

## Brain-wave recognition of sentences

PATRICK SUPPES\*†, BING HAN\*, AND ZHONG-LIN LU‡

\*Center for the Study of Language and Information, Stanford University, Stanford, CA 94305-4115; and ‡Department of Psychology, University of Southern California, Los Angeles, CA 90089

Contributed by Patrick Suppes, October 22, 1998

**ABSTRACT** Electrical and magnetic brain waves of two subjects were recorded for the purpose of recognizing which one of 12 sentences or seven words auditorily presented was processed. The analysis consisted of averaging over trials to create prototypes and test samples, to each of which a Fourier transform was applied, followed by filtering and an inverse transformation to the time domain. The filters used were optimal predictive filters, selected for each subject. A still further improvement was obtained by taking differences between recordings of two electrodes to obtain bipolar pairs that then were used for the same analysis. Recognition rates, based on a least-squares criterion, varied, but the best were above 90%. The first words of prototypes of sentences also were cut and pasted to test, at least partially, the invariance of a word's brain wave in different sentence contexts. The best result was above 80% correct recognition. Test samples made up only of individual trials also were analyzed. The best result was 134 correct of 288 (47%), which is promising, given that the expected recognition number by chance is just 24 (or 8.3%). The work reported in this paper extends our earlier work on brain-wave recognition of words only. The recognition rates reported here further strengthen the case that recordings of electric brain waves of words or sentences, together with extensive mathematical and statistical analysis, can be the basis of new developments in our understanding of brain processing of language.

This paper extends the work reported in ref. 1 on brain-wave recognition of words to such recognition of simple sentences like *Bill sees Susan*. A review of the earlier literature is given in ref. 1 and will not be repeated here. But in that review we did not adequately cover the literature on event-related potentials (ERP) and language comprehension. A good, relatively recent review is that of Osterhout and Holcomb (2), especially the second half that focuses on sentence comprehension. Most of the past work on comprehension has been concerned about establishing a reliable relation between something like the N400 component of the ERP and the semantic or syntactic context. For example, in a classic study Kutas and Hillyard (3, 4) found that semantically inappropriate words elicit a larger N400 component than do semantically appropriate words in the same context.

The research reported here does not focus on event-related potentials, but on the entire filtered brain waves recorded by electroencephalography (EEG) or magnetoencephalography (MEG) methods. The aim of the research is to recognize correctly from brain-wave recordings which simple auditory sentence was heard. Without analyzing semantic aspects yet, we also address the question of identifying the brain waves of individual words of a sentence.

## METHODS

For subjects S8 and S9, EEG and MEG recordings were performed simultaneously in a magnetically shielded room in the Magnetic Source Imaging Laboratory (Biomagnetic Technology, San Diego) housed in Scripps Research Institute. (We number the two subjects consecutively with the seven used in ref. 1, because we continue to apply new methods of analysis to our earlier data.) Sixteen EEG sensors were used. Specifically, the sensors, referenced to the average of the left and the right mastoids, were attached to the scalp of a subject, following the standard 10–20 EEG system: F7, T3, T5, Fp1, F3, C3, P3, Fz, Cz, Fp2, F4, C4, P4, F8, T4, and T6. Two electrooculogram and three electromyograph sensors were used, as in ref. 1. The Magnes 2500 WH Magnetic Source Imaging System (Biomagnetic Technology) with 148 superconductive-quantum-interference sensors was used to record the magnetic field near the scalp. The sensor array is arranged like a helmet that covers the entire scalp of most of the subjects. The recording bandwidth was from 0.1 to 200 Hz with a sampling rate of 678 Hz. For the two subjects, 0.3-s prestimulus baseline was recorded, followed by 3.7-s recording after the onset of the stimulus for S8 and 1.6 s for S9.

An Amiga 2000 computer was used to present the auditory stimuli (digitized speech at 22 kHz) to the subject via airplane earphones with long plastic tube leads. Stimulus onset asynchrony varied from 4.0 to 4.2 s for S8 and from 1.9 to 2.1 s for S9. To reduce the alpha wave in this condition, a scenery picture was placed in front of the subject, who was asked to look at the picture during the recording.

The two subjects, normal, right-handed male native English speakers, aged 32 and 31 years, were run with simultaneous 16-sensor EEG and 148-sensor MEG recordings of brain activity. The observations recorded were of electric (EEG) or magnetic (MEG) field amplitude every 1.47 ms for each sensor.

Both subjects were recorded under the auditory comprehension condition of being presented randomly one of a small set of spoken sentences (S8) or words (S9), 50 trials for each sentence and 80 trials for each word. Subjects were instructed to passively but carefully listen to the spoken sentences or words and try to comprehend them. The seven spoken words used were the four proper names *Bill*, *John*, *Mary*, and *Susan* and the three transitive verbs *hates*, *loves*, and *sees*. The 12 spoken sentences were constructed from the seven words. The list is shown in the first column of Table 1, where the first letter of each word is used to abbreviate the sentence. For example, the entry in the first row of Table 1 is BHJ, which stands for *Bill hates John*.

The verb is always represented by the middle letter, so in this notation there is no ambiguity. The letter S stands for *sees* in the middle position and stands for *Susan* at either end. For subsequent comparison with the corresponding brain wave, the beginning ( $s_i$ ), after the onset of the stimulus trigger, and the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9515861-6\$2.00/0  
PNAS is available online at www.pnas.org.

Abbreviations: EEG, electroencephalography; MEG, magnetoencephalography; FFT, fast Fourier transform.

†To whom reprint requests should be addressed. e-mail: suppes@ockham.stanford.edu.

Table 1. Beginning and end in ms of words in each sentence

	$s_1$	$e_1$	$s_2$	$e_2$	$s_3$	$e_3$
BHJ	125	611	723	1240	1561	1990
BLM	156	614	894	1522	2099	2487
BSS	259	677	959	1645	2020	2489
JHS	99	664	1090	1514	1999	2525
JLM	129	805	1292	1912	2320	2758
JSB	70	650	1055	1712	2009	2351
MHB	30	580	963	1440	1763	2110
MLS	83	637	1029	1639	2049	2544
MSJ	71	614	914	1561	1806	2281
SHB	63	728	992	1458	1683	2004
SLJ	156	859	1385	2011	2473	3012
SSM	112	793	1084	1793	2289	2729
Min	30	580	723	1240	1561	1990
Mean	113	686	1032	1621	2006	2440
Max	259	859	1385	2011	2473	3012

Min, minimum; Max, maximum.

end ( $e_i$ ) in ms,  $i = 1, 2, 3$ , of each word of each sentence are given in Table 1. From a linguistic standpoint the temporal lengths of the spoken words in each sentence seem long, because we used cutoff points of essentially zero energy in the sound spectrum. There are, as can be seen from Table 1, short silences between the words. The sentences were spoken slowly with each word distinctly articulated. A key question, discussed later, is whether or not these breaks are present in the corresponding brain waves.

For each trial, we used the average of 203 observations before the onset of the stimulus as the baseline. After subtracting out the baseline from each trial, to eliminate some noise, we then averaged data, for each sentence or word and each EEG and MEG sensor, over every other trial, starting with trial 2, for half of the total number of trials. Using all of the even trials, this averaging created a prototype wave for each sentence or word. In similar fashion, five test wave forms, using five trials each for the sentences and eight trials each for the words, were produced for each stimulus under each condition by dividing all the odd trials evenly into five groups and averaging within each group. This analysis was labeled E/O. We then reversed the roles of prototypes and test samples by averaging odd trials for prototype and even trials for test. This analysis was labeled O/E. In ref. 1 we did both analyses for all cases and the two did not greatly differ. Here the analyses we computed were results only for E/O. We imposed this limitation because of the larger data files for the sentences as compared with the isolated words studied in ref. 1.

The main additional methods of data analysis were the following. First, we took the complete sequence of observations for each prototype or test, both before and after the stimulus onset, and placed the sequence in the center of a sequence of 4,096 observations for the sentences or 2,048 observations for the words. Next, we filled the beginning part of the longer sequence by mirroring the beginning part of the centered prototype or test sequence, and the ending part by mirroring the ending part of the center sequence. Using a Gaussian function, we then smoothed the two parts filled in by mirroring. The Gaussian function was put at the center of the whole prototype or test sequence with a SD equal to half of the length of the prototype or test sequence. Each observation in the longer sequence that was beyond the centered prototype or test sequence was multiplied by the ratio of the value of the Gaussian function at this observation and the value of the Gaussian function at either end of the centered prototype or test sequence. After mirroring and smoothing, we applied a fast Fourier transform (FFT) to the whole sequence of 4,096 observations for the sentences or 2,048 observations for the words for each sensor. (The FFT algorithm used restricted the

number of observations to a power of two.) We then filtered the result with a fourth-order Butterworth bandpass filter (5) selected optimally for each subject, as described in ref. 1. After the filtering, an inverse-FFT was applied to obtain the filtered wave form in the time domain, whose baseline then was normalized again.

As in ref. 1, the decision criterion for prediction was a standard least-squares one. We first computed the difference between the observed field amplitude of prototype and test sample, for each observation of each sensor after the onset of the stimulus, for a duration whose beginning and end were parameters to be estimated. We next squared this difference and then summed over the observations in the optimal interval. The measure of best fit between prototype and test sample for each sensor was the minimum sum of squares. In other words, a test sample was classified as matching best the prototype having the smallest sum of squares for this test sample.

## RESULTS AND DISCUSSION

**Best EEG Recognition Rate for Sentences.** Following the methods of analysis in ref. 1, we estimated four parameters, first, as in ref. 1, the low frequency ( $L$ ) of the optimal bandpass filter and the width ( $W$ ) of the bandpass. Second, we estimated the starting point ( $s$ ) and ending point ( $e$ ) of the sample sequence of observations yielding the best classification result. The parameters  $s$  and  $e$  were measured in ms from the onset of the stimulus.

After some exploration, we ran a four-dimensional grid of the parameters by using the following increments and ranges: the low frequency  $L$  from 2.5 to 3.5 Hz in step size of 0.5 Hz, the width  $W$  from 6.5 to 7.5 Hz, also in a step size of 0.5 Hz; the starting sample point ( $s$ ) from 0 to 441 ms after the onset of stimulus with a step size of 15 ms, and the ending sample point ( $e$ ) from 1,911 to 3,381 ms with a step size of 15 ms. Note that 10 observations equal 14.7 ms. We show  $s$  and  $e$  only in ms, rather than number of observations. (To convert back to number of observations, divide the temporal interval of observations by 1.47 ms.)

The best EEG result was correct recognition of 52 of the 60 test samples, for a rate of recognition of 86.7%. The parameters were:  $L = 3$  Hz,  $W = 7$  Hz,  $s = 88$  ms,  $e = 2,455$  ms. This result is comparable to the results for the best two of the seven subjects of ref. 1.

The best recognition rate of 86.7% was for sensor C3. In Fig. 1 we show the recognition rate for the 10–20 system of EEG sensors, with isocontour lines of recognition. The variation in sensors is quite large, ranging from nine for Fp1 to the 52 for C3. But, as is well known, inference from location on the scalp of the best sensor to the physical site of processing in the cortex is difficult. What Fig. 1 makes clear is the importance of having a sensor close to the optimal location on the scalp, whatever may be the location of the source in the cortex of the observed electric field.

In Fig. 2 we show a contour map, including the initial exploration, of the recognition-rate surface for filter parameters  $L$  and  $W$  of S8. Again, the recognition-rate sensitivity to parameter variation is high, comparable to that reported in ref. 1. In Fig. 3 we show a recognition-rate contour map for starting and ending parameters  $s$  and  $e$ . Variation of parameters around the best pair of values seems not as sensitive as in the other two figures, but the scales used are not directly comparable.

**Bipolar Version of EEG Recognition.** The virtues and defects of bipolar analysis are discussed rather carefully in ref. 6, but not with respect to problems of recognition, as characterized here. First, a matter of terminology, the standard 10–20 EEG system recordings are physically bipolar, for they record the potential differences between each electrode  $v_i$  on the scalp

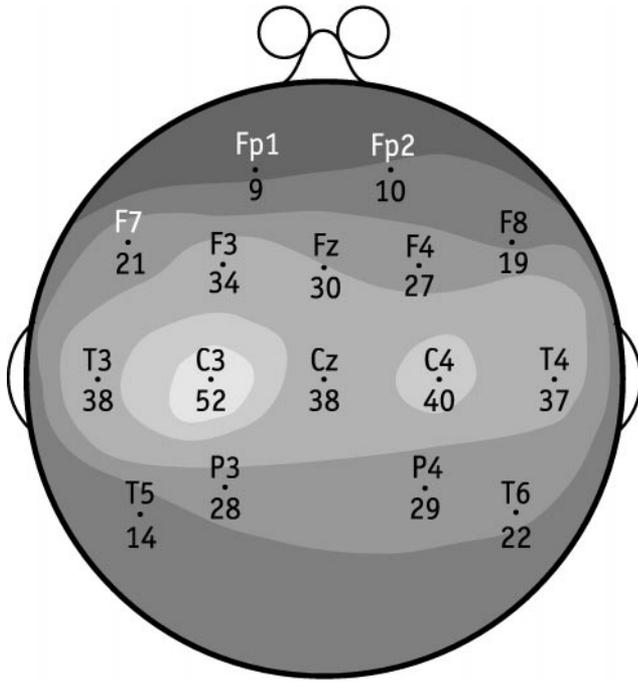


FIG. 1. Shaded contour map of recognition-rate surface for the 10–20 system of EEG sensors for subject S8. The physical surface of the scalp is represented as a plane circle as is standard in representations of the 10–20 system. The recognition rate for each sensor is shown, next to it, as the number of test samples of sentences correctly recognized of a total of 60. The predictions are for the best parameters  $L = 3$  Hz,  $W = 7$  Hz,  $s = 88$  ms, and  $e = 2,455$  ms.

and the reference electrode  $v_r$ . But in EEG research bipolar often refers to potential differences between two nearby electrodes on the scalp. (However, here each bipolar pair was

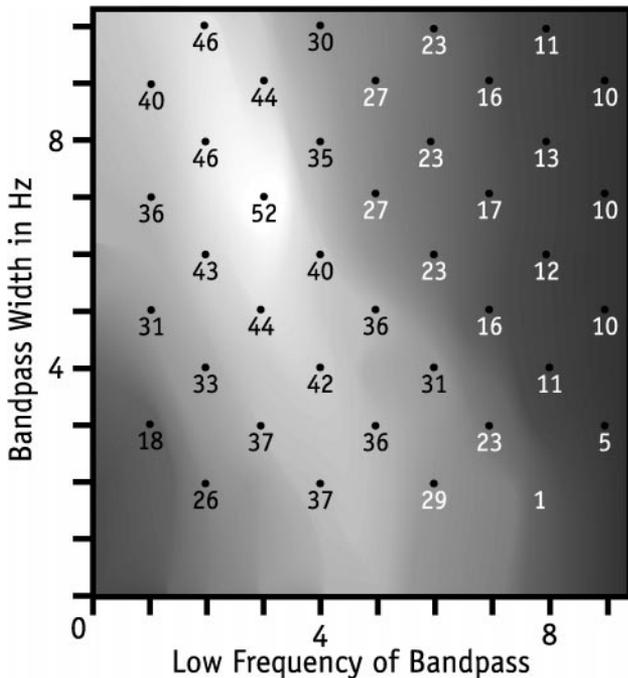


FIG. 2. Contour map of recognition-rate surface for filter parameters  $L$  and  $W$  for subject S8. The  $x$  coordinate is the low frequency ( $L$ ) in Hz and the  $y$  coordinate is the width ( $W$ ) in Hz of a filter. The number plotted as a point on the map is the number of test samples of sentences correctly recognized of a total of 60 by the Butterworth filter with the coordinates of the point. The predictions are for the best temporal parameters  $s = 88$  ms,  $e = 2,455$  ms, and best sensor C3.

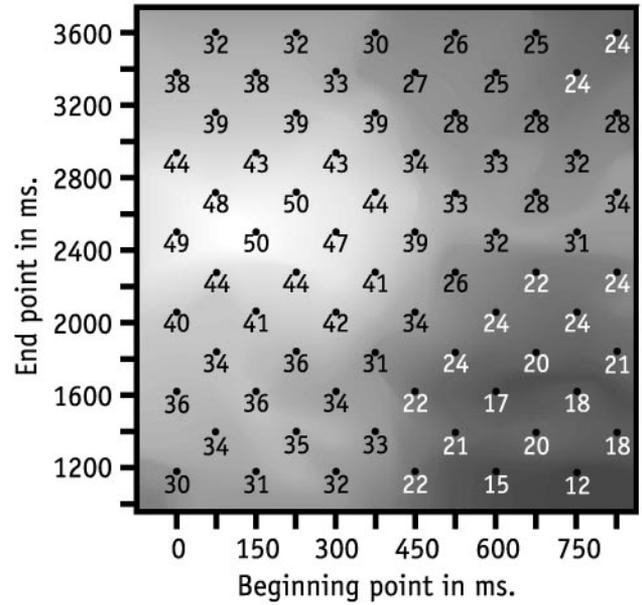


FIG. 3. Contour map of recognition-rate surface for start ( $s$ ) and ending ( $e$ ) parameters for subject S8. The  $x$  coordinate is the start point ( $s$ ), measured in ms, of the interval of observations, used for prediction. The  $y$  coordinate is the ending point ( $e$ ), measured in ms, of the same interval. The recognition-rate numbers have the same meaning as in Figs. 1 and 2. The predictions are for the best filter parameters  $L = 3$  Hz and  $W = 7$  Hz, and best sensor C3.

not recorded as such, but obtained from the data by taking the difference for each individual trial between the recordings for two electrodes.) We define bipolar  $v_{ij}$ , where  $v_i$  and  $v_j$ ,  $i \neq j$ , are scalp electrodes as:

$$v_{ij} = (v_i - v_r) - (v_j - v_r) = v_i - v_j.$$

(Here we explicitly show  $v_r$  to indicate how it is removed in the subtraction.)

Using the same Fourier transform and filter analysis on the bipolar  $v_{ij}$ , as was used above, we obtained slightly better predictive results. Typical, and tied for the best, was 56 of 60 (93.3%) for electrodes C4–T6, with the filter 2.5–9 Hz and the temporal window from 29 to 2,808 ms, i.e.,  $s = 29$  ms and  $e = 2,808$  ms.

Also of interest is the fact that the corresponding pair of electrodes, C3–T5, over the left hemisphere, did almost as well, namely, 52 of 60. Two other sets of pairs displayed considerable hemispheric symmetry in prediction. The results are summarized in Table 2. Of the 120 possible pairs, only one other pair, beyond those listed in Table 2, predicted correctly 40 or more of 60; this pair was an anomalous pair that straddled both hemispheres: C4–T5 had 43 correct predictions.

**Bipolar Revision of S5 Data of Ref. 1.** The worst case of brain-wave recognition for spoken words as stimuli in our earlier article (1) was for subject S5 in the O/E condition, with 13 of 35 (37.1%) correct. Using a 3- to 13-Hz filter, the temporal window 103 to 809 ms, and the bipolar pair Cz–T4, the recognition rate improved to 19 of 35 (54.3%).

**Individual Trials for Sentences.** The excellent predictive results for the two best bipolar pairs encouraged us to see how

Table 2. Bipolar predictions from the two hemispheres

Left hemisphere			Right hemisphere		
Sensor	Sentence	Words	Sensor	Sentence	Words
C3–T5	52	24	C4–T6	56	19
C3–T3	46	12	C4–T4	36	15
C3–P3	43	17	C4–P4	42	15

well these two bipolar pairs would do in classifying test samples consisting of individual trials. When we used a 2.5- to 9-Hz filter and a temporal window from 118 to 2,631 ms just for these two pairs, the recognition results were: for C3–T5 134 of 288 correct (46.5%) and for C4–T6 122 of 288 correct (42.4%). These results are encouraging, given that the chance level is 1/12. But it is also clear that further methods of analysis will be needed to come close to the 90% success level already reached in classifying test samples of averaged trials.

**Best MEG Recognition Rate for Sentences.** When we used the same parameters as for the best EEG result, namely, a 3- to 10-Hz filter and a window from 88 to 2,455 ms, the best MEG sensor had a recognition rate of 32 of 60. With extensive additional search of the filter-parameter space, we raised the recognition rate only to 33 of 60 for a 3- to 11-Hz filter. That none of the MEG sensors came close to the best EEG sensor is not surprising. This result was similar to what was reported in ref. 1. In addition, the electrooculogram and electromyograph sensors were close to chance level in their recognition rates.

For this reason we report only on EEG results in the remainder of this paper.

**Cut-and-Paste Sentences.** For the 12 sentences used as stimuli in the experiment, we grouped three sentences that have the same first word together, which resulted in four groups: (BHJ, BLM, BSS), (JHS, JLM, JSB), (MHB, MLS, MSJ), and (SHB, SLJ, SSM). Within each group, we left the five test samples for each sentence intact, and cut and pasted the first word in the sentences between prototypes of the three sentences, in the following way: 1→2, e.g., cutting *Bill* from BHJ and pasting it to BLM; 2→3, e.g., cutting *Bill* from BLM and pasting it to BSS; 3→1, e.g., cutting *Bill* from BSS and pasting it to BHJ.

The cut-and-paste operations were done before the FFT, filtering, and inverse-FFT. We used a 3- to 10-Hz filter. The start point (*s*) of the cut-and-paste window was fixed at 88 ms after stimulus onset. The end point (*e*) of the cut-and-paste window was run from 221 ms after stimulus onset to 735 ms after stimulus onset with a step increment of 15 ms.

A least-squares analysis then was carried out for the 60 test samples in total against the 12 cut-and-pasted prototypes. There were several local maxima. Some were early, which is hardly surprising because the best cut would be at the start, i.e., at 88 ms. The two later significant local maxima were at 544 ms with 49 of 60, and at 676 ms with 42 of 60. These results do not sharply identify the endpoint of the first word. More detailed investigation will be required to do this.

**Cut and Pasting Among Different Words.** The cut and paste of different wave tokens of the same word suggests a comparison with cut and paste of different first words, e.g., replacing *Bill* by *John*. Instead of within groups, now we did it across groups, in the order of (BHJ, BLM, BSS) → (JHS, JLM, JSB) → (MHB, MLS, MSJ) → (SHB, SLJ, SSM) → (BHJ, BLM, BSS). Between each two groups directly linked by an arrow, we cut and pasted between prototype sentences with the same verb. For example, (BHJ, BLM, BSS) → (JHS, JLM, JSB) represents cutting *Bill* from BHJ and pasting it to JHS, cutting *Bill* from BLM and pasting it to JLM, and cutting *Bill* from BSS and pasting it to JSB.

The best sensor for the filter 3–10 Hz and the observation window 0–544 ms was again C3, which recognized 32 of the 60 test samples (53.3%). That for this condition we got 32 correct of 60 classifications, with a chance level of 1/12, is not surprising, most likely because in many cases the second and third words of the prototypes, taken together, which were unchanged, were sufficient as a basis for correct recognition. In contrast, when we made the cut at 1,176 ms in the middle of the second word as well, as gauged by the data of Table 1, the best results for the same remaining parameters were much worse, 17 correct predictions of 60 (28.3%).

**Totally Artificial Prototypes.** We cut and pasted the prototypes in such a way that the prototype of each sentence was composed of three segments originally from three other different prototypes and the words corresponding to the new segments were such that the three words for the three new segments still composed, even if artificially, the original sentence. We omit the specific details, which follow along the lines laid out for the two cases of cut and paste just analyzed. The best result was 11 of 60, for sensor T3, with the cut after the first spoken word at 882 ms and after the second word at 2,058 ms. The first cut was between the means of  $e_1$  and  $s_2$  shown in Table 1; the second cut was just after the mean of  $s_3$ . The expectation from pure chance is five, so this result is barely significant. Clearly we need to have a better understanding than we now have to create satisfactory totally artificial prototypes.

**Silence Between Words.** Silence in the speech does not mean silence in the brain waves recorded at the same time (compare the two parts of Fig. 4). The shaded part is the speech spectrum for the sentence *Bill sees Susan* and the solid curved line is the prototype filtered bipolar C4–T6 wave of the sentence used in recognizing test samples. The silence between words is not comparably evident in the brain waves, which, if typical, shows that recognizing the beginnings and ends of words is even more difficult in brain waves than it is in continuous speech. (As indicated earlier, our speech recordings of the sentences were not fast enough to qualify as natural continuous speech, but rather as spoken sentences with each word clearly delineated, as is common in dictation.)

**Auditory-Word Condition for S9.** To begin with, we did our standard FFT and filter analysis. We explored the start (*s*) of the signal sequence from 59 to 118 ms after stimulus onset with an increment of 15 ms, the end (*e*) from 441 to 1,176 ms with increment 74 ms, low frequency of the filter from 2 to 4 Hz with increment 1 Hz, and frequency width from 4 to 8 Hz with increment 1 Hz. Classification was done with all possible combinations of values of these variables. The best performance was 23 correct of 35 (65.7%), with a temporal window from 88 to 662 ms after stimulus onset, and with a filter from 2 to 8 Hz. This best performance was achieved by sensor F8.

**Bipolar Version for S9.** The best analysis was for a filter from 2 to 10 Hz and a window from 118 to 882 ms. The best pair was C3–T5 at 24 of 35 (68.6%) correct predictions, and the corresponding pair on the right hemisphere was C4–T6 as second best of all pairs, with 19 of 35 (54.3%) correct predictions. Further data on corresponding bipolar pairs are given in Table 2 and exhibit performance for matching locations in the

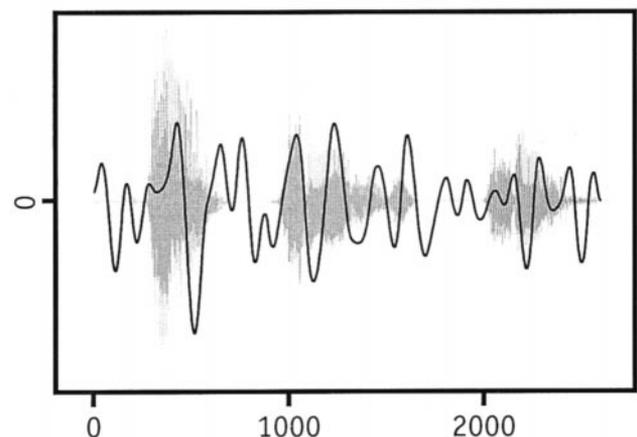


FIG. 4. Speech spectrum and corresponding brain wave of sentence *Bill sees Susan*. The *x* axis is measured in ms after onset of spoken-sentence stimulus. The speech and brain waves have different amplitudes, so no common scale is shown on the *y* axis.

two hemispheres comparable to the bipolar sentence recognition data.

**Using Words of S9 to Recognize Words of S8.** It is natural to ask whether or not the brain wave for the first word in a sentence resembles the brain wave for the word when said alone by a different person. We took the prototypes for both the S8 auditory-sentence condition and the S9 auditory-word condition (all prototypes were generated with half of all the trials according to the way prototypes were generated in the standard E/O case). We used the prototypes for the words *Bill*, *John*, *Mary*, and *Susan* as prototypes in our new classification. For each of these four words, we took the three prototypes for the three sentences that start with the same word in the S8 auditory-sentence case as test samples for the classification using word prototypes from S9. We then did the classification with four prototypes and 12 test samples in total on various windows. The best result obtained was six correct predictions of the 12 possible for sensor T5 with filter of 2–10 Hz and a time window from 59 to 720 ms after stimulus onset. Of these six correct predictions, three were for *John*, i.e., all the test samples of this word, and no other test samples were classified as *John*. In Fig. 5 we show the prototype and best of the three test samples, in terms of least squares, on the same graph. Considering that we are comparing the brain waves of different subjects, the similarity is remarkable.

**Bipolar Version.** The best analysis was correct recognition of 10 of the 12 prototype samples of S8 using prototypes of S9. The parameters and bipolar pair were unusual: filter 9–12 Hz, window 15–368 ms for P4–Fp1. However, many bipolar pairs with many different parameters recognized nine of the 12. We are encouraged to look deeper for brain-wave features invariant between subjects who are processing the same spoken word.

**Analyses of Methodological Interest.** We investigated several different mathematical and conceptual approaches to recognition. Because none of them improved on the recognition results reported above, we describe these alternative approaches only briefly.

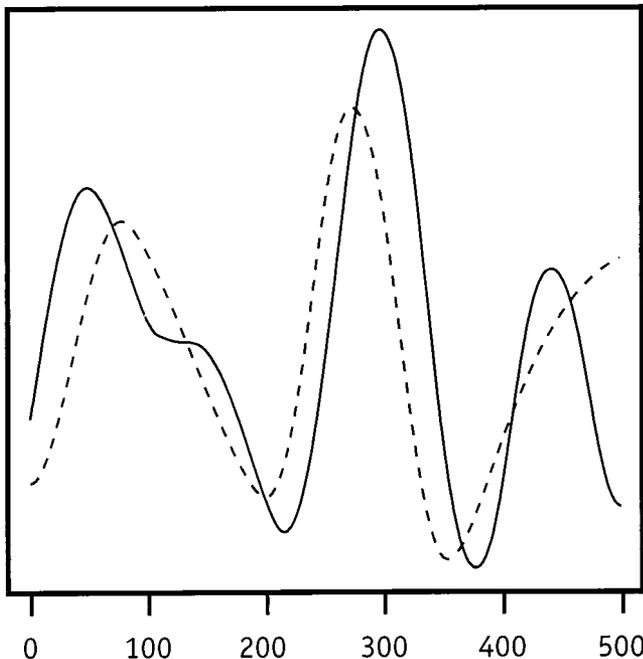


FIG. 5. Brain waves of S8 (dotted curved line) and S9 (solid curved line) for spoken word *John*. In the case of S8, the spoken word occurred as the first word of a sentence. The  $x$  axis is measured in ms after onset of the auditory stimulus.

**Three normalization schemes.** To see whether some rather natural linear transformations would normalize the data across trials in a way that would improve our predictions, we tried the following three schemes on the sentence data of S8, which are described and analyzed above. The transformations were applied to the prototypes and test samples on the hypothesis this might make the prototypes and corresponding test samples more comparable.

The first transformation  $M$  simply subtracted the mean amplitude from the observed amplitude for each observation  $i$ , where  $n$  is the number of observations in each trial:

$$(M) \quad x'_i = x_i - \left( \frac{1}{n} \sum_{k=1}^n x_k \right),$$

the second ( $A$ ) rescaled the amplitude by dividing it by the average absolute value of the amplitude:

$$(A) \quad x'_i = x_i / \left( \frac{1}{n} \sum_{k=1}^n |x_k| \right),$$

and the third ( $E$ ) rescaled the amplitude by dividing it by the square root of the average energy:

$$(E) \quad x'_i = x_i / \sqrt{\left( \frac{1}{n} \sum_{k=1}^n x_k^2 \right)}.$$

We ran several different linear transformations made up of combinations of  $M$ ,  $A$ , and  $E$ . Two of the best were  $A$  alone with correct recognition of 48 of the 60 test samples, and  $E(Mx_i)$  with 49 of 60 correct.

**Exploratory use of wavelets.** We used the discrete wavelet transform (DWT) with symmlet “s8,” introduced by Daubechies (7) (for computational details, see ref. 8). The subbands after DWT were (d1, d2, . . . , d8, d9, s9) on the observation sequence from 0 to 2,048 samples (3,011 ms) after stimulus onset for the prototypes and test samples of all the 12 sentences for sensor C3. For each of the observation sequences, the subbands d1, d2, d3, d4, d5, d9, and s9 were set to zero and inverse-DWT then was taken. By retaining non-zero coefficients corresponding to about 10.6 Hz (d6), 5.3 Hz (d7), and 2.7 Hz (d8), the result is analogous to filtering via a 3- to 10-Hz filter with FFT and inverse-FFT. We then did a least-squares analysis with the wavelet-filtered waveforms for the temporal window from 88 to 2,455 ms after stimulus onset, and obtained 44 of 60 cases correctly recognized. Finally, we explored classification on a variety of temporal windows. The best result was 48 of 60 with five different sets of parameters close together on the grid of beginning and ending points; a

Table 3. Summary of recognition rates

Type of recognition	Result	%
Sentences, S8		
EEG	52 of 60	87
Bipolar EEG	56 of 60	93
Individual trials EEG	134 of 288	47
MEG	33 of 60	55
Cut and paste	49 of 60	82
Artificial prototypes	11 of 60	18
Words, S9		
Bipolar	24 of 35	69
Words (S9) for words (S8)	6 of 12	50
Bipolar	10 of 12	83
Further methods for sentences		
A transform	48 of 60	80
EM transform	49 of 60	82
Wavelets, S8	48 of 60	80
Nearest neighbor	19 of 36	53

typical interval was from 191 to 2,808 ms. So, the wavelet performance was pretty good, in terms of this first exploratory analysis, but not as good as the FFT approach described above.

We also did a recognition analysis using only, in sequence, the individual subbands d6, d7, and d8. Not surprisingly, the results were comparatively poor, the best for d6 was 32 of 60, for d7, 23 of 60, and for d8, 18 of 60.

*Multiple prototypes.* In the current literature of machine learning and adaptive statistics a commonly used method of classification is that of choosing the known classification of the nearest neighbor. In our context we tried a variant of this method by creating for each sentence in the data of S8 five rather than one prototype, so now the prototypes had the merit of being averaged over the same number of trials as the test samples, namely, five trials. To the multiple prototypes we applied the same FFT, filter, inverse-FFT, and least-squares analysis, as in the case of only one prototype. We classified the test sample as the same as that of the prototype, of the many used, which had the best fit. Because of the extra computing required by the many additional Fourier transforms and least squares, we analyzed in blocks of 12 only 36 of the test samples. The best result, which was for sensor C3, window 88 to 2,455 ms and a 3- to 10-Hz filter, was 19 of 36, not nearly as good as the single-prototype result of 52 of 60.

### CONCLUSION

This study and our earlier one (1) show that brain-wave recognition of words and simple sentences being processed auditorily is feasible. In the bipolar analysis we also found surprising symmetry between the two hemispheres for the performing pairs. Moreover, individual words in sentences were identified and successfully cut and pasted to form partly artificial brain-wave representations of sentences. We summarize in Table 3 the main recognition results. (Here and earlier

we have not given standard statistical levels of significance, because the natural null hypothesis of the sensors' being independent and identically distributed random variables is obviously rejected by the regularity of the isocontours of recognition rate, as well as by the large deviations of the best results from chance levels.) We emphasize these recognition results are all for prototypes and test samples averaged over from five to 40 trials, with the one exception of 47% for individual trials. Achieving rates for individual trials close to the best for averaged trials is an important goal for future research of this kind.

We have received useful comments and suggestions for revision from Stanley Peters and Richard F. Thompson. We also thank Biomagnetic Technology, Inc., as well as Scripps Research Institute, for use of the imaging equipment and Barry Schwartz for assistance in running subjects S8 and S9. We are grateful to Paul Dimitre for producing the five figures and to Ann Petersen for preparation of the manuscript.

1. Suppes, P., Lu, Z.-L. & Han, B. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14965–14969.
2. Osterhout, L. & Holcomb, P. J. (1995/1996) in *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, eds. Rugg, M. D. & Coles, M. G. H. (Oxford Univ. Press, New York), pp. 171–215.
3. Kutas, M. & Hillyard, S. A. (1980) *Biol. Psychol.* **11**, 99–116.
4. Kutas, M. & Hillyard, S. A. (1980) *Brain Lang.* **11**, 354–373.
5. Oppenheim, A. V. & Schaffer, R. W. (1975) *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ), pp. 211–218.
6. Nunez, P. L. (1981) *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford Univ. Press, New York).
7. Daubechies, I. (1992) *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, PA).
8. Bruce, A. & Gao, H.-Y. (1996) *Applied Wavelet Analysis with S-PLUS* (Springer, New York).