rights committee of the University of Washington.

The single unit recordings lasted about 30 min. Details of the surgical techniques are described elsewhere (Ojemann 1985), and details of the actual procedure of the microelectrode recording may be found in the preceding communication (Ojemann et al. 1988).

Briefly, after a large opening of the skull and the dura mater over the temporal lobe and the perisylvian cortex under local anesthesia, ECoG-exploration and mapping with electrical stimulation for speech and memory was done as a routine diagnostic procedure. This lasted about 2–3 h. Electrolytically sharpened, lacquer insulated tungsten microelectrodes with a tip diameter of 1–5 μm and a tip resistance of 3–7 $\text{M}\Omega$ were used. The electrodes were remote controlled by a hydraulic micro-manipulator (Wells), the head stage of which was held by a rod clamped to the skull. In 10 patients, two electrodes were mounted together in the manipulator with a vertical tip distance of 2–4 mm. In order to reduce tissue movements due to pulsation and respiration in the region of the microelectrode record-

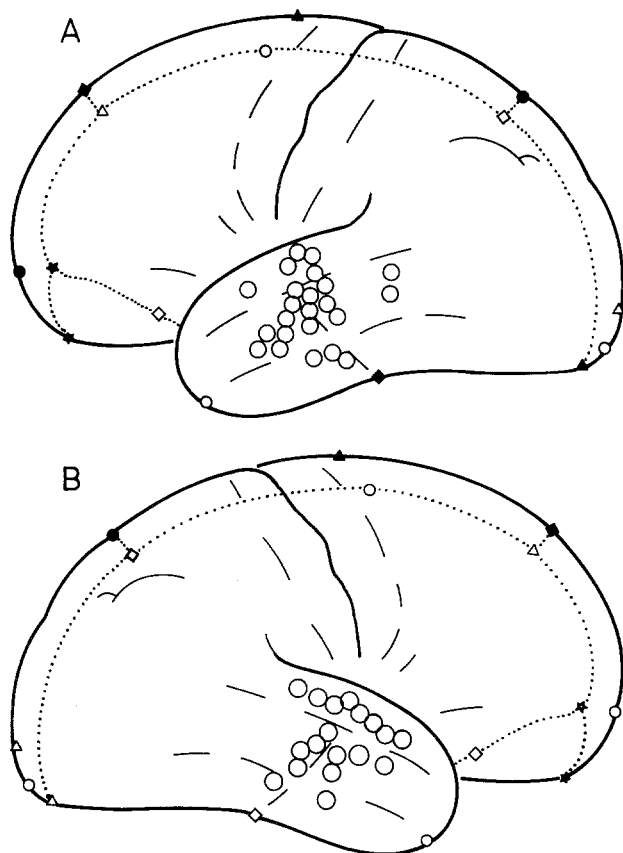


Fig. 1A, B. Schematic drawing of the left (A) and right (B) hemisphere. Sites of microelectrode recordings are marked by circles. The number of recording sites (41) is larger than the number of patients (34), because in some patients recordings from different sites were taken in succession (e.g. middle and superior temporal gyrus, respectively). Successive recordings or simultaneous recordings of several units with one or two microelectrodes at one site are represented as one site

ing, a "pressure foot" made of plexiglass and with a diameter of 1.5 cm was attached to the micromanipulator and lowered onto the cortical surface with slight pressure but without affecting the pial circulation. The microelectrodes were introduced through a 1–2 mm hole in the center of the pressure foot. A silver electrode was attached to the pressure foot close to the central hole for recording the surface ECoG near the microelectrode. Several carbon-silver electrodes were placed on the cortex for recording ECoG and evoked potentials.

When a unit or a cluster of units was identified and found sufficiently stable over a period of 1–2 min, the testing was begun. During recordings from the left hemisphere, visually presented language and memory tests were done first. These tests required 15–20 min. They are described in detail in the preceding paper (Ojemann et al. 1988) and in the fourth report of this series. We shall refer to them here only in the context of speech related activities and for comparison.

Auditory stimuli were then presented in an open field situation. The patients were lying on the side contralateral to the brain operation. The surgical drapes separated the operation field from the face and body, so that the patient could see and directly communicate with the anesthetist and the testing personal, but would neither see the surgeon nor the operation nurse. As the patient was lying on a padded head rest, auditory

stimuli reached him largely through the ear homolateral to the recording site.

Auditory stimuli were presented from prerecorded tapes by a studio cassette recorder placed about 1–1.5 m in front of the patients face. The loudness was adjusted so that it corresponded to the loudness of a normal speaker (~60 dB sound pressure). In addition, activity was recorded during conversation such as specific instructions, information exchange between the surgeon and other persons. Noises and conversations not related to the tests were kept to a minimum, but could not be avoided completely. They were recorded as well and were often useful as control stimuli.

The formal auditory tests always started with a *word list* taken from speech audiometry tests at a loudness corresponding to conversational speech. This list contained 20 words (Corkscrew, Newspaper, Rhinoceros, Rattlesnake, Hotdog, Airplane, Christmas tree, Telephone, Automobile, Alligator, Horseshoe, Spaghetti, Helicopter, Cadillac, Grasshopper, Hamburger, Caterpillar, Toothbrush, Strawberry, Crocodile). They were spoken by a male American speaker. 500 ms before each word a 1000 or 500 Hz tone of 500 ms duration appeared. Each period from one word to the next lasted 5 s. For the last 11 patients, a new word list was designed which consisted of 10 triplets of words, which contained at least one identical formant, but the first word had only one syllable (e.g. Time), the second had 2–3 syllables (e.g. Tiredness) and the third was a compound word of 2–3 syllables (e.g. Pipe-dream). Occasional German words were interspersed the meaning of which was not known to the patient. One patient was German and could not understand English. For him a German wordlist was composed with some English words interspersed, which he could not understand. The patients were first asked to just listen to the words and, in a second run, to repeat each word aloud.

A second tape contained short *sentences* such as "Popcorn balls cost ten cents", "Today is Peters graduation", "The Automobile has two doors". These sentences were presented in various distortions: 1) The peaks of the sound envelop triggered short (0.5 ms) pulses so that only the time structure including the fundamental frequency but not the dynamic structure (loudness, harmonics, envelop and prosody) of the speech signal could be heard; 2) The speech was dissected into several frequency bands corresponding about to the critical frequency bands (Frequenzgruppen, frequency groups, Zwicker 1960, 1961). Their envelops were used to modulate harmonic spectral lines at distances of 100 Hz, which were then resynthesized (vocoder transformation, see Schroeder 1966); 3) Original speech; 4) Same vocoder transformation as in b; 5) The tape recorded sentence was played backwards; 6) The patient was asked to repeat the sentence. Transformations 1 and 2 could not be understood. Transformation 2 sounded like an electronically produced robot voice, while transformation 1 was a monotonous noise in which the time structure of speech could be heard but the sound was very different from a human voice. The vocoder transformation in 4, although physically the same as in 2, could now be understood because the patient had heard and understood the original sentence in 3. The reverse speech in 5 could not at all be recognized. It contained all the spectral elements of the original, but in reversed order, and therefore sounded like a foreign language.

In some patients we also used vocoder transformed single words of the word list. In others we asked the patients to repeat a word or a short sentence only after additional instruction about 5 sec after they had heard the word. Words and sentences were of increasing length and complexity. They were spoken by an American woman.

Microelectrode activity was high pass filtered (1000–3000 Hz) and tape recorded together with the ECoGs from 3 loca-

tions including that of the pressure foot, the electronic stimulus signals and the microphone. After the operation, the tape records were played back on an oscilloscope and photographed with a Grass-camera at a speed of 2.5 or 5 cm/s. Later, all records were computer processed for construction of peristimulus time histograms (PSTH), activity histograms and for averaged evoked potentials. For constructing PSTHs, the single or grouped unit activity was passed through an amplitude detector (Schmitt trigger). Sometimes, PSTHs of the same record were constructed twice at low and at high trigger level, if it was found that both PSTHs were not only quantitatively different but also differed in their time structure suggesting that the activity at low amplitude contained neurons which responded differently from those picked up at the high trigger level.

This is an exploratory study in the setting of ongoing brain surgery, where time for this type of research had to be restricted to 20–30 min or less. Stimuli had to be designed beforehand, with little opportunity to change them during the operation. As no a priori knowledge on the possible responsiveness of neuronal activities in the various locations was available the experimental design had to try to cover a relatively broad range of possibilities. Data were inspected and analyzed only after the operation, since monitoring of on-going activity during the operation was limited to presence of activity free of injury or movement artefacts. Even when the later analysis of responses suggested the application of further tests during recordings, variation of the experimental design were kept to a minimum in order to facilitate comparison between patients and different electrode locations.

Results

Recording sites and general recording situation

We recorded single unit activity from 23 sites in the left and from 18 sites in the right hemisphere during various auditory and speech tests including reading and naming (Fig. 1). Recordings in some patients were not of sufficient quality for detailed analysis, which brings the number of successful recording sites for auditory tests to 36. Often the activity of one single unit could be recorded through one or several tests, but in the majority groups of units were recorded which were usually composed of up to 3–5 single units. Although the same unit or the same cluster of units could be recorded over a considerable time and thus through several parts of the test in some patients, in others the activity was lost intermittently and the microelectrode had to be advanced before the next test. In some sites, we recorded successfully activity from two microelectrodes, and some tests were repeated at different electrode depths or in different locations in the same patient. Therefore, the total number of recordings during speech testing alone comes close to 100 different unit activities.

Single units in the lateral temporal lobe usually do not respond to inadvertent noises in the operation room such as jingling from moving surgical

instruments, sipping noises from the suction apparatus, steps or jingling clicks. Only very few units exhibited slight changes of activity (a slight excitation or suppression) during the 1000 or 500 Hz tone preceding each word. Such small responses could be revealed only by averaging over many trials. Responses to music will be treated in a separate report.

However, conspicuous activity changes related to speech were frequently observed either during listening to the tape recorded words or sentences, to informal speech (instructions) or when a patient spoke himself. In order to give an impression of what sort of activities may be going on in the brain during a recording situation, we show a record from two microelectrodes in the right superior temporal gyrus during conversation directed to the patient or between the investigators (Fig. 2). Both electrodes picked up the activity from several units, but some large (record 1) and smaller action potentials (record 2) could be identified throughout most of the recording situation, especially at higher speed recordings. Figure 2A begins with the last response of a face comparison task, in which the patient was asked to say whether a face shown on the projector was identical to the preceding face or different. When the patient (P.) said “different”, there was some low amplitude excitation visible in record 1(a). However, when speaker Oj. terminated the test and announced the next one (b), record 1 showed much activity and record 2 a slight decrease of discharge rate. This pattern appeared again when speaker OC suggested a different test at low voice (c). The activation in electrode 1 and the suppression in 2 lasted throughout the speaking. Activity in electrode 1 increased again when OC spoke (d in B). During sucking noises (arrows in B) no time locked discharges appeared, and the announcement of some technical information over the loud speaker by speaker HL (e), not directed to the patient, activated only some discharges in 1, which soon wore off, however. Non-verbal vocalization of OC at low voice turned on some activity in 1 (f), but the following jingling noises of instruments did not have much effect. In C and D, the patient received some instruction from speaker OC at low voice (g, i), and several units in Record 1 were activated, while the activity in record 2 decreased. When the patient responded with “hmm” (h) low amplitude activity appeared in record 1. When OC spoke louder (E, j), unit discharges in record 1 were elicited by single syllables.

Such ongoing recordings indicate that 1. nearby electrodes may pick up different activity patterns related to the acoustic signals, 2. speech

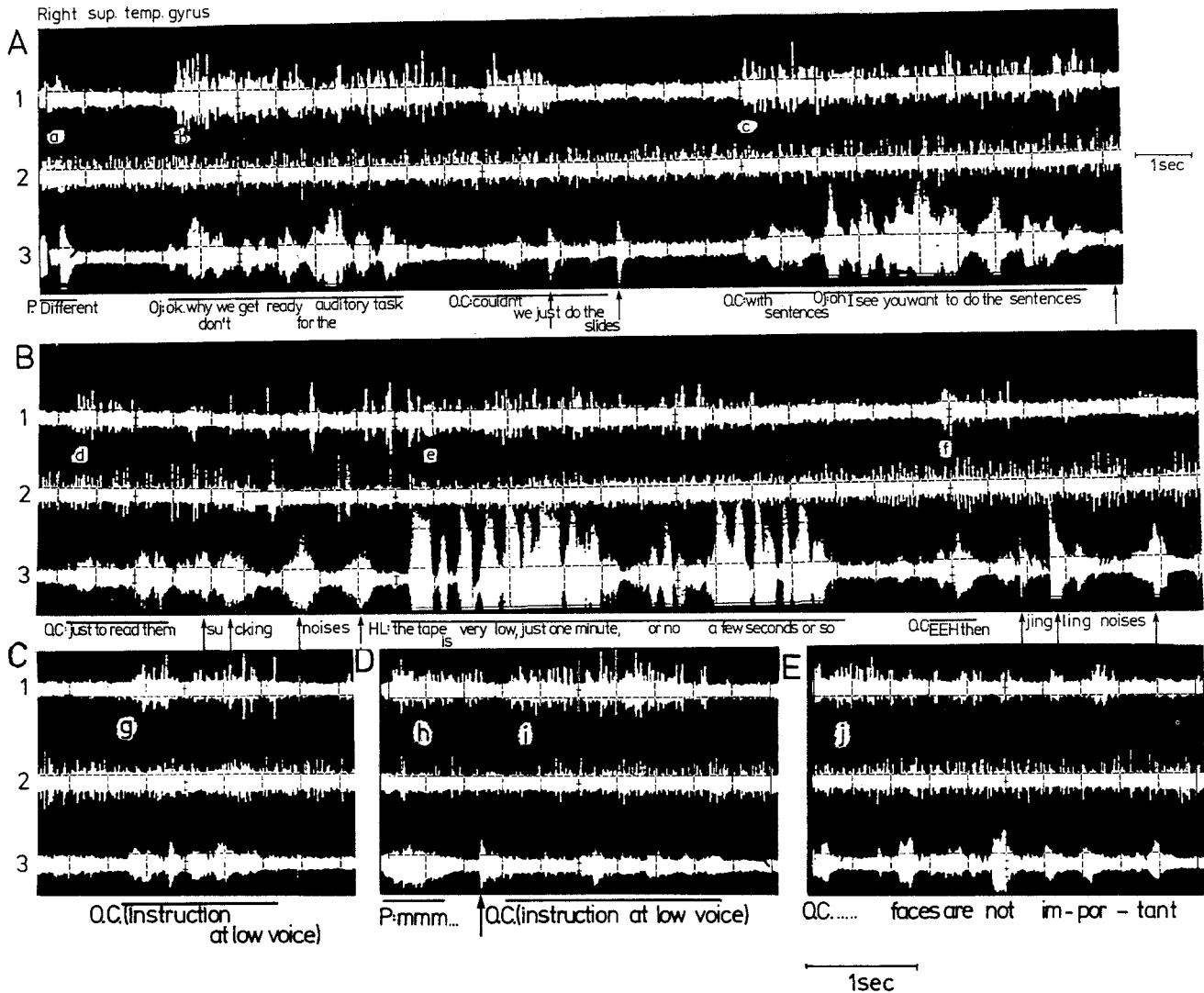


Fig. 2A–E. Simultaneous recording of unit activities with two microelectrodes from the right inferior temporal gyrus during free speech (patient 8533, male). Electrodes were mounted together, the tips were 2 to 3 mm above each other. **A, B** Continuous recording. **C–E** Sections of the recording a few seconds after **A** and **B**. 1 and 2: Recording from the two microelectrodes, 3: Audio signal. The speakers (P.: Patient, Oj., O.C., HL) and what they said is indicated below the recordings. The small letters a–j refer to details mentioned in the text. Time signal (1 sec) after **A** is for records **A** and **B**, that at bottom is for **C–E**.

addressed to the patient or related to his task even at very low voice could produce stronger responses than unrelated loud speech, 3. the patient's own vocalization also may activate neurons responding to exterior speech, and 4. non-speech noises usually do not activate neurons responding to speech. In the following, we shall attempt to break down these neuronal response patterns by demonstrating the responses to various speech related tasks in different locations of the temporal lobe.

Location of speech related neuron responses

Before describing these responses in more detail, we will give a survey of the recording sites (Fig. 1) and their responsiveness to speech signals (Fig. 16).

We recorded successfully during speech testing from 12 sites in the superior temporal gyrus in 11 patients (8 right hemispheres, 4 left hemispheres). All these recordings except 3 were from sites lateral to the pre- and post-Rolandic cortex and 3 were more anteriorly located. 19 recordings (17 patients) were from the middle temporal gyrus (8 right, 11 left hemispheres) and five from the inferior temporal gyrus. All of these recordings were obtained from patients with left hemisphere dominances for speech, based on the amytal test, with one exception. This patient (8531) with left temporal recording, was right hemisphere dominant for speech. Recordings obtained from here are specifically identified as "left, nondominant".

All *superior temporal gyrus* recordings clearly showed speech related responses to words or sentences presented by tape or informally, and when the patient spoke himself. An exception were the three more rostrally located sites in or near the *temporal pole*, of which only one showed a significant excitatory response during informal speech and patient speaking. Responsiveness of *middle temporal gyrus* neuron activity to speech sounds was much less conspicuous. In fact, only one recording showed a clear excitatory response to tape recorded words, two showed a slight excitatory and six a slight inhibitory modulation of discharge rate. In ten middle temporal gyrus recordings (60% of the sites with successful recordings), the unit activity was unresponsive to speech signals from outside. When the patients spoke themselves (repeating words, naming, answers) middle temporal gyrus recordings were more frequently affected (see subsequent report).

Three of the five *inferior temporal gyrus* recordings showed a slight suppression during speech, while two responded with a slight suppression and one with a slight activation, when the patient spoke himself.

Responses to single words

Superior temporal gyrus. Some neurons in this location responded specifically to certain phonemic components of a word. The responses of other neurons distinguished between short and long words or between multisyllable and compound words. Still other neurons showed rebound activation near the end of a word or immediately following it, especially after short words. All excitatory neuronal responses in the superior temporal gyrus were characterized by bursts of activity ranging from a few discharges to discharge trains lasting up to a few hundred ms. Peak discharge rates of single units could reach 100/s and mean rate during bursts were between 25–50/s. Examples are shown in the following figures.

Some units appeared to be specifically responsive to certain phoneme categories. An example is shown in Fig. 3, from the right superior temporal gyrus of a female patient. This unit substantially discharged during some words of the list (words 1 and 7 in this figure), only during some syllables in others (15, 17, 19, 20) and did not respond at all to the remaining words (4, 6, 8, 9, 13, 16). These responses were repeatable as the same pattern of responsiveness was seen when the word list was presented two minutes later in B, when the patient

was asked to repeat each word. With some variations, the activation pattern was similar during the first and second presentation of the same words. Closer inspection shows that the responses were related to certain sounds. Activation typically appeared after the guttural closure consonants *k* and *g*, especially when these were combined with another consonant such as *r* (*cr*, *gr*) or *s* (*s*, *sk*), and after the fricative *s*, in combination with another closure consonant (*st*, *str*). The activation often continued throughout the vowels following these consonant sequences. About three minutes after this word listening and repetition task, the patient was asked to listen to sentences spoken by a female speaker and to repeat them after a delay. Again the neuron responded strongly to the consonant sequences *cr* in the sentence: "The child *is crying again*", both when listening and when repeating the sentence (see Fig. 3 in the following report, Creutzfeldt et al. 1989a). The time that elapsed from the beginning of the consonants (e.g. *chr* in *christmastree* or *cr* in *crocodile*) was 60–100 ms. Although this is tuned to a broad category of phonemes, we cannot decide at this stage whether the consonant category alone or the combination with a subsequent vowel was the trigger feature. The long lasting activation during some syllables suggests the latter.

The response of the neuron of Fig. 4 clearly distinguished between long multisyllabic and compound words. It was also recorded in the right superior temporal gyrus but of another patient. A burst of activity appeared during the second part of compound words, in which the second syllable changed the semantic meaning conveyed by the first syllable (A–D, I), but this activation was reduced or completely absent during the second or third syllables of polysyllabic words of similar length, in which the first syllables were non-words with no semantic meaning (E–H, K). The time that elapsed between the beginning of the first vowel of the second word component and the neuronal activation was between 60 and 100 ms. However, the activation lasted longer than the first syllable of the second word component and could continue throughout following syllables. The one exception in this series is the word "*helicopter*". This is, of course, also a compound word, but this may not have been obvious to the listener. The first part of this anglicized word could exist, however, by itself as an independent word or could be part of another polysyllabic English word (*hell*, *hello*, *help*, *helix*, *helical*), and the syllable combination "*heli*" is used in various words such as *heliport*, *helium*, *heliocentric* etc, which maybe not uncommon to

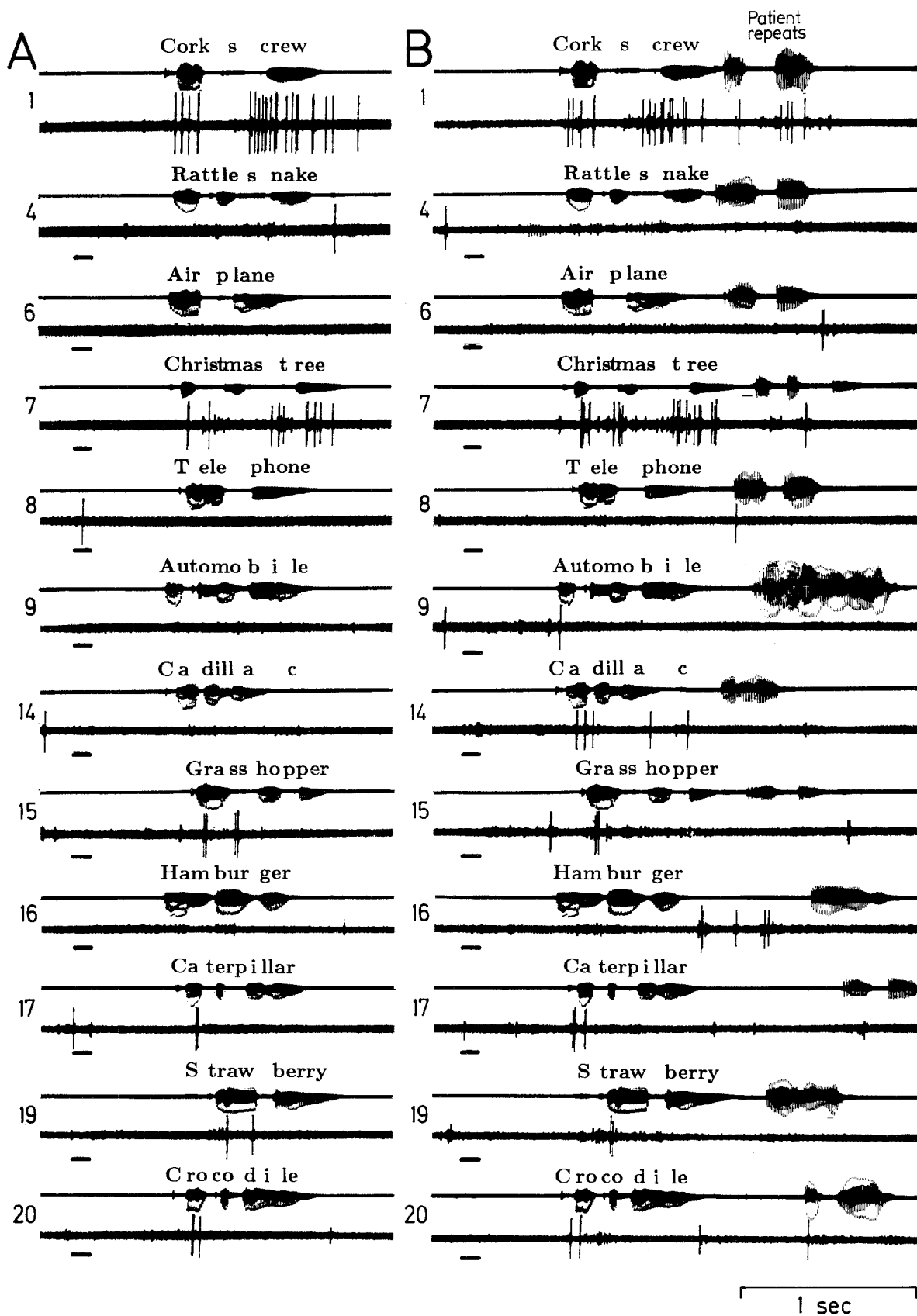


Fig. 3A, B. Responses of a neuron in the right superior temporal gyrus to a restricted class of phonemes (8512, female). The words were tape recorded and played to the patient in an open field situation with 5 seconds between words. The numbers to the left of the records refer to the sequential number of the respective word in the word list. Not all words are shown in this figure. **A** The patient was asked just to listen to the words. **B** 3 min later, the patient was asked to repeat each word. In each record, the audio signal is recorded on top and the single unit activity below. The short line at the left below each record indicates the appearance of the 1000 Hz tone which preceded each word

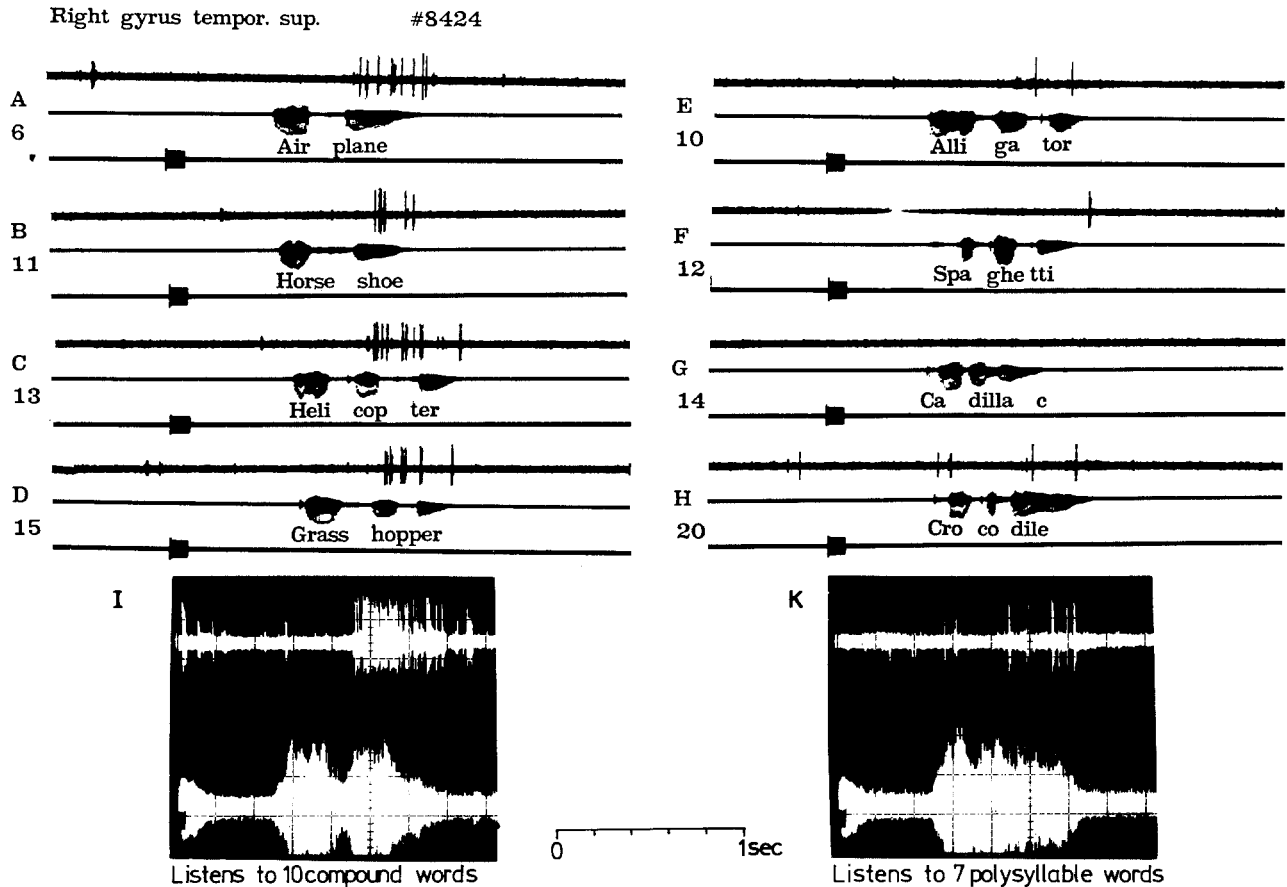


Fig. 4A–K. Selective responses of a neuron in the right superior temporal gyrus to the second part of compound words (8424, female). A–H Selection of original recordings. The sequential number of the words in the tape recorded word list is indicated under each letter (e.g. the word airplane was the sixth word in the word list, the word horseshoe the eleventh etc.). In each recording, the single unit recording is on top, the audio-signal in the middle and the 1000 Hz tone in the third line. I, K Superimposed responses to all ten compound words (I) and to all seven polysyllabic words (K) of the word list

average American speakers. The speaker had, in fact, separated both parts of the word by a long pause as in the other compound words, indicating its compound nature, while in the usual multisyllabic words the pauses between syllables were shorter (see phonograms in Fig. 4). Neuronal responses of this type may have a function in composition of word segments to word entities probably largely based on temporal features, but without specific semantic coding, since the activation was seen in all compound words of this series.

Other neurons only distinguished short and long words, and tended to respond more to long polysyllabic than to short monosyllabic words, irrespective of whether the long words were compound or not. A response of such a neuron is shown in Fig. 5. It was recorded in the left nondominant superior temporal gyrus of a female patient. The large unit in this recording discharged only once or not at all during monosyllabic words (left column), but showed a longer burst of activity from

the second syllable onwards (right column). The activation always started during the second syllable even in those compound words in which the second word component began with the third syllable (Angel-fish, Christmas-tree). This activation appeared also during a long German word (Liebeslied), the meaning of which was not known to the patient. Such responses may be significant for prosody of speech, integrating a syllable into a word, i.e. speech functions that have been particularly related to the nondominant hemisphere (Ross 1981). The first syllable might have increased the excitability, so that excitation from the second syllable onward reached threshold. It may be noted that in this record a second, smaller unit appeared during the last words, which was suppressed during short words (peak) and during the first syllable of multisyllabic words (Liebeslied).

Some units demonstrated both enhanced activity towards the end of compound words and activation specifically related to certain phonemes. An exam-

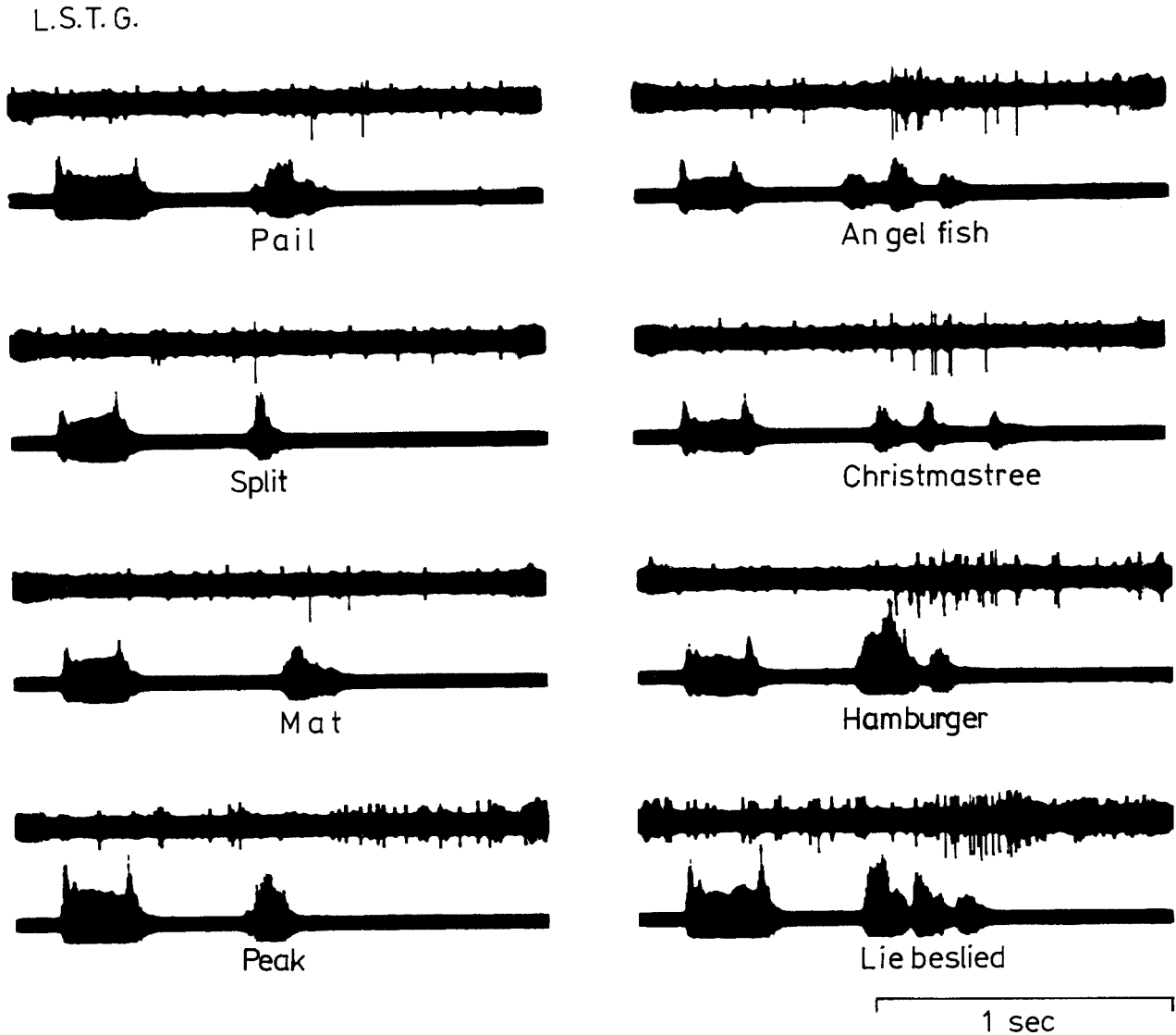
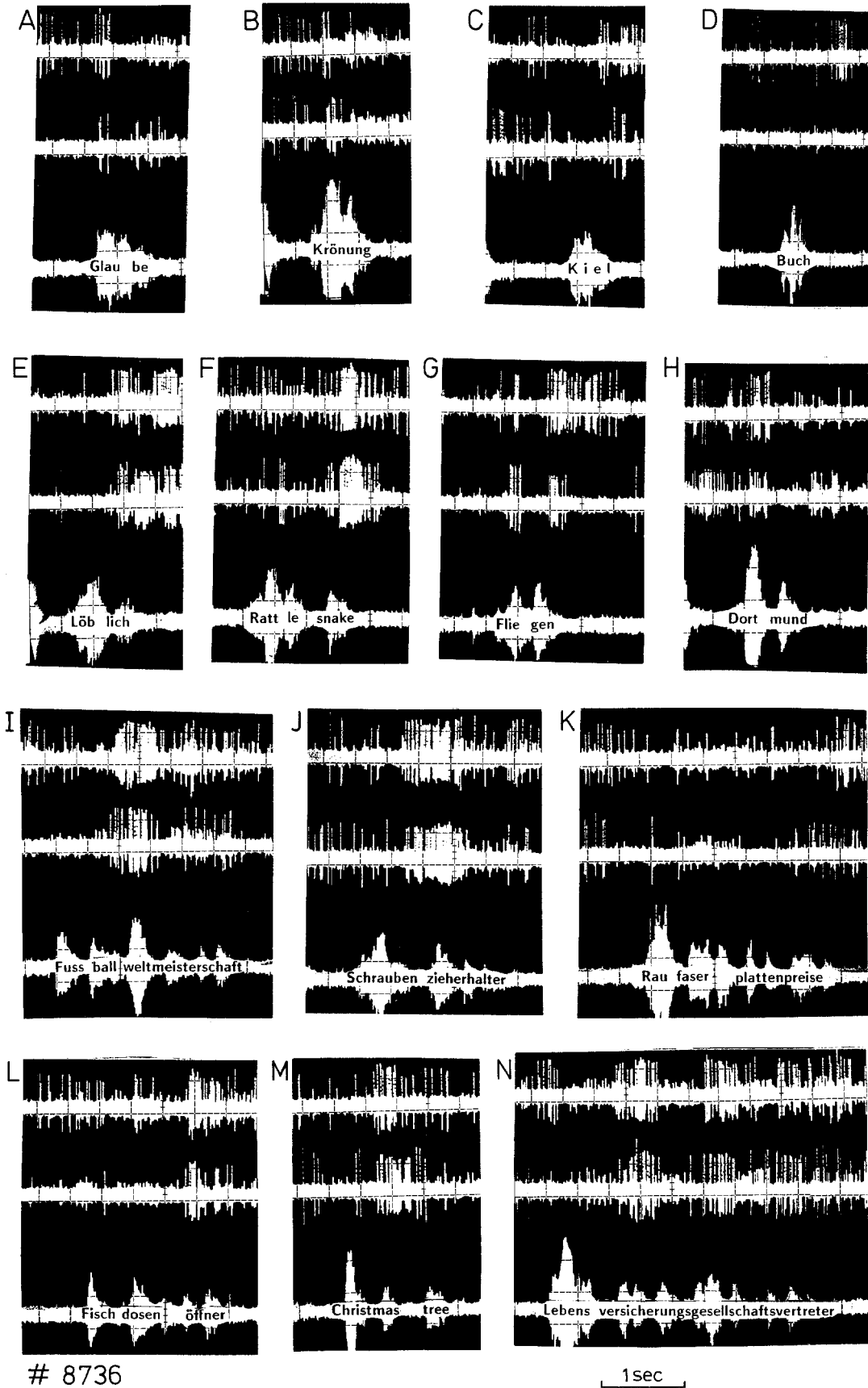


Fig. 5. Strong responses to multisyllable (right) but not to monosyllable words (left) of a neuron in the left nondominant superior temporal gyrus (8531, female). Single unit recording on top, audio recording at bottom (500 Hz tone preceding each word). The large unit could be identified throughout the whole recording session, and an additional unit with small action potentials was picked up by the electrode during the last two words

ple of this is shown in the recordings of Figs. 6 and 7. These records are from the left superior temporal gyrus of a male German speaking patient who had virtually no knowledge of English. The two records in each box of Fig. 6 are responses to the same words presented twice but at an interval of three minutes and in different conditions. During the first presentations, the patient was asked just to listen to the words, and the second time he was asked to repeat each word aloud or to count silently the animals contained in another word list. The record during these auditory tests was dominated by one large unit which could be identified throughout the five minute recording al-

though the amplitude decreased slightly towards the end. A third unit of still larger amplitude was recruited regularly during word repeat (see Fig. 6 in the following paper). The activation pattern elicited by the same words under the different conditions was very similar if not identical, as can be appreciated in Fig. 6. During short words, the unit was usually not active, although there was often a short rebound activation towards the end or following the word (Fig. 6C, D; Fig. 7A). After 2-syllable words, such a late activation occasionally also appeared (Fig. 6E, G), but was less common (Fig. 7B). During long compound words, a strong activation regularly appeared during the second or



8736
Fig. 6

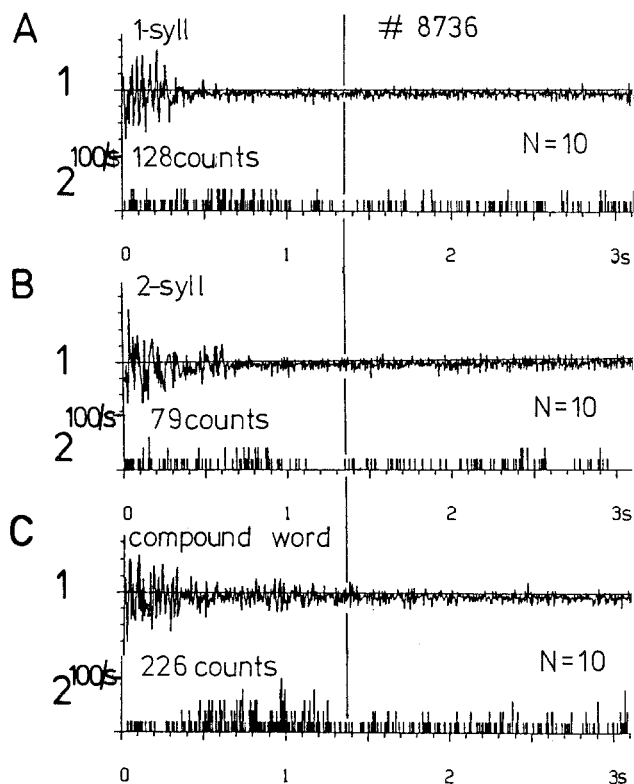


Fig. 7A–C. Different responses to short and long words of a unit in the left superior temporal gyrus (8736, male). The histograms are from the activity of which a selection is shown in Fig. 6. Discharge averages during 10 one syllable **A**, ten two syllable **B**, and ten compound words **C**. The PSTHs were synchronized on the beginning of each word. The numbers written above each histogram are total numbers of discharges during 1.4 s, which correspond to the longest compound word of the list. Binwidth 5 ms

third part of the words (Fig. 6I–N, Fig. 7C). A closer look on these records suggests, in addition, activation by a broad class of phonemes contained in certain compound consonants (*gl* in Fig. 6A, *kr* in B, *bl* in C, *k* in F, *fl* in G, *rt* in H). These activations always lasted longer than the word component that apparently triggered them so that it is difficult, as in the example of Fig. 3 to decide whether only the consonants or a given consonant/vowel combination were responsive.

Suppression of activity during certain parts of words was also seen, but was less conspicuous partly due to the low spontaneous discharge rate of lateral temporal lobe neurons. The activity of some active neurons or neuronal populations may be suppressed throughout words or sentences, usually followed by a rebound at the end of speech.

Middle temporal gyrus. As mentioned above responses to words or sentences were found in only half the recording sites. They were much less obvious and, in the majority inhibitory. Excitatory responses never reached maximum discharge rates above 10/s and were usually less. No close links to phonemic or prosodic features could be identified. The activity averaged in Fig. 8A and B was slightly suppressed during the word (and also during the 1000 Hz tone), but sometimes showed a slight rebound afterwards.

When the patient was asked to repeat the words (B), the suppression during the repeat was more obvious. The neuronal activity of Fig. 8C, from another patient, was also initially suppressed with a slight rebound towards the end of the words, which is seen more clearly when the activity was synchronized with the beginning of the word (E, F). The initial suppression during listening changed into a slight activation when the patient had to repeat the words aloud (D; see also below).

Variability of responses depending on the task

The responses of some neurons varied with attention. Thus, the neurons in Fig. 2 (right superior temporal gyrus) responded strongly to speech addressed to the patient (b, c), but only slightly with speech not directly addressed to him (e). Excitatory responses to word presentation were sometimes enhanced or diminished, or a slight suppression changed into excitation when the patient was asked to repeat each word instead of just listening to them. This could represent increased attention directed to the stimulus or preparatory activity related to the overt response. However, such task relat-

Fig. 6A–N. Responses to phonemic aspects and word length of a neuron from the left superior temporal gyrus (8736, male, German speaking). Selection of short and long compound words from a tape recorded word list. This electrode picked up the activity of several units, but one large action potential could be identified throughout the recording period. In each section of this figure, the response of the unit is shown to a first (top) and second (middle) presentation of the same word (bottom). The first and second recordings were separated by about 3 minutes. The recordings on top were taken while the patient just listened to the words, while during the second recordings in **A–D**, **G**, **H**, the patient was asked to silently count the number of animals in that word list. In the remaining recordings **C**, **E**, **F**, **I–N**, the patient was asked to repeat each word. Note the similar responses to the various words during the first and second presentation. The legend on the audiograms only approximately matches the syllables with the corresponding parts of the envelope. This patient could not understand English words

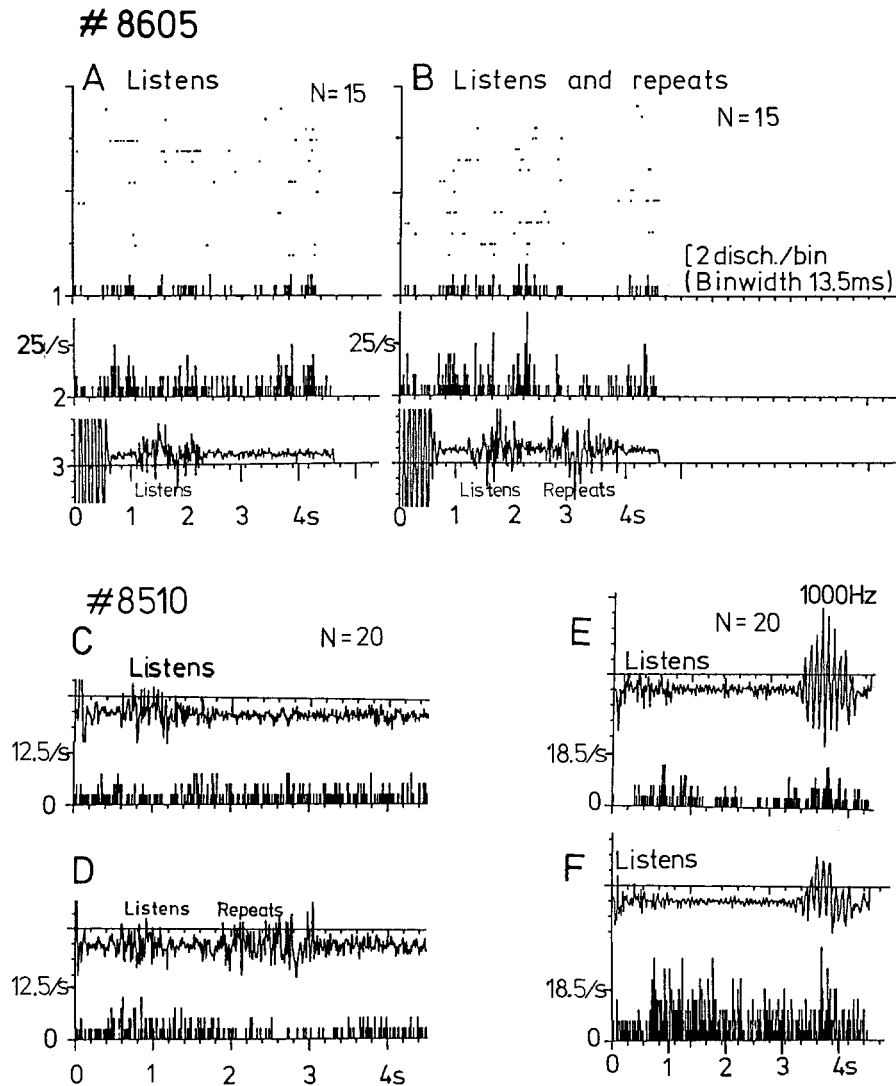


Fig. 8 A–F. Responses to words of different units in the middle temporal gyrus of two different patients (A, B: 8605, right side, C–F: 8510, left side). A, B 1: Single dot display and PSTH at high trigger level (one single unit), 2: PSTH at lower trigger level (multiunit activity), 3: Averaged audio-response. In A, the patient listened to 15 words preceded by a 500 Hz tone. In B, he repeated each word after presentation. Binwidth 13.5 ms. C–E Histograms from one single unit, synchronized with the 1000 Hz tone in C, D, and with the words in E, F. C–E High trigger level (one single unit). F Low trigger (multiunit activity). C, E, F The patient listens to 20 words. D The patient was asked to repeat each word after its presentation. Binwidth 10 ms in C, D, 13.5 ms in E, F. The slow modulations on top of the 1000 Hz tone in this and other averaged records are due to sampling and do not reflect the tone frequency

ed changes of responsiveness were not seen in all neurons, probably in less than 50% (we hesitate to give an exact figure, as each block of the listening and repeat task lasted several minutes, and stability of recording could not always be guaranteed). Task related changes of responses were again most obvious in some *superior temporal gyrus* recordings. The large unit in Fig. 9A–C only occasionally discharged towards the end of the word when the patient just listened to them (A). However, when the patient was asked to repeat each word, the unit became activated already early during the words but it did not respond during the patients overt word repeat (B). When a message was spoken over the audiosystem the neuron became active only after a significant latency as if it was turned on only when the patient concentrated on the auditory signal (C). A similarly strong

effect is shown in the example of Fig. 4 of the following report.

Some units of the superior temporal gyrus decreased activity (down to 50–70%) when listening to words that the patient had been instructed to repeat compared to the same words presented without the instruction. The histograms of Fig. 10 are from the same recordings as those shown in Figs. 6 and 7. The response decrease from the pure listening runs (1) to the listening for repeat runs (2) was strongest in the monosyllable word series (Fig. 10A, 1 and 2, respectively). As described earlier, activation by one-syllable words was, in this recording, predominantly a rebound activation following the word. When the words had to be repeated, this period was the preparation interval for repeating the word. One may therefore argue that the mental process involved in this preparation for

8509 R.S.T.G.

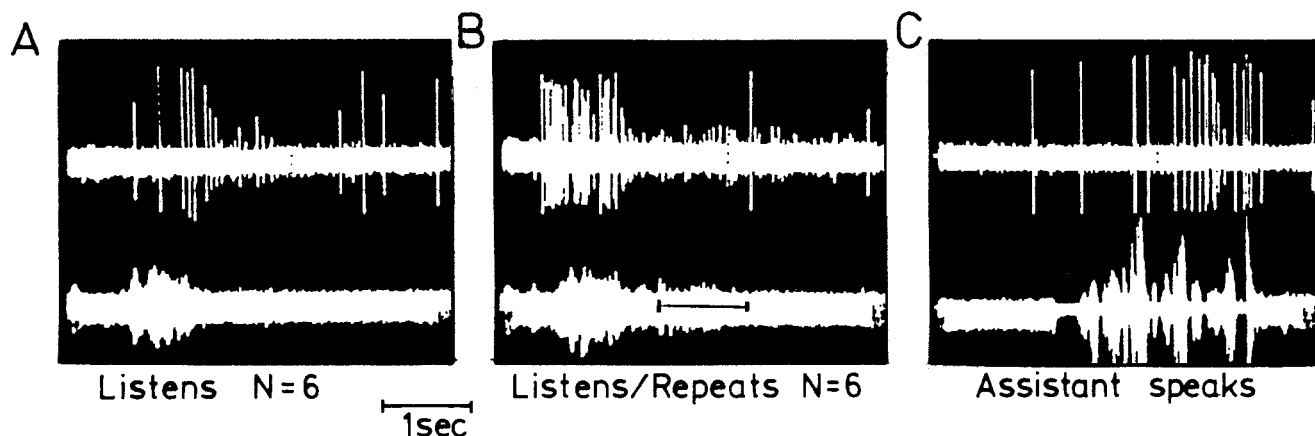


Fig. 9A-C. Task dependant neuronal responses to language in the right superior temporal gyrus (8509, male). **A** When simply listening to words (A) the unit responded only rarely towards the end of the words. Superimposition of responses to 6 words. **B** When the patient was asked to repeat the words, the unit already responded at the beginning of the words, but does not respond to the patients repeat (horizontal line on audio-record). Superimposition of 6 words. **C** When a message is spoken over the audio-communication-system, the unit becomes activated only towards the end of the sentence

repeat contributed to the suppression of the rebound activation.

In *middle temporal gyrus neurons* such changes of responsiveness were less obvious but could be recognized occasionally (see Fig. 8). Thus the small rebound activation in the example of Fig. 8A was slightly enhanced when the patient had to repeat the word (B), and the slight suppression of activity in the example of Fig. 8C turned into a slight activation when the patient was asked to repeat the words (D). Now, the suppression appeared during the word repetition.

During the free speech situation, it was sometimes difficult to decide to what extent the activation of a neuron was an effect (or correlate) of general arousal, of directed attention or whether it was specifically related to the acoustic speech signals. A case where this difficulty is apparent is shown in Fig. 11. In this recording from the left non-dominant *superior temporal gyrus*, the unit(s) discharged in rhythmical bursts with a beat of 3-4/s (A). These bursts were closely related to the slow θ -waves of the EEG which in this patient were the prevailing EEG pattern of this region (not

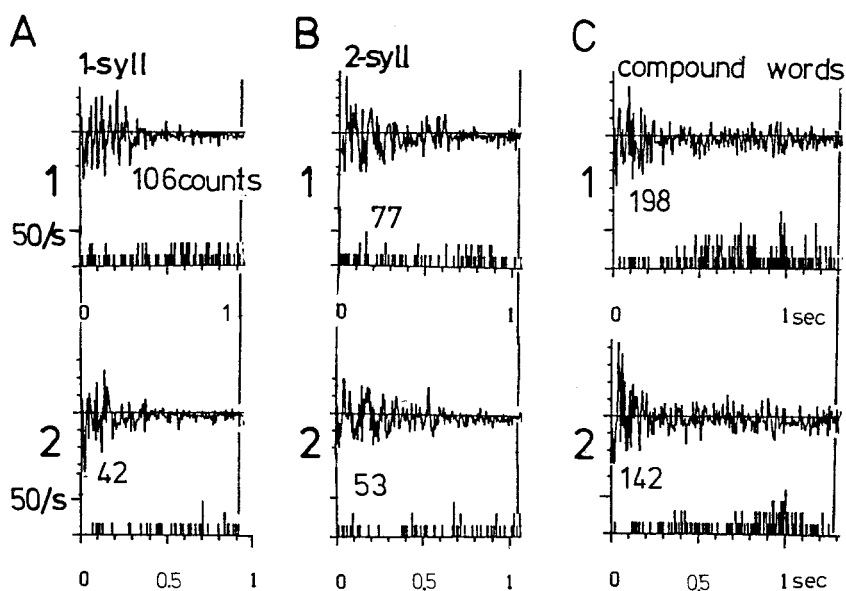


Fig. 10A-C. Decrease of responses to short and long words in the left superior temporal gyrus when words had to be repeated (8736). **A** Responses to one syllable, **B** to two syllable, and **C** to compound words. 1: The patient just listened to the words, 2: He is asked to repeat the words. The numbers refer to the discharges counted during the response periods indicated by the vertical lines. Binwidth 5 ms

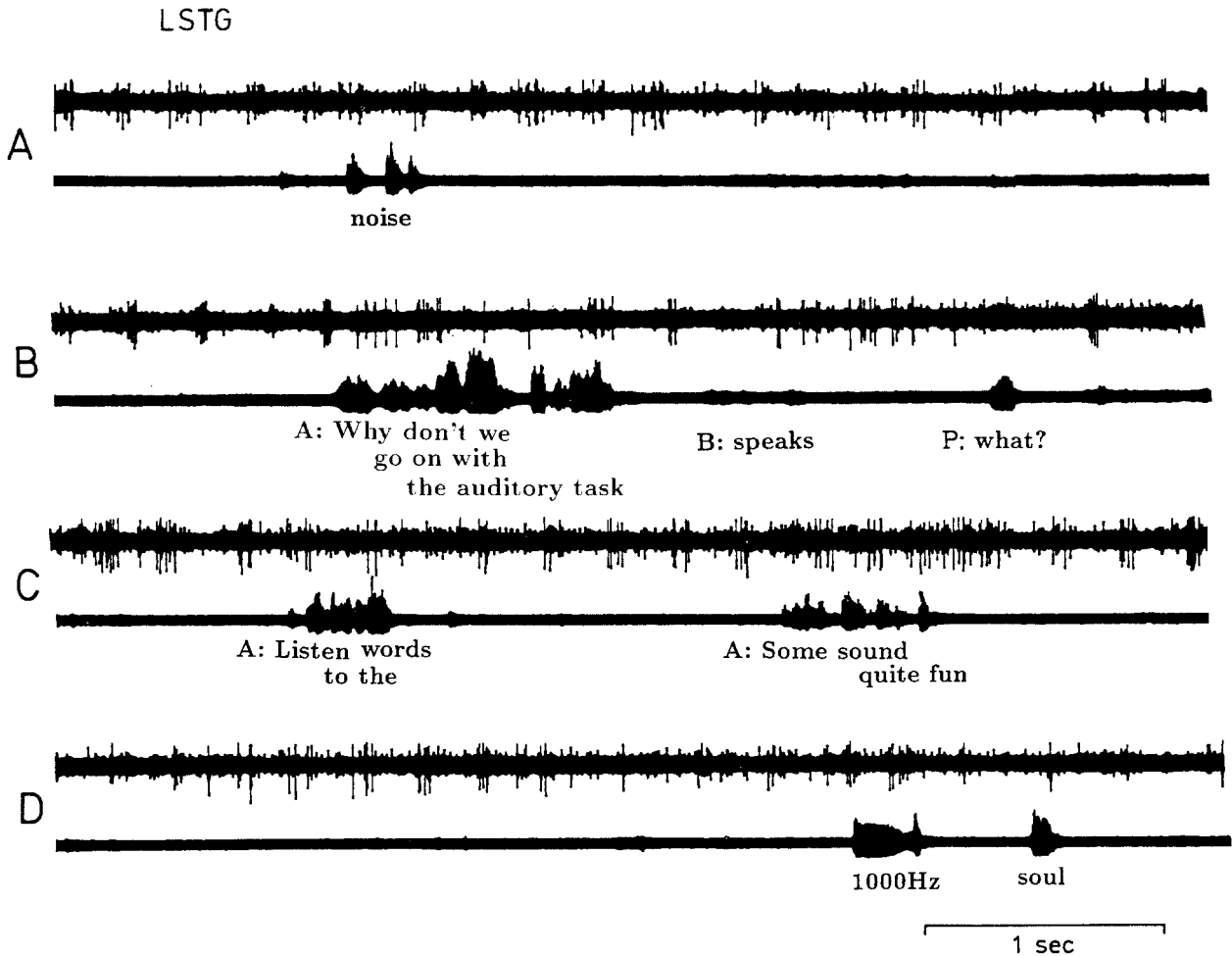


Fig. 11A-D. Change of discharge pattern of single units in the left nondominant superior temporal gyrus during speech addressed to the patient (8531, female). One large unit can be identified throughout the record, in addition to activity of smaller units. **A** During rest the units discharge in short bursts at a burst rate of 3-4 per s. These bursts are synchronous with the θ -waves in the ECoG which are not shown in this record. A jingling noise does not interrupt this activity. **B** When speaker A announces the next auditory task, the rhythmical burst activity is interrupted and the large unit is shortly activated. This is also seen when speaker B speaks at low voice to the patient. When the patient (P) responds, the activity is suppressed. **C** Ten seconds later, the burst activity has been resumed but is to some extent interrupted by the task announcement. **D** While the patient waits for the beginning of the word list, some burst activity is present, but again interrupted, when the task begins (1000 Hz, followed by a vocoder transformed word, which the patient could not understand)

shown in this record). When a task was announced by speaker A, the burst activity changed into a more continuous activation pattern which was even stronger when speaker B gave some additional information at low voice (B). The patient's voice, when she asked "What" did not activate the neuron but rather silenced it for a short while. The next instruction from A, given a few seconds later, did not affect the activity significantly, but a further comment led again to an interruption of the rhythmical burst discharges and a more continuous activation (C). Following this, the activity relapsed into the 3-4/s burst rhythm until the auditory task began. During the formal testing, the same units had only been activated by long words (see Fig. 5).

We may therefore interpret the activation of this unit as a specific response to the speech addressed to the patient but under control of attention.

In the *middle temporal gyrus* a slight arousal activation was seen during ongoing speech in some neurons. Here, the neuron in Fig. 12 may serve as an example. It was recorded in the anterior portion of the right middle temporal gyrus during the diagnostic ECoG-exploration. From time to time, the electroencephalographer reported on his observations and diagnostic evaluation over the audio-communication system. The neuronal activity was low throughout the resting period, but became activated during these announcements without apparent relation to phonetic elements. Activity in

#8603 R.M.T.G.

	Large unit	Multunit
Resting activity	1.1 ± 1.2	2.4 ± 1.5 disch/s
During speaking	2.9 ± 1.2	5.6 ± 1.8 disch/s

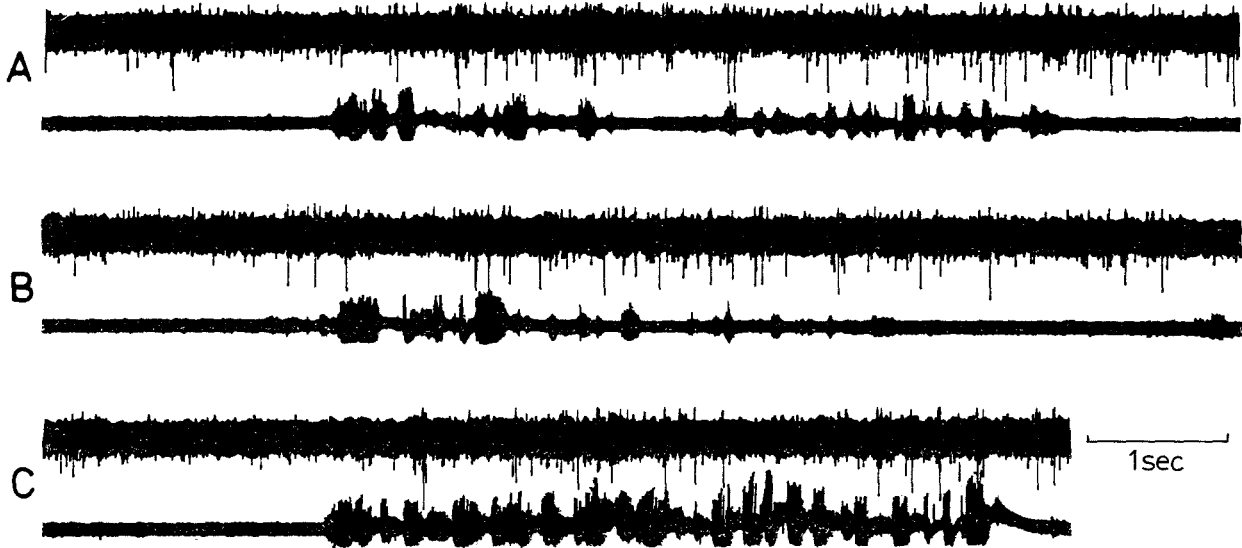


Fig. 12A–C. Non-specific activation of neuronal activity in the right middle temporal gyrus by speech (8603, female). One single unit with large amplitude and multiunit activity with small amplitude can be distinguished. The unit activity is low during rest and increases 2–3 times while the electroencephalographer describes the EEG records to the surgeon over the audio-communication-system (see audio monitoring in second line of records). On top, discharge rates per second during 30 s of resting activity and 20 s of speech (mean rates and standard deviation for each second)

the right superior temporal gyrus of this same patient showed a clear activation predominantly during compound words and with coupling of activity to some phonemic aspects.

Acoustically distorted speech

This part of the study was designed to distinguish between responses to acoustically correct and to distorted, not-understandable speech. As further explained in Methods, language was distorted in different ways: reducing it to pulse frequency modulation (1), vocoder transformation (2, 4) or by playing it backwards (5). The speech samples were short sentences. Each sentence was presented twice in a distorted way, which the patient did not understand (1, 2), then in its original form (3), followed by a repeat of the same vocoder transformation as 2, which could now be understood because the listener knew what it meant (4), and then by reversed speech (5), which contained the same spectral components in reversed order but was not understandable. Each run was completed by asking the patient to repeat the sentence. We have run this test during microelectrode recordings in 4 patients in the superior temporal gyrus, in 7 patients

in the middle and in 2 patients in the inferior temporal gyrus. This test takes half a minute for each sentence, and therefore stability of the recordings was often not sufficient to run through all 10 sentences (5 min).

In only 5 recordings have we observed consistent differences of neuronal activity during the variously distorted sentences. Four neurons changed activity only during undistorted speech, and one other only during two understandable speech representations. An example of this is shown in Fig. 13A–D. This neuron in the left middle temporal gyrus was slightly activated during the real speech sentence (3). As is typical for middle temporal gyrus neurons, the activity was not linked to any phonetic features of the sentences. One may interpret this activation as representing “surprise” (ah, that’s what he said), or selective attention related to language rather than a response to the specific phonetic content or sentence meaning. It should be realized that this neuron also showed periods of activation not related to the sound of the speech such as preceding transformation 1 and 5 in Fig. 13B. Thus, the activations during the normal speech representation can have, per se, only limited significance to the understanding of the

LMTG

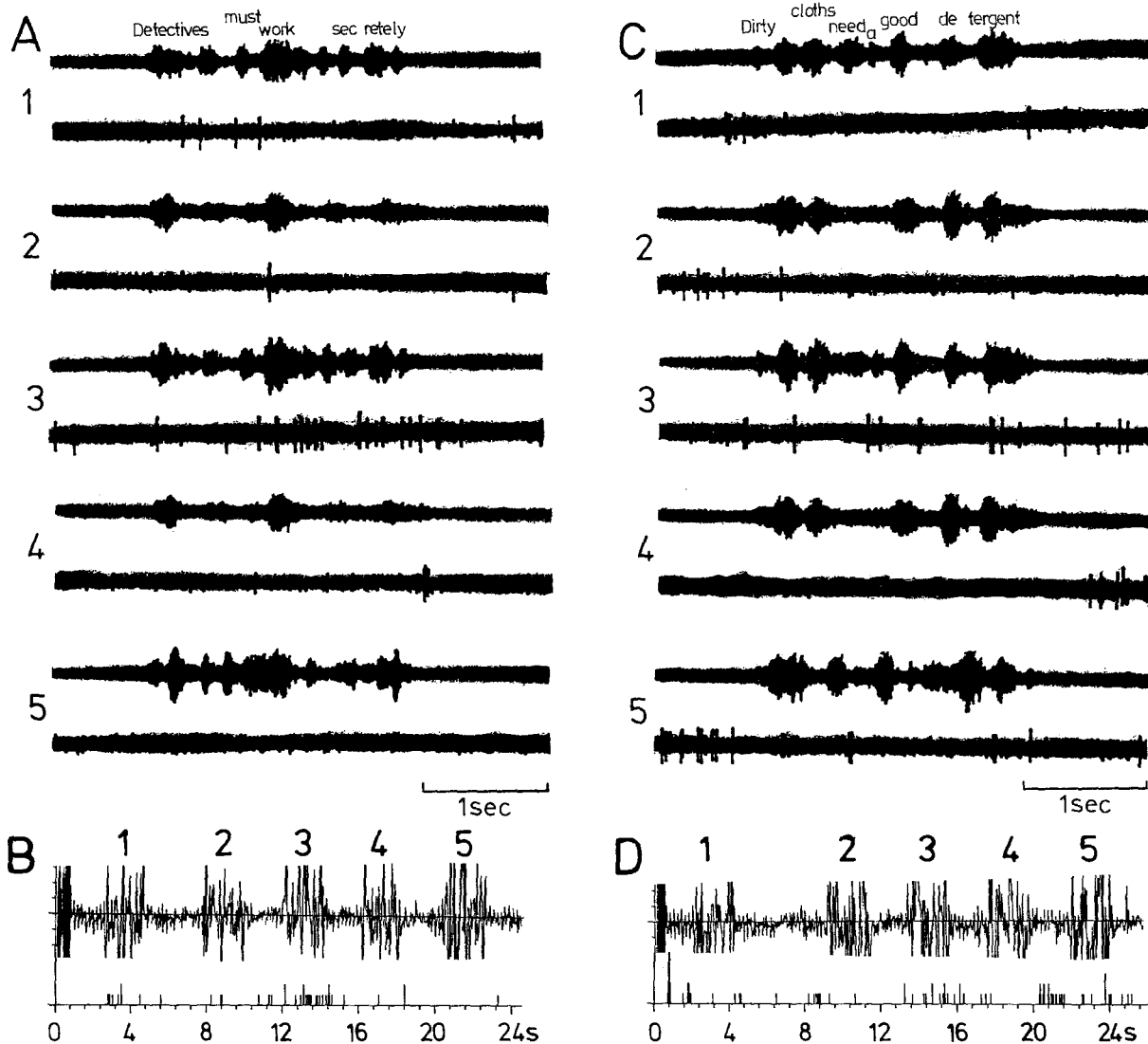


Fig. 13A–D. Activation of a single unit in the left middle temporal gyrus only by undistorted speech (Patient 85 PRO, female). **A, C** Original records. **B, D** Recording of discharge rates. The patient heard two different sentences in **A, B** and **C, D**, respectively, as written on top of the first audio record. The speech was distorted acoustically in different ways (see Methods): 1) Pulse interval transformation, 2) Vocoder transformation, 3) Original speech, 4) Vocoder transformation (as in 2), 5) Back to front replay of the sentence. Only sentences 3 and 4 could be understood. A slight activation of the unit activity is only seen during the presentation of the sentence presented in original speech (record 3) but occasional bursts of activity appear between sentences. Binwidths in **B, D** 40 ms

sentence, if not correlated with activity patterns of other neurons in a rather specific way.

Figure 14 shows discharge counts during and following the presentation of the various sentence transformation of three other neurons. The activity in Fig. 14A is from a neuron in the right superior temporal gyrus. It is clearly suppressed during the presentation of the undistorted (3) and the reversed (5) sentences, and shows a sharp off-activation following the real speech sentence (3) and a smaller

one after the reversed sentence (5). This might indicate, as one might expect from a superior temporal gyrus neuron, a response specifically related to undistorted language sounds. During the real speech sentence, some short activations related to segment borders were present in the original records during two sentences. The activity of the middle temporal gyrus neuron in Fig. 14B showed an increase only during the real speech sentence (3), and that in Fig. 14C appeared to be slightly suppressed

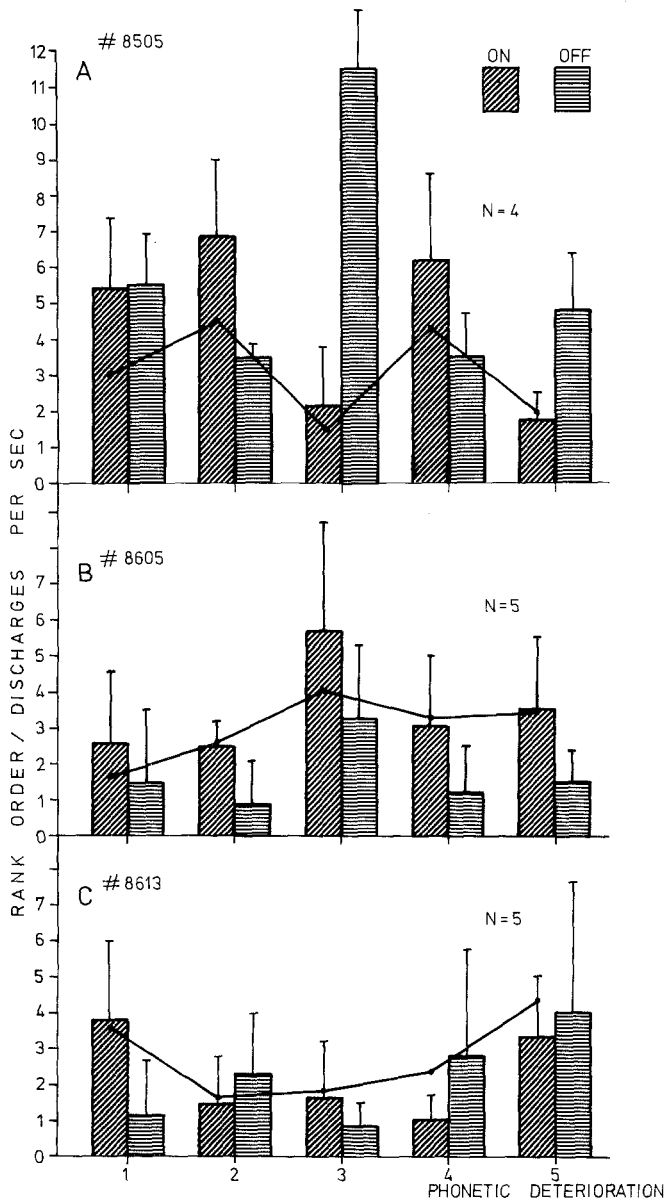


Fig. 14A–C. Discharge rates during acoustically distorted sentences. The numbers on the abscissa refer to the type of speech distortion as explained in the legend to Fig. 13 and in the text. *Diagonally hatched columns:* Activity during sentence presentation (2 s). *Horizontally hatched columns:* Activity during 2 s following sentence presentation. Mean discharge rates with standard deviation (vertical lines on top of each column). *Points connected by lines:* Mean value of rank orders of activity during each sentence. **A** Patient 8505, superior temporal gyrus. **B** Patient 8605, middle temporal gyrus. **C** Patient 8613, middle temporal gyrus. Data are, in each patient, from one single neuron

throughout the period of presentations 2–4, i.e. the just incomprehensible vocoder transformation (2), the real speech sentence (3) and the understandable vocoder transformation (4). One is tempted to relate also these middle temporal gyrus responses

to the process of understanding or trying to understand the sentences.

Despite the limited data so far obtained with the test, the differences between the superior and the middle/inferior temporal gyrus recordings were in the same direction as those found during the normal speech tests. The activity of the superior temporal gyrus neuron showed a distinction between real and acoustically distorted speech with sharp discrimination between the speech and the non-speech period, while the middle temporal gyrus neurons showed only slight activity changes without sharp discrimination between speech and nonspeech periods.

Evoked ECoG-potentials

The electro-corticogram was recorded at the microelectrode recording site with a small silver ball electrode fixed to the pressure foot (see Methods), as well as from various other sites. Details of the electro-corticographic recordings and neuronal unit activity will be described in a later report. One example is illustrated here to show the complex relationship between the activity of single units and the surface evoked potential, and that the latter may give only limited information on the actual distribution and time course of intracortical neuronal activities evoked by a complex auditory signal. The potentials in Fig. 15 were recorded from the same patient and during the same auditory tasks as shown in Figs. 6 and 7. In A and B the patient listened to 40 short words, in C and D to 30 words, 10 of them one-syllable, 10 two-syllable and 10 long compound words. The histograms below each evoked potential are from the largest units recorded at this site. It can be recognized, that the ECoG-potentials elicited by the 500 Hz-tones (Fig. 15A and C) clearly differed from those elicited by the words (B and D). Also the summed unit activity differed in both situations. No prediction can be made from the evoked potential as to the responses of this single unit, however. In A and B the unit activity was slightly suppressed during the initial positive phase of the evoked potential (60–120 ms) and slightly activated towards the end of the subsequent negative phase (120–250 ms), but this is less clear in C and D. We have other examples in our material which show a closer relationship between single or multiple unit activation, and it may be that a grand average from many units might give a better correlation. Yet, a tighter correlation between both events can be expected only, if stimuli are identical, i.e. if responses to single words would be averaged.

8736 L.S.T.G.

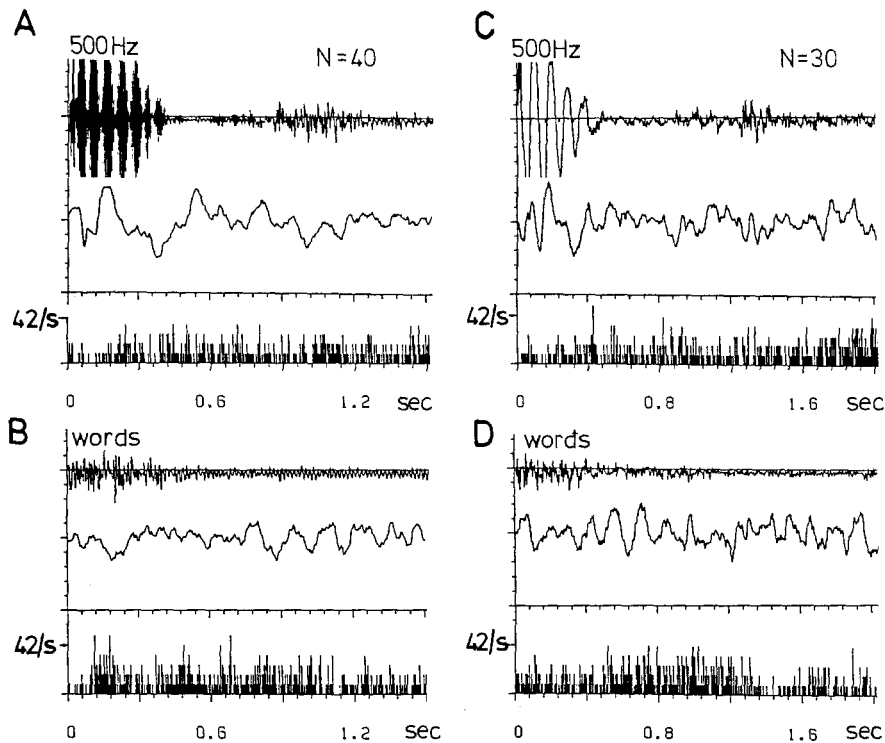


Fig. 15A–D. Evoked ECoG-potentials and single unit responses to a 500 Hz tone **A, C** and to single words **B, D** (Patient 8736, left superior temporal gyrus). The electrocorticogram was recorded from the surface above the microelectrode recording (see Methods). **A, B** Averaged responses to 40 tones **A** and 40 short one and two-syllable words **B**. **C, D** Averaged responses to 30 tones **C** and 30 predominantly long words **D**. Note, different time scales in **A, B** (bin width 3 ms) and **C, D** (bin width 4 ms)

Discussion and conclusions

The observations on single neuron responses in the lateral temporal lobe to speech give some insight into the working of this brain region during listening to linguistic signals. At this stage of a pilot exploration of a new field we considered it appropriate to present examples of original recordings and analysis, in order to give an impression of the type of complexity and variety of neuronal signals elicited by speech sounds. The sample of recordings was collected without regard to any effects of the behavioral measures, and with few other constraints: recordings were limited to the area of the planned resection, the neuronal populations had to be situated so that activity could be held throughout the auditory test (tending to bias the sample toward deeper cortical layers), and activity with frequent high frequency bursting, suggestive of epileptic activity (Calvin et al. 1973), was excluded.

We are aware that some observations reported in this study may be considered incomplete by rigorous standards as they lack the systematic variation of stimulus parameters needed for clearly defining the class of signals represented by maximal excitation of a given unit or in a certain area. When

searching for optimal stimulus parameters in the auditory and periauditory cortex one should be aware, however, that apart from tone and modulation frequency straight forward parameters such as found in the visual cortex (e.g. contrast, contours, colour, orientation, movement etc.) have barely been defined in the auditory domain even in animal studies (Symmes 1981; Creutzfeldt et al. 1980; Goldstein and Abeles 1975; Newman 1977, 1978). Yet, by applying a variety of defined stimuli and by varying the behavioral responses to these stimuli, we were able to test a number of questions related to neuronal responses to spoken language and we were thus able to get at least some ideas in which way and to what a degree the areas and neurons from which we recorded were involved in speech processing. All response types which we have described and documented in this paper have been observed on several patients and often in simultaneous or successive recordings of different neurons in the same patients at the same recording site. They are therefore probably very common, widespread changes. The fact that in our patients the auditory input was essentially restricted to the ipsilateral ear might have had a negative effect on sampling of responsive units since auditory cortex neurons may have a predominant input from either

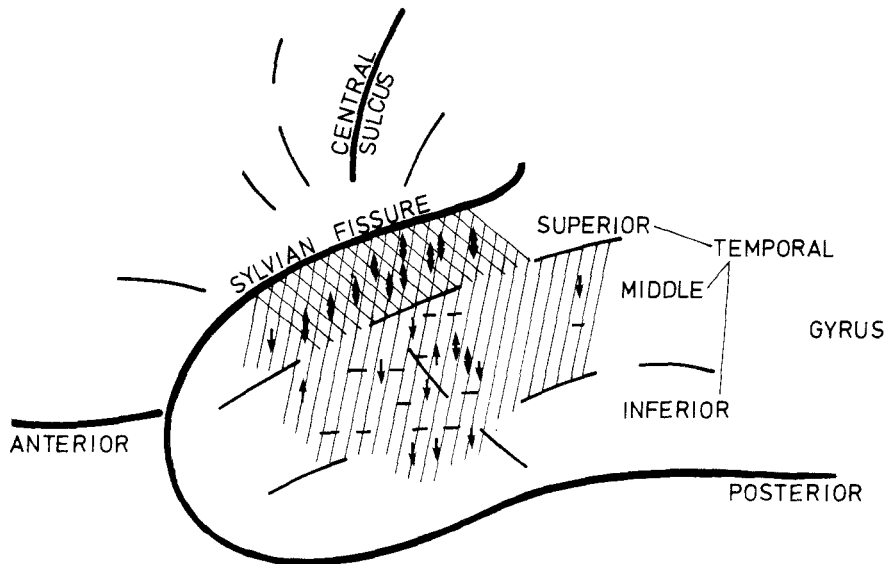


Fig. 16. Schematic representation of response distribution over the temporal lobe to auditory language signals. Data from right and left hemisphere are drawn into one (left) temporal lobe. Symbols indicate the predominant response at one recording site, independent of whether only one single unit, multiunit activity, or several units were recorded simultaneously or successively. Recording sites in which auditory responses to language were not tested or where units could not sufficiently discriminated are not included. Symbols: \uparrow strong and specific modulation of activity by spoken language (for details see text). \downarrow only suppression of spontaneous discharge rate, $-$ no response. The region with predominantly strong responses is cross-hatched, that with weak responses is marked with diagonal lines

the contra- or the ipsilateral ear (Imig and Adrian 1971). We do not expect a qualitative difference, however, neither in response patterns nor in topography due to the predominantly monaural stimulation.

In the following we will discuss different aspects of this study, namely the question of the anatomical distribution of speech related neuronal activity in the lateral temporal lobe (A), the question of the specificity of neuronal responses in relation to auditory, phonemic and linguistic variables (B), and the question of task dependant responses (C).

The topographical distribution of language related activities in the lateral temporal lobe

The responsiveness to auditory speech signals of the lateral temporal cortex does not show obvious differences between the *right and the left hemisphere*. This in itself could indicate that the neuronal signals which we recorded probably do not represent, as such, the semantic elements of the spoken language since otherwise one would have expected clear lateral differences. The detailed analysis of individual responses in other tests (see Ojemann et al. 1988; Creutzfeldt, Ojemann, Lettich, in preparation) also failed to give any indication that semantic meaning as such is represented in the activity of neurons in the middle and anterior

segment of the lateral temporal lobe including the superior temporal gyrus.

The superior and the middle temporal gyrus and the anterior temporal cortex are differently affected by spoken language (see Fig. 16). All neurons in the *superior temporal gyrus* lateral to the sensory and motor cortex showed activity modulations closely related in one way or the other to different aspects of spoken language including the subjects own voice (for the latter see the subsequent report). Excitatory responses were strong and obvious with event related bursts of discharges at a mean rate of 20–50/s and peak rates reaching 70–100/s. In contrast, *middle temporal gyrus* neurons were either not at all or only slightly affected by spoken language with mean rates below 5–10/s. Thus, the superior temporal gyrus is clearly more involved in the representation of auditory linguistic signals than the middle temporal gyrus. This is consistent with the observation that phoneme recognition may be impaired by electrical stimulation of the superior temporal gyrus lateral to the sensory and motor fields (Ojemann 1983). It is also in line with the definition of the superior temporal gyrus as an auditory association cortex (see Creutzfeldt 1983). The lack of obvious pitch or noise sensitivity of its neurons, and the loose – if any – coupling of their activity to harmonic elements of speech signals such as formants (vowels) indicates

that this area is not part of the tonotopically organized auditory cortex. Our recording area in the superior temporal gyrus corresponds approximately to the internal and external parakonio-cortex as defined cytoarchitectonically by Galaburda and Sanides (1980), which may correspond to the non-tonotopically organized lateral or ventral auditory association cortex of other mammals such as the rhesus monkey (Merzenich and Brugge 1973), cat (Merzenich et al. 1979), or guinea pig (Redies et al. 1989).

As to the pathways which carry activation and inhibition into this auditory association area, we cannot say much from our data. Whenever an activation can be related to a well defined phonetic event, the response latency is between 60 and 100 ms. At the sound pressure used, this long latency suggests several relays. It must be considered, however, that the sound pressure of voiceless consonants is much lower than that of the vowels. Responses to and perception of voiceless consonants can therefore be easily suppressed by backwards masking (Spenner and Urbas 1986). The long lasting activations during and following words and often across phoneme boundaries indicate that activities in this region are not just due to feature related selection of primary auditory cortex activities but represent additional labels related to acoustic signals (see section B). They contribute rather than extract specific information to the multiple parallel representation of auditory linguistic signals in the brain in a way similar to that seen in other sensory systems (Creutzfeldt 1983, 1985). Such additional informations could reach these areas via parallel thalamo-cortical projections (through the pulvinar) or could result from a combination of pulvinar-thalamic and transcortical-association inputs. Here, it should be noted that neuronal activities in the thalamus and other subcortical structures are also affected by spoken language in a specific manner (Bechtereva et al. 1979), and that electrical stimulation or lesions in the pulvinar may lead to speech disturbance (Ojemann 1976).

Neurons in the *middle temporal gyrus* were as a whole much less responsive, but not all were completely un-responsive to auditory language signals (see Fig. 16). No specific features of our auditory stimuli could reliably be related to activity changes of units in this area, however (see below). Visual stimuli, written language, pictures or faces also failed to modulate the activity of these neurons in a prominent and reliable manner, although modification of discharge rate was found in some of them during various tasks involving memory (see Ojemann et al. 1988). Possibly other variables

which were not included in our exploration, have to be linked to the auditory stimuli in order to affect the middle temporal gyrus neurons more consistently. Yet, as auditory responsiveness, including responses to the subjects own voice were found in nearly 50% of the recording sites on the lateral surface of the middle temporal gyrus below the sensory and motor cortex but no visual responses, this region may still be considered as part of the auditory rather than the visual association cortex, if one wants to connect it to one of these two senses at all.

The lack of auditory or visual responsiveness of recordings in the anterior lateral temporal cortex (*temporal pole*) leaves the question of the functional significance of this part of the human temporal lobe again unanswered. The *inferior temporal gyrus* does not appear to be much involved in auditory speech analysis either. In three out of the five recordings here, in which we tested speech, we found only a slight suppression of activity during words and no activity changes in the other two. As will be discussed in the last report, neurons in this region did also not respond specifically to complex visual stimuli or reading (see also Ojemann et al. 1988). Thus, our single unit recordings in this segment of the human inferior temporal gyrus indicates so far that it is not functionally homologous to the monkeys infero-temporal cortex which appears to be involved in recognition of visual objects (Mishkin et al. 1983; Ungerleider 1985). The homologous region in the human brain is probably located further posteriorly.

The functional topography of the lateral temporal lobe with respect to the analysis of speech as derived from our single unit study is consistent with some, but not with other conclusions based on observations with other methodological approaches. The changes of RCBF and metabolic rate are distributed more widely over both temporal lobes during listening to speech and during conversation, although – in some studies – heaviest in the region over the superior temporal gyrus (Ingvar 1983; Lassen and Larsen 1980; Raichle 1987; Reivich 1983). Our single unit studies indicate that during listening to speech only the superior temporal gyrus is specifically activated, but during overt speech of the subject we have noted a slightly stronger involvement of the middle and inferior temporal gyrus (see following report). During silent reading and naming we have seen only few if any systematic changes of activity over the whole lateral temporal lobe (see last report of this series, and Ojemann et al. 1988) and also RCBF-changes during silent reading appear to be restricted to the

riphery (Sachs et al. 1982; Delgutte and Kiang 1984) or in fibres of the auditory radiation (Stein-schneider et al. 1982). Most neurons remained in an activated state for up to several hundred ms after the specific sound combination or speech segment had vanished which had turned them on, like an undamped string on a keyboard instrument which goes on ringing after it had been touched. This reverberation may even continue across phonemic segmentations of subsequent syllables and after the end of a word.

As mentioned earlier, our data do not give any hint that responses of superior temporal gyrus neurons are related to semantic categories of spoken language. This can be concluded from responses to foreign words the meaning of which was unknown to the patient, and which affected neuronal activity in the same general manner as familiar words. On the other hand, the meaning of language to the patient, i.e. whether it was directed to him or whether he had to attend to it or not could have some effect on the amplitude of responses (see section C). The closest we have seen as coming to semantics were the recordings in which neurons responded only to the second part of compound words, since these second parts altered the semantic meaning of the first part of the words. Yet, one cannot assign a specific semantic signal to such responses as they appeared with each compound word contained in our word list. Such neurons indicate, however, that speech signals can be meaningfully segmented by neuronal activities already at an early, presemantic level and before conscious access to the vocabulary.

A related but distinct type of response is the activation exclusively or predominantly during the longer words, no matter whether they are compound or just multisyllable. Such activations could "mean" that the phonetic sequence is all belonging to one single word. Although such a category "one word" appears to imply some semantic understanding, the semantics are not contained in these activations as they develop during long and compound words without clear distinctions between meanings. It is more probable that the prosodic and temporal characteristics of multisyllable in contrast to a sequence of short words is the determining factor for such long word-responses.

Another aspect of temporal sequencing of speech signals is suggested by some neurons or multiunit activities which responded indiscriminately to every word of our word lists, or to almost each syllable of a sentence addressed to them. They could thus beat in the rhythm of speech, at least for a limited period (see e.g. Fig. 2E, unit 1). This

beat related activity might also contribute to the segmentation of ongoing speech. Again other neurons were just generally activated by speaking. Such neurons were not or only slightly affected, however, by the 500 or 1000 Hz tone preceding the words, or by non-linguistic noises around the subject. The message of the activation of such neurons could then just be that speech is going on.

Also *inhibition* can be a feature of neuronal responses of superior temporal gyrus neurons during words or speech. In a number of neurons suppression of ongoing activity was the first response to a word presented, which could then be followed by activation or rebound. The activity of other neurons was strongly or completely suppressed during ongoing speech (sentences, instructions), and short off-activation could appear after termination of the sentence. That this inhibition may be language specific is indicated by the observation documented in Fig. 14A, where only the acoustically undistorted sentences suppressed the activity and elicited an off-discharge.

Thus, our observations so far suggest that neuronal activity in the *superior temporal gyrus* during speech may carry information on broadly tuned, mainly voiceless phoneme categories or complexes, on segmentation of phoneme sequences within words and spoken language, on coherence of phoneme sequences in long polysyllabic and in compound words and on the presence of language altogether. This information appears to be distributed across different neurons or sets of neurons each of which is not tuned to just one of these various aspects, and activation patterns elicited by the various triggering elements overlap in time. It is quite possible that in different linguistic contexts neurons may show different patterns, and it is probably the ensemble of activities going on in this cortical network which determines its contribution to the representation of speech signals rather than the specific activation pattern of one single neuron.

An appropriate metaphor might be that of an orchestra, in which each instrument may have a certain function according to its place and timbre, but within a piece of music it may be involved in various aspects of representing the music, such as carrying the tune, responding to it or only accompanying it and thus affecting its sound and quality, beating the rhythm or only accentuating it. None of the instruments is involved only in one aspect of music but changes its function continually. The sound of the instrument is invariantly determined by its built like that of neurons by their connectivity, but its function is not only one and invariant within the piece of music, which express

posterior perisylvian cortex, in addition to patches of increase in the frontal and occipital cortex, while during silent speech, changes of RCBF are restricted to the dorsal frontal cortex (Ingvar 1983). When comparing results of the two approaches one also has to keep in mind the differences in the experimental situations in which the measurements were taken. RCBF- and metabolic measurements are integrated over long time and the tasks may involve more attention and mental concentration over longer periods, whereas during the single unit studies temporally restricted activity changes related to short linguistic signals are tested. Both approaches agree on the bilateral participation of the temporal lobes in speech perception, except for the minor and apparently variable lateral differences in the RCBF- and metabolic studies (see cited references).

Evoked potential studies also demonstrate a bilateral involvement of both temporal lobes in speech analysis but without further dorso-lateral differentiation (Hillyard and Picton 1987). This may be due to the fact that evoked potentials may give relatively little information on the exact location of neuronal processing, as an electrocorticographic and more so an EEG-electrode picks up not only local activity but also far field potentials as amply documented in evoked potential literature (Desmedt 1977). Therefore, the somewhat more widespread distribution of speech related and other auditory evoked potentials does not contradict the conclusion of a more restricted involvement of essentially only the superior temporal gyrus in acoustic representation of speech signals. Interhemispheric differences of evoked potentials elicited by words have been observed by some authors, but not by others, indicating that if differences are real they must be minor (see p. 561 in Hillyard and Picton 1987). Such minor differences could be caused by the different and interindividually variable anatomy of the two temporal lobes with a larger superior temporal gyrus in the speech-dominant hemisphere.

Specificity of responses

Our recordings have shown that responses of *superior temporal gyrus neurons* may be specifically related to certain temporal and phonemic aspects of spoken language, such as the length of a word (longer words may elicit a stronger excitation than short words), some neurons may be excited only by the second part of a compound word, or responses may be restricted to only one phoneme combination or syllable of a word. In other neu-

rons, excitation or inhibition by spoken words or sentences may be unspecific in that activation patterns cannot be related to any consistent variables of the spoken language. In many neurons a rebound activation is seen following short words and towards the end of longer words. If a word or phoneme elicits a specific response, the same response pattern appears again when this word or phoneme is repeated at a later time. Although we have found such reproducible response patterns even at different sound pressure levels, our data do not allow at this point to conclude that such responses were always completely sound pressure invariant. It is interesting to note, that also in sub-human primates (squirrel monkeys) the probability of complex responses increased with the length of a natural call in neurons in A II, whereas neurons in A I were only transiently activated by a variety of stimulus components (Glass and Wollberg 1983).

Some of the responses described in this report could help in *categorization* of certain classes of phonemes. This applies mainly to consonants. Although certain consonant combinations such as *chr* or *stm* in "Christmastree" or *scr* in "corkscrew", the voiceless closure guttural *k* in "caterpillar" or "rattlesnake" or fricatives such as "s" or "f" may lead to strong excitation in some neurons, our data do not indicate that such neurons were exclusively "tuned" to only one of these consonants or consonant combinations. With respect to formants we have not found one neuron that responded exclusively to only one vowel although our second word list was specifically designed for vowel detection (see Methods). The long activations elicited by voiceless consonants usually reached across the vowel phonation and it may therefore be that the consonant/vowel combination or coarticulation (Pols and Schouten 1982) as such was marked. In any case, neurons appear to be relatively broadly tuned to phoneme classes and their activation pattern can therefore only code for broad phonemic and linguistic categories within words.

Auditory neurons in the superior temporal gyrus are not turned on or off by peaks of the sound pressure envelop such as neurons in the primary auditory cortex of guinea pigs, cats and monkeys, which are activated only by sound peaks reaching into their spectral sensitivity range (Creutzfeldt et al. 1980; Sovijärvi 1975; Winter and Funkenstein 1973). There was also no indication that neuronal discharges were phase locked to the basic frequency or comparable higher frequency temporal modulations such as seen in the auditory pe-

over time various aspects of a tune, connecting it with other tunes and bringing it into different harmonic and syntactic contexts.

In the *middle temporal gyrus*, suppression of activity was the only response to words or speech in the majority of those recordings which showed any activity changes (see Fig. 16). In half of the recordings, the neuronal activity was not at all affected by language. In all responsive recordings, the excitatory or suppressive effect of language was weak and activations never went above 10/s, usually less than 5/s discharge rate. In several units the activity was suppressed during the words and showed a slightly increased discharge rate only following the word. In others the activity was slightly depressed following the word. Yet, these responses did not appear to be just unspecific auditory responses as noises or the 1000 or 500 Hz tones affected such neurons much less or not at all. In some units we observed suppression of ongoing activity only during understandable normal speech, but not during incomprehensibly distorted or backward speech. These observations indicate that although some recording sites in the middle temporal gyrus may be affected by spoken language there are no clear indications that they contribute an important signal about phonetic, temporal or semantic aspects of speech. Instead, the changes of middle temporal gyrus activity may be related to understanding of language in a very general way and the prominent suppression may even suggest that activity is being tuned down during listening to speech, as it might interfere with comprehension of spoken language otherwise.

Task dependent changes of responsiveness and activity

Responses to words or short sentences could be stronger or weaker when the task or the situation demanded specific attention to them. Less than half of our recordings were affected by attention, however. In the *superior temporal gyrus* responsive variation related to attention were essentially quantitative, i.e. more or less of the same type of activation, but not a different type of response. To what extent the increased responsiveness reflects gating of afferent signals in the thalamus controlled by reticulo-thalamic circuits (Singer 1977) or at the cortical level through prefrontal mechanisms (Roland 1982) via association fibre connections, must be left open at this stage. Increased as well as decreased responsiveness may indicate that the total excitation pattern in the neuronal network is not only generally elevated or

lowered during directed attention, but that the distribution of relative excitations may also change. Response variations depending on the attention paid to the task have also been observed in monkeys even in the primary auditory cortex but there again, only a minority of neurons showed such effects (Beaton and Miller 1975; Hoehermann et al. 1976; Benson and Heinz 1978). Also in humans, language and other auditory evoked potentials may change depending on the attention paid to the stimulus (Hillyard et al. 1973; Näätänen et al. 1981; for further references see Hillyard and Picton 1987).

The non-specific activation or suppression of ongoing activity by speech, which we observed in the *superior* as well as in the *middle temporal gyrus* may be interpreted as a more general arousal, however, indicating that more extended cortical regions are temporarily alerted by language. These may be reduced again if their contribution is not needed or may be even disturbing for a detailed representation of the signal.

Concluding remarks

We have emphasized in this report the specific involvement of the superior temporal gyrus in the analysis of speech. Neuronal activities in this region may add to the multiple representation of acoustic language, in parallel with other predominantly acoustic and predominantly semantic representations in nearby areas. Lesion of the superior temporal gyrus on both sides produces auditory agnosia, which disables the patient's ability to recognize sounds and to distinguish between language and other sounds (Kleist 1934; Luria 1966). Resections that included those sites in many cases, including all resections in nondominant hemisphere, were followed by no-lasting language disturbance at all. Resections on the speech dominant side were occasionally followed by deficits in comprehending or repeating long words. This defect may last for some time (up to several weeks) and thus outlast the anomia seen during the first days following the operation due to post-operative local edema. Our observations are consistent with the suggestion that this intermediate part of the superior temporal gyrus belongs to perisylvian language cortex and thus plays a role in the mental composition of acoustic noises to a linguistic code but without major lateral specialization. The relatively small deficits after unilateral removal indicate that this region is at least unilaterally disposable and thus not essential for language functions.

Acknowledgements. We express our sincere thanks to the patients who agreed to cooperate in this study. They have contributed to our better understanding of brain functions. We also gratefully acknowledge the help of the Computer-Group of the Max-Planck-Institute, especially that of Mrs. Jutta Tolksdorf in data analysis, Mrs. Tamara Wellem for preparing the figures, and Mrs. Eva-Maria Hölscher for typing the manuscripts. The critical reading of an older version of the manuscript by Dr. Hermann Redies is much appreciated. The work was supported largely by NIH grant NS 21724, 17111 and 20482, as well as by a travel grant to O.C. from the Max-Planck-Society.

References

- Beaton R, Miller JM (1975) Single cell activity in the auditory cortex of the unanesthetized, behaving monkey: correlation with stimulus controlled behavior. *Brain Res* 100:543–562
- Bechtereva N, Bundzen Z, Gogotitsin Y, Malyshev V, Perepelkin P (1979) Neurophysiological codes of words in subcortical structures of the human brain. *Brain Lang* 7:145–163
- Benson DA, Heinz RD (1978) Single unit activity in the auditory cortex of monkeys selectively attending left vs. right ear stimuli. *Brain Res* 159:307–320
- Calvin WH, Ojemann GA, Ward Jr. AA (1973) Human cortical neurons in epileptogenic foci: comparison of interictal firing patterns to those of “epileptic” neurons in animals. *Electroencephalogr Clin Neurophysiol* 34:337–351
- Creutzfeldt OD (1983) *Cortex cerebri*. Springer, Berlin Heidelberg New York
- Creutzfeldt OD (1985) Comparative aspects of representation in the visual system. In: Chagas C, Gattass R, Gross Ch (eds) *Pattern recognition mechanisms*. Proc Study Week Pontif Acad Sci. *Exp Brain Res Suppl* 11:53–81
- Creutzfeldt OD, Hellweg FC, Schreiner C (1980) Thalamo-cortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 39:87–104
- Creutzfeldt OD, Ojemann G, Lettich E (1987) Single neuron activity in the right and left human temporal lobe during listening and speaking. In: Engel Jr J, Ojemann GA, Lüders HO, Williamson P (eds) *Fundamental mechanisms of human brain function*. Raven Press, New York, pp 69–81 and In: Ottoson D (ed) *Duality and unity of the brain*. Proc of an Int Symp, Wenner-Gren Center, Stockholm, pp 295–310
- Creutzfeldt O, Ojemann G, Lettich E (1989a) Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Exp Brain Res* 77:476–489
- Creutzfeldt O, Ojemann G (1989b) Neuronal activity in the human lateral temporal lobe. III. Activity changes during music. *Exp Brain Res* 77:490–498
- Delgutte B, Kiang NYS (1984) Speech coding in the auditory nerve, Parts I–V. *J Acoust Soc Am* 75:866–918
- Desmedt JE (ed) (1977) *Language and hemispheric specialization in man: cerebral event related potentials*. In: *Progr Clin Neurophysiol*, Vol 3. Karger, Basel
- Galaburda A, Sanides F (1980) Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol* 190:597–610
- Glass I, Wollberg Z (1983) Responses of cells in the auditory cortex of awake squirrel monkeys to normal and reversed species specific vocalizations. *Hearing Res* 9:27–33
- Goldstein Jr MH, Abeles M (1975) Single unit activity in the auditory cortex. In: Keidel WE, Neff WD (eds) *Handbook of sensory physiology*, Vol V/2. Springer, Berlin Heidelberg New York, pp 199–218
- Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. *Science* 182:177–180
- Hillyard SA, Picton TW (1987) *Electrophysiology of cognition*. In: Mountcastle VB, Plum F, Geiger SR (eds) *Handbook of physiology: higher functions of the brain*, Sect 1, Vol V/2. Am Physiol Soc, Bethesda Md, pp 519–584 (p 561)
- Hocherman SD, Benson DA, Goldstein Jr MM, Heffner HE, Heinz RD (1976) Evoked unit activity in auditory cortex of monkey performing a selective attention task. *Brain Res* 117:51–68
- Imig TJ, Adrian HO (1977) Binaural columns in the primary field (A1) of cat auditory cortex. *Brain Res* 138:241–257
- Ingvar DH (1983) Serial aspects of language and speech related to prefrontal cortical activity: a selective review. *Hum Neurobiol* 2:177–189
- Jackson JM (1898) Case of epilepsy with tasting movements and “dreamy state” – very small patch of softening in the left uncinata gyrus. *Brain* 21:580–590
- Kleist K (1934) *Gehirnpathologie*. Joh Ambros Barth, Leipzig
- Lassen NA, Larsen B (1980) Cortical activity in the right and left hemispheres during language-related brain functions. *Phonetica* 37:27–37
- Luria AR (1966) *Higher cortical functions in man*, English edition. Basic Books Inc, New York
- Mazziotta JC, Phelps ME, Halgren E (1983) Local cerebral glucose metabolic response to audiovisual stimulation and deprivation: studies in human subjects with positron CT. *Hum Neurobiol* 2:11–23
- Merzenich MM, Andersen RA, Middlebrooks JH (1979) Functional organization of the auditory cortex. In: Creutzfeldt OD, Scheich H, Schreiner C (eds) *Hearing mechanisms and speech*. *Exp Brain Res Suppl* 2:60–75
- Merzenich MM, Brugge JF (1973) Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Res* 50:275–296
- Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. *Trends Neurosci* 6:414–417
- Näätänen R, Graillard AWK, Varcy CA (1981) Attention effects on auditory EPs as a function of inter-stimulus interval. *Biol Psychol* 13:173–187
- Newman JD (1977) Biological filtering and neural mechanisms: group report. In: Bullock TH (ed) *Recognition of complex acoustic signals*. Dahlem Workshop, Berlin, pp 279–306
- Newman JD (1978) Central nervous system processing of sounds in primates. In: Steklis H, Raleigh MJ (eds) *Neurobiology of social communication in primates: an evolutionary perspective*. MJ Ralcigh, New York, pp 69–110
- Ojemann GA (1976) Subcortical language mechanisms. In: Whitaker H, Whitaker HA (eds) *Studies in neurolinguistics*, Vol 1. Academic Press, New York, pp 103–138
- Ojemann GA (1983) Brain organisation for language from the perspective of electrical stimulation mapping. *Behav Brain Sci* 6:189–206
- Ojemann GA (1985) Surgical treatment of epilepsy. In: Wilkens R, Rengachary S (eds) *Neurosurgery*. McGraw Hill, New York, pp 2517–2526
- Ojemann GA, Creutzfeldt OD (1987) Language in humans and animals: contribution of brain stimulation and recording. In: Mountcastle VB, Plum F, Geiger SR (eds) *Handbook of physiology*, Sec 1. The nervous system, Vol V. Higher function of the brain, Part 2. Am Physiol Soc, Bethesda Md, pp 675–700
- Ojemann GA, Creutzfeldt OD, Lettich E, Haglund MM (1988) Neuronal activity in human lateral temporal cortex related to short-term verbal memory, naming and reading. *Brain* 111:1383–1403

- Penfield W, Jasper H (1954) *Epilepsy and the functional anatomy of the human brain*. Little and Brown, Boston
- Penfield W, Perrot P (1963) The brains record of auditory and visual experience. *Brain* 86:595–696
- Pols LCW, Schouten MEH (1982) Perceptual relevance of coarticulation. In: Carlson R, Granström BE (eds) *The representation of speech in the peripheral auditory system*. Elsevier Biomedical Press, Amsterdam New York Oxford, pp 203–208
- Raichle ME (1987) Circulatory and metabolic correlates of brain functions in normal humans. In: Mountcastle VB, Plum F, Geiger S (eds) *Handbook of physiology, Sec. 1. The nervous system, Vol. V/2: Higher functions of the brain*. Am Physiol Soc, Bethesda Md, pp 643–674
- Reivich M, Gur R, Alavi A (1983) Positron emission tomographic studies of sensory stimuli, cognitive processes and anxiety. *Hum Neurobiol* 2:25–33
- Redies H, Sieben U, Creutzfeldt OD (1989) Functional subdivision of the auditory cortex of the guinea pig. *J Comp Neurol* 282:473–488
- Roland PE (1982) Cortical regulation of selective attention in man: a regional cerebral blood flow study. *J Neurophysiol* 48:1059–1078
- Ross ED (1981) The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol* 38:561–569
- Sachs MB, Young ED, Mitter MI (1982) Encoding of speech features in the auditory nerve. In: Carlson R, Granström BE (eds) *The representation of speech in the peripheral auditory system*. Elsevier Biomedical Press, Amsterdam New York Oxford, pp 115–130
- Schroeder MR (1966) Vocoders: analysis and synthesis of speech. *Proc IEEE* 54:720–734
- Singer W (1977) Control of thalamic transmission by corticofugal and ascending reticular pathways in the visual system. *Physiol Rev* 57:386–420
- Sovijärvi ARA (1975) Detection of natural complex sounds in the primary auditory cortex of the cat. *Acta Physiol Scand* 93:318–335
- Spenner H, Urbas JV (1986) Masking and the perception of stop consonants: psychoacoustical and electrophysiological experiments. *Exp Brain Res* 62:607–617
- Steinschneider M, Arezzo J, Vaughan Jr HD (1982) Speech evoked activity in the auditory radiations and cortex of the awake monkey. *Brain Res* 252:353–365
- Symmes D (1981) On the use of natural stimuli in neurophysiological studies of audition. *Hearing Res* 4:203–214
- Ungerleider LG (1985) The corticocortical pathways for object recognition and spatial perception. In: Chagas C, Gattass R, Gross Ch (eds) *Pattern recognition mechanisms*. Proc Study Week Pontif Acad Sci. *Exp Brain Res Suppl* 11:21–37
- Wada J, Rasmussen T (1960) Intracortical injections of sodium amytal for the lateralisation of cerebral speech dominance. *J Neurosurg* 17:266–282
- Winter PA, Funkenstein HH (1973) The effects of species-specific vocalizations on the discharge of auditory cortical cells in the awake squirrel monkey (*Saimiri sciureus*). *Exp Brain Res* 18:489–504
- Zwicker E (1960) Über die Rolle der Frequenzgruppe beim Hören. *Ergebn Biol* 23:187–203
- Zwicker E (1961) Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *J Acoust Soc Am* 83:248

Received December 6, 1988 / Accepted June 9, 1989