

Neuronal activity in the human lateral temporal lobe

II. Responses to the subjects own voice

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. We have recorded neuronal responses in the lateral temporal lobe of man to overt speech during open brain surgery for epilepsy. Tests included overt naming of objects and reading words or short sentences shown on a projector screen, repetition of tape recorded words or sentences presented over a loudspeaker, and free conversation. Neuronal activity in the dominant and non-dominant temporal lobe were about equally affected by overt speech. As during listening to language (see Creutzfeldt et al. 1989), responses differed between recordings from sites in the superior and the middle or inferior temporal gyrus. In the *superior temporal gyrus* all neurons responded clearly and each in a characteristic manner. Activation could be related to phonemic aspects, to segmentation or to the length of spoken words or sentences. However, neurons were mostly differently affected by listening to words and language as compared to overt speaking. In neuronal populations recorded simultaneously with one or two microelectrodes, some neurons responded predominantly to one or the other type of speech. Excitatory responses during overt speaking were always auditory. In the *middle temporal gyrus* more neurons (about 2/3) responded to overt speaking than to listening alone. Activations elicited during overt speech were seen in about 1/3 of our sample, but they were more sluggish than those recorded in the superior gyrus. A prominent feature was suppression of on-going activity, which we found in about 1/3 of middle and in some superior temporal gyrus neurons. This suppression could precede vocalization by up to a few hundred ms, and could outlast it by up to 1 s. Evoked ECoG-potentials to words heard or spoken were different, and those to overt speech were more widespread.

Key words: Single unit recordings – Speaking – Temporal lobe – Man

Introduction

If we speak we also hear our own voice. Yet, we know it is ourselves who speak and not someone else. Is it because other brain areas, such as speech command regions are active at the same time, or that the neuronal activation patterns elicited by our own speech are different from those elicited by other people speaking? Do different neuronal populations and different brain regions participate in the speech analysis if we just listen to others and if we hear ourselves speaking? Or, is it just the fact that bone conduction has high pass filter properties? An answer to these questions could be given if the activity patterns across the brain would be known during listening and speaking.

During our exploration of the lateral temporal lobe of awake patients during epilepsy operations we have recorded neuronal activity during listening to speech and when the subjects spoke themselves, and can thus give at least a partial answer to these questions for the regions explored, i.e. the anterior and intermediate sections of the superior, the middle and the inferior temporal gyrus of both hemispheres. We have shown in the preceding paper that temporal and phonemic aspects of speech affect neuronal activity in the superior temporal gyrus in a specific way during listening to speech signals, whereas neuronal activity in the middle and inferior temporal gyrus was only slightly and unspecifically affected or not at all (Creutzfeldt et al. 1989). In the same patients we have also formally tested neuronal responses to overt speech during naming of objects, repetition of words and

sentences, reading and memorizing. The changes in activity during speech vocalizations recorded during our study are the subject of this report and will be compared with the activity patterns seen during speech perception.

Methods

Patients and recording of single unit activity were identical to those of the preceding study and we refer for details to that report (Creutzfeldt et al. 1989; Ojemann et al. 1988). The present study is based on 41 microelectrode recordings from single units or groups of units in 34 patients. 23 recordings were from the left and 18 from the right hemisphere. The recording sites are shown in Fig. 1 of the preceding paper: 12 recordings were from the superior temporal gyrus (8 right, 4 left), lateral to the somato-sensory cortex. 24 recordings were from the middle temporal gyrus (8 right, 16 left) spread over a slightly wider antero-posterior region than the superior temporal gyrus recordings, and 5 recordings were from the inferior temporal gyrus (2 right, 3 left). As explained in the preceding paper, we recorded in some patients simultaneously from 2 microelectrodes, and recordings were often repeated at different depths. Thus, the total number of single unit and multiunit recordings available for analysis was well over 100. Neuronal activity during the patients own speech was recorded during spontaneous speech, free speech on demand ("tell me your name, "...address", "...the name of your parents", "...of ten animals" etc.), and during formal tests. These included the naming of objects, of face expressions, reading single words or short sentences projected on a translucent screen, and repeating single words or sentences played to the patient from a tape. All naming and reading slides were repeated and the patient was asked to name or read them silently.

Single or grouped neuronal activity was recorded on tape together with the corticographically recorded EEG from 3–4 sites including that of the microelectrode recording, and an audio-record of voices of the patient, the doctors and the pre-recorded tests. The tape recorded neuronal activity was computer analyzed by synchronizing the beginning of the histograms with various signals including the patients voice. For naming and reading, the patient was shown drawings of objects, words or sentences projected from slides on a translucent screen. Usually slides were shown for two seconds and the patient was asked to respond immediately when a slide appeared. Each task typically consisted of six consecutive slides of the same kind, but occasionally a serious included more slides. For the word repetition task, the 20 or 30 words of our tape recorded word lists were played to the patient in an open field situation at a loudness of conversational speech. Patients were always asked first just to listen to the words. Then the same words or another section of the list was played to them and they were instructed before the test to repeat each word immediately after they had heard it. The time between the end of the word and the beginning of the patients response was between 220 and 600 ms, but could be longer (up to 1.2 s) after difficult or foreign words. The interval between words was 5 s, and each new word was preceded by a 1000 or 500 Hz tone. Test words were short (monosyllable), longer (2–3-syllable) and compound words. For details of the word lists see the preceding report and descriptions in the Results of this paper. For sentence repetition, the patient had to wait 2–3 s after the end of each sentence until asked to repeat it. Patients had not heard the tape recorded speech lists before the operation.

Results

Recordings during speaking in the set-up of brain surgery are prone to be confounded with artefacts due to slight respiratory movements of the brain or due to microphonic resonances of the recording circuit. We excluded those recordings in which there was any suspicion of such artefacts. With this precaution it can be clearly stated that the activity of single units maybe strongly affected, either activated or inhibited, by the persons own voice. Such voice related activity changes were recorded in all regions that we studied, i.e. in the superior, middle and inferior temporal gyrus of both sides with some characteristic differences of overall responsiveness and of the specificity of responses. In the *superior temporal gyrus* predominant activation occurred in 8 out of 11 recording sites and predominant suppression in two recordings. One site showed no responses. In the *middle temporal gyrus* only 5 recording sites out of 16 showed a faint activation, 6 showed clear suppression and 5 no response. In the *inferior temporal gyrus*, only one out of five recordings showed a clear activation, two a weak depression and two were unresponsive. When several units could be distinguished in one recording or when two sites 2–4 mm apart were recorded simultaneously with two microelectrodes or successively, responses of both activities could be identical or different. No significant differences between the respective sites of the right and left temporal lobe were found. During silent reading or naming slight changes of discharge rate were sometimes seen, but these were not in close correlation to the presentation of the slide (see Ojemann et al. 1988) and therefore probably related to other variables than reading.

Superior temporal gyrus

As during listening to language, responses related to specific phonemic or temporal aspects of the words spoken by the subject were only found in superior temporal gyrus recordings. In Figs. 1 and 2 we show three different types of responses recorded simultaneously with two microelectrodes. The patient either read words or named the emotional expression on a photographed face shown to him. Electrode (1) recorded relatively small multiunit activity which appeared everytime when the patient spoke. As can be seen in the records with higher time resolution (Fig. 2B–E), the activation of this multiunit activity started 80–100 ms after the first vowel formant of a word began, but typi-

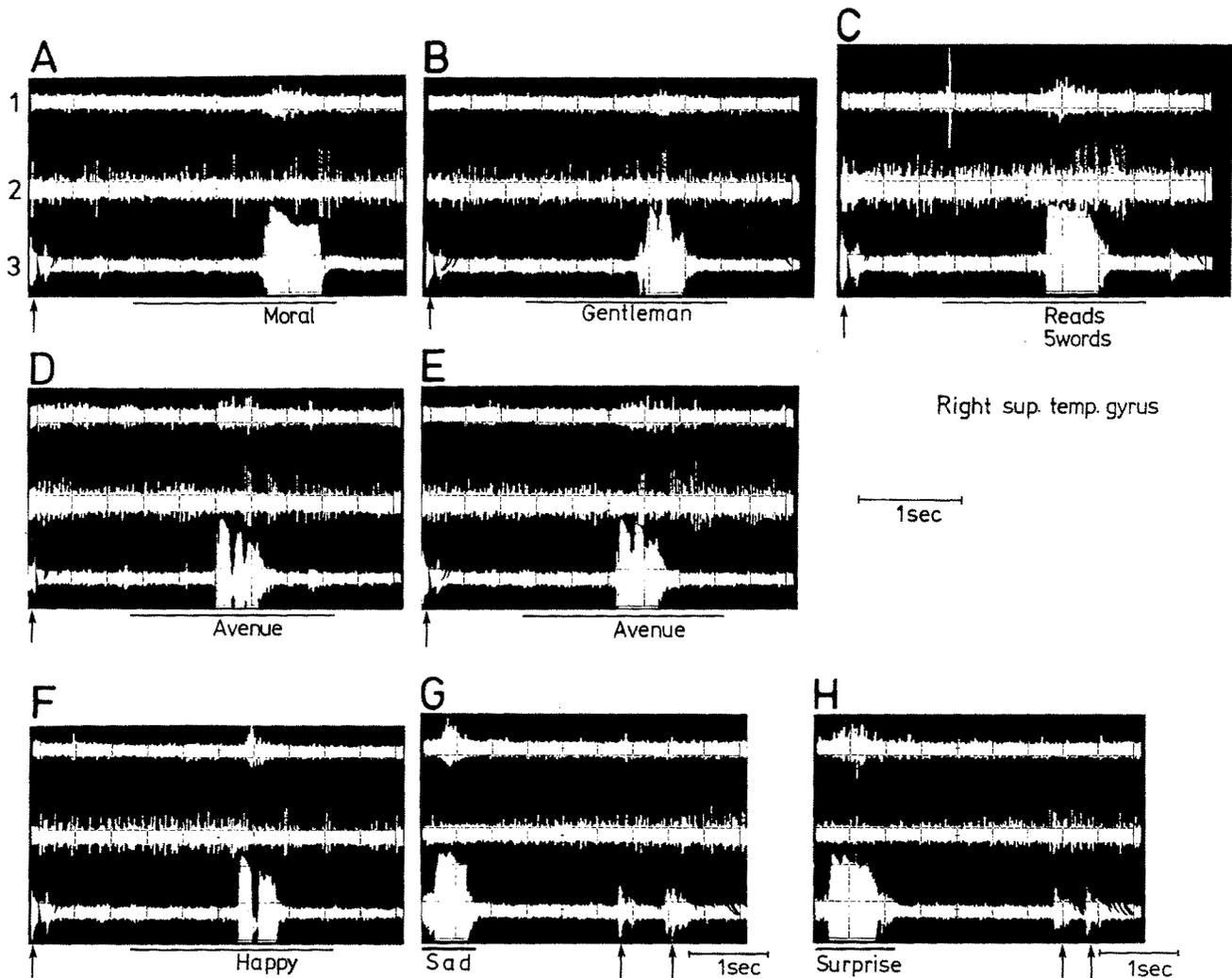


Fig. 1A-H. Responses of single neurones in the right superior temporal gyrus during overt reading and naming (male patient 8533). Simultaneous recording with two micro-electrodes, the tips of which were 3–4 mm above each other (records 1 and 2). 3: Audiorecord. **A-E** The patient was shown a word on the projection screen and was asked to read it aloud. **F-H** The patient was shown the photograph of an actor expressing a certain emotion and had to name that emotion. Vertical arrows below the audio-record indicate the noise produced by the slide falling into the frame. The horizontal lines indicate the time of shutter opening. The words pronounced by the patient are written below the speech records. Note, that records **A-F** are triggered by the slide falling, while records **G, H** are triggered by the beginning of the patients speech. In record 1 only small multiunit activity can be recognized, while in record 2 one unit with large and several with small action potentials can be distinguished

cally did not last throughout the whole word. A prolonged activation was only seen when the patient hummed “mmm” while pondering on the face expression (Fig. 2A and F) or during the closing m of the word “same” (D, F).

In the recordings from electrode (2) small, probably multiunit activity and a larger unit could be distinguished. The large unit tended to respond with a short burst during or/and at the termination of a spoken word (Fig. 1A–F). In the examples shown here it responded either 60–100 ms after the beginning of the second syllable (B, D, E) or towards the end of the word spoken by the patient (A, F). Although it is difficult to assign a specific

phonemic or temporal feature to these short activations, it is apparent that this neuron responded with essentially the same pattern to identical words (“avenue” in Fig. 1D and E) and that it became active only during the second part of all words (see superimposed records in Fig. 1C). It thus appears that some phonetic features, which we cannot define clearly, as well as temporal aspects were involved in the activation of this neuron. It is not clear from this recording, whether the lack of responses to the adjectives designating the facial expression (Fig. 1G, H; Fig. 2) was specifically related to this task as compared to the reading, or whether the neuron became less responsive or even

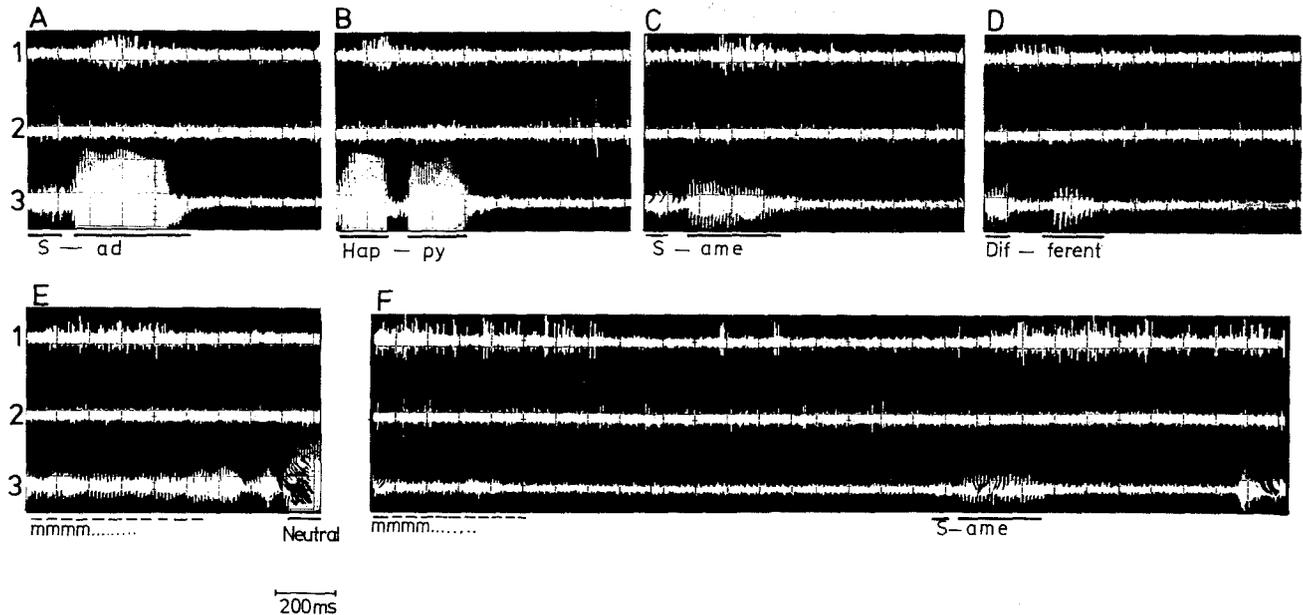


Fig. 2 A–F. Speech related activity in the superior temporal gyrus. (Male, patient 8533). Records at higher time resolution, taken from the same patient as in Fig. 1 during the emotion naming task (for details see legend to Fig. 1). The multiunit activity in record 1 can now be better distinguished and the long delay between speech onset and activation of the unit can be recognized in A–D. The large unit in record 2 is not activated during these words. E, F The multiunit activity in record 1 is increased also during the humming of the patient while he ponders about the expression on the face he is shown

disappeared during this further recording period. As we shall see later, there are some indications that at least some types of responses may be dependent on the context in which a word is spoken.

The third type of response which can be recognized in this recording is seen in the small unit activity picked up by electrode (2). This small unit activity is suppressed during vocalization and returns only 0.5–1 s after the end of the word (see Fig. 1 C, F–H). This suppression sometimes begins up to 250 ms before the overt speech response begins (Fig. 1 D, E).

We were not able to test this neuronal activity during a formal word listening test, but when one of the investigators spoke to the patient, many units of large amplitude were activated in recording (1) and the small unit activity in electrode 2 was suppressed (see Fig. 2 of the preceding paper). Thus, the patterns of activation and inhibition of the three identifiable unit activities was clearly different during listening to speech and speaking.

It was rare to see neurons responding to the same phonetic feature during listening and speaking. We saw it clearly only in two instances. One of these recordings is shown in Fig. 3. This neuron had responded strongly to the broad phoneme category *cr*, *gr*, *k* and *str* in the word list spoken by a male speaker as documented in Fig. 3 of the preceding paper. Also, when the patient listened

to sentences spoken by a female speaker, the neuron responded vigorously to the phoneme combination “*scr*” in the sentence “The child’s crying again” (Fig. 3, A2), and slightly also to other, less well defined but similar features in this and other sentences (The *dog’s* barking; She *is n’t* easy to understand; ...*has his hair cut*). When the patient, also a woman, repeated the sentences, a strong activation appeared again during “is crying” of sentence 2. It may be noted that the patient repeated the sentences much faster than they were spoken to her. Therefore, the activation period in the averaged activity histograms in C and D is much longer during listening (C) than during repeat (D). In both situations the activation is followed by a suppression of ongoing activity. (The fact that the discharge rate during listening and repeat appears to decrease towards the end of the activation period in the PSTH’s is due to the fact that the 10 sentences were not of the same length but became longer from 1–10). Also during the single word repetition task, the test words or consonant combinations which had activated the neuron specifically during listening, were accompanied by discharges during repeat of these same words or phonemes as well, but these responses were weaker during repeat (see Fig. 3 in the preceding paper: *kscr* in *Corkscrew*; *str* in *Christmastree*; *cr* in *crocodile*).

Apart from such identical and specifically

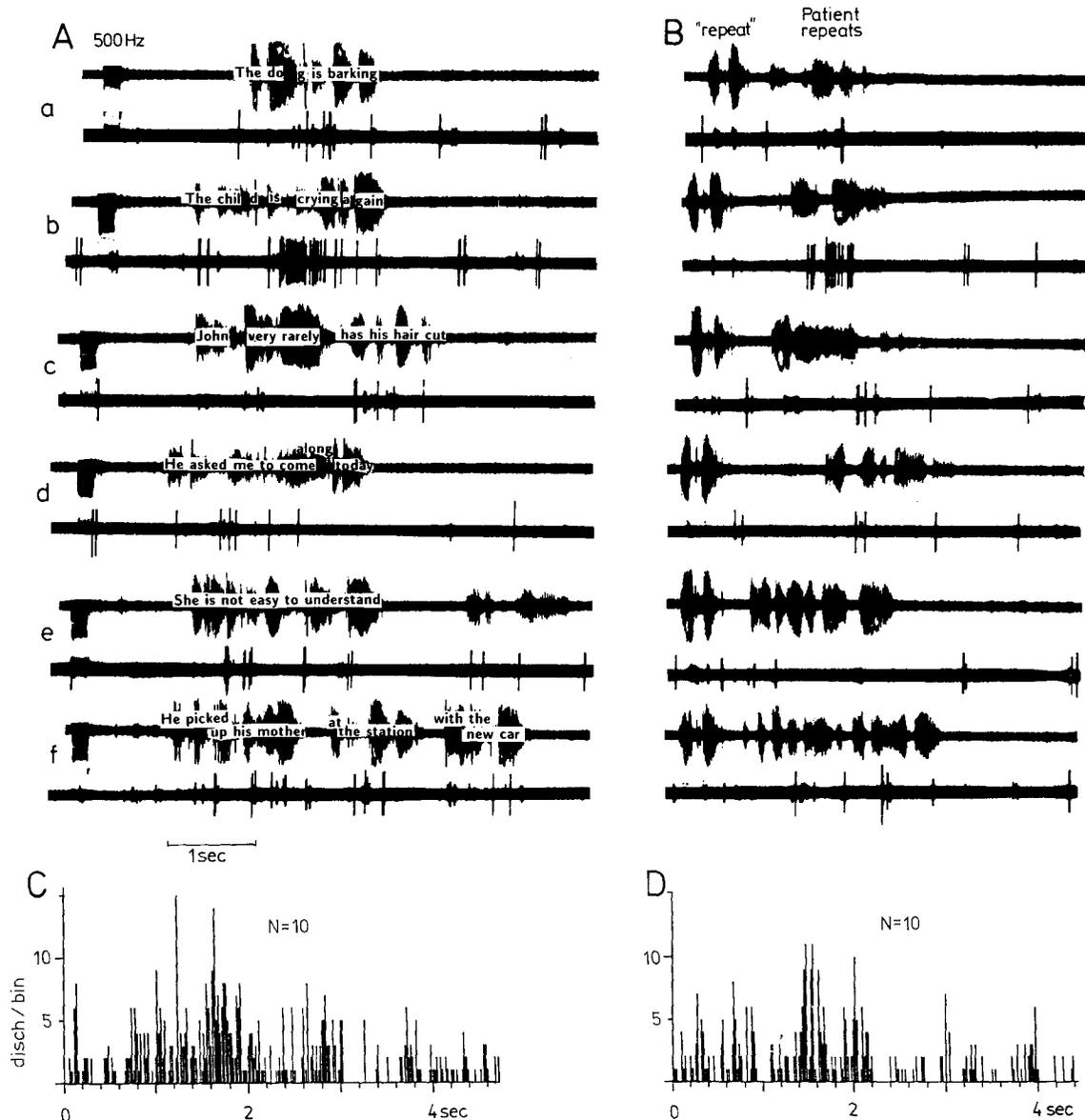


Fig. 3A–D. Activation of a unit in the left superior temporal gyrus during listening and repeating sentences (8512, female). These records are from the same patient and neuron as in Fig. 3 of the preceding report. **A** Audio-record (top) and single unit activity (bottom) when sentences sounded from a tape recorder, spoken by a women speaker. Each sentence is preceded by a short 500 Hz tone. **B** Following the command “repeat” spoken by the female speaker from tape, the patient repeated the sentence just heard (see audio-record). Note, that the patient repeats these sentences faster than they were spoken to her. Note the strong activation by the consonant sequence *scr* when hearing and repeating sentence 2. **C, D** Average discharge rates during listening **C** and repeating the sentences **D**. Bin width in **C** and **D** is 10 ms

phoneme related activation patterns during listening and speaking, the most common picture was that individual neurons were simply affected differently by words or sentences spoken to the patient as compared to the patient repeating them himself. When the activity of several units was recorded simultaneously, different neurons could be differently affected by one or the other situation, resulting in a different activity distribution within that neuron population during listening and repeating.

This is documented in the averaged responses of Fig. 4, which are from two different units in the right superior temporal gyrus recorded simultaneously with two microelectrodes at a tip separation of approximately 3–4 mm (A/B and C–E, respectively). The unit recorded by electrode 1 was slightly and irregularly activated following short words and stronger during and following long words during the listening task (Fig. 4A). This activation by words when the patient was only listen-

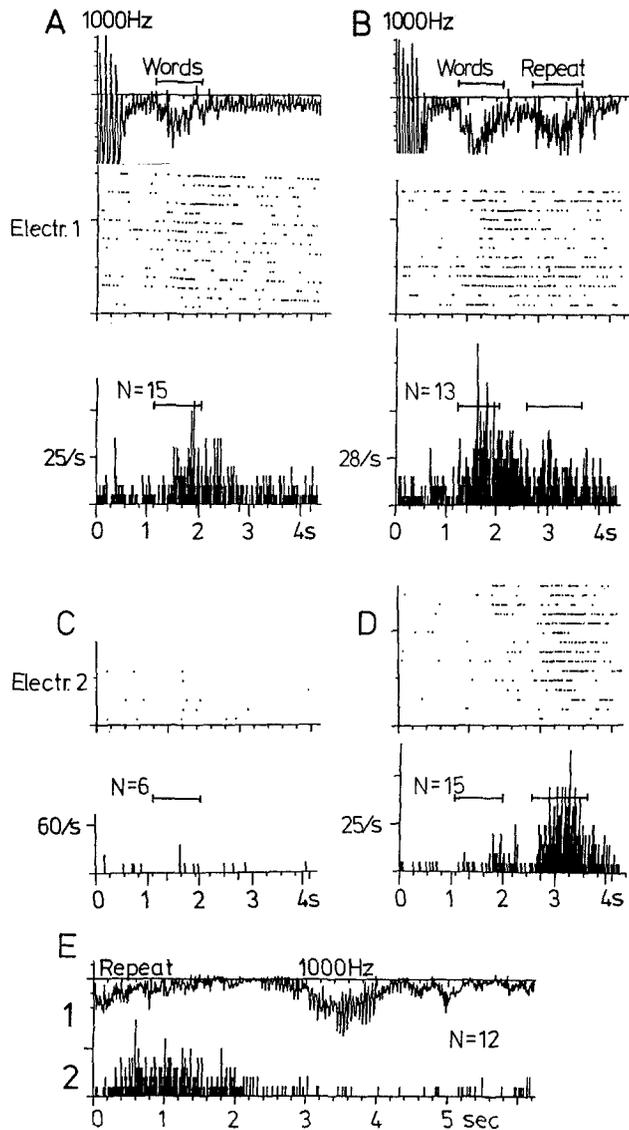


Fig. 4A–E. Differential responses of two simultaneously recorded neuron activities in the right superior temporal gyrus to hearing and repeating the same words. (8603, female). The tips of the recording electrodes were about 3–4 mm above each other. Averaged audiograms on top of dot display of unit discharges and averaged discharge rates during each condition. **A, B** The unit activity in electrode 1 (essentially from one single large unit) is activated by some words (especially longer words) **A** and this activation is enhanced when the patient is asked to repeat words **B**. During phonation of the word the unit is only sometimes activated (see dot display). **C, D** The activity recorded by electrode 2 (one single unit) does not respond to words during listening **C**, but is activated at the end of some words when the patient had to repeat them. The unit responds strongly to the patient's voice, however, when she repeats the words **D**. This activation during repeat has a definite latency after the beginning of the patient's vocalization as shown in **E**, where the averaging is started by the patient's voice. Note that the undulations of the averaged 1000 Hz tone in this and other averaged records are due to sampling and do not correspond to the tone frequency

ing was further increased when the patient had to repeat the words (**B**). Little if any response was seen during the overt repeating itself. The activity recorded by electrode 2 (also predominantly derived from one single unit), on the other hand, did not respond to the word presentation (**Fig. 4C**) except towards the end of some words before repeating them (**D**), but it was strongly activated during the overt repeat (**D**). The histogram in **Fig. 4E** in which the averaging sweep was triggered by the patient's own voice clearly shows that this response was auditory as it started only after a definite latency following the beginning of the patient's speech.

Such differential activation patterns of nearby units by listening and speaking is even seen in units recorded by the same electrode as shown in the example of **Fig. 5**, a recording from the left superior temporal gyrus of a German speaking patient. The words spoken to him from tape (bold letters) activated predominantly a neuron of medium amplitude. This activation was related to some phonemic as well as temporal aspects of the words and was especially strong during the later parts of some compound words (see preceding report). When the patient repeated the words (italic letters) another unit of larger amplitude became active. In contrast to the activation of the smaller unit during listening, the activation of the large unit during speaking was only slightly segmented and was especially strong (25–35 disch/s) and long lasting during long compound words (e.g. **Fig. 5E, F, H**, and histograms and dot displays in **Fig. 6**). It appeared also during repeat of English words whose meanings were not known to the patient (e.g. **Fig. 5C**) (note that the time between the end of the word “rattlesnake” and the beginning of repeat was much longer (850 ms) than that after German words (350–450 ms)). The activation of the large unit during word repeat always started after some latency as the voice triggered response averages of this unit show (**Fig. 6**). Responses related to monosyllable words usually appeared only after the end of the vocalization (**Fig. 5A, 6A**).

Middle temporal gyrus

In the middle temporal gyrus, the neuronal activations during word repeat were weaker and, as during listening, not related to any temporal or phonemic aspects of the vocalization. However, a conspicuous effect was the inhibition of spontaneous discharge activity while the patient spoke. Of the 16 recording sites, 5 showed predominantly an activation, 6 predominantly an inhibition and 5 were

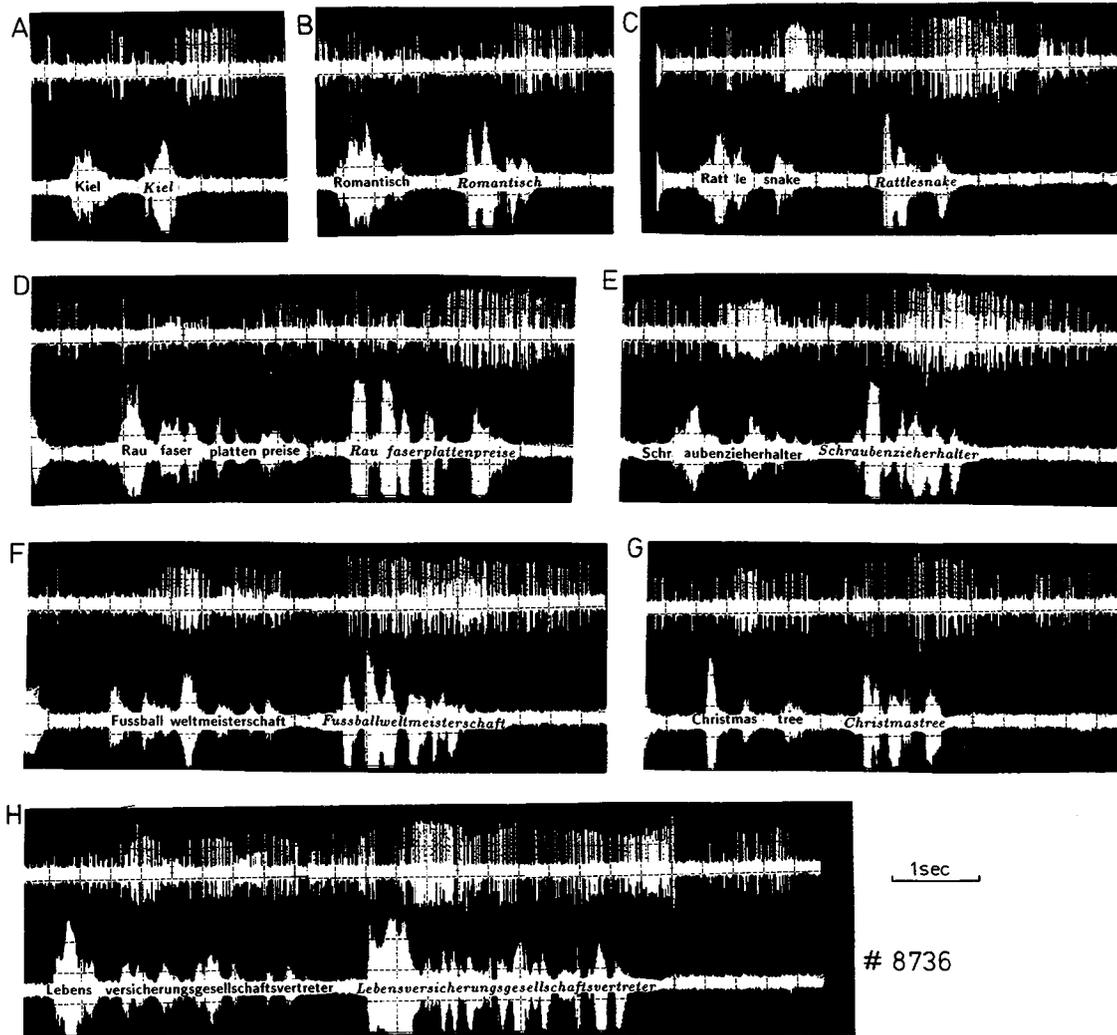


Fig. 5A–H. Recording from the left superior temporal gyrus during the word repeat task (8736, male). This patient could only speak German and did not understand the English words in this series. In the audiograms, the words spoken to the patient from a tape by a male speaker are indicated with bold, the patients repeat with italic letters. Two units can be clearly distinguished, one of very large amplitude and one of medium amplitude. An additional smaller unit activity can be recognized in the background. The unit of middle amplitude is activated during listening by certain phonemic combinations, especially during longer words and may be entrained into the speech rhythm (e.g. **H**) as described in detail in the preceding paper. The large unit is almost exclusively activated only during the repeat and may outlast the patients phonation by several 100 ms (e.g. **B–F**)

unresponsive. Also here, different units recorded simultaneously with one or two adjacent electrodes could respond differently. In the few sites in which units were found to respond to listening, the responses to the patients own voice were stronger, but could be in the same or a different direction.

A typical example of the weak unspecific excitation elicited in middle temporal gyrus recordings during word repeat is shown in Fig. 7. In this record from the right middle temporal gyrus, two units can be distinguished because of different amplitude (the larger is marked by dots). They did not respond to word presentation, but the larger one became more active during repetition. This ac-

tivation was turned on sluggishly and was relatively weak with 7–15 disch/s. This weak and sluggish type of response is in marked contrast to the crisp and sharp responses seen in the superior temporal gyrus (e.g. Figs. 5 and 6).

Examples of strong inhibition during speaking are shown in Fig. 8. The units in Fig. 8A/B and C, respectively, were recorded successively at the same site in the right middle temporal gyrus. They both had an unusually high maintained discharge rate (20 disch/s in Fig. 8 A/B and around 30–50/s in C). The unit in A/B actually always discharged in doublets or triplets (see inset F) and therefore might be considered pathological (“epileptic neu-

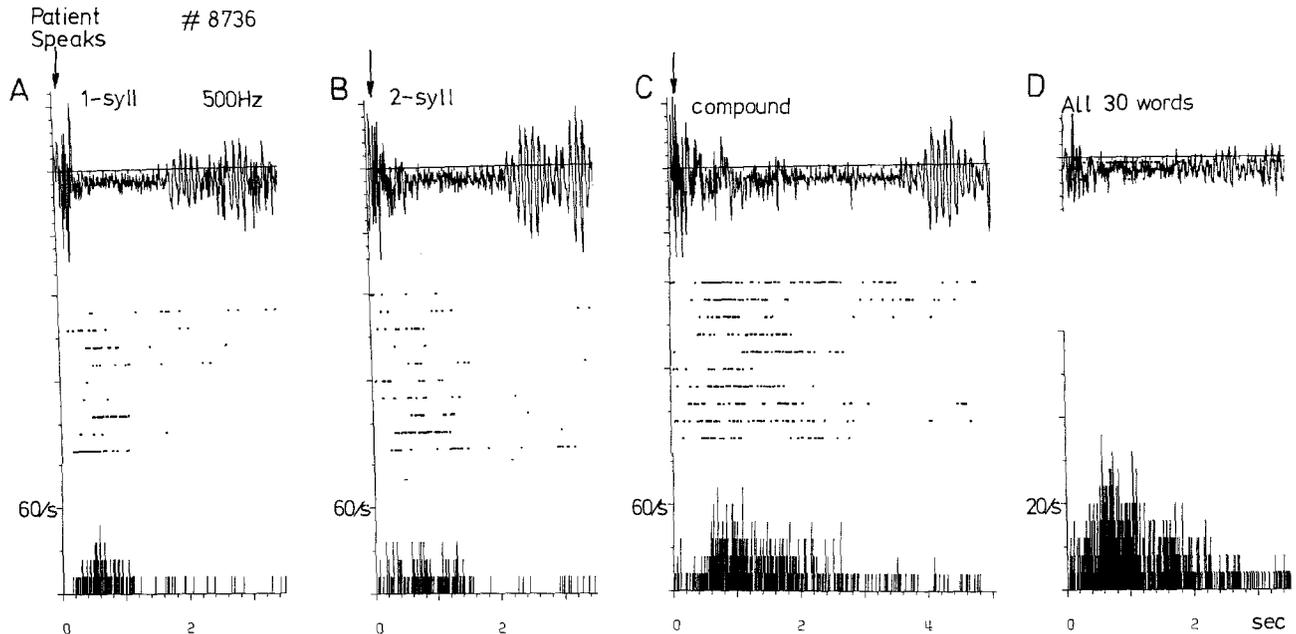


Fig. 6A–D. Averaged discharge rates of the largest unit shown in the selected records of Fig. 5. The averaging was started by the patients voice. *Top*: averaged audiograms, *middle*: single dot displays of the activity during each run, *bottom*: averaged discharge rates. **A** Responses to 10 monosyllable words, **B** 10 two-syllable words, and **C** 10 long compound words. **D** Average of all 30 words. Bin width 10 ms

ron", Wyler and Ward 1980). Both units were not significantly affected by words or sentences spoken to the patient, but their high spontaneous activity was completely suppressed when he started to repeat a word or sentence. This suppression could last for up to nearly one second after the end of the word. The inhibition in both units usually started about at the same time as the vocalization, but could sometimes start before the patient started to speak (A2, B2). This is clearly different from excitatory responses in the superior and middle temporal gyrus which always started at a considerable latency after the beginning of the vocalization. This might suggest that the inhibition is not so much caused by the sound of the patients voice like the excitations, but by the intention to speak. In Fig. 8A the electrode had also picked up the activity of a unit with smaller amplitude which appeared to be activated during speaking. Another unit recorded simultaneously with a second microelectrode at a tip distance of 3–4 mm showed only little if any response to speaking (not shown).

Inferior temporal gyrus

In this location, we have investigated activity changes during speaking only in four patients. Two recording sites were unresponsive. In the other

two, increase and suppression of single neuronal activity during speaking was recorded similar to that observed in the middle temporal gyrus. Activation was only moderate (less than 20 disch/s), started late during the vocalization and lasted for some hundred ms following word utterance (Fig. 9A, B, top record). Suppression of neuronal activity could be recorded in one patient at the same site where the activated unit was recorded (Fig. 9, lower histograms). However, the decrease of activity does not appear to be as sharp and complete as that seen in some middle temporal gyrus recordings.

Discussion and conclusions

Our recordings of single unit activity in the lateral temporal lobe during speaking of the subject have revealed that neuronal activity in this region is differently affected by the patients own voice and by the voice of others, that the neuronal populations affected by speaking have a slightly wider distribution over the lateral temporal lobe than those affected by only listening to language sounds, and that neuronal responses in the superior temporal gyrus were more specifically related to the voice elements than those elicited in other parts of the temporal lobe. As we have not seen such strong and timelocked changes of neuronal activity during

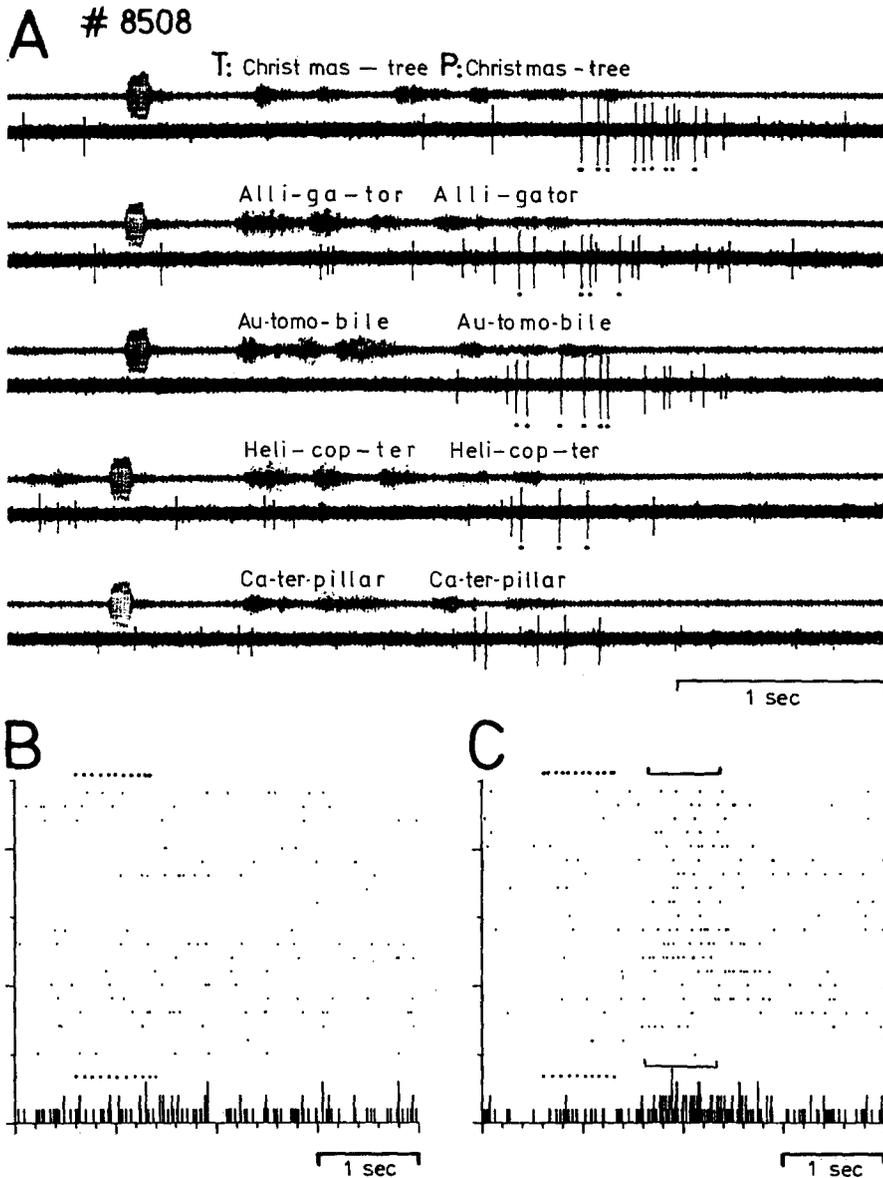


Fig. 7A-C. Slight activation of a unit in the right middle temporal gyrus during word phonation (8508, female). **A** Original recording with audio-record on top. Two units can be distinguished, one of larger (marked by dots) and one of smaller amplitude. The large unit is slightly activated during the second half of word repetition. No response to word listening. **B, C** Dot diagrams and averaged discharge rates. Words spoken to the patient (male speaker from tape) are marked T in **A** and by the dotted horizontal lines in **B, C**, the patient's voice is marked P in **A**, and by the continuous line in **C**.

silent reading or naming, the responses during speaking must be related to the phonation in one way or the other. As during listening to speech, we did not see prominent differences of responses in the right and left hemisphere.

Superior temporal gyrus

In this region, which we had defined in the preceding report as an auditory association area, neuronal activation during overt speech could be quite specifically related to single elements of phonation. These are auditory responses as the definite latency of activations after onset of the respective speech sounds indicate. We have, so far, no indication of activations appearing before phonation. This

excludes input from motor control regions of speech production, in the sense of a corollary discharge. It appears to be rare, on the other hand, that the same phonemic combination which elicits a response during listening to language would also elicit a response during speaking. We saw it clearly only in two recordings. Even in the most impressive example (Fig. 3 of this report and Fig. 3 of the preceding paper), the responses to the phonemic category *scr* in *corkscrew* and *...is crying...* was much weaker when the patient pronounced it than when another male or female speaker said it. This could be due to the different sound spectrum reaching the inner ear from outside or when speaking oneself (air vs. bone conduction). On the other hand, acoustic elements of language are not

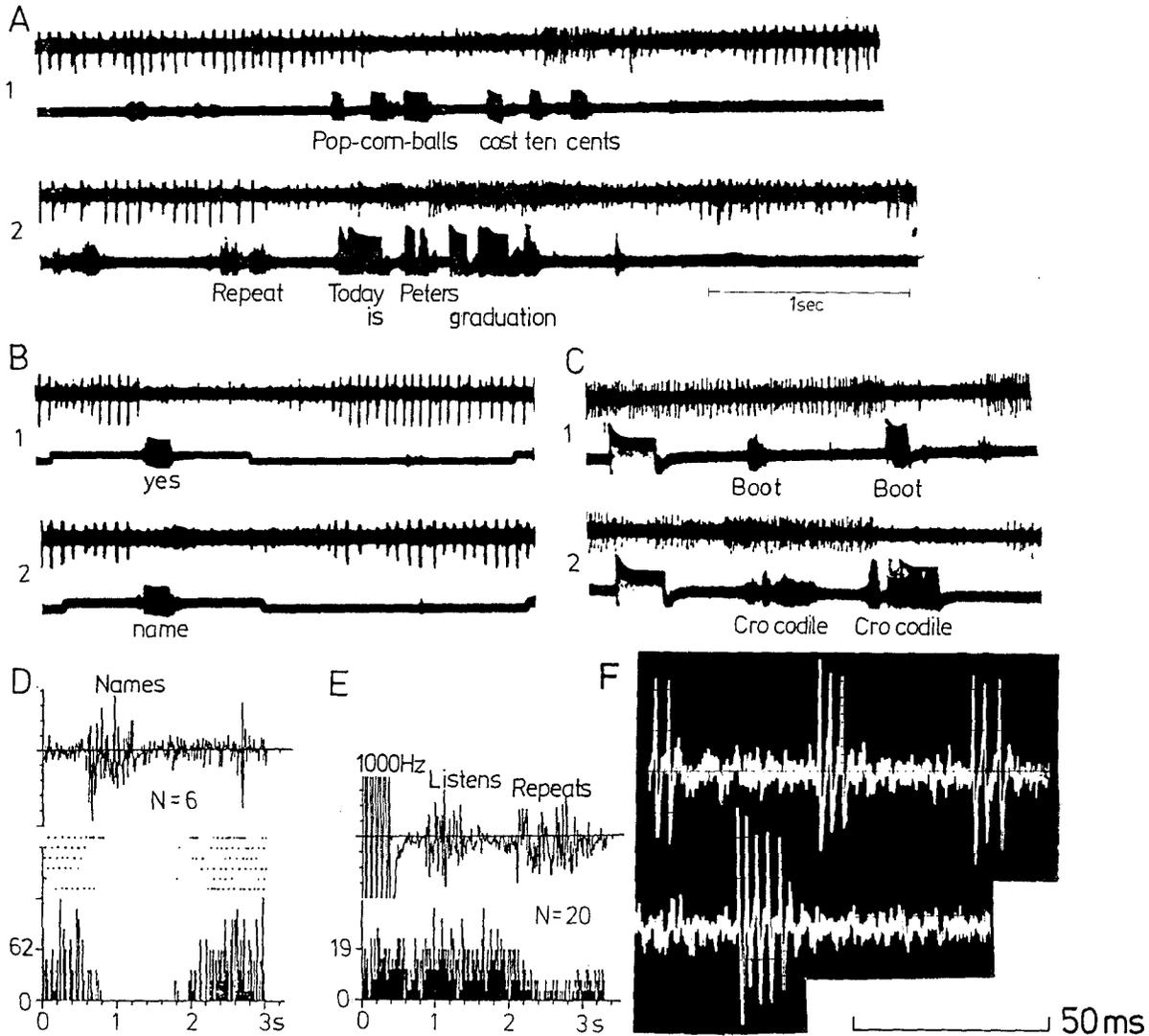


Fig. 8A-F. Suppression of neuronal activity in the right middle temporal gyrus during speaking (8529, male). **A, B, D** This unit always discharged in short high frequency bursts (see **F**), but the activity was completely suppressed when the patient repeated a sentence (**A 1, 2**) or said a word (**B 1**) or named an object (**B 2, D**). Note, that this suppression of activity could begin before the patient started to speak (**A 2, B 2**) and always outlasted the phonation. Low amplitude multiunit activity may be activated during the sentences. **C, E** Another unit recorded earlier at the same site but at a slightly different depth was also suppressed during speaking (repeat of words), but the activity was not changed when the patient heard the words from tape. **F** Recording at faster speed of some discharges of the unit in **A, B, D**

the only and often not the most prominent features represented by neuronal responses in the superior temporal gyrus as demonstrated and discussed in the preceding paper. As during listening, responses are more frequently related to the temporal aspects of spoken language and may thus be connected to segmentation and word length as documented in the example of Fig. 5.

A typical finding in our recordings was that an individual neuron responded predominantly to listening or speaking, or that the response to one was much stronger than to the other. Consequently, the local distribution of activated neurons with-

in the responsive region of the superior temporal gyrus becomes a distinguishing characteristic of representation of language sounds from outside or from the subject himself. We must leave it open at this stage whether acoustic features of heard and self-produced language determine as such that one neuron responds more or less to one or the other type of language, or whether the responsiveness of the whole local cortical circuitry is changed by inputs from other speech related areas, so that different sets of neurons are set free to respond to one or the other type of speech input. We hesitate to conclude, at this point, that there are neu-

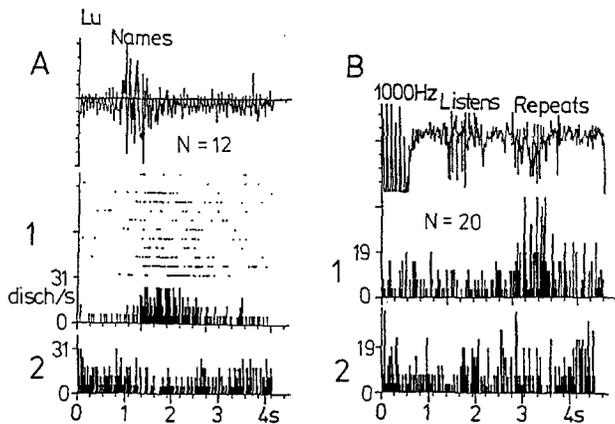


Fig. 9A, B. Activation of one and suppression of another, simultaneously recorded unit from the right inferior temporal gyrus during word utterance (8529). These recordings were done in the same patient as those shown in Fig. 8, but from a different location. During naming of objects **A** and word repeat **B** the activity recorded by electrode 1 was increased, while that recorded by electrode 2 was decreased. No clear response during word listening. The electrode tips were separated by about 3–4 mm and were located above each other

rons in the superior temporal gyrus, which exclusively respond to own or to outside linguistic sounds, although within the framework of our tests some neurons may indeed respond quite exclusively to one or the other.

Another aspect of neuronal responses to voices from outside and to ones own voice is the *inhibition of ongoing spontaneous activity* by one or the other. Thus, for example, the activity of the small unit in Fig. 1, second record, is suppressed when the patient speaks, preceding and outlasting the vocalization, but not if another person speaks (see Fig. 2 in the preceding paper). Such neurons could have the function of gate neurons which are turned on or off by one or the other speech situation, and thus open or close the access of incoming activity to a group of neurons.

Middle and inferior temporal gyrus

In contrast to the little responsiveness of neurons in these regions to language spoken by others, about 2/3 of our recordings revealed clear excitatory or inhibitory neuronal responses during phonation. Yet, the excitatory responses were much less specifically related to phonemic or temporal aspects of the words than in the superior temporal gyrus and were more of the unspecific type as defined in the preceding report. In 1/3 of our population and in about half of the responsive units in the middle and inferior temporal gyrus, spontaneous activity was strongly suppressed during speak-

ing, however. Like in the superior temporal gyrus, this suppression often preceded the actual phonation by up to a few 100 ms and outlasted it by up to a second. This could indicate that the activity is turned off not so much or not only by the auditory signal itself but by other inputs more closely related to speech command. The purpose of such a wide spread suppression of activity in regions not directly related to language analysis or speech command is not clear, but one may speculate that spontaneous activities in these regions could interfere with speaking and therefore have to be cut down during speaking.

We have asked in the Introduction what might distinguish the neuronal representation of speech signals if somebody else speaks to us and if we speak ourselves. We can now give some answers: Regions which are involved in auditory language representation, such as the superior temporal gyrus, are also involved in the auditory representation of our own voice. Representation of the own voice follows similar if not the same general rules as representation of voices from outside, but individual neurons may respond differently and as a consequence the distribution of activities differs in the two situations. This is further supported by the evoked potential pattern such as exemplified in Fig. 10, which shows the potentials recorded just above the recording point of the microelectrode, together with the discharge rates of the units. The surface potentials evoked by the 500 Hz tone (Fig. 10A, C) differ from those elicited by the words sounded to the patient from the tape (B, D), and these again differ from those elicited by the patients own voice when he repeats the words (E). Thus all three conditions evoked potentials in this site, but all three potentials differed. It may be noted that listening to words lead to a synchronized rhythmical afterdischarge of 7/s waves (Fig. 10B) which were almost completely absent when the patient had to repeat the words (D). This is consistent with earlier observations of a reduction of 7–12/s activity in temporo-parietal recordings during overt naming (Fried et al. 1981; Ojemann and Lettich 1989). This suppression of rhythmical EEG-waves may be a consequence of the suppression of ongoing activity during speaking, as we often found a weak correlation between rhythmical EEG-waves and single unit discharges (Creutzfeldt, Gädicke, Ojemann and Chatrian, in preparation).

Our data furthermore answer to the question whether, in addition to the different local activity pattern, the extent of cortical surface involved in speaking and listening to speech differs. This is

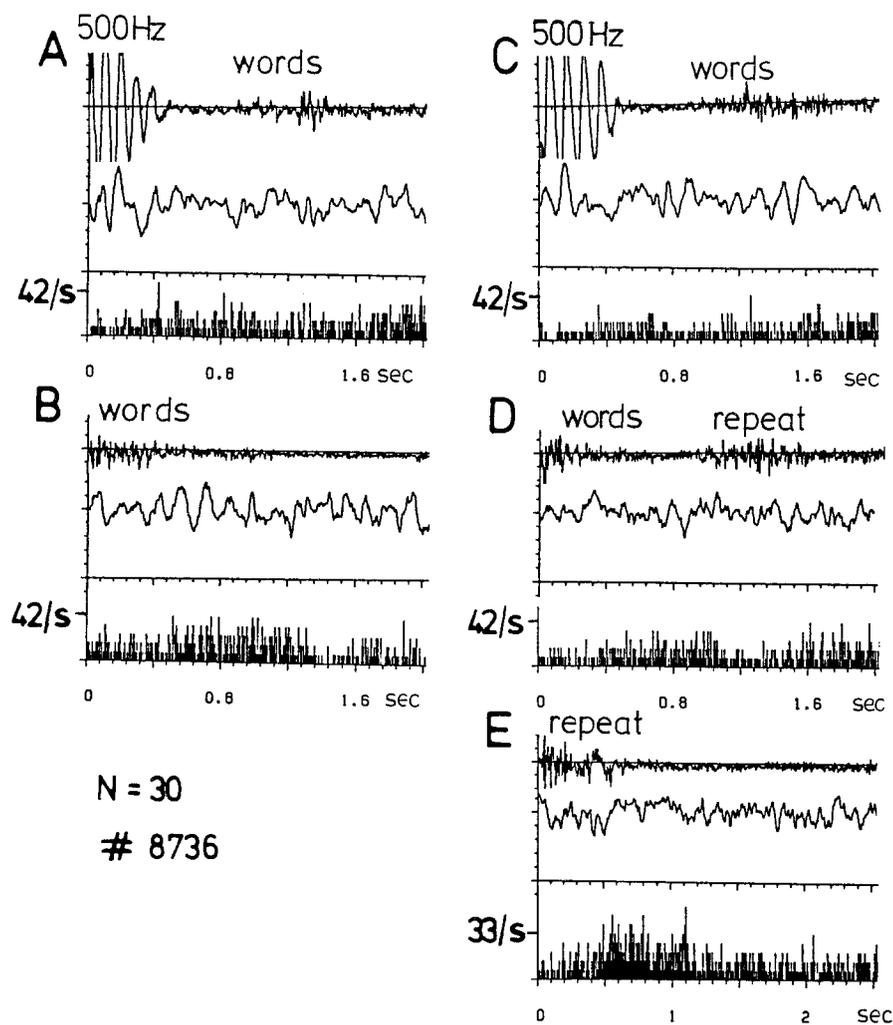


Fig. 10A-E. Averaged evoked potentials in the electrocorticogram and neuronal activity in the left superior temporal gyrus during listening to words and during word repeat (8736, male). Each record is averaged from 30 single presentations of 10 short and 20 longer to very long words. The averaged audiogram is shown on top, that of the electrocorticogram in the second row and the average unit activity in the third row. **A, B** The patient was asked just to listen to the words. In **A**, the averaging was synchronized with the beginning of the tones and in **B** with the speaker's voice. **C-E** The patient was asked to repeat each word. The averaging was synchronized with the 500 Hz tone in **C**, the speaker's voice in **D** and the patient's voice in **E**. Note, slightly different time scale in **A-D, E**, respectively. Since the words were of different length, the repeats of words in **D** begin at different times after the beginning of the averaging evoked potential sweep, and therefore no voice related evoked potential or unit activation is seen. In **E** the sweeps are synchronized on the patient's voice. Histograms are averages of discharge rates of three units recorded by the same microelectrode. The undulations on the averaged 500 Hz tone peeps are not directly related to the tone frequency but caused by sampling. Bin width in **A, B** 8 ms, in **C-E** 10 ms

indeed indicated by the fact, that middle and inferior temporal gyrus activity is slightly more affected by the subjects speaking than by listening to speech. It is more obvious from the distribution of evoked potential. In Fig. 11, we show the averaged surface potentials from five different sites. These recordings are from a patient with a frontal glioma the outlines of which are indicated by the broken line. When only listening to words (**A**), the evoked potential, i.e. a large positive wave at a latency of about 300 ms, was essentially restricted to the anterior part of the superior temporal gyrus. When the patient was asked to repeat the words, a potential evoked by the words was now also recorded over the posterior part of the superior temporal gyrus. During repeat, potential variations of different amplitude and time course appeared over the whole superior temporal gyrus, as well as over the frontal cortex in the region of Broca's area, where a negative motor potential preceded the phonation. It may be noted that in this record as

well, the rhythmical 9/s-activity following the word-evoked positive wave during listening was almost completely wiped out when the patient had to repeat the words. A more widespread involvement of various cortical regions in the frontal, prefrontal and temporal cortex is, of course, also evident from measurements of regional cerebral blood flow during overt speech (Ingvar 1983; Ingvar and Schwartz 1974; Ryding et al. 1985) and of cerebral event related potentials (Desmedt 1977; Hillyard and Picton 1987).

The only animal study comparable to our human recordings has been done on squirrel monkeys (Müller-Preuss 1979). Single units were recorded in the primary auditory cortex during phonation and when the same tape-recorded cry was played back to the animal. In this situation some neurons in the primary auditory cortex responded differently, others nearly identically to the self-produced and the play back sounds. These experiments are, however, not directly comparable to our experi-

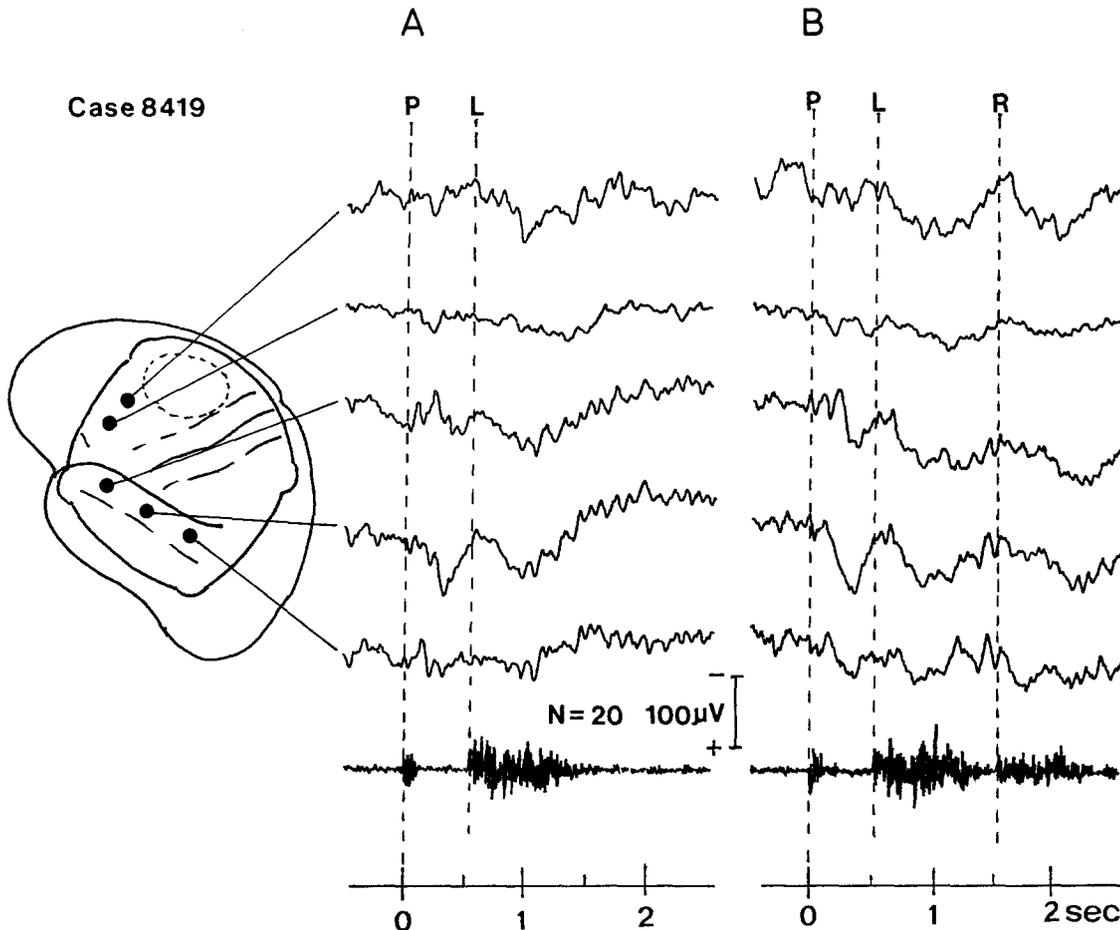


Fig. 11A, B. Auditory evoked potentials recorded from five corticographic electrodes on the left frontal and the superior temporal lobe as indicated (8419, male). **A** Listening to words. **B** Listening and repeat. The potentials are averages from 20 runs. The vertical broken lines indicate the onset of the 1000 Hz tone (P), the word presentation from the tape (L) and the beginning of the patient's voice (R). There was a slight time jitter of up to 15 ms between the beginning of the tone and the beginning of the word presentation, and an even larger jitter (up to 200 ms) to the beginning of the word repeat. Note, different amplitudes and distribution of evoked activities during listening and word repeat (see text)

mental design as the tape recorded words and sentences in our experiments were spoken by different speakers and not by the patient. Interestingly, also in the monkey's primary auditory cortex the activity of some neurons was suppressed during phonation (Müller-Preuss and Ploog 1981).

The lateral temporal cortex including the anterior and middle part of the superior temporal gyrus from which we recorded is not essential and thus disposable to some extent for language comprehension and speaking even in the language dominant hemisphere, at least as long as the other side is intact. When the resection of the superior temporal gyrus in the language dominant hemisphere has to be extended further posteriorly patients may have difficulties for some time to repeat long compound words, and they avoid them during spontaneous speech. We may conclude therefore that this part of the temporal lobe, at least on the language

dominant hemisphere, is involved and to some extent even necessary for the sensory control of speech and the composition of complicated words. Ischemic lesions in the perisylvian language cortex do not only lead to pure sensory aphasias but also to severe disturbances of language production and speech control such as seen in Wernicke's aphasia. These observations in conjunction with our findings thus indicate, that the sensory feedback and its representation in auditory areas during speaking must play an important role for speech control.

References

- Creutzfeldt O, Ojemann G, Lettich E (1989) Neuronal activity in the human lateral temporal lobe. I. Responses to speech. *Exp Brain Res* 77:451-475
- Desmedt JE (ed) (1977) Language and hemispheric specialization in man: cerebral event related potentials. *Progr Clin Neurophysiol*, Vol 3. Karger, Basel

- Fried M, Ojemann GA, Fetz E (1981) Language related potential specific to human language cortex. *Science* 212:353–356
- 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
- Ingvar DH (1983) Serial aspects of language and speech related to prefrontal cortical activity: a selective review. *Hum Neurobiol* 2:177–189
- Ingvar DH, Schwartz MS (1974) Blood flow patterns induced in the dominant hemisphere by speech and reading. *Brain* 97:273–288
- Larsen B, Skinhøj E, Lassen NA (1978) Variation in regional cortical blood flow in the right and left hemispheres during automatic speech. *Brain* 101:193–209
- Müller-Preuss P (1979) Processing of self-produced vocalizations by single neurons in the auditory cortex of the squirrel monkey (*Saimiri sciureus*). In: Creutzfeldt OD, Scheich H, Schreiner C (eds) *Hearing mechanisms and speech*. *Exp Brain Res Suppl* 2:146–151
- Müller-Preuss P, Ploog D (1981) Inhibition of auditory cortical neurons during phonation. *Brain Res* 215:61–76
- Ojemann GA, Fried I, Lettich E (1989) Electrocorticographic (ECoG) correlates of language: desynchronization in temporal language cortex during object naming. *Electroencephal Clin Neurophysiol* (in press)
- 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
- Ryding E, Brådvik B, Ingvar DH (1985) Simultaneous bilateral changes in regional cerebral blood flow (RCBF) from verbal and nonverbal activations. *J Cereb Blood Flow Metabol, Suppl* 1, 5:207–208
- Wyler AR, Ward jr AA (1980) Epileptic neurons. In: Lockard JS, Ward jr AA (eds) *Epilepsy: a window to brain mechanisms*. Raven Press, New York, pp 51–68

Received December 6, 1988 / Accepted June 9, 1989