

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-521768
(P2024-521768A)

(43)公表日 令和6年6月4日(2024.6.4)

(51)国際特許分類
A 6 1 B 5/372(2021.01)

F I
A 6 1 B 5/372

テーマコード (参考)
4 C 1 2 7

審査請求 未請求 予備審査請求 未請求 (全187頁)			
(21)出願番号	特願2023-572722(P2023-572722)	(71)出願人	506115514
(86)(22)出願日	令和4年5月26日(2022.5.26)		ザ リージェンツ オブ ザ ユニバーシテ
(85)翻訳文提出日	令和6年1月18日(2024.1.18)		ィ オブ カリフォルニア
(86)国際出願番号	PCT/US2022/031101		The Regents of the U
(87)国際公開番号	WO2022/251472		niversity of Califo
(87)国際公開日	令和4年12月1日(2022.12.1)		rnia
(31)優先権主張番号	63/193,351		アメリカ合衆国, カリフォルニア州 9
(32)優先日	令和3年5月26日(2021.5.26)		4 6 0 7 - 5 2 0 0, オークランド, フ
(33)優先権主張国・地域又は機関	米国(US)		ランクリン ストリート 1 1 1 1, 1 2
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA	(74)代理人	番 フロア
	,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(		100114188
	AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A	(74)代理人	弁理士 小野 誠
	T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR		100119253
	,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,	(74)代理人	弁理士 金山 賢教
	最終頁に続く		100124855
			最終頁に続く

(54)【発明の名称】 神経活動からのリアルタイム単語及び発話復号のための方法及びデバイス

(57)【要約】

個人のコミュニケーションを支援するための方法、デバイス、及びシステムが提供される。具体的には、個人の神経活動から直接単語及び文を復号するための方法、デバイス、及びシステムが提供される。発話処理に關与する脳の領域からの皮質活動が、個人が単語の発話又はスペルアウトを試行している間に記録される。深層学習計算モデルが、記録された脳活動からの単語を検出及び分類するために使用される。脳活動からの発話の復号は、特定の単語系列がどのように出現する可能性があるかを予測する言語モデルの使用によって補助される。加えて、神経活動からの試行非発話運動の復号を使用して、コミュニケーションを更に支援することができる。このニューロテクノロジーは、発話する能力を失った患者のコミュニケーションを回復させるために使用することができ、自律性及び生活の質を向上させる可能性がある。

【選択図】 図 1

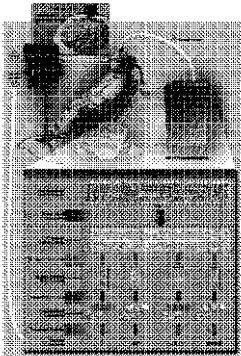


FIG. 1

【特許請求の範囲】

【請求項 1】

対象のコミュニケーションを支援する方法であって、前記方法が、

前記対象による試行発話と関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、前記対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、

コンピューティングデバイスと通信するインターフェースを、前記対象の頭部上のロケーションに位置付けることであって、前記インターフェースが、前記神経記録デバイスに接続されている、位置付けることと、

前記神経記録デバイスを使用して、前記対象による試行発話と関連付けられる前記脳電気信号データを記録することであって、前記インターフェースが、前記神経記録デバイスから前記脳電気信号データを受信し、前記脳電気信号データを前記コンピューティングデバイスのプロセッサに送信する、記録することと、

前記プロセッサを使用して、前記記録された脳電気信号データから単語、句、又は文を復号することと、を含む、方法。

10

【請求項 2】

前記対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために前記コミュニケーションが困難である、請求項 1 に記載の方法。

【請求項 3】

前記対象が、麻痺状態にある、請求項 1 又は 2 に記載の方法。

20

【請求項 4】

前記神経記録デバイスの前記ロケーションが、腹側感覚運動皮質内にある、請求項 1 ～ 3 のいずれか一項に記載の方法。

【請求項 5】

前記電極が、前記感覚運動皮質領域の表面上又は前記感覚運動皮質領域内に位置付けられる、請求項 1 ～ 4 のいずれか一項に記載の方法。

【請求項 6】

前記電極が、硬膜下腔内の前記脳の前記感覚運動皮質領域の表面上に位置付けられる、請求項 5 に記載の方法。

【請求項 7】

前記神経記録デバイスが、脳侵入電極アレイを含む、請求項 1 ～ 6 のいずれか一項に記載の方法。

30

【請求項 8】

前記神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、請求項 1 ～ 7 のいずれか一項に記載の方法。

【請求項 9】

前記電極が、深部電極又は表面電極である、請求項 1 ～ 8 のいずれか一項に記載の方法。

【請求項 10】

前記電気信号データが、高ガンマ周波数成分特徴を含む、請求項 1 ～ 9 のいずれか一項に記載の方法。

40

【請求項 11】

前記電気信号データが、70 Hz ～ 150 Hz の範囲内の神経振動を含む、請求項 10 に記載の方法。

【請求項 12】

前記脳電気信号データを前記記録することが、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの前記脳電気信号データを記録することを含む、請求項 1 ～ 11 のいずれか一項に記載の方法。

【請求項 13】

50

前記対象による前記試行発話と関連付けられる前記脳電気信号を記録するために前記電極を位置付けるための最適なロケーションを識別するために、前記対象の前記脳をマッピングすることを更に含む、請求項 1 ～ 1 2 のいずれか一項に記載の方法。

【請求項 1 4】

前記インターフェースが、前記対象の頭蓋に取り付けられた経皮台座コネクタを備える、請求項 1 ～ 1 3 のいずれか一項に記載の方法。

【請求項 1 5】

前記インターフェースが、前記経皮台座コネクタに接続されたヘッドステージを更に備える、請求項 1 4 に記載の方法。

【請求項 1 6】

前記プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、請求項 1 ～ 1 5 のいずれか一項に記載の方法。

【請求項 1 7】

前記ハンドヘルドデバイスが、携帯電話又はタブレットである、請求項 1 6 に記載の方法。

【請求項 1 8】

前記プロセッサが、試行単語生成と関連付けられる前記記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、発話検出、単語分類、及び文復号を自動化するようにプログラムされている、請求項 1 ～ 1 7 のいずれか一項に記載の方法。

【請求項 1 9】

前記プロセッサが、発話検出、単語分類、及び文復号のために機械学習アルゴリズムを使用するようにプログラムされている、請求項 1 8 に記載の方法。

【請求項 2 0】

人工ニューラルネットワーク（ANN）モデルが、前記発話検出及び前記単語分類に使用され、隠れマルコフモデル（HMM）、ビタビ復号モデル、又は自然言語処理技法が、前記文復号に使用される、請求項 1 9 に記載の方法。

【請求項 2 1】

前記プロセッサが、前記対象による前記試行発話中の単語生成の開始及び終了の検出を自動化するようにプログラムされている、請求項 1 ～ 2 0 のいずれか一項に記載の方法。

【請求項 2 2】

準備、発話、及び休止のための発話イベントラベルを前記脳電気信号データの前記記録中の時点に割り当てることを更に含む、請求項 2 1 に記載の方法。

【請求項 2 3】

前記プロセッサが、前記単語分類の検出された開始の周りのタイムウィンドウ内に前記記録された脳電気信号データを使用するようにプログラムされている、請求項 2 1 又は 2 2 に記載の方法。

【請求項 2 4】

前記対象が、前記試行発話について指定の単語セットに制限される、請求項 1 ～ 2 3 のいずれか一項に記載の方法。

【請求項 2 5】

前記プロセッサが、前記単語セットの単語が、前記対象が前記試行発話中に生成しようとした、意図された単語である確率を算出するようにプログラムされている、請求項 2 4 に記載の方法。

【請求項 2 6】

前記プロセッサが、前記単語セットの全ての単語について、前記単語セットの単語が、前記対象が前記試行発話中に生成しようとした、意図された単語である前記確率を算出するようにプログラムされている、請求項 2 5 に記載の方法。

【請求項 2 7】

前記単語セットが、am（～である）、are（～である）、bad（悪い）、bring（運ぶ）、clean（きれい、きれいにする）、closer（より近い）、co

10

20

30

40

50

m f o r t a b l e (快適)、c o m i n g (来ている)、c o m p u t e r (コンピュータ)、d o (行う)、f a i t h (信仰)、f a m i l y (家族)、f e e l (感じる)、g l a s s e s (眼鏡)、g o i n g (行っている)、g o o d (良い)、g o o d b y e (さようなら)、h a v e (有する)、h e l l o (こんにちは)、h e l p (手助け、手伝って)、h e r e (ここ)、h o p e (願う)、h o w (どのように)、h u n g r y (空腹)、I (私)、i s (〜である)、i t (それ)、l i k e (好き)、m u s i c (音楽)、m y (私の)、n e e d (必要とする)、n o (いいえ)、n o t (〜でない)、n u r s e (看護師)、o k a y (オーケー)、o u t s i d e (外側)、p l e a s e (どうか)、r i g h t (すぐ、正しい)、s u c c e s s (成功)、t e l l (伝える)、t h a t (あれ)、t h e y (それら)、t h i r s t y (喉が渴いている)、t i r e d (疲れている)、u p (上)、v e r y (とても)、w h a t (何)、w h e r e (どこ)、y e s (はい)、及び y o u (あなた)を含む、請求項 24～26 のいずれか一項に記載の方法。

10

【請求項 28】

前記対象が、文を作成するために、制限なしに前記単語セットの前記単語を使用し得る、請求項 1～27 のいずれか一項に記載の方法。

【請求項 29】

前記プロセッサが、単語系列が、前記対象が前記試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされている、請求項 28 に記載の方法。

【請求項 30】

前記プロセッサが、予測単語系列確率を決定することによって前記復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされている、請求項 1～29 のいずれか一項に記載の方法。

20

【請求項 31】

前記言語モデルに従って、より頻繁に出現する単語が、それほど頻繁に出現しない単語よりも大きい重みを割り当てられる、請求項 30 に記載の方法。

【請求項 32】

前記プロセッサが、前記試行発話と関連付けられる前記脳電気信号データ、前記機械学習アルゴリズムを使用した単語分類モデルからの前記予測単語確率、及び前記言語モデルを使用した前記単語系列確率を所与として、前記対象の前記意図された発話内の最も可能性の高い前記単語系列を決定するためにビタビ復号モデルを使用するようにプログラムされている、請求項 30 又は 31 に記載の方法。

30

【請求項 33】

前記対象の試行非発話運動と関連付けられる脳電気信号データを記録することであって、前記対象が、前記試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、前記試行非発話運動を実施する、記録することと、

前記試行非発話運動と関連付けられる記録された前記脳電気信号データ内の電気信号のパターンを識別し、前記対象が前記非発話運動を試行した確率を算出する非発話運動分類モデルを使用して前記脳電気信号データを分析することと、を更に含む、請求項 1～32 のいずれか一項に記載の方法。

40

【請求項 34】

前記試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、請求項 33 に記載の方法。

【請求項 35】

前記手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、請求項 34 に記載の方法。

【請求項 36】

前記プロセッサが、前記試行非発話運動と関連付けられる前記記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、前記対象による前記試行発話の終了をシグナリングする前記対象の試行非発話運動の検出を自動化するように更にプログ

50

ラムされている、請求項 33～35 のいずれか一項に記載の方法。

【請求項 37】

前記プロセッサが、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てるように更にプログラムされている、請求項 36 に記載の方法。

【請求項 38】

前記方法が、前記復号の正確度を評価することを更に含む、請求項 1～37 のいずれか一項に記載の方法。

【請求項 39】

対象による試行発話と関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法であって、前記コンピュータが、

a) 前記対象による前記試行発話と関連付けられる前記記録された脳電気信号データを受信するステップと、

b) 前記脳電気信号データの記録中の任意の時点で試行発話が発生している確率を算出し、前記対象による前記試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、前記記録された脳電気信号データを分析するステップと、

c) 前記対象による試行単語生成と関連付けられる前記記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、前記脳電気信号データを分析するステップと、

d) 前記単語分類モデルからの算出された前記単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した前記文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、前記単語分類モデル及び前記言語モデルを使用して決定された前記予測単語確率に基づいて、前記文内の最も可能性の高い前記単語系列を決定するステップと、

e) 前記記録された脳電気信号データから復号された前記文を表示するステップと、を含むステップを実施する、コンピュータ実装方法。

【請求項 40】

機械学習アルゴリズムが、発話検出、単語分類、及び文復号に使用される、請求項 39 に記載のコンピュータ実装方法。

【請求項 41】

人工ニューラルネットワーク (ANN) モデルが、前記発話検出及び前記単語分類に使用され、隠れマルコフモデル (HMM)、ビタビ復号モデル、又は自然言語処理技法が、前記文復号に使用される、請求項 40 に記載のコンピュータ実装方法。

【請求項 42】

前記対象が、前記試行発話について指定の単語セットに制限される、請求項 39～41 のいずれか一項に記載のコンピュータ実装方法。

【請求項 43】

前記単語セットの全ての単語について、前記単語セットの単語が、前記対象が前記試行発話中に生成しようとした、意図された単語である確率を算出することと、前記対象が試行発話中に生成しようとした、前記意図された単語である確率が最も高い前記単語セットの前記単語を選択することと、を更に含む、請求項 42 に記載のコンピュータ実装方法。

【請求項 44】

前記対象が、文を作成するために、制限なしに前記単語セットの前記単語を使用し得るか、又は、前記試行発話について指定の文セットに制限される、請求項 39～43 のいずれか一項に記載のコンピュータ実装方法。

【請求項 45】

単語系列が、前記対象が前記試行発話中に生成しようとした、意図された文である確率を算出することを更に含む、請求項 39～44 のいずれか一項に記載のコンピュータ実装方法。

【請求項 46】

最も可能性の高い文及び1つ以上のより可能性の低い文を維持することと、各単語の復号後に、単語系列が、前記試行発話中に前記対象が生成しようとした、意図された文である前記確率を再算出することと、を更に含む、請求項45に記載のコンピュータ実装方法。

【請求項47】

前記最も可能性の高い文及び前記1つ以上のより可能性の低い文が、前記試行発話のために前記対象によって使用される前記単語セットからの単語のみから構成される、請求項46に記載のコンピュータ実装方法。

【請求項48】

準備、発話、及び休止のための発話イベントラベルを前記脳電気信号データの前記記録中の時点に割り当てることを更に含む、請求項39～47のいずれか一項に記載のコンピュータ実装方法。

【請求項49】

前記単語分類の検出された開始の周りのタイムウィンドウ内で前記記録された脳電気信号データのみが使用される、請求項48に記載のコンピュータ実装方法。

【請求項50】

前記言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みが割り当てられる、請求項39～49のいずれか一項に記載のコンピュータ実装方法。

【請求項51】

前記対象による試行単語生成と関連付けられる前記記録された脳電気信号データ内の前記電気信号のパターンに関する情報を含む前記対象のユーザプロファイルを格納することを更に含む、請求項39～50のいずれか一項に記載のコンピュータ実装方法。

【請求項52】

前記対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、前記対象が、前記試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、前記試行非発話運動を実施する、受信することと、

前記試行非発話運動と関連付けられる前記記録された脳電気信号データ内の電気信号のパターンを識別し、前記対象が前記非発話運動を試行した確率を算出する分類モデルを使用して前記脳電気信号データを分析することと、を更に含む、請求項39～51のいずれか一項に記載のコンピュータ実装方法。

【請求項53】

前記試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、請求項52に記載のコンピュータ実装方法。

【請求項54】

前記手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、請求項53に記載のコンピュータ実装方法。

【請求項55】

前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てることを更に含む、請求項52～54のいずれか一項に記載のコンピュータ実装方法。

【請求項56】

コンピュータ内のプロセッサによって実行されるとき、前記プロセッサに、請求項39～55のいずれか一項に記載の方法を実行させるプログラム命令を含む、非一時的コンピュータ可読媒体。

【請求項57】

請求項56に記載の非一時的コンピュータ可読媒体と、対象による試行発話と関連付けられる脳電気信号データを復号するための命令と、を備える、キット。

【請求項58】

対象のコミュニケーションを支援するためのシステムであって、前記システムが、

10

20

30

40

50

前記対象による試行発話又は試行非発話運動と関連付けられる脳電気信号データを記録するために、前記対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、

請求項 39～55 のいずれか一項に記載のコンピュータ実装方法に従って、前記記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、

コンピューティングデバイスと通信するインターフェースであって、前記インターフェースが、前記対象の頭部上のロケーションに位置するように適合されており、前記インターフェースが、前記神経記録デバイスから前記脳電気信号データを受信し、前記脳電気信号データを前記プロセッサに送信する、インターフェースと、

前記記録された脳電気信号データから復号された前記文を表示するためのディスプレイコンポーネントと、を備える、システム。

10

【請求項 59】

前記対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために前記コミュニケーションが困難である、請求項 58 に記載のシステム。

【請求項 60】

前記神経記録デバイスの前記ロケーションが、腹側感覚運動皮質内にある、請求項 58 又は 59 に記載のシステム。

【請求項 61】

前記電極が、前記感覚運動皮質領域の表面上又は前記感覚運動皮質領域内に位置するように適合されている、請求項 58～60 のいずれか一項に記載のシステム。

20

【請求項 62】

前記電極が、硬膜下腔内の前記脳の前記感覚運動皮質領域の表面上に位置するように適合されている、請求項 61 に記載のシステム。

【請求項 63】

前記神経記録デバイスが、脳侵入電極アレイを含む、請求項 58～62 のいずれか一項に記載のシステム。

【請求項 64】

前記神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、請求項 58～63 のいずれか一項に記載のシステム。

【請求項 65】

前記電極が、深部電極又は表面電極である、請求項 58～64 のいずれか一項に記載のシステム。

30

【請求項 66】

前記電気信号データが、高ガンマ周波数成分特徴を含む、請求項 58～65 のいずれか一項に記載のシステム。

【請求項 67】

前記電気信号データが、70 Hz～150 Hz の範囲内の神経振動を含む、請求項 66 に記載のシステム。

【請求項 68】

前記インターフェースが、前記対象の頭蓋に取り付けられた経皮台座コネクタを備える、請求項 58～67 のいずれか一項に記載のシステム。

40

【請求項 69】

前記インターフェースが、前記経皮台座コネクタに接続可能であるヘッドステージを更に備える、請求項 68 に記載のシステム。

【請求項 70】

前記プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、請求項 58～69 のいずれか一項に記載のシステム。

【請求項 71】

前記ハンドヘルドデバイスが、携帯電話又はタブレットである、請求項 70 に記載のシステム。

50

【請求項 7 2】

機械学習アルゴリズムが、発話検出、単語分類、及び文復号に使用される、請求項 5 8 ~ 7 1 のいずれか一項に記載のシステム。

【請求項 7 3】

人工ニューラルネットワーク (ANN) モデルが、前記発話検出及び前記単語分類に使用され、隠れマルコフモデル (HMM)、ビタビ復号モデル、又は自然言語処理技法が、前記文復号に使用される、請求項 7 2 に記載のシステム。

【請求項 7 4】

前記プロセッサが、準備、発話、及び休止のための発話イベントラベルを前記脳電気信号データの前記記録中の時点に割り当てるように更にプログラムされている、請求項 5 8 ~ 7 3 のいずれか一項に記載のシステム。

【請求項 7 5】

前記プロセッサが、前記単語分類の検出された開始の周りのタイムウィンドウ内で前記記録された脳電気信号データを使用するように更にプログラムされている、請求項 7 4 に記載のシステム。

【請求項 7 6】

前記対象が、前記試行発話について指定の単語セットに制限される、請求項 5 8 ~ 7 5 のいずれか一項に記載のシステム。

【請求項 7 7】

前記プロセッサが、前記単語セットの全ての単語について、前記単語セットの単語が、前記対象が前記試行発話中に生成しようとした、意図された単語である確率を算出し、前記対象が前記試行発話中に生成しようとした、前記意図された単語である確率が最も高い前記単語セットの前記単語を選択するように更にプログラムされている、請求項 7 6 に記載のシステム。

【請求項 7 8】

前記単語セットが、am、are、bad、bring、clean、closer、comfortable、coming、computer、do、faith、family、feel、glasses、going、good、goodbye、have、hello、help、here、hope、how、hungry、I、is、it、like、music、my、need、no、not、nurse、okay、outside、please、right、success、tell、that、they、thirsty、tired、up、very、what、where、yes、及びyouを含む、請求項 7 6 又は 7 7 に記載のシステム。

【請求項 7 9】

前記対象が、選択された前記単語セットの任意の選ばれた単語系列を使用し得る、請求項 7 6 ~ 7 8 のいずれか一項に記載のシステム。

【請求項 8 0】

前記プロセッサが、単語系列が、前記対象が前記試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされている、請求項 7 9 に記載のシステム。

【請求項 8 1】

前記プロセッサが、最も可能性の高い文及び 1 つ以上のより可能性の低い文を維持し、各単語の復号後に、単語系列が、前記試行発話中に前記対象が生成しようとした、意図された文である前記確率を再算出するようにプログラムされている、請求項 8 0 に記載のシステム。

【請求項 8 2】

前記最も可能性の高い文及び前記 1 つ以上のより可能性の低い文が、前記試行発話のために前記対象によって使用される前記単語セットからの単語のみから構成される、請求項 8 1 に記載のシステム。

【請求項 8 3】

10

20

30

40

50

前記プロセッサが、前記試行非発話運動と関連付けられる前記記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、前記対象による前記試行発話の開始又は終了をシグナリングする前記対象の試行非発話運動の検出を自動化するように更にプログラムされている、請求項 58～82 のいずれか一項に記載のシステム。

【請求項 84】

前記プロセッサが、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てるように更にプログラムされている、請求項 83 に記載のシステム。

【請求項 85】

請求項 58～84 のいずれか一項に記載のシステムと、対象による試行発話と関連付けられる脳電気信号データを記録し、復号するために、前記システムを使用するための命令と、を備える、キット。

【請求項 86】

対象のコミュニケーションを支援する方法であって、前記方法が、
前記対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、前記対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、

コンピューティングデバイスと通信するインターフェースを、前記対象の頭部上のロケーションに位置付けることであって、前記インターフェースが、前記神経記録デバイスに接続されている、位置付けることと、

前記神経記録デバイスを使用して、前記対象による前記試行スペリングと関連付けられる前記脳電気信号データを記録することであって、前記インターフェースが、前記神経記録デバイスから前記脳電気信号データを受信し、前記脳電気信号データを前記コンピューティングデバイスのプロセッサに送信する、記録することと、

前記プロセッサを使用して、記録された前記脳電気信号データから前記意図された文のスペリングされた前記単語を復号することと、を含む、方法。

【請求項 87】

前記対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために前記コミュニケーションが困難である、請求項 86 に記載の方法。

【請求項 88】

前記対象が、麻痺状態にある、請求項 86 又は 87 に記載の方法。

【請求項 89】

前記神経記録デバイスの前記ロケーションが、腹側感覚運動皮質内にある、請求項 86～88 のいずれか一項に記載の方法。

【請求項 90】

前記電極が、前記感覚運動皮質領域の表面上又は前記感覚運動皮質領域内に位置付けられる、請求項 86～89 のいずれか一項に記載の方法。

【請求項 91】

前記電極が、硬膜下腔内の前記脳の前記感覚運動皮質領域の表面上に位置付けられる、請求項 90 に記載の方法。

【請求項 92】

前記神経記録デバイスが、脳侵入電極アレイを含む、請求項 86～91 のいずれか一項に記載の方法。

【請求項 93】

前記神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、請求項 86～92 のいずれか一項に記載の方法。

【請求項 94】

前記電極が、深部電極又は表面電極である、請求項 86～93 のいずれか一項に記載の方法。

【請求項 95】

前記電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、請求項 86～94 のいずれか一項に記載の方法。

【請求項 96】

前記電気信号データが、70 Hz～150 Hz の高ガンマ周波数範囲、及び 0.3 Hz～100 Hz の低周波数範囲内の神経振動を含む、請求項 95 に記載の方法。

【請求項 97】

前記脳電気信号データの前記記録が、中心前回領域、中心後回領域、後中前頭回領域、後上前頭回領域、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの前記脳電気信号データを記録することを含む、請求項 86～96 のいずれか一項に記載の方法。

10

【請求項 98】

前記対象による単語の前記試行スペリング又は前記試行非発話運動試行発話と関連付けられる前記脳電気信号を記録するために前記電極を位置付けるための最適なロケーションを識別するために、前記対象の前記脳をマッピングすることを更に含む、請求項 86～97 のいずれか一項に記載の方法。

【請求項 99】

前記インターフェースが、前記対象の頭蓋に取り付けられた経皮台座コネクタを備える、請求項 86～98 のいずれか一項に記載の方法。

【請求項 100】

前記インターフェースが、前記経皮台座コネクタに接続されたヘッドステージを更に備える、請求項 99 に記載の方法。

20

【請求項 101】

前記プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、請求項 86～100 のいずれか一項に記載の方法。

【請求項 102】

前記ハンドヘルドデバイスが、携帯電話又はタブレットである、請求項 101 に記載の方法。

【請求項 103】

前記プロセッサが、前記対象による単語の前記試行スペリングと関連付けられる前記記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、前記試行スペリング、文字分類、単語分類、及び文復号の検出を自動化するようにプログラムされている、請求項 86～102 のいずれか一項に記載の方法。

30

【請求項 104】

前記プロセッサが、前記発話検出、文字分類、単語分類、及び文復号のために機械学習アルゴリズムを使用するようにプログラムされている、請求項 103 に記載の方法。

【請求項 105】

前記プロセッサが、前記対象による単語の試行スペリングと関連付けられる神経活動から復号される文字系列からの単語分類を、前記対象によって使用される言語の語彙内の単語のみに制約するように更にプログラムされる、請求項 104 に記載の方法。

【請求項 106】

40

前記プロセッサが、前記準備、試行スペリング、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされている、請求項 86～105 のいずれか一項に記載の方法。

【請求項 107】

前記プロセッサが、前記対象による文字の試行スペリングの検出された前記開始の周りのタイムウィンドウ内で前記記録された脳電気信号データを使用するようにプログラムされている、請求項 106 に記載の方法。

【請求項 108】

前記対象が前記意図された文の前記単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを前記対象に提供することを更に含む、請求項 86～107

50

のいずれか一項に記載の方法。

【請求項 1 0 9】

前記一連のゴーキューが、ディスプレイ上に視覚的に提供される、請求項 1 0 8 に記載の方法。

【請求項 1 1 0】

各ゴーキューに先行して、前記ゴーキューの提示へのカウントダウンが行われ、次にスペリングされる文字の前記カウントダウンが、前記ディスプレイ上に視覚的に提供され、各ゴーキューの後に自動的に開始される、請求項 1 0 9 に記載の方法。

【請求項 1 1 1】

前記一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される、請求項 1 0 8 ~ 1 1 0 のいずれか一項に記載の方法。

【請求項 1 1 2】

前記対象が、各ゴーキュー間の前記設定時間間隔を制御することができる、請求項 1 1 1 に記載の方法。

【請求項 1 1 3】

前記プロセッサが、前記ゴーキューに後続するタイムウィンドウ内で前記記録された脳電気信号データを使用するようにプログラムされている、請求項 1 0 8 ~ 1 1 2 のいずれか一項に記載の方法。

【請求項 1 1 4】

前記プロセッサが、復号された文字の系列からの復号された単語の系列が、前記対象による意図された文の単語の文字の前記試行スペリング中に前記対象が生成しようとした、意図された文である確率を算出するようにプログラムされている、請求項 8 6 ~ 1 1 3 のいずれか一項に記載の方法。

【請求項 1 1 5】

前記プロセッサが、予測単語系列確率を決定することによって前記復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされている、請求項 8 6 ~ 1 1 4 のいずれか一項に記載の方法。

【請求項 1 1 6】

前記言語モデルに従って、より頻繁に出現する単語が、それほど頻繁に出現しない単語よりも大きい重みを割り当てられる、請求項 1 1 5 に記載の方法。

【請求項 1 1 7】

前記プロセッサが、予測文字確率の系列を使用して、潜在的な文候補を計算し、前記文候補内の予測される単語間の文字系列に空白を自動的に挿入するように更にプログラムされている、請求項 8 6 ~ 1 1 6 のいずれか一項に記載の方法。

【請求項 1 1 8】

前記対象の試行非発話運動と関連付けられる脳電気信号データを記録することであって、前記対象が、前記意図された文の単語の前記試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、前記試行非発話運動を実施する、記録することと、

前記試行非発話運動と関連付けられる記録された前記脳電気信号データ内の電気信号のパターンを識別し、前記対象が前記非発話運動を試行した確率を算出する分類モデルを使用して前記脳電気信号データを分析することと、を更に含む、請求項 8 6 ~ 1 1 7 のいずれか一項に記載の方法。

【請求項 1 1 9】

前記試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、請求項 1 1 8 に記載の方法。

【請求項 1 2 0】

前記手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、請求項 1 1 9 に記載の方法。

【請求項 1 2 1】

10

20

30

40

50

前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てることを更に含む、請求項 1 1 8 ～ 1 2 0 のいずれか一項に記載の方法。

【請求項 1 2 2】

前記復号の正確度を評価することを更に含む、請求項 8 6 ～ 1 2 1 のいずれか一項に記載の方法。

【請求項 1 2 3】

前記神経記録デバイスを使用して、前記対象による試行発話と関連付けられる脳電気信号データを記録することであって、前記インターフェースが、前記神経記録デバイスから前記脳電気信号データを受信し、前記脳電気信号データを前記コンピューティングデバイスの前記プロセッサに送信する、記録することと、

前記プロセッサを使用して、前記対象による試行発話と関連付けられる記録された前記脳電気信号データから単語、句、又は文を復号することと、を更に含む、請求項 8 6 ～ 1 2 2 のいずれか一項に記載の方法。

【請求項 1 2 4】

対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法であって、前記コンピュータが、

a) 前記対象による意図された文の単語の文字の前記試行スペリングと関連付けられる前記記録された脳電気信号データを受信するステップと、

b) 前記電気信号データの前記記録中の任意の時点で試行スペリングが発生している確率を算出し、前記対象による前記試行スペリング中の文字生成の開始及び終了を検出するために、発話検出モデルを使用して、前記記録された脳電気信号データを分析するステップと、

c) 前記対象による試行文字生成と関連付けられる前記記録された脳電気信号データ内の電気信号のパターンを識別し、予測文字確率の系列を算出する文字分類モデルを使用して、前記脳電気信号データを分析するステップと、

d) 前記予測文字確率の系列に基づいて潜在的な文候補を計算し、前記文候補内の予測単語間の前記文字系列に空白を自動的に挿入するステップであって、前記文字系列内の復号される単語は、前記対象によって使用される言語の語彙内の単語のみに制約される、計算し、挿入するステップと、

e) 予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用して前記潜在的な文候補を分析し、文内の最も可能性の高い前記単語系列を決定するステップと、

f) 前記記録された脳電気信号データから復号された前記文を表示するステップと、を含むステップを実施する、コンピュータ実装方法。

【請求項 1 2 5】

前記記録された脳電気信号データが、前記対象による文字の試行スペリングの検出された前記開始の周りのタイムウィンドウ内でのみ使用される、請求項 1 2 4 に記載のコンピュータ実装方法。

【請求項 1 2 6】

前記対象が前記意図された文の前記単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを前記対象に表示することを更に含む、請求項 1 2 4 又は 1 2 5 に記載のコンピュータ実装方法。

【請求項 1 2 7】

各ゴーキューに先行して、前記ゴーキューの提示へのカウントダウンが表示され、次にスペリングされる文字の前記カウントダウンが、各ゴーキューの後に自動的に開始される、請求項 1 2 6 に記載のコンピュータ実装方法。

【請求項 1 2 8】

前記一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される、請求項 1 2 6 又は 1 2 7 に記載のコンピュータ実装方法。

10

20

30

40

50

【請求項 1 2 9】

前記対象が、各ゴーキュー間の前記設定時間間隔を制御することができる、請求項 1 2 8 に記載のコンピュータ実装方法。

【請求項 1 3 0】

前記ゴーキューに後続するタイムウィンドウ内で前記記録された脳電気信号データが文字分類に使用される、請求項 1 2 2 ～ 1 2 7 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 1】

前記対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、前記対象が、前記意図された文の単語の前記試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、前記試行非発話運動を実施する、受信することと、

10

前記試行非発話運動と関連付けられる前記記録された脳電気信号データ内の電気信号のパターンを識別し、前記対象が前記非発話運動を試行した確率を算出する分類モデルを使用して前記脳電気信号データを分析することと、を更に含む、請求項 1 2 4 ～ 1 3 0 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 2】

前記試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、請求項 1 3 1 に記載の方法。

【請求項 1 3 3】

20

前記手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、請求項 1 3 2 に記載の方法。

【請求項 1 3 4】

機械学習アルゴリズムが、試行スペリング若しくは試行非発話運動の検出又は文字分類に使用される、請求項 1 2 4 ～ 1 3 3 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 5】

前記言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みを割り当てることを更に含む、請求項 1 2 4 ～ 1 3 4 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 6】

30

前記対象による試行スペリング中の文字生成と関連付けられる前記記録された脳電気信号データ内の前記電気信号のパターンに関する情報を含む前記対象のユーザプロファイルを格納することを更に含む、請求項 1 2 4 ～ 1 3 5 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 7】

前記電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、請求項 1 2 4 ～ 1 3 6 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 3 8】

前記電気信号データが、70 Hz ～ 150 Hz の高ガンマ周波数範囲、及び 0.3 Hz ～ 100 Hz の低周波数範囲内の神経振動を含む、請求項 1 3 7 に記載のコンピュータ実装方法。

40

【請求項 1 3 9】

前記復号の正確度を評価することを更に含む、請求項 1 2 4 ～ 1 3 8 のいずれか一項に記載のコンピュータ実装方法。

【請求項 1 4 0】

前記対象による試行発話と関連付けられる記録された脳電気信号データから文を復号することを更に含み、前記コンピュータが、

a) 前記対象による前記試行発話と関連付けられる前記記録された脳電気信号データを受信するステップと、

b) 任意の時点で試行発話が発生している確率を算出し、前記対象による前記試行発話

50

中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、前記記録された脳電気信号データを分析するステップと、

c) 前記対象による試行単語生成と関連付けられる前記記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、前記脳電気信号データを分析するステップと、

d) 前記単語分類モデルからの算出された前記単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した前記文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、前記単語分類モデル及び前記言語モデルを使用して決定された前記予測単語確率に基づいて、前記文内の最も可能性の高い前記単語系列を決定するステップと、

e) 前記記録された脳電気信号データから復号された前記文を表示するステップと、を含むステップを実施する、請求項 1 2 4 ~ 1 3 9 のいずれか一項にコンピュータ実装方法

10

【請求項 1 4 1】

機械学習アルゴリズムが、発話検出及び単語分類、並びに文復号に使用される、請求項 1 4 0 に記載のコンピュータ実装方法。

【請求項 1 4 2】

人工ニューラルネットワーク (ANN) モデルが、前記発話検出及び前記単語分類に使用され、隠れマルコフモデル (HMM)、ピタビ復号モデル、又は自然言語処理技法が、前記文復号に使用される、請求項 1 4 1 に記載のコンピュータ実装方法。

20

【請求項 1 4 3】

コンピュータ内のプロセッサによって実行されるとき、前記プロセッサに、請求項 1 2 4 ~ 1 4 2 のいずれか一項に記載の方法を実行させるプログラム命令を含む、非一時的コンピュータ可読媒体。

【請求項 1 4 4】

請求項 1 4 3 に記載の非一時的コンピュータ可読媒体と、対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを復号するための命令と、を備える、キット。

【請求項 1 4 5】

対象のコミュニケーションを支援するためのシステムであって、前記システムが、

前記対象による試行発話、意図された文の単語の文字の試行スペリング、若しくは試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録するために、前記対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、

30

請求項 1 2 4 ~ 1 4 2 のいずれか一項に記載のコンピュータ実装方法に従って、前記記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、

コンピューティングデバイスと通信するインターフェースであって、前記インターフェースが、前記対象の頭部上のロケーションに位置するように適合されており、前記インターフェースが、前記神経記録デバイスから前記脳電気信号データを受信し、前記脳電気信号データを前記プロセッサに送信する、インターフェースと、

40

前記記録された脳電気信号データから復号された前記文を表示するためのディスプレイコンポーネントと、を備える、システム。

【請求項 1 4 6】

前記対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために前記コミュニケーションが困難である、請求項 1 4 5 に記載のシステム。

【請求項 1 4 7】

前記神経記録デバイスの前記ロケーションが、腹側感覚運動皮質内にある、請求項 1 4 5 又は 1 4 6 に記載のシステム。

【請求項 1 4 8】

前記電極が、前記感覚運動皮質領域の表面上又は前記感覚運動皮質領域内に位置するよ

50

うに適合されている、請求項 1 4 5 ～ 1 4 7 のいずれか一項に記載のシステム。

【請求項 1 4 9】

前記電極が、硬膜下腔内の前記脳の前記感覚運動皮質領域の表面上に位置するように適合されている、請求項 1 4 8 に記載のシステム。

【請求項 1 5 0】

前記神経記録デバイスが、脳侵入電極アレイを含む、請求項 1 4 5 ～ 1 4 9 のいずれか一項に記載のシステム。

【請求項 1 5 1】

前記神経記録デバイスが、皮質脳波記録（E C o G）電極アレイを含む、請求項 1 4 5 ～ 1 5 0 のいずれか一項に記載のシステム。

【請求項 1 5 2】

前記電極が、深部電極又は表面電極である、請求項 1 4 5 ～ 1 5 1 のいずれか一項に記載のシステム。

【請求項 1 5 3】

前記電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、請求項 1 4 5 ～ 1 5 2 のいずれか一項に記載のシステム。

【請求項 1 5 4】

前記電気信号データが、7 0 H z ～ 1 5 0 H z の高ガンマ周波数範囲、及び 0 . 3 H z ～ 1 0 0 H z の低周波数範囲内の神経振動を含む、請求項 1 5 3 に記載のシステム。

【請求項 1 5 5】

前記インターフェースが、前記対象の頭蓋に取り付けられた経皮台座コネクタを備える、請求項 1 4 5 ～ 1 5 4 のいずれか一項に記載のシステム。

【請求項 1 5 6】

前記インターフェースが、前記経皮台座コネクタに接続可能であるヘッドステージを更に備える、請求項 1 5 5 に記載のシステム。

【請求項 1 5 7】

前記プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、請求項 1 4 5 ～ 1 5 6 のいずれか一項に記載のシステム。

【請求項 1 5 8】

前記ハンドヘルドデバイスが、携帯電話又はタブレットである、請求項 1 5 7 に記載のシステム。

【請求項 1 5 9】

請求項 1 4 5 ～ 1 5 8 のいずれか一項に記載のシステムと、対象による試行発話、単語の試行スペリング、若しくは試行試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録し、復号するために、前記システムを使用するための命令と、を備える、キット。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願の相互参照

本出願は、2 0 2 1 年 5 月 2 6 日に提出された仮出願第 6 3 / 1 9 3 , 3 5 1 号の米国特許法第 1 1 9 条（e）に基づく利益を主張し、当該仮出願は、参照によりその全体が本明細書に組み込まれる。

【0 0 0 2】

政府支援に関する陳述

本発明は、米国国立衛生研究所（N I H）によって付与された助成金番号 U 0 1 N S 0 9 8 9 7 1 - 0 1 の下で政府支援を受けて行われた。政府は、本発明に特定の権利を有する。

【0 0 0 3】

序論

10

20

30

40

50

構語障害とは、発話能力の喪失である。構語障害は、脳卒中、外傷性脳損傷、及び筋萎縮性側索硬化症を含む様々な状態に起因し得る（Beukelman et al. (2007) *Augmentative and Alternative Communication* 23 (3) : 230-242）。重度の運動障害を有する麻痺した個人の場合、構語障害は、家族、友人、及び介護者とのコミュニケーションを妨げ、自己申告の生活の質を低下させる（Felgoise et al. (2016) *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 17 (3-4) : 179-183）。発話する能力を失った麻痺した患者のコミュニケーションを回復させるために設計されたニューロテクノロジーは、自律性及び生活の質を向上させる可能性がある。しかしながら、既存のアプローチのほとんどは、自然な発話に比べて遅く、退屈である。したがって、構語障害を有する患者にコミュニケーション能力を回復させるためのより良好な方法に対する必要性が依然として存在する。

10

【発明の概要】

【0004】

個人のコミュニケーションを支援するための方法、デバイス、及びシステムが提供される。具体的には、個人の神経活動から直接単語及び文を復号するための方法、デバイス、及びシステムが提供される。開示されている方法において、発話処理に関与する脳の領域からの皮質活動が、個人が単語の発語又はスペルアウトを試行している間に（単語又は綴られた文字が発声されていなくても）記録される。深層学習計算モデルが、記録された脳活動からの単語を検出及び分類するために使用される。脳活動からの発話の復号は、特定の単語系列がどのように出現する可能性があるかを予測する言語モデルの使用によって補助される。加えて、神経活動からの試行非発話運動の復号を使用して、コミュニケーションを更に支援することができる。本明細書に記載されているニューロテクノロジーは、発話する能力を失った患者のコミュニケーションを回復させるために使用することができ、自律性及び生活の質を向上させる可能性がある。

20

【0005】

一態様において、対象のコミュニケーションを支援する方法が提供され、本方法が、対象による試行発話と関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、コンピューティングデバイスと通信するインターフェースを、対象の頭部上のロケーションに位置付けることであって、インターフェースが、神経記録デバイスに接続されている、位置付けることと、神経記録デバイスを使用して、対象による試行発話と関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、記録することと、プロセッサを使用して、記録された脳電気信号データから単語、句、又は文を復号することと、を含む。

30

【0006】

特定の実施形態において、対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のためにコミュニケーションが困難である。いくつかの実施形態において、対象が、麻痺状態にある。

40

【0007】

特定の実施形態において、神経記録デバイスのロケーションが、腹側感覚運動皮質内にある。例えば、電極が、感覚運動皮質領域の表面上に、又は感覚運動皮質領域内に位置付けられ得る。いくつかの実施形態において、電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置付けられる。

【0008】

特定の実施形態において、本方法が、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの脳電気信号データを記録することを含む。

50

【0009】

特定の実施形態において、神経記録デバイスが、脳侵入電極アレイ又は皮質脳波記録（E C o G）電極アレイを含む。

【0010】

特定の実施形態において、電極が、深部電極又は表面電極である。

【0011】

特定の実施形態において、プロセッサによって使用される特徴が、電気信号データに含まれる高ガンマ周波数成分特徴である。いくつかの実施形態において、高ガンマ周波数電気信号データが、70Hz～150Hzの範囲内の神経振動を含み得る。

【0012】

特定の実施形態において、本方法が、対象による試行発話と関連付けられる脳電気信号を記録するために電極を位置付けるための最適なロケーションを識別するために、対象の脳をマッピングすることを更に含む。

【0013】

特定の実施形態において、インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える。いくつかの実施形態において、インターフェースが、経皮台座コネクタに接続された取り外し可能ヘッドステージを更に備える。

【0014】

特定の実施形態において、プロセッサが、コンピュータ又はハンドヘルドデバイス（例えば、携帯電話又はタブレット）によって提供される。

【0015】

特定の実施形態において、プロセッサが、対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、機械学習アルゴリズムを使用して、発話検出、単語分類、及び文復号を自動化するようにプログラムされる。いくつかの実施形態において、機械学習アルゴリズムが、発話検出のために人工ニューラルネットワーク（ANN）モデルを使用し、文復号のために、限定ではないが、隠れマルコフモデル（HMM）又はビタビ復号モデルなどの単語分類及び自然言語処理技法を使用する。

【0016】

特定の実施形態において、プロセッサが、対象による試行発話中の単語生成の開始及び終了の検出を自動化するようにプログラムされる。いくつかの実施形態において、本方法が、準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む。いくつかの実施形態において、プロセッサが、単語分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされる。

【0017】

特定の実施形態において、対象が、試行発話について指定の単語セットに制限される。

【0018】

特定の実施形態において、プロセッサが、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出し、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択するようにプログラムされる。

【0019】

特定の実施形態において、単語セットが、a m（～である）、a r e（～である）、b a d（悪い）、b r i n g（運ぶ）、c l e a n（きれい、きれいにする）、c l o s e r（より近い）、c o m f o r t a b l e（快適）、c o m i n g（来ている）、c o m p u t e r（コンピュータ）、d o（行う）、f a i t h（信仰）、f a m i l y（家族）、f e e l（感じる）、g l a s s e s（眼鏡）、g o i n g（行っている）、g o o

10

20

30

40

50

d (良い)、goodbye (さようなら)、have (有する)、hello (こんにちは)、help (手助け、手伝って)、here (ここ)、hope (願う)、how (どのように)、hungry (空腹)、I (私)、is (～である)、it (それ)、like (好き)、music (音楽)、my (私の)、need (必要とする)、no (いいえ)、not (～でない)、nurse (看護師)、okay (オーケー)、outside (外側)、please (どうか)、right (すぐ、正しい)、success (成功)、tell (伝える)、that (あれ)、they (それら)、thirsty (喉が渴いている)、tired (疲れている)、up (上)、very (とても)、what (何)、where (どこ)、yes (はい)、及び you (あなた) を含む。

10

【0020】

特定の実施形態において、対象が、文を作成するために、制限なしに単語セットの単語を使用してもよい。他の実施形態において、対象が、試行発話の指定の文セットに制限される。

【0021】

特定の実施形態において、プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、文セットの全ての文について、文セットの文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、指定の単語セットからの単語から全体が構成される多くの可能性のある文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、対象が試行発話中に生成しようとした可能性が最も高い、指定の単語セットからの単語から全体が構成される文、及びその可能性がより低い他のそのような文を維持するようにプログラムされる。いくつかの実施形態において、プロセッサが、任意の所与の時点における、1 番目、2 番目、及び 3 番目に可能性の高い文の確率を追跡するようにプログラムされる。新しい単語イベントが処理されるとき、最も可能性の高い文が、変化し得る。例えば、ある単語イベントの処理に基づく 2 番目に可能性の高い文が、1 つ以上の追加の単語イベントが処理された後に、最も可能性の高い文になり得る。

20

【0022】

特定の実施形態において、文セットが、対象が介護者が実施することを所望するタスクに関して、介護者とコミュニケーションするために選択することができる文を含む。いくつかの実施形態において、指定の単語セットからの単語から全体が構成され得る文が、対象が介護者が実施することを所望するタスクに関して、介護者とコミュニケーションするために使用することができる文を含む。

30

【0023】

特定の実施形態において、文セットが、Are you going outside (あなたは外出していますか)、Are you tired (あなたは疲れていますか)、Bring my glasses here (私の眼鏡をここに持ってきてください)、Bring my glasses please (どうか私の眼鏡を持ってきてください)、Do not feel bad (気を悪くしないでください)、Do you feel comfortable (あなたは快適ですか)、Faith is good (信仰はよいことです)、Hello how are you (こんにちは、元気ですか)、Here is my computer (ここにあるのは私のコンピュータです)、How do you feel (気分はどうですか)、How do you like my music (私の音楽はあなたの気に入りましたか)、I am going outside (私は外出しています)、I am not going (私は外出していません)、I am not hungry (私は空腹ではありません)、I am not okay (私は体調が良くありません)、I am okay (私は体調が良いです)、I am outside (私は外にいます)、I am thi

40

50

r s t y (私は喉が渴いています)、I d o n o t f e e l c o m f o r t a b
 l e (私は快適ではありません)、I f e e l v e r y c o m f o r t a b l e (私は非常に快適です)、I f e e l v e r y h u n g r y (私は非常に空腹です)、I h o p e i t i s c l e a n (私はそれが清潔であると思います)、I l i k e m y n u r s e (私は私の看護師が好きです)、I n e e d m y g l a s s e s (私は私の眼鏡がほしいです)、I n e e d y o u (私にはあなたが必要です)、I t i s c o m f o r t a b l e (それは快適です)、I t i s g o o d (それは良いです)、I t i s o k a y (それは満足のいくものです)、I t i s r i g h t h e r e (それはちょうどここにあります)、M y c o m p u t e r i s c l e a n (私のコンピュータは清潔です)、M y f a m i l y i s h e r e (私の家族はここにいます)、M y f a m i l y i s o u t s i d e (私の家族は外にいます)、M y f a m i l y i s v e r y c o m f o r t a b l e (私の家族は非常に快適です)、M y g l a s s e s a r e c l e a n (私の眼鏡は清潔です)、M y g l a s s e s a r e c o m f o r t a b l e (私の眼鏡は快適です)、M y n u r s e i s o u t s i d e (私の看護師は外にいます)、M y n u r s e i s r i g h t o u t s i d e (私の看護師はちょうど外にいます)、N o (いいえ)、P l e a s e b r i n g m y g l a s s e s h e r e (どうか私の眼鏡をここに持ってきてください)、P l e a s e c l e a n i t (どうかそれを清潔にしてください)、P l e a s e t e l l m y f a m i l y (どうか私の家族に電話してください)、T h a t i s v e r y c l e a n (あれは非常に清潔です)、T h e y a r e c o m i n g h e r e (彼らはここにきています)、T h e y a r e c o m i n g o u t s i d e (彼らは外にきています)、T h e y a r e g o i n g o u t s i d e (彼らは外出しています)、T h e y h a v e f a i t h (彼らには信仰があります)、W h a t d o y o u d o (あなたは何をしていますか)、W h e r e i s i t (それはどこにありますか)、Y e s (はい)、及び Y o u a r e n o t r i g h t (あなたは正しくありません)を含む。

10

20

30

40

50

【0024】

特定の実施形態において、プロセッサが、予測単語系列確率を決定することによって復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされる。例えば、言語モデルに従って、より頻

【0025】

特定の実施形態において、プロセッサが、試行発話と関連付けられる脳電気信号データ、機械学習アルゴリズムを使用した単語分類からの予測単語確率、及び言語モデルを使用した単語系列確率を所与として、対象の意図された発話内の最も可能性の高い単語系列を決定するために、隠れマルコフモデル(HMM)又はビタビ復号モデルを使用するようにプログラムされる。

【0026】

特定の実施形態において、本方法が、対象の試行非発話運動と関連付けられる脳電気信号データを記録することであって、対象が、試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する非発話運動分類モデルを使用して脳電気信号データを分析することと、を更に含む。いくつかの実施形態において、試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む。

【0027】

特定の実施形態において、プロセッサが、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、対象の試行非発話運動の検出を自動化するように更にプログラムされる。いくつかの実施形態において、プ

ロセッサが、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てるように更にプログラムされる。

【0028】

特定の実施形態において、本方法が、復号の正確度を評価することを更に含む。

【0029】

別の態様において、対象による試行発話と関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法が提供され、コンピュータが、a) 記録された脳電気信号データを対象から受信するステップと、b) 脳電気信号データの記録中の任意の時点で試行発話が発生している確率を算出し、対象による試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、c) 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、脳電気信号データを分析するステップと、d) 単語分類モデルからの算出された単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、単語分類モデル及び言語モデルを使用して決定された予測単語確率に基づいて、文内の最も可能性の高い単語系列を決定するステップと、e) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する。

【0030】

特定の実施形態において、プロセッサが、対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、機械学習アルゴリズムを使用して、発話検出、単語分類、及び文復号を自動化するようにプログラムされる。いくつかの実施形態において、機械学習アルゴリズムが、発話検出のために人工ニューラルネットワーク (ANN) モデルを使用し、文復号のために、限定ではないが、隠れマルコフモデル (HMM) 又はビタビ復号モデルなどの単語分類及び自然言語処理技法を使用する。

【0031】

特定の実施形態において、対象が、試行発話について指定の単語セットに制限される。いくつかの実施形態において、プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出し、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択するように更にプログラムされる。

【0032】

特定の実施形態において、対象が、文を作成するために、制限なしに単語セットの単語を使用してもよい。他の実施形態において、対象が、試行発話の指定の文セットに制限される。いくつかの実施形態において、プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するように更にプログラムされる。いくつかの実施形態において、プロセッサが、文セットの文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するように更にプログラムされる。

【0033】

特定の実施形態において、コンピュータ実装方法が、準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む。

【0034】

特定の実施形態において、コンピュータ実装方法が、単語分類の検出された開始の周りのタイムウィンドウ内に (例えば、単語分類の検出された開始の1秒前から、検出された開始の3秒後まで) 記録された脳電気信号データを分析することを更に含む。

【0035】

特定の実施形態において、コンピュータ実装方法が、言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みを割り当てることを更

10

20

30

40

50

に含む。

【0036】

特定の実施形態において、コンピュータ実装方法が、対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、対象が、試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する非発話運動分類モデルを使用して脳電気信号データを分析することと、を更に含む。いくつかの実施形態において、試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む。いくつかの実施形態において、コンピュータ実装方法が、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てることを更に含む。

10

【0037】

特定の実施形態において、コンピュータ実装方法が、対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンに関する情報を含む対象のユーザプロファイルを格納することを更に含む。

【0038】

別の態様において、コンピュータ内のプロセッサによって実行されると、プロセッサに、対象による試行発話と関連付けられる記録された脳電気信号データから文を復号するための、本明細書に記載されたコンピュータ実装方法を実施させるプログラム命令を含む、非一時的コンピュータ可読媒体が提供される。

20

【0039】

別の態様において、非一時的コンピュータ可読媒体と、対象による試行発話と関連付けられる脳電気信号データを復号するための命令と、を備える、キットが提供される。

【0040】

別の態様において、対象のコミュニケーションを支援するためのシステムが提供され、システムが、対象による試行発話と関連付けられる脳電気信号データを記録するために、対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、本明細書に記載されているコンピュータ実装方法に従って、記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、対象の頭部上のロケーションに位置するように適合されている、コンピューティングデバイスと通信するインターフェースであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、記録された脳電気信号データから復号された文を表示するためのディスプレイコンポーネントと、を備える。

30

【0041】

特定の実施形態において、対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のためにコミュニケーションが困難である。

【0042】

特定の実施形態において、神経記録デバイスのロケーションが、腹側感覚運動皮質内にある。

40

【0043】

特定の実施形態において、電極が、感覚運動皮質領域の表面上に、又は感覚運動皮質領域内に位置するように適合されている。いくつかの実施形態において、電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置するように適合されている。

【0044】

特定の実施形態において、神経記録デバイスが、脳侵入電極アレイ又は皮質脳波記録（ECOG）電極アレイを含む。

【0045】

特定の実施形態において、電極が、深部電極又は表面電極である。

【0046】

50

特定の実施形態において、電気信号データが、高ガンマ周波数成分特徴を含む。いくつかの実施形態において、高ガンマ周波数電気信号データが、70Hz～150Hzの範囲内の神経振動を含む。

【0047】

特定の実施形態において、インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える。いくつかの実施形態において、インターフェースが、経皮台座コネクタに接続可能であるヘッドステージを更に備える。

【0048】

特定の実施形態において、プロセッサが、コンピュータ又はハンドヘルドデバイス（例えば、携帯電話又はタブレット）によって提供される。

10

【0049】

特定の実施形態において、プロセッサが、対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、機械学習アルゴリズムを使用して、発話検出、単語分類、及び文復号を自動化するようにプログラムされる。いくつかの実施形態において、機械学習アルゴリズムが、発話検出のために人工ニューラルネットワーク（ANN）モデルを使用し、文復号のために、限定ではないが、隠れマルコフモデル（HMM）又はビタビ復号モデルなどの単語分類及び自然言語処理技法を使用する。

【0050】

特定の実施形態において、プロセッサが、準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされる。いくつかの実施形態において、プロセッサが、単語分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するように更にプログラムされる。

20

【0051】

特定の実施形態において、対象が、試行発話について指定の単語セットに制限される。いくつかの実施形態において、プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出し、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択するように更にプログラムされる。

【0052】

特定の実施形態において、単語セットが、am（～である）、are（～である）、bad（悪い）、bring（運ぶ）、clean（きれい、きれいにする）、closer（より近い）、comfortable（快適）、coming（来ている）、computer（コンピュータ）、do（行う）、faith（信仰）、family（家族）、feel（感じる）、glasses（眼鏡）、going（行っている）、good（良い）、goodbye（さようなら）、have（有する）、hello（こんにちは）、help（手助け、手伝って）、here（ここ）、hope（願う）、how（どのように）、hungry（空腹）、I（私）、is（～である）、it（それ）、like（好き）、music（音楽）、my（私の）、need（必要とする）、no（いいえ）、not（～でない）、nurse（看護師）、okay（オーケー）、outside（外側）、please（どうか）、right（すぐ、正しい）、success（成功）、tell（伝える）、that（あれ）、they（それら）、thirsty（喉が渴いている）、tired（疲れている）、up（上）、very（とても）、what（何）、where（どこ）、yes（はい）、及びyou（あなた）を含む。

30

【0053】

特定の実施形態において、対象が、文を作成するために、制限なしに単語セットの単語を使用してもよい。他の実施形態において、対象が、試行発話の指定の文セットに制限される。いくつかの実施形態において、プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するように更にプログラムされる。いく

40

50

つかの実施形態において、プロセッサが、文セットの文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するように更にプログラムされる。いくつかの実施形態において、文セットが、対象が介護者が実施することを所望するタスクに関して、介護者とコミュニケーションするために選択することができる文を含む。

【0054】

特定の実施形態において、文セットが、Are you going outside (あなたは外出していますか)、Are you tired (あなたは疲れていますか)、Bring my glasses here (私の眼鏡をここに持ってきてください)、Bring my glasses please (どうか私の眼鏡を持ってきてください)、Do not feel bad (気を悪くしないでください)、Do you feel comfortable (あなたは快適ですか)、Faith is good (信仰はよいことです)、Hello how are you (こんにちは、元気ですか)、Here is my computer (ここにあるのは私のコンピュータです)、How do you feel (気分はどうですか)、How do you like my music (私の音楽はあなたの気に入りましたか)、I am going outside (私は外出しています)、I am not going (私は外出していません)、I am not hungry (私は空腹ではありません)、I am not okay (私は体調が良くありません)、I am okay (私は体調が良いです)、I am outside (私は外にいます)、I am thirsty (私は喉が渴いています)、I do not feel comfortable (私は快適ではありません)、I feel very comfortable (私は非常に快適です)、I feel very hungry (私は非常に空腹です)、I hope it is clean (私はそれが清潔であると思います)、I like my nurse (私は私の看護師が好きです)、I need my glasses (私は私の眼鏡がほしいです)、I need you (私にはあなたが必要です)、It is comfortable (それは快適です)、It is good (それは良いです)、It is okay (それは満足のいくものです)、It is right here (それはちょうどここにあります)、My computer is clean (私のコンピュータは清潔です)、My family is here (私の家族はここにいます)、My family is outside (私の家族は外にいます)、My family is very comfortable (私の家族は非常に快適です)、My glasses are clean (私の眼鏡は清潔です)、My glasses are comfortable (私の眼鏡は快適です)、My nurse is outside (私の看護師は外にいます)、My nurse is right outside (私の看護師はちょうど外にいます)、No (いいえ)、Please bring my glasses here (どうか私の眼鏡をここに持ってきてください)、Please clean it (どうかそれを清潔にしてください)、Please tell my family (どうか私の家族に電話してください)、That is very clean (あれは非常に清潔です)、They are coming here (彼らはここに来ています)、They are coming outside (彼らは外に来ています)、They are going outside (彼らは外出しています)、They have faith (彼らには信仰があります)、What do you do (あなたは何をしていますか)、Where is it (それはどこにありますか)、Yes (はい)、及びYou are not right (あなたは正しくありません)を含む。

【0055】

特定の実施形態において、プロセッサが、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、対象の試行非発話運動の検出を自動化するように更にプログラムされる。いくつかの実施形態において、プロセッサが、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中

10

20

30

40

50

の時点に割り当てるように更にプログラムされる。

【0056】

別の態様において、対象のコミュニケーションを支援するための本明細書に記載されているシステムと、非一時的コンピュータ可読媒体と、対象による試行発話と関連付けられる脳電気信号データを記録し、復号するために、システムを使用するための命令と、を備える、キットが提供される。

【0057】

別の態様において、対象のコミュニケーションを支援する方法が提供され、本方法が、対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、コンピューティングデバイスと通信するインターフェースを、対象の頭部上のロケーションに位置付けることであって、インターフェースが、神経記録デバイスに接続されている、位置付けることと、神経記録デバイスを使用して、対象による上記試行スペリングと関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをコンピューティングデバイスのプロセッサに送信する、記録することと、プロセッサを使用して、記録された脳電気信号データから意図された文のスペリングされた単語を復号することと、を含む。

【0058】

特定の実施形態において、電気信号データが、高ガンマ周波数成分特徴（例えば、70 Hz ~ 150 Hz）及び低周波数成分特徴（例えば、0.3 Hz ~ 100 Hz）を含む。

【0059】

特定の実施形態において、脳電気信号データを記録することが、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの脳電気信号データを記録することを含む。

【0060】

特定の実施形態において、本方法が、対象による単語の試行スペリングと関連付けられる脳の電気信号を記録するために電極を位置付けるための最適なロケーションを識別するために、対象の脳をマッピングすることを更に含む。

【0061】

特定の実施形態において、プロセッサが、対象による単語の試行スペリングと関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、試行スペリングと関連付けられる脳活動の検出、文字分類、単語分類、及び文復号を自動化するようにプログラムされる。

【0062】

特定の実施形態において、プロセッサが、発話検出、文字分類、単語分類、及び文復号のために機械学習アルゴリズムを使用するようにプログラムされる。いくつかの実施形態において、機械学習アルゴリズムが、自然言語処理技法を使用してもよい。

【0063】

特定の実施形態において、プロセッサが、対象による単語の試行スペリングと関連付けられる神経活動から復号される文字系列からの単語分類を、対象によって使用される言語の語彙内の単語のみに制約するように更にプログラムされる。

【0064】

特定の実施形態において、プロセッサが、対象による試行スペリング中の文字生成の開始及び終了の検出を自動化するようにプログラムされる。

【0065】

特定の実施形態において、プロセッサが、準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされる。

【0066】

特定の実施形態において、プロセッサが、対象による文字の試行スペリングの検出され

10

20

30

40

50

た開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされる。

【0067】

特定の実施形態において、本方法が、対象が意図された文の単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを対象に提供することを更に含む。いくつかの実施形態において、一連のゴーキューが、ディスプレイ上に視覚的に提供される。いくつかの実施形態において、各ゴーキューに先行して、ゴーキューの提示へのカウントダウンが行われ、次にスペリングされる文字のカウントダウンが、ディスプレイ上に視覚的に提供され、各ゴーキューの後に自動的に開始される。いくつかの実施形態において、一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される。いくつかの実施形態において、対象が、各ゴーキュー間の設定時間間隔を制御することができる。いくつかの実施形態において、プロセッサが、ゴーキューに後続するタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされる。

10

【0068】

特定の実施形態において、プロセッサが、復号された文字の系列からの復号された単語の系列が、対象による意図された文の単語の文字の試行スペリング中に対象が生成しようとした、意図された文である確率を算出するようにプログラムされる。

【0069】

特定の実施形態において、プロセッサが、予測単語系列確率を決定することによって復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされる。いくつかの実施形態において、言語モデルに従って、より頻繁に出現する単語が、それほど頻繁に出現しない単語よりも大きい重みを割り当てられる。

20

【0070】

特定の実施形態において、プロセッサが、予測文字確率の系列を使用して、潜在的な文候補を計算し、文候補内の予測される単語間の文字系列に空白を自動的に挿入するように更にプログラムされる。

【0071】

特定の実施形態において、本方法が、対象の試行非発話運動と関連付けられる脳電気信号データを記録することであって、対象が、意図された文の単語の試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する分類モデルを使用して脳電気信号データを分析することと、を更に含む。

30

【0072】

特定の実施形態において、試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む。いくつかの実施形態において、手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む。

【0073】

特定の実施形態において、プロセッサが、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、対象による試行スペリングの終了をシグナリングする対象の試行非発話運動の検出を自動化するようにプログラムされる。いくつかの実施形態において、プロセッサが、前記試行非発話運動のイベントラベルを前記脳電気信号データの前記記録中の時点に割り当てるように更にプログラムされる。

40

【0074】

一態様において、本方法が、神経記録デバイスを使用して、対象による試行発話と関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをコンピューティングデバイスのプロセッサに送信する、記録することと、本明細書に記載されているように、プロセッ

50

サを使用して、対象による試行発話と関連付けられる記録された脳電気信号データから単語、句、又は文を復号することと、を更に含む。

【0075】

特定の実施形態において、本方法が、復号の正確度を評価することを更に含む。

【0076】

別の態様において、対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法が提供され、コンピュータが、a) 対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データを受信するステップと、b) 任意の時点で試行スペリングが発生している確率を算出し、対象による試行スペリング中の文字生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、c) 対象による試行文字生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測文字確率の系列を算出する文字分類モデルを使用して、脳電気信号データを分析するステップと、d) 予測文字確率の系列に基づいて潜在的な文候補を計算し、文候補内の予測される単語間の文字系列に空白を自動的に挿入するステップであって、文字系列内の復号される単語は、対象によって使用される言語の語彙内の単語のみに制約される、計算し、挿入するステップと、e) 予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用して潜在的な文候補を分析し、文内の最も可能性の高い単語系列を決定するステップと、f) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する。

【0077】

特定の実施形態において、記録された脳電気信号データが、対象による文字の試行スペリングの検出された開始の周りのタイムウィンドウ内でのみ使用される。

【0078】

特定の実施形態において、本方法が、対象が意図された文の単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを対象に表示することを更に含む。いくつかの実施形態において、各ゴーキューに先行して、ゴーキューの提示へのカウントダウンが表示され、次にスペリングされる文字のカウントダウンが、各ゴーキューの後に自動的に開始される。いくつかの実施形態において、一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される。いくつかの実施形態において、対象が、各ゴーキュー間の設定時間間隔を制御することができる。いくつかの実施形態において、ゴーキューに後続するタイムウィンドウ内で記録された脳電気信号データが、文字分類に使用される。

【0079】

特定の実施形態において、コンピュータ実装方法が、対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、対象が、意図された文の単語の試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する運動分類モデルを使用して脳電気信号データを分析することと、を更に含む。いくつかの実施形態において、試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む。いくつかの実施形態において、手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む。

【0080】

特定の実施形態において、機械学習アルゴリズムが、発話検出及び文字分類に使用される。

【0081】

特定の実施形態において、コンピュータ実装方法が、言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みを割り当てることを更に含む。

【0082】

特定の実施形態において、コンピュータ実装方法が、対象による試行スペリング中の文字生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンに関する情報を含む対象のユーザプロファイルを格納することを更に含む。

【0083】

特定の実施形態において、電気信号データが、高ガンマ周波数成分特徴（例えば、70 Hz ~ 150 Hz）及び低周波数成分特徴（例えば、0.3 Hz ~ 100 Hz）を含む。

【0084】

特定の実施形態において、コンピュータ実装方法が、復号の正確度を評価することを更に含む。

【0085】

特定の実施形態において、コンピュータ実装方法が、対象による試行発話と関連付けられる記録された脳電気信号データから文を復号することを更に含み、コンピュータが、a) 対象による試行発話と関連付けられる記録された脳電気信号データを受信するステップと、b) 任意の時点で試行発話が発生している確率を算出し、対象による試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、c) 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、脳電気信号データを分析するステップと、d) 単語分類モデルからの算出された単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、単語分類モデル及び言語モデルを使用して決定された予測単語確率に基づいて、文内の最も可能性の高い単語系列を決定するステップと、e) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する。いくつかの実施形態において、機械学習アルゴリズムが、発話検出、単語分類、及び文復号に使用される。いくつかの実施形態において、人工ニューラルネットワーク（ANN）モデルが、発話検出及び単語分類に使用され、隠れマルコフモデル（HMM）、ビタビ復号モデル、又は他の自然言語処理技法が、文復号に使用される。

【0086】

別の態様において、非一時的コンピュータ可読媒体は、提供され、非一時的コンピュータ可読媒体は、コンピュータ内のプロセッサによって実行されるとき、プロセッサに、本明細書に記載されたコンピュータ実装方法を実行させるプログラム命令を含む。

【0087】

別の態様において、非一時的コンピュータ可読媒体と、対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを復号するための命令と、を備える、キットが提供される。

【0088】

別の態様において、対象のコミュニケーションを支援するためのシステムが提供され、システムが、対象による試行発話、意図された文の単語の文字の試行スペリング、若しくは試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録するために、対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている、電極を備える神経記録デバイスと、本明細書に記載されているコンピュータ実装方法に従って、記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、コンピューティングデバイスと通信するインターフェースであって、上記インターフェースが、対象の頭部上のロケーションに位置するように適合されており、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、記録された脳電気信号データから復号された文を表示するためのディスプレイコンポーネントと、を備える。

【0089】

特定の実施形態において、電極が、感覚運動皮質領域の表面上に、又は感覚運動皮質領域内に位置するように適合されている。

【0090】

特定の実施形態において、電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置するように適合されている。

【0091】

特定の実施形態において、神経記録デバイスが、脳侵入電極アレイを含む。

【0092】

特定の実施形態において、神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む。

【0093】

特定の実施形態において、電極が、深部電極又は表面電極である。

【0094】

特定の実施形態において、電気信号データが、高ガンマ周波数成分特徴（例えば、70 Hz～150 Hz）及び低周波数成分特徴（例えば、0.3 Hz～100 Hz）を含む。

【0095】

特定の実施形態において、インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える。

【0096】

特定の実施形態において、インターフェースが、経皮台座コネクタに接続可能であるヘッドステージを更に備える。

【0097】

特定の実施形態において、プロセッサが、コンピュータ又はハンドヘルドデバイス（例えば、携帯電話又はタブレット）によって提供される。

【0098】

別の態様において、本明細書に記載されているシステムと、非一時的コンピュータ可読媒体と、対象による試行発話、単語の試行スペリング、若しくは試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録し、復号するために、システムを使用するための命令と、を備える、キット。

【0099】

試行発話、単語の試行スペリング、又は試行非発話運動と関連付けられる神経活動の復号を通じて対象のコミュニケーションを支援する方法は、組み合わせることができる。これらの技法は相補的である。いくつかの事例において、試行スペリングの復号が、試行発話の復号よりも大きい語彙が使用されることを可能にし得る。しかしながら、試行発話の復号は、より速く、直接的な単語復号を可能にするため、対象にとってより容易であり、より便利であり得、これが、頻繁に使用される単語を表現するのに好ましい場合がある。復号を支援するために、試行非発話運動を使用して、対象が試行発話又は意図されたメッセージのスペルアウトを開始又は終了していることをシグナリングすることができる。

【図面の簡単な説明】

【0100】

【図1】直接発話BCIの概略図。重度の麻痺を有する臨床試験参加者に埋め込まれた治療用皮質脳波記録（ECOG）電極アレイから取得された神経活動が使用されて、リアルタイムで単語及び文が直接復号される。会話型のデモンストレーションにおいて、参加者は質問によって視覚的に促され（A）、事前定義された50語の語彙からの単語を使用して応答を試行するように指示される。同時に、皮質信号が、ECOGデバイスを介して脳の表面から取得され（B）、リアルタイムで処理される（C）。発話検出モデルが、処理された神経信号をサンプルごとに分析して、参加者の発話する試行を検出する（D）。分類器が、関連する神経活動の各検出された窓から（50の可能な単語にわたって）単語確率を計算する（E）。ビタビ復号アルゴリズムが、これらの確率を、別個に訓練された言語モデルからの単語系列確率と併せて使用して、ECOGデータを所与として最も可能性

10

20

30

40

50

の高い文を復号する（F）。単語が復号されるたびに更新される、予測される文が、参加者へのフィードバックとして表示される（G）。

【図2A】神経信号処理及び言語モデリングが、様々な文のリアルタイムでの復号を可能にする。図2Aは、文タスクブロック中に参加者の皮質活動から復号された単語系列の単語誤り率を示す。単語誤り率は、復号誤りが行われた頻度を定量化する（単語誤り率が低いほど、パフォーマンスがより良好であることを示す）。言語モデル（LM）を用いて、及び用いずに単語を復号する場合、単語の誤り率は偶然よりも有意に低く、復号中にLMを使用する場合、パフォーマンスは有意に改善された（*全て $P < 0.001$ 、3方向ホルム・ボンフェローニ補正）。図2Bは、誤って復号された単語を含むか又は除外するかのいずれかの場合の、全ての試験にわたる復号された単語毎分値を示す。各バイオリン分布は、スコット帯域幅推定を用いたカーネル密度推定を使用して作成されており、中央値を示す太い水平線、及び範囲を示すより小さい水平線（平均を下回るか又は上回る4つの標準偏差を超える外れ値を除く）を伴う。図2Cは、各試験において検出された単語の数と実際の単語の数との間の差の要約を示し、正しい文の長さを有する試験の割合は黒色で示され、誤った文の長さは暗赤色で示される。図2Dは、全ての試験及び50全ての文標的にわたるLMを用いた、及び用いていない復号された文の編集距離（行われた復号誤りの数）を示し、LMによる予測についての編集距離の昇順でソートされる（より低い編集距離はより良好なパフォーマンスを示す）。各小さい垂直ダッシュは、1つの試験の編集距離を表す（標的文ごとに3つの試験がある。同一の編集距離のマークは、視覚化のために水平にずらして表示される）。各点は、その標的文の平均編集距離を表す。下部のヒストグラムは、全ての試験にわたる編集距離カウントを示している。図2Eは、7つの異なる試行の標的文並びにLMを使用した、及び使用していない、復号された文を示す。正しく復号された単語は黒色で示され、誤った単語は赤色で示される。

【図2B】神経信号処理及び言語モデリングが、様々な文のリアルタイムでの復号を可能にする。図2Aは、文タスクブロック中に参加者の皮質活動から復号された単語系列の単語誤り率を示す。単語誤り率は、復号誤りが行われた頻度を定量化する（単語誤り率が低いほど、パフォーマンスがより良好であることを示す）。言語モデル（LM）を用いて、及び用いずに単語を復号する場合、単語の誤り率は偶然よりも有意に低く、復号中にLMを使用する場合、パフォーマンスは有意に改善された（*全て $P < 0.001$ 、3方向ホルム・ボンフェローニ補正）。図2Bは、誤って復号された単語を含むか又は除外するかのいずれかの場合の、全ての試験にわたる復号された単語毎分値を示す。各バイオリン分布は、スコット帯域幅推定を用いたカーネル密度推定を使用して作成されており、中央値を示す太い水平線、及び範囲を示すより小さい水平線（平均を下回るか又は上回る4つの標準偏差を超える外れ値を除く）を伴う。図2Cは、各試験において検出された単語の数と実際の単語の数との間の差の要約を示し、正しい文の長さを有する試験の割合は黒色で示され、誤った文の長さは暗赤色で示される。図2Dは、全ての試験及び50全ての文標的にわたるLMを用いた、及び用いていない復号された文の編集距離（行われた復号誤りの数）を示し、LMによる予測についての編集距離の昇順でソートされる（より低い編集距離はより良好なパフォーマンスを示す）。各小さい垂直ダッシュは、1つの試験の編集距離を表す（標的文ごとに3つの試験がある。同一の編集距離のマークは、視覚化のために水平にずらして表示される）。各点は、その標的文の平均編集距離を表す。下部のヒストグラムは、全ての試験にわたる編集距離カウントを示している。図2Eは、7つの異なる試行の標的文並びにLMを使用した、及び使用していない、復号された文を示す。正しく復号された単語は黒色で示され、誤った単語は赤色で示される。

【図2C】神経信号処理及び言語モデリングが、様々な文のリアルタイムでの復号を可能にする。図2Aは、文タスクブロック中に参加者の皮質活動から復号された単語系列の単語誤り率を示す。単語誤り率は、復号誤りが行われた頻度を定量化する（単語誤り率が低いほど、パフォーマンスがより良好であることを示す）。言語モデル（LM）を用いて、及び用いずに単語を復号する場合、単語の誤り率は偶然よりも有意に低く、復号中にLMを使用する場合、パフォーマンスは有意に改善された（*全て $P < 0.001$ 、3方向ホ

10

20

30

40

50

ルム・ボンフェローニ補正)。図2Bは、誤って復号された単語を含むか又は除外するかのいずれかの場合の、全ての試験にわたる復号された単語毎分値を示す。各バイオリン分布は、スコット帯域幅推定を用いたカーネル密度推定を使用して作成されており、中央値を示す太い水平線、及び範囲を示すより小さい水平線（平均を下回るか又は上回る4つの標準偏差を超える外れ値を除く）を伴う。図2Cは、各試験において検出された単語の数と実際の単語の数との間の差の要約を示し、正しい文の長さを有する試験の割合は黒色で示され、誤った文の長さは暗赤色で示される。図2Dは、全ての試験及び50全ての文標的にわたるLMを用いた、及び用いていない復号された文の編集距離（行われた復号誤りの数）を示し、LMによる予測についての編集距離の昇順でソートされる（より低い編集距離はより良好なパフォーマンスを示す）。各小さい垂直ダッシュは、1つの試験の編集距離を表す（標的文ごとに3つの試験がある。同一の編集距離のマークは、視覚化のために水平にずらして表示される）。各点は、その標的文の平均編集距離を表す。下部のヒストグラムは、全ての試験にわたる編集距離カウントを示している。図2Eは、7つの異なる試行の標的文並びにLMを使用した、及び使用していない、復号された文を示す。正しく復号された単語は黒色で示され、誤った単語は赤色で示される。

【図2D】神経信号処理及び言語モデリングが、様々な文のリアルタイムでの復号を可能にする。図2Aは、文タスクブロック中に参加者の皮質活動から復号された単語系列の単語誤り率を示す。単語誤り率は、復号誤りが行われた頻度を定量化する（単語誤り率が低いほど、パフォーマンスがより良好であることを示す）。言語モデル（LM）を用いて、及び用いずに単語を復号する場合、単語の誤り率は偶然よりも有意に低く、復号中にLMを使用する場合、パフォーマンスは有意に改善された（*全て $P < 0.001$ 、3方向ホルム・ボンフェローニ補正）。図2Bは、誤って復号された単語を含むか又は除外するかのいずれかの場合の、全ての試験にわたる復号された単語毎分値を示す。各バイオリン分布は、スコット帯域幅推定を用いたカーネル密度推定を使用して作成されており、中央値を示す太い水平線、及び範囲を示すより小さい水平線（平均を下回るか又は上回る4つの標準偏差を超える外れ値を除く）を伴う。図2Cは、各試験において検出された単語の数と実際の単語の数との間の差の要約を示し、正しい文の長さを有する試験の割合は黒色で示され、誤った文の長さは暗赤色で示される。図2Dは、全ての試験及び50全ての文標的にわたるLMを用いた、及び用いていない復号された文の編集距離（行われた復号誤りの数）を示し、LMによる予測についての編集距離の昇順でソートされる（より低い編集距離はより良好なパフォーマンスを示す）。各小さい垂直ダッシュは、1つの試験の編集距離を表す（標的文ごとに3つの試験がある。同一の編集距離のマークは、視覚化のために水平にずらして表示される）。各点は、その標的文の平均編集距離を表す。下部のヒストグラムは、全ての試験にわたる編集距離カウントを示している。図2Eは、7つの異なる試行の標的文並びにLMを使用した、及び使用していない、復号された文を示す。正しく復号された単語は黒色で示され、誤った単語は赤色で示される。

【図2E】神経信号処理及び言語モデリングが、様々な文のリアルタイムでの復号を可能にする。図2Aは、文タスクブロック中に参加者の皮質活動から復号された単語系列の単語誤り率を示す。単語誤り率は、復号誤りが行われた頻度を定量化する（単語誤り率が低いほど、パフォーマンスがより良好であることを示す）。言語モデル（LM）を用いて、及び用いずに単語を復号する場合、単語の誤り率は偶然よりも有意に低く、復号中にLMを使用する場合、パフォーマンスは有意に改善された（*全て $P < 0.001$ 、3方向ホルム・ボンフェローニ補正）。図2Bは、誤って復号された単語を含むか又は除外するかのいずれかの場合の、全ての試験にわたる復号された単語毎分値を示す。各バイオリン分布は、スコット帯域幅推定を用いたカーネル密度推定を使用して作成されており、中央値を示す太い水平線、及び範囲を示すより小さい水平線（平均を下回るか又は上回る4つの標準偏差を超える外れ値を除く）を伴う。図2Cは、各試験において検出された単語の数と実際の単語の数との間の差の要約を示し、正しい文の長さを有する試験の割合は黒色で示され、誤った文の長さは暗赤色で示される。図2Dは、全ての試験及び50全ての文標的にわたるLMを用いた、及び用いていない復号された文の編集距離（行われた復号誤り

10

20

30

40

50

の数)を示し、LMによる予測についての編集距離の昇順でソートされる(より低い編集距離はより良好なパフォーマンスを示す)。各小さい垂直ダッシュは、1つの試験の編集距離を表す(標的文ごとに3つの試験がある。同一の編集距離のマークは、視覚化のために水平にずらして表示される)。各点は、その標的文の平均編集距離を表す。下部のヒストグラムは、全ての試験にわたる編集距離カウントを示している。図2Eは、7つの異なる試行の標的文並びにLMを使用した、及び使用していない、復号された文を示す。正しく復号された単語は黒色で示され、誤った単語は赤色で示される。

【図3A】異なる神経活動パターンが、単語生成試行の基礎となる。図3Aは、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度に対する訓練データの量の影響を示す。各点は、10交差検証フォールドにわたる平均±標準偏差を示す。偶然正確度が、水平破線として描かれている。図3Bは、埋め込まれた電極の位置並びに発話検出モデル及び単語分類モデルへのそれらの寄与と重ねられた参加者の脳再構築を示す。プロットされた電極のサイズ(面積)及び不透明度は、相対的な寄与によってスケールされる(重要な電極は他の電極よりも大きく不透明に見える)。各寄与セットは、合計が1になるように正規化されている。解剖学的参照のために、中心前回は明青色で強調表示されている。図3Cは、分類結果からの単語の混乱を示し、分類器が標的単語の同一性を所与として参加者が発語を試行していた50個の単語の各々を予測した頻度を示す(対角線に沿った値は正しい分類に対応する)。

10

【図3B】異なる神経活動パターンが、単語生成試行の基礎となる。図3Aは、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度に対する訓練データの量の影響を示す。各点は、10交差検証フォールドにわたる平均±標準偏差を示す。偶然正確度が、水平破線として描かれている。図3Bは、埋め込まれた電極の位置並びに発話検出モデル及び単語分類モデルへのそれらの寄与と重ねられた参加者の脳再構築を示す。プロットされた電極のサイズ(面積)及び不透明度は、相対的な寄与によってスケールされる(重要な電極は他の電極よりも大きく不透明に見える)。各寄与セットは、合計が1になるように正規化されている。解剖学的参照のために、中心前回は明青色で強調表示されている。図3Cは、分類結果からの単語の混乱を示し、分類器が標的単語の同一性を所与として参加者が発語を試行していた50個の単語の各々を予測した頻度を示す(対角線に沿った値は正しい分類に対応する)。

20

【図3C】異なる神経活動パターンが、単語生成試行の基礎となる。図3Aは、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度に対する訓練データの量の影響を示す。各点は、10交差検証フォールドにわたる平均±標準偏差を示す。偶然正確度が、水平破線として描かれている。図3Bは、埋め込まれた電極の位置並びに発話検出モデル及び単語分類モデルへのそれらの寄与と重ねられた参加者の脳再構築を示す。プロットされた電極のサイズ(面積)及び不透明度は、相対的な寄与によってスケールされる(重要な電極は他の電極よりも大きく不透明に見える)。各寄与セットは、合計が1になるように正規化されている。解剖学的参照のために、中心前回は明青色で強調表示されている。図3Cは、分類結果からの単語の混乱を示し、分類器が標的単語の同一性を所与として参加者が発語を試行していた50個の単語の各々を予測した頻度を示す(対角線に沿った値は正しい分類に対応する)。

30

【図4A】発話試行中に記録された神経活動は、長期的な安定性を呈する。図4Aは、18ヶ月にわたる記録に及ぶ、孤立単語タスク中に「Goodbye(さようなら)」という単語を発語する参加者の全ての試行にわたる単一の電極からの神経活動を示す。図4Bは、4つの重複しない日付範囲からサンプリングされた、孤立単語データのサブセットに対する検出器及び分類器の訓練及びテストからの単語分類結果を示す。各サブセットには、各単語の20の試行された生成からのデータが含まれている。各実線バーは、単一のサブセット内の交差検証された評価からの結果を示し、各点線バーは、評価されているサブセットを除く全てのサブセットからのデータに対する訓練からの結果を示す。各バーは、10評価フォールドにわたる平均±標準誤差を表す。偶然正確度が、水平破線として描かれている。また、各試験サブセットについての4つの同サブセット評価間(* $P < 0.0$

40

50

1、両側フィッシャー正確確率検定、10方向ホルム・ボンフェローニ補正)、及び2つの評価間(* $P < 0.01$ 、両側正確マクネマー検定、10方向ホルム・ボンフェローニ補正)の有意差も示されている。単一のサブセット内の交差検証された評価中に計算された電極寄与が、上部に示される(右上隅に最も背面及び後方の電極を配向される)。プロットされた電極のサイズ(面積)及び不透明度は、相対的な寄与によってスケールされる。各寄与セットは、合計が1になるように正規化されている。

【図4B】発話試行中に記録された神経活動は、長期的な安定性を呈する。図4Aは、18ヶ月にわたる記録に及ぶ、孤立単語タスク中に「Goodbye(さようなら)」という単語を発語する参加者の全ての試行にわたる単一の電極からの神経活動を示す。図4Bは、4つの重複しない日付範囲からサンプリングされた、孤立単語データのサブセットに対する検出器及び分類器の訓練及びテストからの単語分類結果を示す。各サブセットには、各単語の20の試行された生成からのデータが含まれている。各実線バーは、単一のサブセット内の交差検証された評価からの結果を示し、各点線バーは、評価されているサブセットを除く全てのサブセットからのデータに対する訓練からの結果を示す。各バーは、10評価フォールドにわたる平均±標準誤差を表す。偶然正確度が、水平破線として描かれている。また、各試験サブセットについての4つの同サブセット評価間(* $P < 0.01$ 、両側フィッシャー正確確率検定、10方向ホルム・ボンフェローニ補正)、及び2つの評価間(* $P < 0.01$ 、両側正確マクネマー検定、10方向ホルム・ボンフェローニ補正)の有意差も示されている。単一のサブセット内の交差検証された評価中に計算された電極寄与が、上部に示される(右上隅に最も背面及び後方の電極を配向される)。プロットされた電極のサイズ(面積)及び不透明度は、相対的な寄与によってスケールされる。各寄与セットは、合計が1になるように正規化されている。

【図5A】参加者のMRI結果。図5Aは、脳軟化症及びポンティーン脳卒中(赤色でラベル付けされている)によって引き起こされる脳幹萎縮(青色でラベル付けされている)を有する参加者の矢状MRIを示す。図5Bは、大脳萎縮が存在しないことを示す2つの追加のMRIスキャンを示し、皮質ニューロン集団(本研究で記録されたものを含む)が参加者の病理によって比較的影響を受けていないはずであることを示唆している。

【図5B】参加者のMRI結果。図5Aは、脳軟化症及びポンティーン脳卒中(赤色でラベル付けされている)によって引き起こされる脳幹萎縮(青色でラベル付けされている)を有する参加者の矢状MRIを示す。図5Bは、大脳萎縮が存在しないことを示す2つの追加のMRIスキャンを示し、皮質ニューロン集団(本研究で記録されたものを含む)が参加者の病理によって比較的影響を受けていないはずであることを示唆している。

【図6】リアルタイム神経データ取得ハードウェアインフラストラクチャ。埋め込まれたアレイ及び経皮台座コネクタから取得された皮質脳波記録(ECoG)データが処理され、Neuropoortデジタル信号プロセッサ(DSP)に送信される。同時に、マイクロフォンデータが取得され、増幅され、DSPに送信される。DSPからの信号は、リアルタイムコンピュータに送信される。リアルタイムコンピュータは、フィードバックとしてリアルタイムで提供される任意の復号された文を含む、参加者に表示されるタスクを制御する。リアルタイムコンピュータからのスピーカ出力もDSPに送信され、神経信号(図示せず)と同期される。以前のセッション中に、台座に接続されたヒト患者ケーブルがECoG信号を取得しており、これはその後、DSPに送信される前にフロントエンド増幅器によって処理された(ヒト患者ケーブル及びフロントエンド増幅器はここに示されていないが、使用時にこのパイプライン内のデジタルヘッドステージ及びデジタルハブと置き換わっていた)。

【図7】リアルタイム神経信号処理パイプライン。データ取得ヘッドステージ及びリグを使用して、参加者の皮質脳波記録(ECoG)信号が30kHzで取得され、広帯域フィルタによってフィルタリングされ、ソフトウェアベースの回線ノイズキャンセル技法によって調節され、500Hzでローパスフィルタリングされ、1kHzでリアルタイムコンピュータにストリーミングされた。リアルタイムコンピュータ上で、カスタムソフトウェアが使用されて、ECoG信号に対する共通平均参照、マルチバンド高ガンマバンドパス

フィルタリング、分析振幅推定、マルチバンド平均化、及び z スコアリングの実行が実施された。次いで、得られた信号が、残りの分析のための高ガンマ活性の測度として使用された。

【図 8】データ収集タイムライン。1 日に複数のデータタイプが収集された場合、バーは垂直に積み重ねられる（任意の所与の日の積み重ねられたバーの高さは、その日に収集された試験の総数に等しい）。データ収集スケジュールの不規則性は、埋め込まれたデバイスとは無関係の外部及び臨床的な時間的制約によって部分的に引き起こされている。55～88 週の隙間は、COVID-19 パンデミックに関する臨床ガイドラインに起因している。

【図 9】発話検出モデル概略図。全ての電極にわたる z スコア高ガンマ活性は、3 つの長・短期記憶層（LSTM）及び単一の高密度（全結合）層のスタックからなる人工ニューラルネットワークによって時点ごとに処理される。高密度層は、最後の LSTM 層の潜在次元を、発話、準備、及び休止の 3 つのイベントクラスの確率空間に投影する。予測発話イベント確率時系列は、平滑化され、次いで、検出された発話イベントの開始時間（ t^* ）及び終了時間をもたらすための確率及び時間閾値で閾値化される。文復号中、発話イベントが検出されるたびに、検出された開始（ t^* ）に対して -1～3 秒に及ぶ神経活動の窓が単語分類器に渡された。示されている神経活動、予測発話確率時系列（右上）、及び検出された発話イベント（右下）は、参加者が単語「family」の生成を試行した、孤立単語試験の 7 秒のタイムウィンドウにわたる実際の神経データ及び検出結果である。

【図 10】単語分類モデル概略図。各分類について、高ガンマ活性の 4 秒のタイムウィンドウが、10 個の人工ニューラルネットワーク（ANN）モデルからなるアンサンブルによって処理される。各 ANN 内で、高ガンマ活性は、時間的畳み込み、及び後続する 2 つの双方向ゲート付き再帰ユニット（GRU）層によって処理される。高密度層は、最終 GRU 層から確率空間に潜在次元を投影し、これは、神経タイムウィンドウと関連付けられる発話生成試行中に、50 単語セットからの単語の各単語が標的単語である確率を含む。アンサンブルされた ANN モデルからの 10 の確率分布が、予測単語確率の最終ベクトルを得るためにともに平均化される。

【図 11】文復号隠れマルコフモデル。この隠れマルコフモデル（HMM）は、参加者が生成を試行する単語（隠れ状態 q_i ）と、関連付けられる検出された神経活動のタイムウィンドウ（観察された状態 y_i ）との間の関係を記述する。HMM 放出確率 $p(y_0 | q_0)$ は、 $p(w_i | y_i)$ （単語分類器によって提供される単語尤度）に単純化することができ、HMM 遷移確率 $p(q_i | q_{i-1})$ は、 $p(w_i | c_i)$ （言語モデルによって提供される単語系列の事前確率）に単純化することができる。

【図 12 A】孤立単語データを用いた補助モデリング結果。図 12 A は、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度（左）及び交差エントロピー損失（右）に対する訓練データの量の影響を示す。交差エントロピーが低いほど、パフォーマンスがより良好であることが指示される。各点は、10 交差検証フォールドにわたる平均±標準偏差を示す（交差エントロピープロットにおける誤差バーは、典型的には、円形マーカーと並んで見ることができないほど小さかった）。偶然パフォーマンスが、各プロットにおいて水平破線として示されている（偶然交差エントロピー損失は、単語標的の数の逆数の負の対数（基数 2）として計算される）。パフォーマンスは、訓練データの最初の 4 時間はより急速に改善し、その後の 5 時間はそれほど急速ではなかったが、プラトーではなかった。利用可能な全ての孤立単語データを使用する場合、情報転送速度は 25.1 ビット／分であった（図示されていない）。図 12 B は、孤立単語データを用いた発話検出及び検出されたイベントキュレーション中の検出誤りの頻度に対する訓練データの量の影響を示す。誤り率が低いほど、パフォーマンスがより良好であることを示す。偽陽性は、単語生成試行に関連付けられていない検出されたイベントであり、偽陰性は、検出されたイベントに関連付けられていない単語生成試行である。各点は、10 交差検証フォールドにわたる平均±標準偏差を示す。利用可能な訓練データの全てを、各発話検出モデルに適合させるために使用したわけではないが、各モデルは、常に 47～83 分間のデ

10

20

30

40

50

ータを使用した（図示しない）。図12Cは、ゴーキュー（100msヒストグラムビンサイズ）に対する9000の孤立単語試行にわたる神経活動から検出された開始の分布を示す。このヒストグラムは、学習曲線スキーム（利用可能な全ての試験が交差検証評価に含まれていた）の最終分析セットの結果を使用して作成した。検出された発話開始の分布は、関連付けられるゴーキューの後の平均が308msであり、標準偏差が1017msであった。この分布は、参加者の応答時間の挙動変動によってある程度影響を受けた可能性が高い。検出されたイベントキュレーション中に、429の試験が複数の候補から検出されたイベントを選択するためにキュレーションを必要とした（420の試験には2つの候補があり、9つの試験には3つの候補があった）。

【図12B】孤立単語データを用いた補助モデリング結果。図12Aは、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度（左）及び交差エントロピー損失（右）に対する訓練データの量の影響を示す。交差エントロピーが低いほど、パフォーマンスがより良好であることが指示される。各点は、10交差検証フォールドにわたる平均±標準偏差を示す（交差エントロピープロットにおける誤差バーは、典型的には、円形マーカーと並んで見ることができないほど小さかった）。偶然パフォーマンスが、各プロットにおいて水平破線として示されている（偶然交差エントロピー損失は、単語標的の数の逆数の負の対数（基数2）として計算される）。パフォーマンスは、訓練データの最初の4時間はより急速に改善し、その後の5時間はそれほど急速ではなかったが、プラトーではなかった。利用可能な全ての孤立単語データを使用する場合、情報転送速度は25.1ビット／分であった（図示されていない）。図12Bは、孤立単語データを用いた発話検出及び検出されたイベントキュレーション中の検出誤りの頻度に対する訓練データの量の影響を示す。誤り率が低いほど、パフォーマンスがより良好であることを示す。偽陽性は、単語生成試行に関連付けられていない検出されたイベントであり、偽陰性は、検出されたイベントに関連付けられていない単語生成試行である。各点は、10交差検証フォールドにわたる平均±標準偏差を示す。利用可能な訓練データの全てを、各発話検出モデルに適合させるために使用したわけではないが、各モデルは、常に47～83分間のデータを使用した（図示しない）。図12Cは、ゴーキュー（100msヒストグラムビンサイズ）に対する9000の孤立単語試行にわたる神経活動から検出された開始の分布を示す。このヒストグラムは、学習曲線スキーム（利用可能な全ての試験が交差検証評価に含まれていた）の最終分析セットの結果を使用して作成した。検出された発話開始の分布は、関連付けられるゴーキューの後の平均が308msであり、標準偏差が1017msであった。この分布は、参加者の応答時間の挙動変動によってある程度影響を受けた可能性が高い。検出されたイベントキュレーション中に、429の試験が複数の候補から検出されたイベントを選択するためにキュレーションを必要とした（420の試験には2つの候補があり、9つの試験には3つの候補があった）。

【図12C】孤立単語データを用いた補助モデリング結果。図12Aは、参加者の孤立単語生成試行中に記録された皮質活動を使用した単語分類正確度（左）及び交差エントロピー損失（右）に対する訓練データの量の影響を示す。交差エントロピーが低いほど、パフォーマンスがより良好であることが指示される。各点は、10交差検証フォールドにわたる平均±標準偏差を示す（交差エントロピープロットにおける誤差バーは、典型的には、円形マーカーと並んで見ることができないほど小さかった）。偶然パフォーマンスが、各プロットにおいて水平破線として示されている（偶然交差エントロピー損失は、単語標的の数の逆数の負の対数（基数2）として計算される）。パフォーマンスは、訓練データの最初の4時間はより急速に改善し、その後の5時間はそれほど急速ではなかったが、プラトーではなかった。利用可能な全ての孤立単語データを使用する場合、情報転送速度は25.1ビット／分であった（図示されていない）。図12Bは、孤立単語データを用いた発話検出及び検出されたイベントキュレーション中の検出誤りの頻度に対する訓練データの量の影響を示す。誤り率が低いほど、パフォーマンスがより良好であることを示す。偽陽性は、単語生成試行に関連付けられていない検出されたイベントであり、偽陰性は、検出されたイベントに関連付けられていない単語生成試行である。各点は、10交差検証フ

10

20

30

40

50

ールドにわたる平均±標準偏差を示す。利用可能な訓練データの全てを、各発話検出モデルに適合させるために使用したわけではないが、各モデルは、常に47～83分間のデータを使用した（図示しない）。図12Cは、ゴーキュー（100msヒストグラムビンサイズ）に対する9000の孤立単語試行にわたる神経活動から検出された開始の分布を示す。このヒストグラムは、学習曲線スキーム（利用可能な全ての試験が交差検証評価に含まれていた）の最終分析セットの結果を使用して作成した。検出された発話開始の分布は、関連付けられるゴーキューの後の平均が308msであり、標準偏差が1017msであった。この分布は、参加者の応答時間の挙動変動によってある程度影響を受けた可能性が高い。検出されたイベントキュレーション中に、429の試験が複数の候補から検出されたイベントを選択するためにキュレーションを必要とした（420の試験には2つの候補があり、9つの試験には3つの候補があった）。

10

【図13】音響汚染調査。各青色曲線は、周波数の関数としての、単一の電極からのスペクトログラムと、時間位置合わせされたマイクロフォン信号からの対応するスペクトログラムとの間の平均相関を示す。赤色曲線は、マイクロフォン信号の平均パワースペクトル密度（PSD）を示す。垂直破線は、60Hz回線ノイズ周波数及びその高調波をマークしている。緑色で強調表示されているのは、復号中に使用される神経特徴を抽出した周波数帯域であった高ガンマ周波数帯（70～150Hz）である。全ての周波数にわたって、電極とマイクロフォン信号との間の相関は小さい。高ガンマ周波数範囲の下限では相関がわずかに増大するが、この相関の増大はマイクロフォンPSDが減少するにつれて生じる。相関は低く、マイクロフォンPSDとともに増大又は減少はしないため、観察された相関は、共有された電気ノイズなどの音響汚染以外の要因に起因する可能性が高い。これらの結果を、音響汚染を記述した研究（ここで使用した汚染分析の情報を提供した）[39]で観察された結果と比較した後、本発明者らは、本発明の復号パフォーマンスが本発明の電気生理学的記録の音響汚染によって人為的に改善されなかったと結論づけた。

20

【図14A】発話誘発信号の長期安定性。図14Aは、81週の記録に及ぶ、孤立単語タスク中に「Goodbye（さようなら）」という単語を発語する参加者の全ての試行にわたる単一の電極からの神経活動を示す。図14Bは、電極位置と重ねられた参加者の脳再構築を示す。パネルAに示されている電極は黒色で埋められている。解剖学的参照のために、中心前回は明青色で強調表示されている。図14Cは、4つの重複しない日付範囲からサンプリングされた、孤立単語データのサブセットに対する検出器及び分類器の訓練及びテストからの単語分類結果を示す。各サブセットには、各単語の20の試行された生成からのデータが含まれている。各実線バーは、単一のサブセット内の交差検証された評価からの結果を示し、各点線バーは、評価されているサブセットを除く全てのサブセットからのデータに対する訓練からの結果を示す。各誤差バーは、交差検証フォールドにわたって計算された、平均の95%信頼区間を示す。偶然正確度が、水平破線として描かれている。単一のサブセット内の交差検証された評価中に計算された電極寄与が、上部に示される（右上隅に最も背面及び後方の電極を配向される）。プロットされた電極のサイズ（面積）及び不透明度は、相対的な寄与によってスケールされる。各寄与セットは、合計が1になるように正規化されている。これらの結果は、発話誘発皮質応答が研究期間全体を通して比較的安定したままであったことを示唆しているが、2～3ヶ月ごとのモデル再較正は、依然として復号パフォーマンスにとって有益であり得る。

30

40

【図14B】発話誘発信号の長期安定性。図14Aは、81週の記録に及ぶ、孤立単語タスク中に「Goodbye（さようなら）」という単語を発語する参加者の全ての試行にわたる単一の電極からの神経活動を示す。図14Bは、電極位置と重ねられた参加者の脳再構築を示す。パネルAに示されている電極は黒色で埋められている。解剖学的参照のために、中心前回は明青色で強調表示されている。図14Cは、4つの重複しない日付範囲からサンプリングされた、孤立単語データのサブセットに対する検出器及び分類器の訓練及びテストからの単語分類結果を示す。各サブセットには、各単語の20の試行された生成からのデータが含まれている。各実線バーは、単一のサブセット内の交差検証された評価からの結果を示し、各点線バーは、評価されているサブセットを除く全てのサブセット

50

からのデータに対する訓練からの結果を示す。各誤差バーは、交差検証フォールドにわたって計算された、平均の95%信頼区間を示す。偶然正確度が、水平破線として描かれている。単一のサブセット内の交差検証された評価中に計算された電極寄与が、上部に示される（右上隅に最も背面及び後方の電極を配向される）。プロットされた電極のサイズ（面積）及び不透明度は、相対的な寄与によってスケールされる。各寄与セットは、合計が1になるように正規化されている。これらの結果は、発話誘発皮質応答が研究期間全体を通して比較的安定したままであったことを示唆しているが、2〜3ヶ月ごとのモデル再較正は、依然として復号パフォーマンスにとって有益であり得る。

【図14C】発話誘発信号の長期安定性。図14Aは、81週の記録に及ぶ、孤立単語タスク中に「Goodbye（さようなら）」という単語を発語する参加者の全ての試行にわたる単一の電極からの神経活動を示す。図14Bは、電極位置と重ねられた参加者の脳再構築を示す。パネルAに示されている電極は黒色で埋められている。解剖学的参照のために、中心前回が明青色で強調表示されている。図14Cは、4つの重複しない日付範囲からサンプリングされた、孤立単語データのサブセットに対する検出器及び分類器の訓練及びテストからの単語分類結果を示す。各サブセットには、各単語の20の試行された生成からのデータが含まれている。各実線バーは、単一のサブセット内の交差検証された評価からの結果を示し、各点線バーは、評価されているサブセットを除く全てのサブセットからのデータに対する訓練からの結果を示す。各誤差バーは、交差検証フォールドにわたって計算された、平均の95%信頼区間を示す。偶然正確度が、水平破線として描かれている。単一のサブセット内の交差検証された評価中に計算された電極寄与が、上部に示される（右上隅に最も背面及び後方の電極を配向される）。プロットされた電極のサイズ（面積）及び不透明度は、相対的な寄与によってスケールされる。各寄与セットは、合計が1になるように正規化されている。これらの結果は、発話誘発皮質応答が研究期間全体を通して比較的安定したままであったことを示唆しているが、2〜3ヶ月ごとのモデル再較正は、依然として復号パフォーマンスにとって有益であり得る。

【図15】スペリングパイプラインの概略図。A. 文スペリング試験の開始時に、参加者は、自発的にスペラを作動させるために、黙って単語を発語することを試行する。B. 神経特徴（高ガンマ活性及び低周波信号）が、タスク全体を通して記録された皮質データからリアルタイムで抽出される。単一の電極（図19Aに示される電極0）からの特徴が示されている。視覚化のために、トレースは、150ミリ秒の標準偏差を有するガウスカーネルを用いた畳み込みを介して平滑化された。マイクロフォン信号は、タスク中に音声出力がないことを示している。C. 再帰ニューラルネットワーク（RNN）及び閾値化演算から構成される発話検出モデルが、神経特徴をサンプルごとに処理して、沈黙発話試行を検出する。試行が検出されると、検出モデルは非アクティブになり、スペリング手順が開始される。D. スペリング手順中、参加者は、2.5秒ごとに発生する文字復号サイクル全体を通じて、意図されたメッセージをスペルアウトする。各サイクルにおいて、参加者にカウントダウンが視覚的に提示され、最終的にゴーキューが提示される。ゴーキューにおいて、参加者は所望の文字を表すコードワードを黙って発語することを試行する。E. 高ガンマ活性及び低周波数信号が、全ての電極チャンネルについてスペリング手順を通して計算され、文字復号サイクルに対応する2.5秒の重複しないタイムウィンドウに分割される。F. RNNベースの文字分類モデルが、これらの神経タイムウィンドウの各々を処理して、参加者が26個の可能性のあるコードワードの各々を黙って発語することを試行していたか、又は手動コマンド（Gを参照）の実施を試行していた確率を予測する。参加者が少なくとも80%の確率で手動コマンドを実施していたと分類器が予測した場合、スペリング手順は終了し、文は確定される（Iを参照）。それ以外の場合、予測文字確率が、リアルタイムでビーム探索アルゴリズムによって処理され、最も可能性の高い文が参加者に表示される。G. 参加者が、意図したメッセージをスペルアウトした後、次の文字復号サイクル中に右手を握ってスペリング手順を終了し、文を完成させることを試行する。H. 手動コマンドと関連付けられる神経タイムウィンドウが、分類モデルに渡される。I. 参加者が手動コマンドを試行したことを分類器が確認した場合、ニューラルネットワー

10

20

30

40

50

クベースの言語モデル（「D i s t i l G P T - 2」）が、完全な単語のみで構成された文を再スコアリングし、システムが、最終予測として再スコアリング後に最も可能性の高い文を使用する。

【図16A】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A．リアルタイム文スペリング（「+LM（リアルタイム結果）」と表記）及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率（CER）。「偶然（Chance）」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ（Only neural decoding）」条件において、文は、（空白文字を一切含まない）文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab．制約（+Vocab． constraints）」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1, 152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM（リアルタイム結果）（+LM（Real-time results）」）とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モデリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した（*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図16B．図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率（WER）。図16C．リアルタイムテスト中に1分間あたりに復号された文字数。図16D．リアルタイムテスト中に1分間あたりに復号された単語数。図16A～16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており（各ブロックにおいて、参加者は2～5文のスペリングを試行した）、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～5文のスペリングを試行した）。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～4の会話応答のスペリングを試行した）。図16E．各復号された文の超過文字の数。超過文字が0文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド（スペラからディスエンゲージするための）が参加者の神経活動から首尾よく識別されたことを示している。図16F．各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1及び2は、リアルタイムで復号された文が少なくとも1つの誤りを含んでいた試験をマークしている。これら2つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図16B】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A．リアルタイム文スペリング（「+LM（リアルタイム結果）」と表記）及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率（CER）。「偶然（Chance）」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ（Only neural decoding）」条件において、文は、（空白文字を一切含まない）文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab．制約（+Vocab． constraints）」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1, 152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM（リアルタイム結果）（+LM（Real-time results）」）とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モ

デリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した（*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図16B・図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率（WER）。図16C・リアルタイムテスト中に1分間あたりに復号された文字数。図16D・リアルタイムテスト中に1分間あたりに復号された単語数。図16A～16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており（各ブロックにおいて、参加者は2～5文のスペリングを試行した）、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～5文のスペリングを試行した）。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～4の会話応答のスペリングを試行した）。図16E・各復号された文の超過文字の数。超過文字が0文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド（スペラからディスエンゲージするための）が参加者の神経活動から首尾よく識別されたことを示している。図16F・各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1及び2は、リアルタイムで復号された文が少なくとも1つの誤りを含んでいた試験をマークしている。これら2つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図16C】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A・リアルタイム文スペリング（「+LM（リアルタイム結果）」と表記）及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率（CER）。「偶然（Chance）」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ（Only neural decoding）」条件において、文は、（空白文字を一切含まない）文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab・制約（+Vocab・constraints）」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1,152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM（リアルタイム結果）（+LM（Real-time results）」とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モデリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した（*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図16B・図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率（WER）。図16C・リアルタイムテスト中に1分間あたりに復号された文字数。図16D・リアルタイムテスト中に1分間あたりに復号された単語数。図16A～16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており（各ブロックにおいて、参加者は2～5文のスペリングを試行した）、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～5文のスペリングを試行した）。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～4の会話応答のスペリングを試行した）。図16E・各復号された文の超過文字の数。超過文字が0文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド（スペラからディスエンゲージするための）が参加者

の神経活動から首尾よく識別されたことを示している。図16F．各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1及び2は、リアルタイムで復号された文が少なくとも1つの誤りを含んでいた試験をマークしている。これら2つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図16D】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A．リアルタイム文スペリング（「+LM（リアルタイム結果）」と表記）及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率（CER）。「偶然（Chance）」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ（Only neural decoding）」条件において、文は、（空白文字を一切含まない）文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab．制約（+Vocab． constraints）」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1, 152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM（リアルタイム結果）（+LM（Real-time results）」とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モデリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した（*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図16B．図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率（WER）。図16C．リアルタイムテスト中に1分間あたりに復号された文字数。図16D．リアルタイムテスト中に1分間あたりに復号された単語数。図16A～16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており（各ブロックにおいて、参加者は2～5文のスペリングを試行した）、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～5文のスペリングを試行した）。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する（これらのブロックの各々において、参加者は、2～4の会話応答のスペリングを試行した）。図16E．各復号された文の超過文字の数。超過文字が0文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド（スペラからディスエンゲージするための）が参加者の神経活動から首尾よく識別されたことを示している。図16F．各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1及び2は、リアルタイムで復号された文が少なくとも1つの誤りを含んでいた試験をマークしている。これら2つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図16E】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A．リアルタイム文スペリング（「+LM（リアルタイム結果）」と表記）及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率（CER）。「偶然（Chance）」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ（Only neural decoding）」条件において、文は、（空白文字を一切含まない）文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab．制約（+Vocab． constraints）」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1, 152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM（リアルタイム結果）（+L

M (Real-time results)」とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モデリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した(*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図16B、図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率(WER)。図16C、リアルタイムテスト中に1分間あたりに復号された文字数。図16D、リアルタイムテスト中に1分間あたりに復号された単語数。図16A~16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており(各ブロックにおいて、参加者は2~5文のスペリングを試行した)、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する(これらのブロックの各々において、参加者は、2~5文のスペリングを試行した)。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する(これらのブロックの各々において、参加者は、2~4の会話応答のスペリングを試行した)。図16E、各復号された文の超過文字の数。超過文字が0文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド(スペラからディスエンゲージするための)が参加者の神経活動から首尾よく識別されたことを示している。図16F、各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1及び2は、リアルタイムで復号された文が少なくとも1つの誤りを含んでいた試験をマークしている。これら2つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図16F】コピータイピングタスク中のスペリングシステムのパフォーマンス概要。図16A、リアルタイム文スペリング(「+LM(リアルタイム結果)」と表記)及びスペリングシステムの部分が省略されたオフラインシミュレーション中に観察された文字誤り率(CER)。「偶然(Chance)」条件において、文は、スペリングパイプラインの残り部分を変更することなく、神経分類器からの出力をランダムに生成された文字確率で置き換えることによって作成された。「神経復号のみ(Only neural decoding)」条件において、文は、(空白文字を一切含まない)文試験中の分類器の予測の各々から最も可能性の高い文字を連結することによってのみ作成された。「+Vocab.制約(+Vocab. constraints)」条件において、神経分類器からの予測された文字確率が、予測された文字系列を制約して1,152語の語彙内から単語を形成するビーム探索とともに使用された。「+LM(リアルタイム結果)(+LM(Real-time results))」とラベル付けされた最終条件は、参加者によるテスト中のリアルタイム結果を示し、ビーム探索中及び文が確定した後の言語モデリングを組み込んでいる。完全なシステムによってリアルタイムに復号された文は、他の条件で復号された文よりも低いCERを呈した(*** $P < 0.0001$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図16B、図16Aからのリアルタイム結果及び対応するオフライン省略シミュレーションの単語誤り率(WER)。図16C、リアルタイムテスト中に1分間あたりに復号された文字数。図16D、リアルタイムテスト中に1分間あたりに復号された単語数。図16A~16Dにおいて、分布は、各ボックスプロットが $n = 34$ 個のリアルタイムブロックにわたって計算されたことを記述しており(各ブロックにおいて、参加者は2~5文のスペリングを試行した)、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図16A及び16Bにおいて、各ボックスプロットは、 $n = 34$ 個のブロックに対応する(これらのブロックの各々において、参加者は、2~5文のスペリングを試行した)。図16Cにおいて、各ボックスプロットは、 $n = 9$ 個のブロックに対応する(これらのブロックの各々において、参加者は、2~4の会話応答のスペリングを試行した)。図16E、各復号された文

の超過文字の数。超過文字が 0 文字の復号された文は、参加者がその文の最後の文字をスペリングした直後に、手動コマンド（スペラからディスエンゲージするための）が参加者の神経活動から首尾よく識別されたことを示している。図 16 F、各非偶然条件からの復号された文を用いた例示的な文スペリング試験。誤った文字は赤色に着色されている。1 及び 2 は、リアルタイムで復号された文が少なくとも 1 つの誤りを含んでいた試験をマークしている。これら 2 つの試験の標的文は、パネルの下部に与えられている。他の全ての例示的な文は、リアルタイム復号誤りを一切含まなかった。

【図 17 A】沈黙発話試行中の高ガンマ活性（HGA）及び低周波数信号（LFS）の特徴付け。図 17 A、HGA 単独を使用した場合、LFS 単独を使用した場合、及び HGA + LFS の両方を同時に使用した場合の、黙って試行された NATO コードワードの 10 フォールド交差検証分類正確度を示す図である。LFS のみを使用した分類正確度は、HGA のみを使用した分類正確度よりも有意に高く、HGA + LFS 結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる（** $P < 0.001$ 、3 方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。偶然正確度は 3.7% である。各ボックスプロットは、四分位間範囲の 1.5 倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図 17 B、HGA 特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図 17 C、結合 HGA + LFS 特徴セットを使用して訓練された分類モデルからの HGA 特徴と関連付けられる電極寄与を示す図である。図 17 D、LFS 特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図 17 E、結合 HGA + LFS 特徴セットを使用して訓練された分類モデルからの LFS 特徴と関連付けられる電極寄与を示す図である。図 17 B ~ 17 E において、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図 17 F、100 回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の 80% 超を説明するために必要な最小限の主成分（PC）数を示す図である。必要な PC の数は、特徴セットごとに有意に異なっていた（** $P < 0.0001$ 、3 方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3 方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図 17 G、100 回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の 80% 超を説明するために必要な最小限の PC 数を示す図である。図 17 F 及び 17 G において、各特徴セットに必要な PC の数は、ヒストグラムとして示され、ここで、x 軸は、特定の数の PC を必要としたブートストラップ反復の割合である。図 17 H、時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの 99% 信頼区間を表す。

【図 17 B】沈黙発話試行中の高ガンマ活性（HGA）及び低周波数信号（LFS）の特徴付け。図 17 A、HGA 単独を使用した場合、LFS 単独を使用した場合、及び HGA + LFS の両方を同時に使用した場合の、黙って試行された NATO コードワードの 10 フォールド交差検証分類正確度を示す図である。LFS のみを使用した分類正確度は、HGA のみを使用した分類正確度よりも有意に高く、HGA + LFS 結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる（** $P < 0.001$ 、3 方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。偶然正確度は 3.7% である。各ボックスプロットは、四分位間範囲の 1.5 倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図 17 B、HGA 特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図 17

C. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのHGA特徴と関連付けられる電極寄与を示す図である。図17D. LFS特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17E. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのLFS特徴と関連付けられる電極寄与を示す図である。図17B~17Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分(PC)数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17C】沈黙発話試行中の高ガンマ活性(HGA)及び低周波数信号(LFS)の特徴付け。図17A. HGA単独を使用した場合、LFS単独を使用した場合、及びHGA+LFSの両方を同時に使用した場合の、黙って試行されたNATOコードワードの10フォールド交差検証分類正確度を示す図である。LFSのみを使用した分類正確度は、HGAのみを使用した分類正確度よりも有意に高く、HGA+LFS結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n=10$ 交差検証フォールドに対応する。図17B. HGA特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17C. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのHGA特徴と関連付けられる電極寄与を示す図である。図17D. LFS特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17E. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのLFS特徴と関連付けられる電極寄与を示す図である。図17B~17Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分(PC)数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17D】沈黙発話試行中の高ガンマ活性(HGA)及び低周波数信号(LFS)の特徴付け。図17A. HGA単独を使用した場合、LFS単独を使用した場合、及びHGA+LFSの両方を同時に使用した場合の、黙って試行されたNATOコードワードの10

フォールド交差検証分類正確度を示す図である。L F Sのみを使用した分類正確度は、H G Aのみを使用した分類正確度よりも有意に高く、H G A + L F S結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる（** $P < 0.001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図17 B. H G A特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17 C. 結合H G A + L F S特徴セットを使用して訓練された分類モデルからのH G A特徴と関連付けられる電極寄与を示す図である。図17 D. L F S特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17 E. 結合H G A + L F S特徴セットを使用して訓練された分類モデルからのL F S特徴と関連付けられる電極寄与を示す図である。図17 B～17 Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17 F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分（PC）数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた（*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。図17 G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17 F及び17 Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17 H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17 E】沈黙発話試行中の高ガンマ活性（H G A）及び低周波数信号（L F S）の特徴付け。図17 A. H G A単独を使用した場合、L F S単独を使用した場合、及びH G A + L F Sの両方を同時に使用した場合の、黙って試行されたN A T Oコードワードの10フォールド交差検証分類正確度を示す図である。L F Sのみを使用した分類正確度は、H G Aのみを使用した分類正確度よりも有意に高く、H G A + L F S結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる（** $P < 0.001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図17 B. H G A特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17 C. 結合H G A + L F S特徴セットを使用して訓練された分類モデルからのH G A特徴と関連付けられる電極寄与を示す図である。図17 D. L F S特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17 E. 結合H G A + L F S特徴セットを使用して訓練された分類モデルからのL F S特徴と関連付けられる電極寄与を示す図である。図17 B～17 Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17 F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分（PC）数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた（*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位

10

20

30

40

50

和検定、* $P < .01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17F】沈黙発話試行中の高ガンマ活性(HGA)及び低周波数信号(LFS)の特徴付け。図17A. HGA単独を使用した場合、LFS単独を使用した場合、及びHGA + LFSの両方を同時に使用した場合の、黙って試行されたNATOコードワードの10フォールド交差検証分類正確度を示す図である。LFSのみを使用した分類正確度は、HGAのみを使用した分類正確度よりも有意に高く、HGA + LFS結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる(** $P < 0.001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図17B. HGA特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17C. 結合HGA + LFS特徴セットを使用して訓練された分類モデルからのHGA特徴と関連付けられる電極寄与を示す図である。図17D. LFS特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17E. 結合HGA + LFS特徴セットを使用して訓練された分類モデルからのLFS特徴と関連付けられる電極寄与を示す図である。図17B~17Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分(PC)数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた(** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < .01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17G】沈黙発話試行中の高ガンマ活性(HGA)及び低周波数信号(LFS)の特徴付け。図17A. HGA単独を使用した場合、LFS単独を使用した場合、及びHGA + LFSの両方を同時に使用した場合の、黙って試行されたNATOコードワードの10フォールド交差検証分類正確度を示す図である。LFSのみを使用した分類正確度は、HGAのみを使用した分類正確度よりも有意に高く、HGA + LFS結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる(** $P < 0.001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。図17B. HGA特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17

10

20

30

40

50

C. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのHGA特徴と関連付けられる電極寄与を示す図である。図17D. LFS特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17E. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのLFS特徴と関連付けられる電極寄与を示す図である。図17B~17Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分(PC)数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図17H】沈黙発話試行中の高ガンマ活性(HGA)及び低周波数信号(LFS)の特徴付け。図17A. HGA単独を使用した場合、LFS単独を使用した場合、及びHGA+LFSの両方を同時に使用した場合の、黙って試行されたNATOコードワードの10フォールド交差検証分類正確度を示す図である。LFSのみを使用した分類正確度は、HGAのみを使用した分類正確度よりも有意に高く、HGA+LFS結果を使用した分類正確度は、いずれかの特徴タイプのみを使用した分類正確度よりも有意に高くなる(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。偶然正確度は3.7%である。各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。各ボックスプロットは、 $n=10$ 交差検証フォールドに対応する。図17B. HGA特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供する。図17C. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのHGA特徴と関連付けられる電極寄与を示す図である。図17D. LFS特徴のみを使用して訓練された分類モデルからの電極寄与を示す図である。図17E. 結合HGA+LFS特徴セットを使用して訓練された分類モデルからのLFS特徴と関連付けられる電極寄与を示す図である。図17B~17Eにおいて、プロットされた電極のサイズ及び不透明度は、相対的な寄与によってスケールされ、より大きく、より不透明に見える電極は、より重要な特徴を分類モデルに提供した。図17F. 100回以上のブートストラップ反復にわたる、各特徴セットの空間次元の分散の80%超を説明するために必要な最小限の主成分(PC)数を示す図である。必要なPCの数は、特徴セットごとに有意に異なっていた(*** $P < 0.0001$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、* $P < 0.01$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。図17G. 100回以上のブートストラップ反復にわたる、各特徴セットの時間次元の分散の80%超を説明するために必要な最小限のPC数を示す図である。図17F及び17Gにおいて、各特徴セットに必要なPCの数は、ヒストグラムとして示され、ここで、x軸は、特定の数のPCを必要としたブートストラップ反復の割合である。図17H. 時間的平滑化が分類正確度に及ぼす影響。各点は、中央値を表し、誤差バーは、中央値のブートストラップされた推定の周りの99%信頼区間を表す。

【図18A】英語文字及びNATOコードワードを黙って発語することを試行している間の神経信号の比較。図18A. HGA+LFS特徴によって訓練されたモデルを使用した分類正確度($n=10$ 交差検証フォールド)は、英語文字よりもNATOコードワードに

10

20

30

40

50

ついて有意に高い（** $P < 0.001$ 、両側ウィルコクソン順位和検定）。点線の水平線は偶然正確度を表す。図18B．結合HGA+LFS特徴セットの最近クラス距離は、文字よりもNATOコードワードの方が有意に大きい（ボックスプロットは、 $n = 26$ 個のコードワード又は文字にわたる値を示す。 $*P < 0.01$ 、両側ウィルコクソン順位和検定）。図18A及び18Bにおいて、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図18C．最近クラス距離は、対応する文字よりもコードワードの大部分で大きい。図18B及び図18Cにおいて、最近クラス距離は、試験平均HGA+LFS特徴間のフロベニウスノルムとして計算される。

【図18B】英語文字及びNATOコードワードを黙って発語することを試行している間の神経信号の比較。図18A．HGA+LFS特徴によって訓練されたモデルを使用した分類正確度（ $n = 10$ 交差検証フォールド）は、英語文字よりもNATOコードワードについて有意に高い（** $P < 0.001$ 、両側ウィルコクソン順位和検定）。点線の水平線は偶然正確度を表す。図18B．結合HGA+LFS特徴セットの最近クラス距離は、文字よりもNATOコードワードの方が有意に大きい（ボックスプロットは、 $n = 26$ 個のコードワード又は文字にわたる値を示す。 $*P < 0.01$ 、両側ウィルコクソン順位和検定）。図18A及び18Bにおいて、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図18C．最近クラス距離は、対応する文字よりもコードワードの大部分で大きい。図18B及び図18Cにおいて、最近クラス距離は、試験平均HGA+LFS特徴間のフロベニウスノルムとして計算される。

【図18C】英語文字及びNATOコードワードを黙って発語することを試行している間の神経信号の比較。図18A．HGA+LFS特徴によって訓練されたモデルを使用した分類正確度（ $n = 10$ 交差検証フォールド）は、英語文字よりもNATOコードワードについて有意に高い（** $P < 0.001$ 、両側ウィルコクソン順位和検定）。点線の水平線は偶然正確度を表す。図18B．結合HGA+LFS特徴セットの最近クラス距離は、文字よりもNATOコードワードの方が有意に大きい（ボックスプロットは、 $n = 26$ 個のコードワード又は文字にわたる値を示す。 $*P < 0.01$ 、両側ウィルコクソン順位和検定）。図18A及び18Bにおいて、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図18C．最近クラス距離は、対応する文字よりもコードワードの大部分で大きい。図18B及び図18Cにおいて、最近クラス距離は、試験平均HGA+LFS特徴間のフロベニウスノルムとして計算される。

【図19A】公然の発話試行と沈黙発話試行との間の神経信号及び分類パフォーマンスの差。図19A．埋め込まれた電極の位置で覆われた参加者の脳のMRI再構成。図19B及び図19Cにおいて使用される電極の位置は、太字になっており、オーバーレイに番号が付されている。図19B．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「k i l o（キロ）」の発語を試行する。図19C．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「t a n g o（タンゴ）」の発語を試行する。図19B及びCの誘発された応答は、時間0において垂直の破線としてマークされているゴーキューに位置合わせされる。各曲線は、 $n = 100$ の発話試行にわたる平均±標準誤差を示す。図19D．様々なモデル訓練スキームによるコードワード分類正確度（10交差検証フォールドにわたる）。全ての比較は、「n s」としてマークされたものを除いて、結果ペア間の有意差を明らかにした（ $P < 0.01$ 、28方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和）。各ボックスプロットは、 $n = 10$ 交差検証フォールドに対応する。偶然正確度は3.84%である。

【図19B】公然の発話試行と沈黙発話試行との間の神経信号及び分類パフォーマンスの差。図19A．埋め込まれた電極の位置で覆われた参加者の脳のMRI再構成。図19B及び図19Cにおいて使用される電極の位置は、太字になっており、オーバーレイに番号

10

20

30

40

50

が付されている。図19B．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「k i l o（キロ）」の発語を試行する。図19C．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「t a n g o（タンゴ）」の発語を試行する。図19B及びCの誘発された応答は、時間0において垂直の破線としてマークされているゴーキューに位置合わせされる。各曲線は、 $n=100$ の発話試行にわたる平均±標準誤差を示す。図19D．様々なモデル訓練スキームによるコードワード分類正確度（10交差検証フォールドにわたる）。全ての比較は、「ns」としてマークされたものを除いて、結果ペア間の有意差を明らかにした（ $P<0.01$ 、28方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和）。各ボックスプロットは、 $n=10$ 交差検証フォールドに対応する。偶然正確度は3.84%である。

10

【図19C】公然の発話試行と沈黙発話試行との間の神経信号及び分類パフォーマンスの差。図19A．埋め込まれた電極の位置で覆われた参加者の脳のMRI再構成。図19B及び図19Cにおいて使用される電極の位置は、太字になっており、オーバーレイに番号が付されている。図19B．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「k i l o（キロ）」の発語を試行する。図19C．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「t a n g o（タンゴ）」の発語を試行する。図19B及びCの誘発された応答は、時間0において垂直の破線としてマークされているゴーキューに位置合わせされる。各曲線は、 $n=100$ の発話試行にわたる平均±標準誤差を示す。図19D．様々なモデル訓練スキームによるコードワード分類正確度（10交差検証フォールドにわたる）。全ての比較は、「ns」としてマークされたものを除いて、結果ペア間の有意差を明らかにした（ $P<0.01$ 、28方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和）。各ボックスプロットは、 $n=10$ 交差検証フォールドに対応する。偶然正確度は3.84%である。

20

【図19D】公然の発話試行と沈黙発話試行との間の神経信号及び分類パフォーマンスの差。図19A．埋め込まれた電極の位置で覆われた参加者の脳のMRI再構成。図19B及び図19Cにおいて使用される電極の位置は、太字になっており、オーバーレイに番号が付されている。図19B．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「k i l o（キロ）」の発語を試行する。図19C．沈黙（オレンジ色）と公然（緑色）との間の高ガンマ活性（HGA）イベント関連電位は、NATOのコードワード「t a n g o（タンゴ）」の発語を試行する。図19B及びCの誘発された応答は、時間0において垂直の破線としてマークされているゴーキューに位置合わせされる。各曲線は、 $n=100$ の発話試行にわたる平均±標準誤差を示す。図19D．様々なモデル訓練スキームによるコードワード分類正確度（10交差検証フォールドにわたる）。全ての比較は、「ns」としてマークされたものを除いて、結果ペア間の有意差を明らかにした（ $P<0.01$ 、28方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和）。各ボックスプロットは、 $n=10$ 交差検証フォールドに対応する。偶然正確度は3.84%である。

30

【図20A】スペリングアプローチは、より大きい語彙及び会話の設定に一般化することができる。図20A．リアルタイム復号中に使用されている元の語彙を含む、異なる語彙によるコピータイピングタスクからのシミュレートされた文字誤り率。図20B．図20Aの対応するシミュレーションからの単語誤り率。図20C．会話タスク条件中にリアルタイムで復号された意志的に選択された応答及びメッセージにわたる文字及び単語の誤り率。図20A～20Cにおいて、各ボックスプロットは、四分位間範囲の1.5倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図20D．参加者の脳活動から復号された対応する応答（右）を伴う、会話タスク条件の試験から提示された質問の例（左）。最後の例において、参加者は質問によって促されることなく、意図したメッセージをスペルアウトした。

40

【図20B】スペリングアプローチは、より大きい語彙及び会話の設定に一般化すること

50

ができる。図 20 A . リアルタイム復号中に使用されている元の語彙を含む、異なる語彙によるコピータイピングタスクからのシミュレートされた文字誤り率。図 20 B . 図 20 A の対応するシミュレーションからの単語誤り率。図 20 C . 会話タスク条件中にリアルタイムで復号された意志的に選択された応答及びメッセージにわたる文字及び単語の誤り率。図 20 A ~ 20 C において、各ボックスプロットは、四分位間範囲の 1 . 5 倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図 20 D . 参加者の脳活動から復号された対応する応答（右）を伴う、会話タスク条件の試験から提示された質問の例（左）。最後の例において、参加者は質問によって促されることなく、意図したメッセージをスペルアウトした。

【図 20 C】スプリングアプローチは、より大きい語彙及び会話の設定に一般化することができる。図 20 A . リアルタイム復号中に使用されている元の語彙を含む、異なる語彙によるコピータイピングタスクからのシミュレートされた文字誤り率。図 20 B . 図 20 A の対応するシミュレーションからの単語誤り率。図 20 C . 会話タスク条件中にリアルタイムで復号された意志的に選択された応答及びメッセージにわたる文字及び単語の誤り率。図 20 A ~ 20 C において、各ボックスプロットは、四分位間範囲の 1 . 5 倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図 20 D . 参加者の脳活動から復号された対応する応答（右）を伴う、会話タスク条件の試験から提示された質問の例（左）。最後の例において、参加者は質問によって促されることなく、意図したメッセージをスペルアウトした。

【図 20 D】スプリングアプローチは、より大きい語彙及び会話の設定に一般化することができる。図 20 A . リアルタイム復号中に使用されている元の語彙を含む、異なる語彙によるコピータイピングタスクからのシミュレートされた文字誤り率。図 20 B . 図 20 A の対応するシミュレーションからの単語誤り率。図 20 C . 会話タスク条件中にリアルタイムで復号された意志的に選択された応答及びメッセージにわたる文字及び単語の誤り率。図 20 A ~ 20 C において、各ボックスプロットは、四分位間範囲の 1 . 5 倍であるデータポイントを除く、分布の残りの部分を示すように伸びるひげを有するデータの四分位数を示す。図 20 D . 参加者の脳活動から復号された対応する応答（右）を伴う、会話タスク条件の試験から提示された質問の例（左）。最後の例において、参加者は質問によって促されることなく、意図したメッセージをスペルアウトした。

【図 21】データ収集タイムライン。各バーは、記録の各日に収集された試行の総数を示している。参加者及び埋め込みの日付は、本発明者らの以前の研究と同じである [2]。1 日に複数のタイプのデータセットが収集された場合、バーは収集された各データセットの割合によって色付けされる。各色は（凡例で指定されているような）特定のデータセットを表す。データセットは、タスクタイプ（孤立標的又はリアルタイム文スプリング）、発言セット（英語文字、N A T O コードワード（試行された手の握りを含む）、コピータイピング文、又は会話文）、及びリアルタイム文スプリングデータセットの場合、データの目的（ハイパーパラメータ最適化又はパフォーマンス評価のための）が異なる。全ての発話関連試験は、その凡例ラベルに「（公然）」があるデータセットを除いて、沈黙発話試行と関連付けられた。追加的に、この公然データセットの試行の 3 . 0 6 % は、参加者がコードワードの公然の生成を試行したバージョンのコピータイピング文スプリングタスクの中に実際に記録された（更なる詳細については、セクション S 3 を参照）。データセットは、神経埋め込みとは無関係であった外部及び臨床時間の制約に起因して不規則なスケジュールで収集された。5 5 ~ 8 8 週の隙間は、特に、対面での記録セッションを制限又は妨げた C O V I D - 1 9 パンデミックの開始中の臨床ガイドラインに起因した。

【図 22】リアルタイム信号処理パイプライン。経皮台座コネクタに取り付けられた取り外し可能データ取得ヘッドステージ（C e r e P l e x E , B l a c k r o c k M i c r o s y s t e m s）が、ハードウェアベースのワイドバンドバターワースフィルタ（0 . 3 H z ~ 7 . 5 k H z）を E C o G 信号に適用し、1 6 ビット、2 5 0 - n V / ビット分解能でデジタル化し、3 0 k H z で追加の接続を通じて N e u r o p o r t システム（B l a c k r o c k M i c r o s y s t e m s）に送信した。N e u r o p o r t シ

10

20

30

40

50

システムは、ソフトウェアベースの回線ノイズキャンセル及びアンチエイリアシングローパスフィルタ（500 Hz）を使用して信号を処理した。その後、処理された信号が、更なるリアルタイム処理及び分析のために1 kHzで別個のコンピュータにストリーミングされ、そこで、ECOGデータの各時間サンプルに（全ての電極チャンネルにわたって）共通平均基準を適用した。次いで、再参照された信号が2つの並列ストリームで処理されて、高ガンマ活性（HGA）及び低周波数信号（LFS）特徴が抽出された。HGA特徴を計算するために、再参照された信号に8つの390次バンドパス有限インパルス応答（FIR）フィルタを適用した（フィルタ中心周波数は、72.0、79.5、87.8、96.9、107.0、118.1、130.4、及び144.0 Hzの高ガンマ帯域内であった）。次いで、各チャンネル及びバンドについて、ヒルベルト変換を近似するために170次FIRフィルタを使用した。具体的には、各チャンネル及びバンドについて、分析信号の実成分を85サンプル（フィルタ次数の半分）遅延された元の信号に等しく設定し、虚成分を元の信号のヒルベルト変換（このFIRフィルタによって近似）に等しく設定した[25]。次いで、5つごとの時間サンプルにおける各分析信号の大きさを計算し、200 Hzの分析振幅信号を生成した。各チャンネルについて、各時点で8つのバンドにわたって分析振幅値を平均化して、そのチャンネルの単一の高ガンマ分析振幅測度を取得した。LFS特徴を計算するために、カットオフ周波数100 Hzの130次アンチエイリアシングローパスFIRフィルタを適用した後、再参照信号を200 Hzにダウンサンプリングした。次いで、2つの特徴ストリーム（高ガンマ分析振幅及びダウンサンプリング信号）からの時間同期値を単一の特徴ストリームに結合した。次に、30秒のスライドウィンドウでWelchの方法を使用して各チャンネル及び各特徴タイプの値をzスコアリングした[26]。最後に、大きいzスコアの大きさが非常に大きいサンプルが実行中のzスコア統計又はダウンストリーム復号プロセスに干渉するのを防ぐために、単純なアーチファクト除去手法を実施した。高ガンマ特徴を計算するために同様の前処理パイプラインを実装した以前の研究[2、27]からのこの図を適応させた。

【図23】発話検出モデル概略図。リアルタイム文スペリング中に参加者の神経活動から沈黙発話試行を検出するために、まず、各電極のzスコア低周波数信号（LFS）及び高ガンマ活性（HGA）が、3つの長・短期記憶（LSTM）層のスタックによって連続的に処理される。次に、単一の高密度（全結合）層が、最終的なLSTMの潜在次元を、発話、発話準備、休止、及び運動の4つの可能なクラスに投影する。次いで、発話確率のストリームが、時間的に平滑化され、確率閾値化され、時間閾値化されて、完全な発話イベントの開始及び終了をもたらす。参加者が黙って何かの発語を試行し、その発話試行が検出されると、スペリングシステムがエンゲージされ、ペーシングされたスペリング手順が開始される。描写されている神経特徴、予測発話確率時系列（右上）、及び検出された発話イベント（右下）は、リアルタイム文コピータイピングタスクの試験の開始時の5秒のタイムウィンドウにわたる実際の神経データ及び検出結果である。この図は、同様の発話検出アーキテクチャを実装した本発明者らの以前の研究[2]から適応したものである。

【図24A】コードワード分類正確度に対する特徴選択の影響。図24A、HGA特徴のみではなく、高ガンマ活性（HGA）及び低周波数信号（LFS）をともに使用する場合（結合HGA+LFS特徴セット）、各コードワードの分類正確度が向上する。図24B、LFS単独ではなくHGA+LFSを使用すると、ほぼ全てのコードワードの分類正確度が向上する。図24A及び図24Bの両方において、コードワードは、小文字として表され、スピアマン順位相関が示される。関連するp値が、並べ替え検定を介して計算され、観察の一方のグループ（HGA、LFS、又はHGA+LFSのいずれかのコードワード正確度）が、その観察グループと他方のグループとの間の相関を再計算する前にシャッフルされた。2つの比較の各々について、並べ替え検定中に2000回の反復が使用された。

【図24B】コードワード分類正確度に対する特徴選択の影響。図24A、HGA特徴のみではなく、高ガンマ活性（HGA）及び低周波数信号（LFS）をともに使用する場合（結合HGA+LFS特徴セット）、各コードワードの分類正確度が向上する。図24B

・ L F S 単独ではなく H G A + L F S を使用すると、ほぼ全てのコードワードの分類正確度が向上する。図 2 4 A 及び図 2 4 B の両方において、コードワードは、小文字として表され、スピアマン順位相関が示される。関連する p 値が、並べ替え検定を介して計算され、観察の一方のグループ（H G A、L F S、又は H G A + L F S のいずれかのコードワード正確度）が、その観察グループと他方のグループとの間の相関を再計算する前にシャッフルされた。2 つの比較の各々について、並べ替え検定中に 2 0 0 0 回の反復が使用された。

【図 2 5】孤立標的試験分類からの混同行列。（高ガンマ活性及び低周波数信号の両方を使用して）孤立標的試験中に記録された神経データのオフライン分類中に計算された混同値が、各 N A T O コードワード及び試行された手の握りについて示される。各行は、標的コードワード又は試行された手の握りに対応し、その行の各列内の値は、標的として正しく分類された（値が対角線に沿っている場合）、又は別の潜在的な標的として誤って分類された（「混同された」）（値が対角線に沿っていない場合）孤立標的タスク試験の割合に対応する。各行の値の合計は 1 0 0 % になる。一般に、沈黙発話と手の握りの試行は確実に分類された。

【図 2 6 A】公然及び沈黙発話試行中の神経活性化特性。図 2 6 A、各画像は、電極位置、並びに、最大ピークコードワード平均大きさとして測定された、各電極、発話試行のタイプ（公然又は沈黙）、及び特徴タイプ（高ガンマ活性（H G A）又は低周波数信号（L F S））についての最大神経活性化と重ね合わされた参加者の脳の M R I 再構成を示す。これらの値を算出するために、孤立標的データセットを使用して、各コードワード、電極、発話試行のタイプ、及び特徴タイプについて、試験平均化神経特徴時系列を計算した（各試行について、ゴーキューが使用された後の 2 . 5 秒のタイムウィンドウ）。次いで、これらの試験平均化時系列の各々のピーク大きさ（絶対値の最大値）を決定した。次いで、各電極、発話試行の種類、及び特徴タイプの最大ピークコードワード平均大きさが、各組み合わせのコードワードにわたるこれらのピーク大きさの最大値として計算された。2 つの列は、各タイプの発話試行（公然、次いで沈黙）の値を示し、2 つの列は、各特徴タイプ（H G A、次いで L F S）の値を示す。図 2 6 B、ピークコードワード平均大きさの標準偏差。ここで、（図 2 6 A で使用される最大値の代わりに）各電極、発話試行のタイプ、及び特徴タイプについてコードワードにわたるピーク平均大きさの標準偏差が計算され、プロットされており、その組み合わせについて発話標的にわたって大きさがどの程度変化したかを示す。図 2 6 A 及び図 2 6 B について、各プロットされた電極の色は、その電極の真の関連する値を示し、各電極のサイズは、（所与のタイプの発話試行及び特徴タイプの）他の電極の値と比較してのその電極の関連する値を示す。

【図 2 6 B】公然及び沈黙発話試行中の神経活性化特性。図 2 6 A、各画像は、電極位置、並びに、最大ピークコードワード平均大きさとして測定された、各電極、発話試行のタイプ（公然又は沈黙）、及び特徴タイプ（高ガンマ活性（H G A）又は低周波数信号（L F S））についての最大神経活性化と重ね合わされた参加者の脳の M R I 再構成を示す。これらの値を算出するために、孤立標的データセットを使用して、各コードワード、電極、発話試行のタイプ、及び特徴タイプについて、試験平均化神経特徴時系列を計算した（各試行について、ゴーキューが使用された後の 2 . 5 秒のタイムウィンドウ）。次いで、これらの試験平均化時系列の各々のピーク大きさ（絶対値の最大値）を決定した。次いで、各電極、発話試行の種類、及び特徴タイプの最大ピークコードワード平均大きさが、各組み合わせのコードワードにわたるこれらのピーク大きさの最大値として計算された。2 つの列は、各タイプの発話試行（公然、次いで沈黙）の値を示し、2 つの列は、各特徴タイプ（H G A、次いで L F S）の値を示す。図 2 6 B、ピークコードワード平均大きさの標準偏差。ここで、（図 2 6 A で使用される最大値の代わりに）各電極、発話試行のタイプ、及び特徴タイプについてコードワードにわたるピーク平均大きさの標準偏差が計算され、プロットされており、その組み合わせについて発話標的にわたって大きさがどの程度変化したかを示す。図 2 6 A 及び図 2 6 B について、各プロットされた電極の色は、その電極の真の関連する値を示し、各電極のサイズは、（所与のタイプの発話試行及び特徴タ

イブの) 他の電極の値と比較してのその電極の関連する値を示す。

【発明を実施するための形態】

【0101】

対象のコミュニケーションを支援するための方法、デバイス、及びシステムが提供される。具体的には、個人の神経活動から直接単語及び文を復号するための方法、デバイス、及びシステムが提供される。開示されている方法において、発話処理に關与する脳の領域からの皮質活動が、個人が文の単語の発語又はスペルアウトを試行している間に記録される。深層学習計算モデルが、記録された脳活動からの単語を検出及び分類するために使用される。脳活動からの発話の復号は、特定の単語系列がどのように出現する可能性があるかを予測する言語モデルの使用によって補助される。加えて、神経活動からの試行非発話運動の復号を使用して、コミュニケーションを更に支援することができる。

10

【0102】

本明細書に開示される方法、デバイス、及びシステムは、脳卒中、外傷性脳損傷、脳腫瘍、筋萎縮性側索硬化症、多発性硬化症、ハンチントン病、ニーマン-ピック病、フリードライヒ運動失調症、ウィルソン病、脳性麻痺、ギラン-バレー症候群、テイ-サックス病、脳症、橋中心髄鞘崩壊症、及び構語障害をもたらす頭部、頸部、又は胸部の筋肉の機能障害又は麻痺を引き起こす他の状態を含むが、これらに限定されない状態及び疾患によって引き起こされるコミュニケーションに困難を有する個人を支援するために使用され得る。本明細書に開示される方法は、そのような個人に対するコミュニケーションを回復し、自律性及び生活の質を改善するために使用され得る。

20

【0103】

本発明の例示的な実施形態を説明する前に、本発明は記載された特定の実施形態に限定されず、それ自体は勿論、変化し得ることを理解されたい。また、本発明の範囲は、添付の特許請求の範囲によってのみ限定されることになるため、本明細書で使用される用語は、特定の実施形態のみを説明する目的のためのものであり、限定することが意図されるものではないことも理解されるべきである。

【0104】

値の範囲が提供される場合、文脈上別段の指示のない限り、その範囲の上限と下限との間の各介在値も、下限の単位の10分の1まで具体的に開示されていることを理解されたい。規定範囲における任意の規定値又は介在値と、その規定範囲における任意の他の規定値又は介在値との間のより小さい各範囲が、本発明に包含される。これらのより小さい範囲の上限及び下限は、独立してその範囲に含まれても、又は除外されてもよく、より小さい範囲に限度のいずれかが含まれる範囲、どちらも含まれない範囲、又は両方の限度が含まれる範囲の各々も、規定範囲内の任意の具体的に除外されている限度に従って本発明に包含される。記載された範囲が限界の一方又は両方を含む場合、それらの含まれる限界のいずれか又は両方を除外する範囲も、同様に本発明に含まれる。

30

【0105】

別途定義されない限り、本明細書で使用される全ての技術用語及び科学用語は、本発明が属する当業者によって一般に理解される意味と同じ意味を有する。本明細書に記載される方法及び材料と類似又は同等の任意の方法及び材料は、本発明の実施又は試験に使用することができるが、ここでは、いくつかの潜在的かつ例示的な方法及び材料が説明され得る。本明細書で言及される任意の及び全ての刊行物は、刊行物が引用される方法及び／又は材料を開示及び説明するために、参照により本明細書に組み込まれる。矛盾がある場合、本開示が、組み込まれた刊行物の任意の開示に優先することを理解されたい。

40

【0106】

本明細書及び添付の特許請求の範囲で使用される場合、単数形「a」、「an」、及び「the」は、文脈が明確に別段の指示をしない限り、複数の指示対象を含むことに留意されたい。このように、例えば、「電極(an electrode)」又は「その電極(the electrode)」への参照は複数のそのような電極を含み、「信号(a signal)」又は「その信号(the signal)」への言及は1つ以上の信号

50

への言及を含む、などとなる。

【0107】

特許請求の範囲は、任意であり得る要素を除外するために起草され得ることに更に留意されたい。したがって、本明細書は、特許請求の範囲の要素の列挙に関連して、「専ら」、「のみ」などの排他的用語を使用すること、又は「否定的な」制限を使用することの先行的な根拠として機能することが意図される。

【0108】

本明細書で論じられる刊行物は、本出願の出願日より前のそれらの開示についてのみ提供される。本明細書のいかなる内容も、本発明が先行発明のために、そのような刊行物に先行する権利がないことの承認として解釈されるべきではない。更に、提供された公開日は、実際の公開日とは異なる場合があり、個別に確認される必要がある場合がある。そのような刊行物が、本開示の明示的又は黙示的な定義と矛盾する用語の定義を定めている限り、本開示の定義が支配する。

【0109】

本開示を読むと当業者には明らかであるように、本明細書に記載及び例証される別個の実施形態の各々は、本発明の範囲又は趣旨から逸脱することなく、他のいくつかの実施形態のいずれかの特徴から容易に分離され得るか、又はこれらと組み合わせられ得る別個の構成要素及び特徴を有する。任意の列挙された方法は、列挙された事象の順序、又は論理的に可能な任意の他の順序で実行され得る。

【0110】

定義

「コミュニケーション障害」という用語は、本明細書においては、対象の発話する能力に影響を及ぼす状態の群を指すために使用される。コミュニケーション障害は、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、筋萎縮性側索硬化症、多発性硬化症、ハンチントン病、ニーマンピーク病、フリードライヒ運動失調症、ウィルソン病、脳性麻痺、ギランバレー症候群、テイサックス病、脳症、橋中心髄鞘崩壊症、及び構語障害をもたらす頭部、頸部、又は胸部の筋肉の機能障害又は麻痺を引き起こす他の状態を含むが、これらに限定されない。

【0111】

「コミュニケーション」という用語は、話し言葉、単語のスペリング、及びテキストの生成（例えば、発話する試行を介してパーソナルデバイスを制御して電子メール又はテキストを生成すること）を含む口頭コミュニケーションなどの単語ベースのコミュニケーション、並びに非発話運動の試行などの行動ベースのコミュニケーションを含む。試行発話は、理解可能であってもよく若しくは理解可能でなくともよい発声される発話、又は発声されない発話を含んでもよい。沈黙発話試行は、発声せずに発話を明確に発することを試行する意志的試行である。沈黙スペリング試行は、発声せずにアルファベット又は数字をスペリングする意志的な試行である。試行非発話運動は、検出可能な物理的な動きが一切伴わない想像上の動きを含み得る。試行非発話運動は、これに限定されないが、想像上の頭部、腕、手、足、及び脚の動きを含み得る。試行非発話運動は、試行された発話若しくはスペリングの開始若しくは終了を示すために、又は外部デバイスを制御するために（例えば、パーソナルデバイス若しくはソフトウェアアプリケーションとの通信のために、又はデバイスをオン若しくはオフにするために）使用され得る。開示されている方法において、神経活動が、個人が任意の音声出力又は検出可能な運動を生成するか否かにかかわらず、コミュニケーションの試行中に記録される。

【0112】

「対象」、「個人」、「患者」、及び「参加者」という用語は、本明細書中で互換的に使用され、コミュニケーション障害を有する患者を指す。患者は、好ましくは、コミュニケーションを回復させるための本明細書に開示されるシステム、デバイス、及び方法から利益を得ることができる、ヒト、例えば、小児、青年、成人、例えば、若年、中年、又は高齢のヒトである。患者は、構語障害を患っていると診断されている場合がある。

【0113】

本明細書において使用されるものとしての「ユーザ」という用語は、本開示の方法の1つ以上のステップを実施するために、本明細書に開示されるデバイス及び／又はシステムと相互作用する人物を指す。ユーザは、処置を受ける患者であってもよい。ユーザは、患者の医師などの医療従事者であってもよい。

【0114】

方法

本開示は、対象のコミュニケーションを支援するための方法を提供する。個人の神経活動から直接単語及び文を復号するための方法が提供される。開示されている方法において、発話処理に関与する脳の領域からの皮質活動が、個人が文の単語の発語又はスペルアウトを試行している間に記録される。言葉を発語又はスペルアウトする試行は、発声を含むか又は除外し得る。すなわち、神経活動が、個人が任意の音声出力を生成するか否かにかかわらず、単語を発語又はスペルアウトする試行中に記録される。いくつかの場合、音声出力は、個人が単語の発語又はスペルアウトを試行するとき、理解不能であり得る。深層学習計算モデルが、記録された脳活動からの単語及び／又はスペリングされた文字を検出及び分類するために使用される。脳活動からの発話の復号は、特定の単語系列がどのように出現するかを予測する言語モデルの使用によって補助される。本明細書に記載されているニューロテクノロジーは、発話する能力を失った患者のコミュニケーションを回復させるために使用することができ、自律性及び生活の質を向上させる可能性がある。ここで、本方法の様々なステップ及び態様について、下記により詳細に説明する。

【0115】

本方法は、対象による試行発話及び／又は試行スペリングと関連付けられる脳電気信号データを記録するために、1つ以上の電極を備える神経記録デバイスを、対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、コンピューティングデバイスと通信するインターフェースを、対象の頭部上のロケーションに位置付けることと、を含む。対象による試行発話及び／又は試行スペリングと関連付けられる脳電気信号データは、神経記録デバイスを使用して記録され、インターフェースは、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信し、プロセッサは、対象による試行発話及び／又はスペリングを検出し、記録された脳電気信号データからスペリングされた文字、単語、句、又は文を復号するようにプログラムされている。

【0116】

記録デバイスは、非脳侵入表面電極又は脳侵入深部電極を含んでもよい。電気信号は、単一の電極、電極対、又は電極アレイを使用して記録されてもよい。いくつかの実施形態において、脳活動は、2つ以上の部位から記録される。特定の実施形態において、脳電気信号データは、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせなどの、発話処理に関与する脳の感覚運動皮質領域から記録される。いくつかの実施形態において、電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置付けられる。

【0117】

脳の特定の領域での脳活動を記録するための電極の位置付けは、頭蓋内電極の配置のための標準的な外科的処置を使用して実行されてもよい。本明細書において使用される場合、「電極 (an electrode)」又は「その電極 (the electrode)」という表現は、単一の電極又は電極アレイなどの複数の電極を指す。本明細書において使用される場合、脳の領域と接触する電極の文脈で使用されるものとしての「接触」という用語は、電極と領域との間の物理的な関連付けを指す。言い換えれば、脳の領域と接触している電極は、脳の領域に物理的に接している。脳の領域と接触している電極を使用して、試行発話及び／又はスペリングと関連付けられる神経活動に対応する電気信号を検出することができる。本明細書に開示される方法において使用される電極は、単極（カソード又はアノード）又はバイポーラ（例えば、アノード及びカソードを有する）であってもよい。

【0118】

特定の実施形態において、1つ以上の電極が、1つ以上の脳領域における試行発話及び／又はスプリングと関連付けられる神経活動のための電気信号を記録するために使用される。電極は、例えば、脳の中心前回、中心後回、後中前頭回、後上前頭回、又は後下前頭回領域などの、発話処理に参与する感覚運動皮質の領域内に配置されてもよい。場合によっては、電極を配置することは、脳の指定された領域の表面上に電極を位置付けることを伴い得る。例えば、電極は、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせにおいて、脳の表面上に配置されてもよい。電極は、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域において、脳の表面の少なくとも一部分に接触してもよい。いくつかの実施形態において、電極は、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域において、実質的に表面領域全体に接触してもよい。いくつかの実施形態において、電極は、追加的に、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域に隣接する領域に接触してもよい。

10

【0119】

いくつかの実施形態において、平面支持基板上に配置された電極アレイが、本明細書で指定される脳領域のうちの1つ以上からの神経活動の電気信号を検出するために使用されてもよい。電極アレイの表面積は、電極アレイと脳との間の所望の接触面積によって決定され得る。表面電極又は表面電極アレイなどの、脳表面に埋め込むための電極は、商業供給元から入手することができる。市販されている電極／電極アレイを、所望の接触面積を達成するように修正してもよい。場合によっては、本明細書に開示される方法において使用され得る非脳侵入電極（表面電極とも称される）は、皮質脳波記録（ECOG）電極又は脳波記録（EEG）電極であってもよい。

20

【0120】

場合によっては、電極を標的領域又は部位に位置付けることは（例えば、神経記録デバイス電極）、脳の特定の領域内に脳侵入電極（深部電極とも称される）を位置付けることを伴い得る。例えば、深部電極は、発話処理に参与する感覚運動皮質の選択された領域（例えば、中心前回、中心後回、後中前頭回、後上前頭回、又は後下前頭回領域）内に配置されてもよい。いくつかの実施形態において、電極は、追加的に、発話処理に参与する感覚運動皮質の選択された領域に隣接する（例えば、中心前回、中心後回、後中前頭回、後上前頭回、又は後下前頭回領域に隣接する）領域に接触してもよい。いくつかの実施形態において、電極アレイが、本明細書において指定されているような、発話処理に参与する感覚運動皮質の選択された領域（例えば、中心前回、中心後回、後中前頭回、後上前頭回、又は後下前頭回領域）において電気信号を記録するために使用されてもよい。

30

【0121】

電極が脳に挿入される深さは、電極アレイと脳との間の所望の接触レベル、及び電極が電気信号を記録するためにアクセスできるであろう神経集団のタイプによって決定され得る。脳侵入電極アレイは、商業供給元から入手され得る。市販されている電極アレイを、脳組織への所望の挿入深さを達成するように修正してもよい。

【0122】

電極アレイに含まれる電極の正確な数（例えば、試行発話と関連付けられる神経活動の記録のための）は、変化してもよい。特定の態様において、電極アレイは、3個以上、10個以上、50個以上、100個以上、200個以上、500個以上などの2つ以上の電極を含んでもよく、4個以上、例えば、約3～6個の電極、約6～12個の電極、約12～18個の電極、約18～24個の電極、約24～30個の電極、約30～48個の電極、約48～72個の電極、約72～96個の電極、約96～128個の電極、約128～196個の電極、約196～294個の電極、又はより多くの電極を含む。電極は、規則的な繰り返しパターン（例えば、電極間に約1cmの間隔を有するグリッドなどのグリッド）に配置されてもよく、又はパターンがなくてもよい。対象による試行発話及び／又はスプリングと関連付けられる神経活動からの電気信号の最適な記録のために標的部位を適

40

50

合させる電極を使用することができる。そのような例の1つは、 $21/2\text{ mm}$ だけ分離された8つの接点を有する単一の多接点電極である。各接点は、約 2 mm のスパンを有することになる。別の例は、 2 mm の介在する隙間を有する2つの 1 cm 接点を有する電極である。また更に、本方法に使用することができる電極の別の例は、標的部位をカバーするための2又は3分岐電極である。これらの3叉電極の各々は、 $2\sim 2.5\text{ mm}$ の中心間分離、及び 1.5 mm のスパンを有する4つの $1\sim 2\text{ mm}$ 接点を有する。

【0123】

いくつかの実施形態において、高密度ECOG電極アレイが、対象による試行発話及び／又はスペリングと関連付けられる神経活動からの電気信号を記録するために使用される。例えば、高密度ECOG電極アレイは、少なくとも100個の電極、少なくとも128個の電極、少なくとも196個の電極、少なくとも256個の電極、少なくとも294個の電極、少なくとも500個の電極、又は少なくとも1000個の電極、又はより多くの電極を含んでもよい。いくつかの実施形態において、高密度ECOG電極アレイにおける電極中心間間隔は、 $250\text{ mm}\sim 4\text{ mm}$ に及び 250 mm 、 300 mm 、 350 mm 、 400 mm 、 500 mm 、 550 mm 、 600 mm 、 650 mm 、 700 mm 、 800 mm 、 900 mm 、 1 mm 、 1.5 mm 、 2 mm 、 2.5 mm 、 3 mm 、 3.5 mm 、又は 4 mm などの、この範囲内の任意の電極中心間間隔を含む。いくつかの実施形態において、高密度ECOG微小電極アレイが使用される。ECOG微小電極アレイは、 150 、 160 、 170 、 180 、 190 、 200 、 210 、 220 、 230 、 240 、又は 250 mm などのこの範囲内の任意の直径を含む、 $150\text{ mm}\sim 250\text{ mm}$ に及ぶ直径を有する電極を含む、 250 mm 以下、 230 mm 以下、又は 200 mm 以下の直径を有する電極を含んでもよい。高密度ECOG電極アレイ及び微小電極アレイの説明については、例えば、参照により本明細書に組み込まれる、Muller et al. (2015) Annu Int Conf IEEE Eng Med Biol Soc 2016:1528-1531、Chiang et al (2020) J. Neural Eng. 17:046008、Escabi et al. (2014) J. Neurophysiol. 112(6):1566-1583を参照されたい。

【0124】

各電極のサイズはまた、アレイ内の電極の数、電極のロケーション、材料、患者の年齢、及び他の要因などの要因に応じて変化してもよい。特定の態様において、各電極は、 $4\text{ mm}\sim 0.25\text{ mm}$ 、 $3\text{ mm}\sim 0.25\text{ mm}$ 、 $2\text{ mm}\sim 0.25\text{ mm}$ 、 $1\text{ mm}\sim 0.25\text{ mm}$ 、又は約 3 mm 、約 2 mm 、約 1 mm 、約 0.5 mm 、又は約 0.25 mm を含む、約 4 mm 以下などの、約 5 mm 以下のサイズ（例えば、直径）を有する。

【0125】

特定の実施形態において、本方法は、電極の位置付けを最適化するために対象の脳をマッピングすることを更に含む。電極の位置付けは、対象による試行発話と関連付けられる脳活動特徴を検出し、試行発話の最適な復号を達成するように最適化される。例えば、特定の周波数範囲（例えば、アルファ、デルタ、ベータ、ガンマ、及び／又は高ガンマ）内の電気信号のパターンが、対象によって意図される単語、句、又は文の試行発話及び／又はスペリングの検出及び復号に使用されてもよい。したがって、電極は、コミュニケーション障害を有する対象へのコミュニケーションを回復させるために、特定の周波数範囲における脳活動の検出及び／又は復号を最適化するように位置付けることができる。

【0126】

特定の態様において、本開示の方法及びシステムは、脳活動、例えば、腹側感覚運動皮質における電気活動を記録することを含み得、ここで、試行発話の単語、句、及び文と関連付けられるガンマ周波数神経活動のパターンが検出され得る。特定の事例において、腹側感覚運動皮質内の複数のロケーションからの電気活動が測定されてもよい。いくつかの実施形態において、高ガンマ周波数範囲（ $70\text{ Hz}\sim 150\text{ Hz}$ など）又は低周波数範囲（ $0.3\text{ Hz}\sim 100\text{ Hz}$ など）の電気活動は、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから測定されてもよい

。いくつかの実施形態において、高ガンマ周波数範囲（70 Hz～150 Hzなど）及び低周波数範囲（0.3 Hz～100 Hzなど）の電気活動は、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから測定されてもよい。

【0127】

脳活動の検出は、当該技術分野で知られている任意の方法によって実施されてもよい。例えば、神経活動の機能的脳イメージングは、皮質脳波記録（ECOG）、脳波記録（EEG）、定位的頭蓋内脳波（sEEG）、脳磁図（MEG）、単光子放出コンピュータ断層撮影（SPECT）などの電気的方法、並びに機能的磁気共鳴イメージング（fMRI）、陽電子放出断層撮影（PET）、機能的近赤外線分光法（fNIRS）、及び時間領域機能的近赤外線分光法などの代謝及び血流研究によって実行されてもよい。いくつかの実施形態において、中心前回、中心後回、後中前頭回、後上前頭回、又は後下前頭回領域が、試行発話及び／又は試行スペリングと関連付けられる神経活動を検出するための電極のための最適な位置付けを決定するためにマッピングされる。これらの領域のうちの1つ以上は、試行発話及び／又は試行スペリングと関連付けられる神経活動からの電気信号を測定するための電極を備える神経記録デバイスを埋め込むことができる。

【0128】

いくつかの事例において、脳内の1つ以上のロケーションにおける電気的活動は、試行発話又は試行スペリング中だけでなく、発話試行又はスペリング試行の直前（すなわち、発話又はスペリングの準備の期間）から試行発話又はスペリングの直後の期間（すなわち、試行発話又はスペリング後の休止期間）までに及ぶ期間中にも測定され得る。特定の部位における神経活動からの発話又はスペリングの復号の正確度の評価は、復号された単語を患者の意図された単語と比較することによって決定され得る。例えば、患者は、支援タイピングデバイスを使用して、正しい意図された単語を伝達することができる。発話イベントの開始及び終了の検出、並びに神経活動の復号からの単語／文字分類正確度の両方を評価することができる。偽陽性は、真の単語又は文字生成試行に関連付けられていない検出された発話イベントを含み、偽陰性は、検出された発話イベントに関連付けられていない単語／文字生成試行を含む。発話イベントの検出、及び神経活動からの単語又はスペリングされた文字の復号におけるより低い誤り率は、より良好なパフォーマンスを示す。場合によっては、電極の配置又は電極の数は、電気信号の検出並びに対象による試行発話及び／又はスペリングの復号を改善するために変更されてもよい。

【0129】

本方法の適用は、コミュニケーション障害の重症度の臨床評価及びコミュニケーションの支援への要求によって決定される必要性に基づいて、神経記録デバイスを埋め込むための患者を選択する事前のステップを含み得、また、認知評価、解剖学的評価、行動評価及び／又は神経生理学的評価を含み得る。コミュニケーションに困難を有する患者は、本明細書に記載のように、コミュニケーションを支援するために神経記録デバイスを埋め込むことができる。

【0130】

コンピューティングデバイスと通信することができるインターフェースが、頭蓋内に埋め込まれるか、又は対象の頭部に配置されて、脳電気信号を神経記録デバイスから取得することができ、復号のためにデータプロセッサに送信することができる、外部からアクセス可能なプラットフォームを提供する。いくつかの実施形態において、インターフェースは、対象の頭蓋内に固定された経皮台座コネクタを備える。インターフェースは、例えば、取り外し可能なデジタルコネクタ及びケーブルによって、コンピュータ又はハンドヘルドコンピューティングデバイス（例えば、携帯電話又はタブレット）などのコンピューティングデバイスに接続され得る。代替的に、インターフェースは、コンピューティングデバイスに無線で接続されてもよい。いくつかの実施形態において、インターフェースは、第2の無線通信ユニットを備えるコンピューティングデバイスと通信する第1の無線通信ユニットを備える。いくつかの実施形態において、第1の無線通信ユニットは、電磁搬送

波（例えば、無線波、マイクロ波、又は赤外線搬送波）又は超音波を使用する無線通信プロトコルを利用して、インターフェースから第2の無線通信ユニットを備えるコンピューティングデバイスにデータを転送する。脳コンピュータインターフェースは、Blackrock Microsystems (Salt Lake City, Utah) のNeuroport（商標）システムを含む市販のものであり、例えば、参照により本明細書に組み込まれるWeiss et al. (2019) Brain-Computer Interfaces 6: 106-117も参照されたい。

【0131】

プロセッサは、記録された脳電気信号データから試行発話及び／又は試行スペリングを復号するようにプログラムされたコンピュータ又はハンドヘルドコンピューティングデバイス（例えば、携帯電話又はタブレット）によって提供されてもよい。

10

【0132】

記録された脳電気活動を分析することは、アルゴリズム又は分類器の使用を含んでもよい。いくつかの実施形態において、機械学習アルゴリズムが、試行発話又はスペリング中の記録された脳活動の分析からの発話検出、文字分類（試行スペリングの場合）、単語分類、及び文復号を自動化するために使用される。機械学習アルゴリズムは、教師あり学習アルゴリズムを含んでもよい。教師あり学習アルゴリズムの例は、平均1依存性推定量（AODE）、人工ニューラルネットワーク（例えば、長・短期記憶（LSTM）層のスタックを含む人工ニューラルネットワーク）、ベイズ統計（例えば、ナイーブベイズ分類器、ベイズネットワーク、ベイズ知識ベース）、事例ベース推論、決定木、帰納論理プログラミング、ガウス過程回帰、データ処理のグループ法（GMDH）、学習オートマトン、学習ベクトル量子化、最小メッセージ長（決定木、決定グラフなど）、怠惰学習、インスタンスベース学習最近傍アルゴリズム、類推モデリング、確率的で近似的に正しい（PAC）学習、リップルダウンルール、知識取得方法論、シンボリック機械学習アルゴリズム、サブシンボリック機械学習アルゴリズム、サポートベクタマシン、ランダムフォレスト、分類器のアンサンブル、ブートストラップアグリゲーティング（バギング）、及びブーストを含んでもよい。教師あり学習は、回帰分析及び情報ファジーネットワーク（IFN）などの順序分類を含んでもよい。代替的に、教師あり学習方法は、AODE、線形分類器（例えば、フィッシャー線形判別子、ロジスティック回帰、ナイーブベイズ分類器、パーセプトロン、及びサポートベクタマシン）、二次分類器、k近傍、ブースト、決定木（例えば、C4.5、ランダムフォレスト）、ベイズネットワーク、及び隠れマルコフモデルなどの統計的分類を含んでもよい。

20

30

【0133】

機械学習アルゴリズムはまた、教師なし学習アルゴリズムを含んでもよい。教師なし学習アルゴリズムの例は、人工ニューラルネットワーク、データクラスタリング、期待値最大化アルゴリズム、自己組織化マップ、動径基底関数ネットワーク、ベクトル量子化、生成地形図、情報ボトルネック法、及びIBSEADを含んでもよい。教師なし学習はまた、アプリアリアルゴリズム、Eclatアルゴリズム、及びFP成長アルゴリズムなどの相関ルール学習アルゴリズムを含んでもよい。単一リンクークラスタリング及び概念的クラスタリングなどの階層的クラスタリングも使用されてもよい。代替的に、教師なし学習は、K平均アルゴリズム及びファジークラスタリングなどの分割クラスタリングを含んでもよい。

40

【0134】

いくつかの事例において、機械学習アルゴリズムは、強化学習アルゴリズムを含む。強化学習アルゴリズムの例は、時間差分学習、Q学習、及び学習オートマトンを含むが、これらに限定されない。代替的に、機械学習アルゴリズムは、データ前処理を含んでもよい。

【0135】

場合によっては、機械学習アルゴリズムは、深層学習を使用してもよい。深層学習（例えば、深層ニューラルネットワーク、深層信念ネットワーク、グラフニューラルネットワ

50

ーク、再帰ニューラルネットワーク、及び畳み込みニューラルネットワーク）は、教師あり、半教師あり、又は教師なしであってもよい。

【0136】

いくつかの実施形態において、機械学習アルゴリズムが、発話検出のために人工ニューラルネットワーク（ANN）モデルを使用し、文復号のために、限定ではないが、隠れマルコフモデル（HMM）又はビタビ復号モデルなどの単語／文字分類及び自然言語処理技法を使用する。

【0137】

いくつかの実施形態において、プロセッサは、発話検出モデルを使用して、神経活動の記録中の任意の時点で試行発話又はスペリングが発生している確率を決定するように、及び／又は神経活動の記録中の試行発話又はスペリングの開始及び終了を検出するようにプログラムされる。線形モデル又は非線形（例えば、人工ニューラルネットワーク（ANN））モデルが、発話検出を自動化するために使用されてもよい。いくつかの実施形態において、深層学習モデルが、具体的には、被験者による試行発話中の単語生成又は被験者による試行スペリング中の文字生成の開始及び終了の検出を自動化するために、発話検出のために使用される。プロセッサは、更に、準備、発話／スペリング、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるようにプログラムされてもよい。いくつかの実施形態において、試行発話／スペリングの検出された開始の周りのタイムウィンドウ内（例えば、発話の検出された開始の1秒前から、発話の検出された開始の3秒後まで）の記録された脳電気信号データが、単語分類又は文字分類のために使用される。

【0138】

単語分類は、機械学習アルゴリズムを使用して、対象による試行発話中の試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別を自動化することができる。文字分類は、機械学習アルゴリズムを使用して、対象による試行スペリング中の試行文字生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別を自動化することができる。

【0139】

特定の実施形態において、対象が意図された文の単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューが、対象に提供される。いくつかの実施形態において、一連のゴーキューが、ディスプレイ上に視覚的に提供される。各ゴーキューに先行して、ゴーキューの提示へのカウントダウンが行われてもよく、次にスペリングされる文字のカウントダウンが、ディスプレイ上に視覚的に提供され、各ゴーキューの後に自動的に開始される。例えば、スペリング手順中、参加者は、文字復号サイクル全体を通じて、意図されたメッセージをスペルアウトする。各サイクルにおいて、参加者にカウントダウンが視覚的に提示され、最終的にゴーキューが提示される。ゴーキューにおいて、参加者は所望の文字を黙って発語することを試行する。いくつかの実施形態において、一連のゴーキューは、ユーザによって調整可能であり得る、各ゴーキュー間の設定時間間隔を提供される。特定の実施形態において、プロセッサは、ゴーキューに後続するタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされる。

【0140】

いくつかの実施形態において、プロセッサは、単語分類モデルを使用して、神経活動の検出されたタイムウィンドウ（例えば、発話検出モデルによって、試行発話又はスペリング中に発生していると識別されるタイムウィンドウ）内の単語を復号するようにプログラムされる。単語分類モデルは、対象が、可能性のある発話／テキスト標的にわたって、試行発話内の特定の単語を意図した確率を決定するために使用される。例えば、ユーザが発語することができる可能性のある単語の語彙内の各単語について、単語分類モデルは、ユーザがその単語の発語を試行したときに神経活動が収集された確率を決定する。単語分類モデルは、線形モデル又は非線形（例えば、ANN）モデルを使用してもよい。

【0141】

10

20

30

40

50

いくつかの実施形態において、プロセッサは、文字分類モデルを使用して、対象によって使用される言語の全ての可能な文字（すなわち、アルファベット又は数字の文字）にわたって、試行スペリング中に対象が特定の文字を意図した確率を決定するようにプログラムされる。特定の実施形態において、プロセッサが、対象による単語の試行スペリングと関連付けられる神経活動から復号される文字系列からの単語分類を、対象によって使用される言語の語彙内の単語のみに制約するように更にプログラムされる。

【0142】

いくつかの実施形態において、プロセッサは、単語系列復号モデルを使用して、単語系列確率に基づいて文を復号し、試行発話又はスペリング中の対象の対応する神経活動から検出された発話イベントと関連付けられる最も可能性の高い単語系列を決定するようにプログラムされる。単語系列復号モデルは、分類モデルからの確率の系列を使用して、復号された系列を構築する。これは、言語モデルを使用して、先験的な文字系列又は単語系列確率を神経復号パイプラインに組み込むことを伴い得る。これはまた、言語モデルからの確率の組み込みを処理するための隠れマルコフモデリング（HMM）又はビタビ復号モデルも伴い得る。これは、線形モデル又は非線形（例えば、ANN）モデルを使用することができる。いくつかの実施形態では、プロセッサはまた、予測された単語シーケンスの確率を決定することによって復号を支援するために、単語のシーケンス内の前の単語またはフレーズが与えられた次の単語の確率を提供する言語モデルを使用するようにプログラムされ、より頻繁に発生する単語は、言語モデルに従ってより頻繁に発生しない単語よりも多くの重みが割り当てられる。加えて、前に検出された発話イベントからの復号された情報、復号を補助するために使用されてもよい。神経活動から試行発話を復号するために使用される発話検出モデル、単語分類モデル、及び言語モデルの詳細な議論については、実施例を参照されたい。

【0143】

対象には、試行発話を、事前定義された語彙（すなわち、単語セット）からの単語に制限するように指示することができる。含まれる単語の数は、意味のある様々な文を作成するのに十分に大きい、十分な神経ベースの分類パフォーマンスを可能にするのに十分に小さいことが好ましい。神経活動からの単語分類のために、対象は、各単語と関連付けられる電気信号のパターンを決定するために、単語セットに含まれる各単語の生成を試行するように指示される。本明細書に記載される方法によって容易に復号され、通信を支援するために使用され得る単語の選択及び単語セットのサイズを評価するために、デバイスの埋め込みに後続する、対象に関する探索的な予備的評価を使用することができる。

【0144】

いくつかの実施形態において、単語セットは、最大50語、最大100語、最大200語、最大300語、最大400語、又は最大500語、又はより多くの単語を含む。例えば、単語セットは、50語、55語、60語、65語、70語、75語、80語、85語、90語、95語、100語、125語、150語、175語、200語、225語、250語、275語、300語、325語、350語、375語、400語、500語、600語、700語、800語、900語、1000語、又はその間の任意の語数を含んでもよい。いくつかの実施形態において、単語セットは、am、are、bad、bring、clean、closer、comfortable、coming、computer、do、faith、family、feel、glasses、going、good、goodbye、have、hello、help、here、hope、how、hungry、I、is、it、like、music、my、need、no、not、nurse、okay、outside、please、right、success、tell、that、they、thirsty、tired、up、very、what、where、yes、及びyouを含む。

【0145】

いくつかの実施形態において、対象の試行発話は、選択された単語セットの任意の選ばれた単語系列を含んでもよい。他の実施形態において、対象の試行発話は、選択された単語

語セットの単語のみを使用する事前定義された文セットに更に制限される。単語セット及び文セットは、対象が介護者が実施することを所望するタスクに関して、介護者とコミュニケーションするために使用することができる文を含むように選択されてもよい。神経活動からの文分類のために、対象は、対象の神経活動が処理され、テキストに復号される間、文セットに含まれる各文の生成を試行するように指示される。インターフェースに接続されているプロセッサは、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、指定の単語セットからの単語から全体が構成される多くの可能性のある文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、対象が試行発話中に生成しようとした可能性が最も高い、指定の単語セットからの単語から全体が構成される文、及びその可能性がより低い他のそのような文を維持するようにプログラムされる。いくつかの実施形態において、プロセッサは、任意の所与の時点における、1番目、2番目、及び3番目に可能性の高い文の確率を維持するようにプログラムされる。新しい単語イベントが処理されるとき、最も可能性の高い文が、変化し得る。例えば、ある単語イベントの処理に基づく2番目に可能性の高い文が、1つ以上の追加の単語イベントが処理された後に、最も可能性の高い文になり得る。

10

20

30

40

50

【0146】

いくつかの実施形態において、文セットは、最大25文、最大50語、最大100文、最大200文、最大300文、最大400文、又は最大500文、又はより多くの文を含む。例えば、文セットは、50文、100文、200文、300文、400文、500文、600文、700文、800文、900文、1000文、又はその間の任意の文数を含んでもよい。いくつかの実施形態において、文セットは、Are you going outside、Are you tired、Bring my glasses here、Bring my glasses please、Do not feel bad、Do you feel comfortable、Faith is good、Hello how are you、Here is my computer、How do you feel、How do you like my music、I am going outside、I am not going、I am not hungry、I am not okay、I am okay、I am outside、I am thirsty、I do not feel comfortable、I feel very comfortable、I feel very hungry、I hope it is clean、I like my nurse、I need my glasses、I need you、It is comfortable、It is good、It is okay、It is right here、My computer is clean、My family is here、My family is outside、My family is very comfortable、My glasses are clean、My glasses are comfortable、My nurse is outside、My nurse is right outside、No、Please bring my glasses here、Please clean it、Please tell my family、That is very clean、They are coming here、They are coming outside、They are going outside、They have faith、What do you do、Where is it、Yes、及びYou are not rightを含む。

【0147】

いくつかの実施形態において、対象の試行発話は、意図されたメッセージの単語のスペルアウトを含む。試行発話標的は、任意の言語（英語など）のアルファベット、及び／又はアルファベットの文字を表すコードワード（例えば、alpha（アルファ）、bra

v o (ブラボー) などの NATO コードワード) を含んでもよい。文字確率は、発話標的 (線形又は非線形 (例えば、ANN) モデルを使用することができる) の分類によって決定することができ、系列復号技法 (例えば、言語モデリング、隠れマルコフモデリング、ビタビ復号など) を使用して処理されて、脳活動から完全な文を復号することができる。

【0148】

特定の実施形態において、方法は、記録された神経活動から試行非発話運動を復号することを更に含むことができる。非発話運動は、これに限定されないが、想像上の頭部、腕、手、足、及び脚の動きを含み得る。非発話運動は、ユーザに有益な任意の様式で 사용할ことができる。例えば、神経活動からの非発話運動の復号は、マウスカーソルを制御するか、若しくは他の様態で他のデバイスと相互作用するか、テキスト復号インターフェース内の誤り訂正方法を制御するか、又はシステムを制御するための高レベルコマンド (「文末」又は「メインメニューに戻る」コマンドなど) を選択するために使用され得る。分類モデルを使用して、運動コマンド (例えば、想像上の手運動) を識別することができ、これは、ユーザが意図されたメッセージの試行発話又はスペルアウトを開始又は終了していることをシステムに示すために使用することができる。

【0149】

試行発話、単語の試行スペリング、又は試行非発話運動と関連付けられる神経活動の復号を通じて対象のコミュニケーションを支援する方法は、組み合わせることができる。これらの技法は相補的である。いくつかの事例において、試行スペリングの復号が、試行発話の復号よりも大きい語彙が使用されることを可能にし得る。しかしながら、試行発話の復号は、より速く、直接的な単語復号を可能にするため、対象にとってより容易であり、より便利であり得、これが、頻繁に使用される単語を表現するのに好ましい場合がある。復号を支援するために、試行非発話運動を使用して、対象が試行発話又は意図されたメッセージのスペルアウトを開始又は終了していることをシグナリングすることができる。

【0150】

脳活動からの試行発話、試行スペリング、及び／又は試行非発話運動を復号するためのシステム及びコンピュータ実装方法

本開示はまた、主題の方法の実践に用途を見出すシステムも提供する。いくつかの実施形態において、システムは、a) 対象による試行発話及び／又は試行スペリング及び／又は試行非発話運動と関連付けられる脳電気信号データを記録するために、対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、b) 記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、c) コンピューティングデバイスと通信するインターフェースであって、当該インターフェースは、対象の頭部上のロケーションに位置するように適合されており、インターフェースは、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、d) 記録された脳電気信号データから復号された文を表示するためのディスプレイコンポーネントと、を含むことができる。

【0151】

例えば、中央前回、中央後回、後中前方回、後上方前方回、または後方下方前方回領域からの高ガンマ周波数範囲 (70 Hz ~ 150 Hz など) 及び／または低周波数範囲 (例えば、0.3 Hz ~ 100 Hz) における電気活動、またはそれらの任意の組み合わせは、このシステムを使用する神経記録デバイスによって記録され得、インターフェースは、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する。プロセッサは、本明細書に記載されるように、1つ以上のアルゴリズムを使用して、記録された脳電気信号データから文字、単語、句、又は文を復号するためのプログラミングを実行してもよい。

【0152】

いくつかの実施形態において、コンピュータ実装方法が、対象による試行発話と関連付けられる記録された脳電気信号データから文を復号するために使用される。プロセッサは

、a) 対象による試行発話と関連付けられる記録された脳電気信号データを受信するステップと、b) 任意の時点で試行発話が発生している確率を算出し、対象による試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、c) 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、脳電気信号データを分析するステップと、d) 単語分類モデルからの算出された単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、単語分類モデル及び言語モデルを使用して決定された予測単語確率に基づいて、文内の最も可能性の高い単語系列を決定するステップと、e) 記録された脳電気信号データから復号された文を表示するステップと、を含む、コンピュータ実装方法のステップを実施するようにプログラムすることができる。

10

【0153】

いくつかの実施形態において、コンピュータ実装方法が、対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データから文を復号するために使用される。プロセッサは、a) 対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データを受信するステップと、b) 任意の時点で試行スペリングが発生している確率を算出し、対象による試行スペリング中の文字生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、c) 対象による試行文字生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測文字確率の系列を算出する文字分類モデルを使用して、脳電気信号データを分析するステップと、d) 予測文字確率の系列に基づいて潜在的な文候補を計算し、文候補内の予測される単語間の文字系列に空白を自動的に挿入するステップであって、文字系列内の復号される単語は、対象によって使用される言語の語彙内の単語のみに制約される、計算し、挿入するステップと、e) 予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用して潜在的な文候補を分析し、文内の最も可能性の高い単語系列を決定するステップと、f) 記録された脳電気信号データから復号された文を表示するステップと、を含む、コンピュータ実装方法のステップを実施するようにプログラムすることができる。

20

30

【0154】

いくつかの実施形態において、コンピュータ実装方法が、対象による試行発話及び試行スペリングと関連付けられる記録された脳電気信号データから文を復号するために使用される。

【0155】

特定の実施形態において、システムは、試行発話又は試行スペリング中に収集される神経活動から発話又はスペリング情報を復号するためだけでなく、記録された神経活動から試行非発話運動を復号するためにも使用され得る。非発話運動は、これに限定されないが、想像上の頭部、腕、手、足、及び脚の動きを含み得る。非発話運動は、ユーザに有益な任意の様式で 사용할ことができる。例えば、神経活動からの非発話運動の復号は、マウスカーソルを制御するか、若しくは他の状態で他のデバイスと相互作用するか、テキスト復号インターフェース内の誤り訂正方法を制御するか、又はシステムを制御するための高レベルコマンド（「文末」又は「メインメニューに戻る」コマンドなど）を選択するために使用され得る。分類モデルを使用して、運動コマンド（例えば、想像上の手運動）を識別することができ、これは、ユーザが意図されたメッセージの試行発話又はスペルアウトを開始又は終了していることをシステムに示すために使用することができる。

40

【0156】

いくつかの実施形態において、コンピュータ実装方法が、対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、対象が、試行発話又は

50

意図された文の単語の試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する分類モデルを使用して脳電気信号データを分析することと、を更に含む。

【0157】

特定の実施形態において、コンピュータ実装方法が、対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンに関する情報を含む対象のユーザプロファイルを格納することを更に含む。

【0158】

いくつかの実施形態において、人工ニューラルネットワーク（ANN）モデルが、発話検出のために使用され、限定ではないが、隠れマルコフモデル（HMM）又はビタビ復号モデルなどの文字／単語分類及び自然言語処理技法が、文復号のために使用される。

【0159】

特定の実施形態において、対象が、試行発話について指定の単語セットに制限される。いくつかの実施形態において、プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出し、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択するように更にプログラムされる。いくつかの実施形態において、対象の試行発話は、選択された単語セットの任意の選ばれた単語系列を含んでもよい。他の実施形態において、対象が、試行発話の指定の文セットに制限される。

【0160】

いくつかの実施形態において、プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するように更にプログラムされる。いくつかの実施形態において、プロセッサが、指定の単語セットからの単語から全体が構成される多くの可能性のある文が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされる。いくつかの実施形態において、プロセッサが、対象が試行発話中に生成しようとした可能性が最も高い、指定の単語セットからの単語から全体が構成される文、及びその可能性がより低い1つ以上のそのような文を維持するようにプログラムされる。いくつかの実施形態において、プロセッサが、任意の所与の時点における、1番目、2番目、及び3番目に可能性の高い文の確率を追跡するようにプログラムされる。新しい単語イベントが処理されるとき、最も可能性の高い文が、変化し得る。例えば、前の回の単語イベントの処理に基づく2番目に可能性の高い文が、1つ以上の追加の単語イベントが処理された後に、最も可能性の高い文になり得る。

【0161】

特定の実施形態において、プロセッサが、準備、発話／スペリング（完全な単語、文字、又は任意の他の発話標的）、非発話運動、及び休止のためのイベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされる。いくつかの実施形態において、プロセッサが、単語又は文字分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するように更にプログラムされる。例えば、プロセッサは、単語又は文字分類の検出された開始の1秒前から、検出された開始の3秒後までに記録された脳電気信号データを使用するようにプログラムされてもよい。

【0162】

特定の実施形態において、プロセッサは、言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みを割り当てるように更にプログラムされる。

【0163】

記録された脳電気信号データは、復号する前に様々な方法で処理されてもよい。例えば、データ処理は、限定されないが、神経特徴ストリームのリアルタイムのサンプルごとの処理、個々の電極チャンネルにわたる共通平均参照の使用、デジタル信号フィルタリングを

10

20

30

40

50

実施するための有限インパルス応答（FIR）フィルタの使用、例えばWelfordの方法を使用したスライドウィンドウ正規化手順の実行、自動アーチファクト除去、並びに計算効率を改善するための並列化及び線形パイプライン化を含んでもよい。神経特徴の処理は、発話／テキスト復号中に使用するための1つ以上の特徴ストリームを抽出するためにリアルタイムで実施されてもよい。データ処理方法の説明については、例えば、参照により全体が本明細書に組み込まれる、Moses et al. (2018) J. Neural. Eng. 15 (3) : 036005, Moses et al. (2019) Nat. Commun. 2019 10 (1) : 3096, Moses et al. (2021) N. Engl. J. Med. 385 (3) : 217-227, Sun et al. (2020) J. Neural. Eng. 17 (6)、及びMakin et al. (2020) Nature Neuroscience 23 : 575-582を参照されたい。

10

【0164】

本明細書に記載されている方法は、デジタル電子回路において、若しくはコンピュータソフトウェア、ファームウェア、又はハードウェアにおいて実装されることができる。開示された実施形態及び他の実施形態は、1つ以上のコンピュータプログラム製品、すなわち、データ処理装置によって実行されるか、又はデータ処理装置の動作を制御するために、コンピュータ可読媒体上に符号化されたコンピュータプログラム命令の1つ以上のモジュールとして実装されることができる。コンピュータ可読媒体は、機械可読格納デバイス、機械可読格納基板、メモリデバイス、機械可読伝播信号に影響を与える物質の組成物、又はこれらの任意の組み合わせであることができる。

20

【0165】

コンピュータプログラム（プログラム、ソフトウェア、ソフトウェアアプリケーション、スクリプト、又はコードとしても知られる）は、コンパイルされた言語又は解釈された言語を含む任意の形式のプログラミング言語で記述されることができ、それは、スタンドアロンプログラムとして、若しくはモジュール、コンポーネント、サブルーチン、又はコンピューティング環境において使用に好適な他のユニットなどを含む任意の形式で展開されることができる。コンピュータプログラムは、必ずしもファイルシステム内のファイルに対応するとは限らない。プログラムは、他のプログラム若しくはデータ（例えば、マークアップ言語文書に格納された1つ以上のスクリプト）を保持するファイルの一部に、問題のプログラム専用の単一のファイルに、又は複数のコーディネートされたファイル（例えば、1つ以上のモジュール、サブプログラム、又はコードの一部を格納するファイル）に格納されることができる。コンピュータプログラムは、1つのサイトに位置するか、又は複数のサイトにわたって分散され、通信ネットワークによって相互接続される1つのコンピュータ又は複数のコンピュータ上で実行されるように展開されることができる。

30

【0166】

更なる態様において、コンピュータ実装方法を実施するためのシステムは、記載されているように、プロセッサ、格納コンポーネント（すなわち、メモリ）、ディスプレイコンポーネント、及び汎用コンピュータに典型的に存在する他のコンポーネントを含むコンピュータを含んでもよい。格納コンポーネントは、プロセッサによって実行され得る命令と、プロセッサによって取り出され得る、操作され得る、又は格納され得るデータとを含む、プロセッサによってアクセス可能な情報を格納する。

40

【0167】

格納コンポーネントは、命令を含む。例えば、格納コンポーネントは、コンピュータ実装方法が、対象による試行発話及び／又は試行スペリングと関連付けられる記録された脳電気信号データから文を復号するための命令を含む。コンピュータプロセッサは、本明細書に記載されたように、対象による試行発話と関連付けられる脳電気信号データを受信し、1つ以上のアルゴリズムに従ってデータを分析するために、格納コンポーネントに結合され、格納コンポーネントに格納された命令を実行するように構成される。ディスプレイコンポーネントは、記録された脳電気信号データから復号された文を表示する。

50

【0168】

格納コンポーネントは、ハードドライブ、メモリカード、ROM、RAM、DVD、CD-ROM、USBフラッシュドライブ、書き込み可能メモリ、及び読み取り専用メモリなど、プロセッサによってアクセス可能な情報を格納することができる任意のタイプであり得る。プロセッサは、Intel Corporationからのプロセッサなどの任意の周知のプロセッサであり得る。代替的に、プロセッサは、ASIC又はFPGAなどの専用コントローラであり得る。

【0169】

命令は、プロセッサによって直接的に（マシンコードなどの）、又は間接的に（スクリプトなどの）実行される命令の任意のセットであり得る。その点において、「命令」、「ステップ」、及び「プログラム」という用語は、本明細書において互換的に使用され得る。命令は、プロセッサによる直接処理のためのオブジェクトコード形式で、又はオンデマンドで解釈されるか、又は事前にコンパイルされる独立ソースコードモジュールのスクリプト若しくはコレクションを含む任意の他のコンピュータ言語で格納され得る。

【0170】

データは、命令に従ってプロセッサによって取り出され、格納され、又は修正され得る。例えば、システムは、任意の特定のデータ構造によって制限されないが、データは、複数の異なるフィールド及びレコード、XMLドキュメント、又はフラットファイルを有するテーブルとして、コンピュータレジスタ、関係データベースに格納され得る。データはまた、限定されないが、バイナリ値、ASCII、又はUnicodeなどの、任意のコンピュータ可読フォーマットでフォーマットされ得る。更に、データは、数字、記述テキスト、独自コード、ポインタ、他のメモリ（他のネットワーク位置を含む）に格納されたデータへの参照、又は関連データを計算するために関数によって使用される情報など、関連情報を識別するのに十分な任意の情報を含み得る。

【0171】

特定の実施形態では、プロセッサ及び格納コンポーネントは、同一の物理ハウジング内に格納されてもよく、又は格納されなくてもよい複数のプロセッサ及び格納コンポーネントを備え得る。例えば、命令及びデータのいくつかは、リムーバブルCD-ROM上に、及び読み取り専用コンピュータチップ内の他のものに格納され得る。命令及びデータの一部又は全ては、プロセッサから物理的に離れているが、依然としてプロセッサによってアクセス可能な場所に格納され得る。同様に、プロセッサは、並列動作し得る、又は並列動作しないプロセッサのコレクションを備え得る。

【0172】

システムはまた、コンピューティングデバイスと通信することができるインターフェースを含む。インターフェースは、頭蓋内に埋め込まれるか、又は対象の頭部に配置されて、脳電気信号を神経記録デバイスから取得することができ、復号のためにコンピューティングデバイスに送信することができる、外部からアクセス可能なプラットフォームを提供することができる。いくつかの実施形態において、インターフェースは、対象の頭蓋内に固定された経皮台座コネクタを備える。インターフェースは、例えば、取り外し可能なデジタルコネクタ及びケーブルによって、コンピュータ又はハンドヘルドコンピューティングデバイス（例えば、携帯電話又はタブレット）などのコンピューティングデバイスに接続され得る。代替的に、インターフェースは、コンピューティングデバイスに無線で接続されてもよい。いくつかの実施形態において、インターフェースは、第2の無線通信ユニットを備えるコンピューティングデバイスと通信する第1の無線通信ユニットを備える。いくつかの実施形態において、第1の無線通信ユニットは、電磁搬送波（例えば、無線波、マイクロ波、又は赤外線搬送波）又は超音波を使用する無線通信プロトコルを利用して、インターフェースから第2の無線通信ユニットを備えるコンピューティングデバイスにデータを転送する。脳コンピュータインターフェースは、Blackrock Microsystems（ユタ州ソルトレークシティ）のNeuroport（商標）システムを含む市販のものであり、例えば、参照により本明細書に組み込まれるWeisset

10

20

30

40

50

al. (2019) Brain-Computer Interfaces 6:106-117 も参照されたい。

【0173】

本開示の方法を実行するためのシステムのコンポーネントは、下記の実施例で更に説明される。

【0174】

キット

本明細書に記載の方法を実行するためのキットも提供される。いくつかの実施形態において、キットは、本明細書に記載されているように、対象による試行発話及び／又は試行スペリングと関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法を実行するためのソフトウェアを備える。いくつかの実施形態において、キットは、本明細書に記載されるように、対象のコミュニケーションを支援するためのシステムを備える。そのようなシステムは、対象による試行発話及び／又は試行スペリング及び／又は非発話運動と関連付けられる脳電気信号データを記録するために、対象の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、本明細書に記載されているコンピュータ実装方法に従って、記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、コンピューティングデバイスと通信することが可能なインターフェースであって、当該インターフェースは、対象の頭部上のロケーションに位置するように適合されており、インターフェースは、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、記録された脳電気信号データから復号された文を表示するためのディスプレイコンポーネントと、を備えることができる。

10

20

【0175】

加えて、キットは、（特定の実施形態では）主題の方法を実施するための命令を更に含み得る。これらの命令は、様々な形態で主題のキット中に存在し得、そのうちの1つ以上は、キット中に存在し得る。例えば、命令は、適切な媒体又は基材、例えば、情報が印刷された1枚又は複数枚の紙、キットのパッケージにおいて、パッケージ挿入物において、などに印刷情報として存在し得る。これらの命令の別の形式は、情報が記録されたコンピュータ可読媒体、例えば、ディスク、コンパクトディスク（CD）、フラッシュドライブなどである。存在し得るこれらの命令の更に別の形態は、削除されたサイトで情報にアクセスするためにインターネットを介して使用され得るウェブサイトアドレスである。

30

【0176】

ユーティリティ

本開示の方法、デバイス、及びシステムは、個人のコミュニケーションの支援に用途を見出す。具体的には、個人の神経活動から直接単語及び文を復号するための方法、デバイス、及びシステムが提供される。開示されている方法において、発話処理に關与する脳の領域からの皮質活動が、個人が意図された文の単語の発語又はスペルアウトを試行している間に記録される。深層学習計算モデルが、記録された脳活動からの文字／単語を検出及び分類するために使用される。脳活動からの発話の復号は、特定の単語系列がどのように出現する可能性があるかを予測する言語モデルの使用によって補助される。加えて、神経活動からの試行非発話運動の復号を使用して、コミュニケーションを更に支援することができる。

40

【0177】

本明細書に開示される方法、デバイス、及びシステムは、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、筋萎縮性側索硬化症、多発性硬化症、ハンチントン病、ニーマン-ピック病、フリードライヒ運動失調症、ウィルソン病、脳性麻痺、ギラン-バレー症候群、テイ-サックス病、脳症、橋中心髄鞘崩壊症、及び構語障害をもたらす頭部、頸部、又は胸部の筋肉の機能障害又は麻痺を引き起こす他の状態を含むが、これらに限定されない状態及び疾患によって引き起こされるコミュニケーションに困難を有する個人を支援するために使用され得る。本明細書に開示される方法は、そのような個人に対するコミュニケーショ

50

ンを回復し、自律性及び生活の質を改善するために使用され得る。

【0178】

本開示の非限定的な態様の例

上記の本主題の実施形態を含む態様は、単独で、又は1つ以上の他の態様若しくは実施形態との組み合わせで有益であり得る。上記の説明を限定することなく、番号1～159の本開示の特定の非限定的な態様を以下に提供する。本開示を読めば当業者には明らかであろうように、個々に番号付けされた態様の各々は、先行する又は後に続く個々に番号付けされた態様のいずれかとともに使用し又は組み合わせてもよい。これは、態様の全てのそのような組み合わせにサポートを提供することを意図しており、以下に明示的に提供される態様の組み合わせに限定されない。

10

1. 対象のコミュニケーションを支援する方法であって、方法が、

対象による試行発話と関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、

コンピューティングデバイスと通信するインターフェースを、対象の頭部上のロケーションに位置付けることであって、インターフェースが、神経記録デバイスに接続されている、位置付けることと、

神経記録デバイスを使用して、対象による試行発話と関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをコンピューティングデバイスのプロセッサに送信する、記録することと、

20

プロセッサを使用して、記録された脳電気信号データから単語、句、又は文を復号することと、を含む、方法。

2. 対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために当該コミュニケーションが困難である、態様1に記載の方法。

3. 対象が、麻痺状態にある、態様1又は2に記載の方法。

4. 神経記録デバイスのロケーションが、腹側感覚運動皮質内にある、態様1～3のいずれか1つに記載の方法。

5. 電極が、感覚運動皮質領域の表面上又は感覚運動皮質領域内に位置付けられる、態様1～4のいずれか1つに記載の方法。

30

6. 電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置付けられる、態様5に記載の方法。

7. 神経記録デバイスが、脳侵入電極アレイを含む、態様1～6のいずれか1つに記載の方法。

8. 神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、請求項1～7のいずれか一項に記載の方法。

9. 電極が、深部電極又は表面電極である、態様1～8のいずれか1つに記載の方法。

10. 電気信号データが、高ガンマ周波数成分特徴を含む、態様1～9のいずれか1つに記載の方法。

11. 電気信号データが、70Hz～150Hzの範囲内の神経振動を含む、態様10に記載の方法。

40

12. 当該脳電気信号データを記録することが、中心前回、中心後回、後中前頭回、後上前頭回、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの脳電気信号データを記録することを含む、態様1～11のいずれか1つに記載の方法。

13. 対象による試行発話と関連付けられる脳電気信号を記録するために電極を位置付けるための最適なロケーションを識別するために、対象の脳をマッピングすることを更に含む、態様1～12のいずれか1つに記載の方法。

14. インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える、態様1～13のいずれか1つに記載の方法。

50

15. インターフェースが、経皮台座コネクタに接続されたヘッドステージを更に備える、態様14に記載の方法。

16. プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、態様1～15のいずれか1つに記載の方法。

17. ハンドヘルドデバイスが、携帯電話又はタブレットである、態様16に記載の方法。

18. プロセッサが、試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、発話検出、単語分類、及び文復号を自動化するようにプログラムされている、態様1～17のいずれか1つに記載の方法。

19. プロセッサが、発話検出、単語分類、及び文復号のために機械学習アルゴリズムを使用するようにプログラムされている、態様18に記載の方法。

20. 人工ニューラルネットワーク (ANN) モデルが、発話検出及び単語分類に使用され、隠れマルコフモデル (HMM)、ビタビ復号モデル、又は自然言語処理技法が、文復号に使用される、態様19に記載の方法。

21. プロセッサが、対象による試行発話中の単語生成の開始及び終了の検出を自動化するようにプログラムされている、態様1～20のいずれか1つに記載の方法。

22. 準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む、態様21に記載の方法。

23. プロセッサが、単語分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされている、態様21又は22に記載の方法。

24. 対象が、試行発話について指定の単語セットに制限される、態様1～23のいずれか1つに記載の方法。

25. プロセッサが、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出するようにプログラムされている、態様24に記載の方法。

26. プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出するようにプログラムされている、態様25に記載の方法。

27. 単語セットが、am、are、bad、bring、clean、closer、comfortable、coming、computer、do、faith、family、feel、glasses、going、good、goodbye、have、hello、help、here、hope、how、hungry、I、is、it、like、music、my、need、no、not、nurse、okay、outside、please、right、success、tell、that、they、thirsty、tired、up、very、what、where、yes、及びyouを含む、態様24～26のいずれか1つに記載の方法。

28. 対象が、文を作成するために、制限なしに単語セットの単語を使用し得る、態様1～27のいずれか1つに記載の方法。

29. プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされている、態様28に記載の方法。

30. プロセッサが、予測単語系列確率を決定することによって復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされている、態様1～29のいずれか1つに記載の方法。

31. 言語モデルに従って、より頻繁に出現する単語が、それほど頻繁に出現しない単語よりも大きい重みを割り当てられる、態様30に記載の方法。

32. プロセッサが、試行発話と関連付けられる脳電気信号データ、機械学習アルゴリズムを使用した単語分類モデルからの予測単語確率、及び言語モデルを使用した単語系列確率を所与として、対象の意図された発話内の最も可能性の高い単語系列を決定するためにビタビ復号モデルを使用するようにプログラムされている、態様30又は31に記載の

10

20

30

40

50

方法。

33．対象の試行非発話運動と関連付けられる脳電気信号データを記録することによって、対象が、試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、

試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する非発話運動分類モデルを使用して脳電気信号データを分析することと、を更に含む、態様1～32のいずれか1つに記載の方法。

34．試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、態様33に記載の方法。

35．手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、態様34に記載の方法。

36．プロセッサが、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、対象による試行発話の終了をシグナリングする対象の試行非発話運動の検出を自動化するように更にプログラムされている、態様33～35のいずれか1つに記載の方法。

37．プロセッサが、試行非発話運動のイベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされている、態様36に記載の方法。

38．方法が、復号の正確度を評価することを更に含む、態様1～37のいずれか1つに記載の方法。

39．対象による試行発話と関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法であって、コンピュータが、

a) 対象による試行発話と関連付けられる記録された脳電気信号データを受信するステップと、

b) 脳電気信号データの記録中の任意の時点で試行発話が発生している確率を算出し、対象による試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、

c) 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、脳電気信号データを分析するステップと、

d) 単語分類モデルからの算出された単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、単語分類モデル及び言語モデルを使用して決定された予測単語確率に基づいて、文内の最も可能性の高い単語系列を決定するステップと、

e) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する、コンピュータ実装方法。

40．機械学習アルゴリズムが、発話検出、単語分類、及び文復号に使用される、態様39に記載のコンピュータ実装方法。

41．人工ニューラルネットワーク(ANN)モデルが、発話検出及び単語分類に使用され、隠れマルコフモデル(HMM)、ビタビ復号モデル、又は自然言語処理技法が、文復号に使用される、態様40に記載のコンピュータ実装方法。

42．対象が、試行発話について指定の単語セットに制限される、態様39～41のいずれか1つに記載のコンピュータ実装方法。

43．単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出することと、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択することと、を更に含む、態様42に記載のコンピュータ実装方法。

44．対象が、文を作成するために、制限なしに単語セットの単語を使用し得るか、又は、試行発話について指定の文セットに制限される、態様39～43のいずれか1つに記

10

20

30

40

50

載のコンピュータ実装方法。

45. 単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出することを更に含む、態様39~44のいずれか1つに記載のコンピュータ実装方法。

46. 最も可能性の高い文及び1つ以上のより可能性の低い文を維持することと、各単語の復号後に、単語系列が、試行発話中に対象が生成しようとした、意図された文である確率を再算出することと、を更に含む、態様45に記載のコンピュータ実装方法。

47. 最も可能性の高い文及び1つ以上のより可能性の低い文が、試行発話のために対象によって使用される単語セットからの単語のみから構成される、態様46に記載のコンピュータ実装方法。

48. 準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む、態様39~47のいずれか1つに記載のコンピュータ実装方法。

49. 単語分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データのみが使用される、態様48に記載のコンピュータ実装方法。

50. 言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みが割り当てられる、態様39~49のいずれか1つに記載のコンピュータ実装方法。

51. 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンに関する情報を含む対象のユーザプロファイルを格納することを更に含む、態様39~50のいずれか1つに記載のコンピュータ実装方法。

52. 対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、対象が、試行発話の開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、受信することと、

試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する分類モデルを使用して脳電気信号データを分析することと、を更に含む、態様39~51のいずれか1つに記載のコンピュータ実装方法。

53. 試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、態様52に記載のコンピュータ実装方法。

54. 手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、態様53に記載のコンピュータ実装方法。

55. 試行非発話運動のイベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む、態様52~54のいずれか1つに記載のコンピュータ実装方法。

56. コンピュータ内のプロセッサによって実行されるとき、プロセッサに、態様39~55のいずれか1つに記載の方法を実行させるプログラム命令を含む、非一時的コンピュータ可読媒体。

57. 態様56に記載の非一時的コンピュータ可読媒体と、対象による試行発話と関連付けられる脳電気信号データを復号するための命令と、を備える、キット。

58. 対象のコミュニケーションを支援するためのシステムであって、システムが、対象による試行発話又は試行非発話運動と関連付けられる脳電気信号データを記録するために、対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、

態様39~55のいずれか1つに記載のコンピュータ実装方法に従って、記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、

コンピューティングデバイスと通信するインターフェースであって、当該インターフェースが、対象の頭部上のロケーションに位置するように適合されており、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、

記録された脳電気信号データから復号された文を表示するためのディスプレイコンポー

10

20

30

40

50

ネットと、を備える、システム。

59. 対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために当該コミュニケーションが困難である、態様58に記載のシステム。

60. 神経記録デバイスのロケーションが、腹側感覚運動皮質内にある、態様58又は59に記載のシステム。

61. 電極が、感覚運動皮質領域の表面上又は感覚運動皮質領域内に位置するように適合されている、態様58～60のいずれか1つに記載のシステム。

62. 電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置するように適合されている、態様61に記載のシステム。

63. 神経記録デバイスが、脳侵入電極アレイを含む、態様58～62のいずれか1つに記載のシステム。 10

64. 神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、態様58～63のいずれか1つに記載のシステム。

65. 電極が、深部電極又は表面電極である、態様58～64のいずれか1つに記載のシステム。

66. 電気信号データが、高ガンマ周波数成分特徴を含む、態様58～65のいずれか1つに記載のシステム。

67. 電気信号データが、70Hz～150Hzの範囲内の神経振動を含む、態様66に記載のシステム。

68. インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える、態様58～67のいずれか1つに記載のシステム。 20

69. インターフェースが、経皮台座コネクタに接続可能であるヘッドステージを更に備える、態様68に記載のシステム。

70. プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、態様58～69のいずれか1つに記載のシステム。

71. ハンドヘルドデバイスが、携帯電話又はタブレットである、態様70に記載のシステム。

72. 機械学習アルゴリズムが、発話検出、単語分類、及び文復号に使用される、態様58～71のいずれか1つに記載のシステム。

73. 人工ニューラルネットワーク（ANN）モデルが、発話検出及び単語分類に使用され、隠れマルコフモデル（HMM）、ビタビ復号モデル、又は自然言語処理技法が、文復号に使用される、態様72に記載のシステム。 30

74. プロセッサが、準備、発話、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされている、態様58～73のいずれか1つに記載のシステム。

75. プロセッサが、単語分類の検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するように更にプログラムされている、態様74に記載のシステム。

76. 対象が、試行発話について指定の単語セットに制限される、態様58～75のいずれか1つに記載のシステム。 40

77. プロセッサが、単語セットの全ての単語について、単語セットの単語が、対象が試行発話中に生成しようとした、意図された単語である確率を算出し、対象が試行発話中に生成しようとした、意図された単語である確率が最も高い単語セットの単語を選択するように更にプログラムされている、態様76に記載のシステム。

78. 単語セットが、am、are、bad、bring、clean、closer、comfortable、coming、computer、do、faith、family、feel、glasses、going、good、goodbye、have、hello、help、here、hope、how、hungry、I、is、it、like、music、my、need、no、not、nurse、okay、outside、please、right、success、tell、that、th 50

ey、thirsty、tired、up、very、what、where、yes、及びyouを含む、態様76又は77に記載のシステム。

79．対象が、選択された単語セットの任意の選ばれた単語系列を使用し得る、態様76～78のいずれか1つに記載のシステム。

80．プロセッサが、単語系列が、対象が試行発話中に生成しようとした、意図された文である確率を算出するようにプログラムされている、態様79に記載のシステム。

81．プロセッサが、最も可能性の高い文及び1つ以上のより可能性の低い文を維持し、各単語の復号後に、単語系列が、試行発話中に対象が生成しようとした、意図された文である確率を再算出するようにプログラムされている、態様80に記載のシステム。

82．最も可能性の高い文及び1つ以上のより可能性の低い文が、試行発話のために対象によって使用される単語セットからの単語のみから構成される、態様81に記載のシステム。 10

83．プロセッサが、試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、対象による試行発話の開始又は終了をシグナリングする対象の試行非発話運動の検出を自動化するように更にプログラムされている、態様58～82のいずれか1つに記載のシステム。

84．プロセッサが、試行非発話運動のイベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされている、態様83に記載のシステム。

85．態様58～84のいずれか1つに記載のシステムと、対象による試行発話と関連付けられる脳電気信号データを記録し、復号するために、システムを使用するための命令と、を備える、キット。 20

86．対象のコミュニケーションを支援する方法であって、方法が、

対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを記録するために、電極を備える神経記録デバイスを、対象の脳の感覚運動皮質領域内のロケーションに位置付けることと、

コンピューティングデバイスと通信するインターフェースを、対象の頭部上のロケーションに位置付けることであって、インターフェースが、神経記録デバイスに接続されている、位置付けることと、

神経記録デバイスを使用して、対象による当該試行スペリングと関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをコンピューティングデバイスのプロセッサに送信する、記録することと、 30

プロセッサを使用して、記録された脳電気信号データから意図された文のスペリングされた単語を復号することと、を含む、方法。

87．対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために当該コミュニケーションが困難である、態様86に記載の方法。

88．対象が、麻痺状態にある、態様86又は87に記載の方法。

89．神経記録デバイスのロケーションが、腹側感覚運動皮質内にある、態様86～88のいずれか1つに記載の方法。

90．電極が、感覚運動皮質領域の表面上又は感覚運動皮質領域内に位置付けられる、態様86～89のいずれか1つに記載の方法。 40

91．電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置付けられる、態様90に記載の方法。

92．神経記録デバイスが、脳侵入電極アレイを含む、態様86～91のいずれか1つに記載の方法。

93．神経記録デバイスが、皮質脳波記録（ECOG）電極アレイを含む、態様86～92のいずれか1つに記載の方法。

94．電極が、深部電極又は表面電極である、態様86～93のいずれか1つに記載の方法。

95．電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、態様 50

86～94のいずれか1つに記載の方法。

96．電気信号データが、70Hz～150Hzの高ガンマ周波数範囲、及び0.3Hz～100Hzの低周波数範囲内の神経振動を含む、態様95に記載の方法。

97．当該脳電気信号データの記録が、中心前回領域、中心後回領域、後中前頭回領域、後上前頭回領域、若しくは後下前頭回領域、又はこれらの任意の組み合わせから選択される感覚運動皮質領域からの脳電気信号データを記録することを含む、態様86～96のいずれか1つに記載の方法。

98．対象による単語の試行スペリング又は試行非発話運動試行発話と関連付けられる脳電気信号を記録するために電極を位置付けるための最適なロケーションを識別するために、対象の脳をマッピングすることを更に含む、態様86～97のいずれか1つに記載の方法。

99．インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える、態様86～98のいずれか1つに記載の方法。

100．インターフェースが、経皮台座コネクタに接続されたヘッドステージを更に備える、態様99に記載の方法。

101．プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、態様86～100のいずれか1つに記載の方法。

102．ハンドヘルドデバイスが、携帯電話又はタブレットである、態様101に記載の方法。

103．プロセッサが、対象による単語の試行スペリングと関連付けられる記録された脳電気信号データ内の電気信号の神経活動パターンの識別に基づいて、試行スペリング、文字分類、単語分類、及び文復号の検出を自動化するようにプログラムされている、態様86～102のいずれか1つに記載の方法。

104．プロセッサが、発話検出、文字分類、単語分類、及び文復号のために機械学習アルゴリズムを使用するようにプログラムされている、態様103に記載の方法。

105．プロセッサが、対象による単語の試行スペリングと関連付けられる神経活動から復号される文字系列からの単語分類を、対象によって使用される言語の語彙内の単語のみに制約するように更にプログラムされる、態様104に記載の方法。

106．プロセッサが、準備、試行スペリング、及び休止のための発話イベントラベルを脳電気信号データの記録中の時点に割り当てるように更にプログラムされている、態様86～105のいずれか1つに記載の方法。

107．プロセッサが、対象による文字の試行スペリングの検出された開始の周りのタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされている、態様106に記載の方法。

108．対象が意図された文の単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを対象に提供することを更に含む、態様86～107のいずれか1つに記載の方法。

109．一連のゴーキューが、ディスプレイ上に視覚的に提供される、態様108に記載の方法。

110．各ゴーキューに先行して、ゴーキューの提示へのカウントダウンが行われ、次にスペリングされる文字のカウントダウンが、ディスプレイ上に視覚的に提供され、各ゴーキューの後に自動的に開始される、態様109に記載の方法。

111．一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される、態様108～110のいずれか1つに記載の方法。

112．対象が、各ゴーキュー間の設定時間間隔を制御することができる、態様111に記載の方法。

113．プロセッサが、ゴーキューに後続するタイムウィンドウ内で記録された脳電気信号データを使用するようにプログラムされている、態様108～112のいずれか1つに記載の方法。

114．プロセッサが、復号された文字の系列からの復号された単語の系列が、対象に

10

20

30

40

50

よる意図された文の単語の文字の試行スペリング中に対象が生成しようとした、意図された文である確率を算出するようにプログラムされている、態様 86～113 のいずれか 1 つに記載の方法。

115. プロセッサが、予測単語系列確率を決定することによって復号を補助するために、単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用するようにプログラムされている、態様 86～114 のいずれか 1 つに記載の方法。

116. 言語モデルに従って、より頻繁に出現する単語が、それほど頻繁に出現しない単語よりも大きい重みを割り当てられる、態様 115 に記載の方法。

117. プロセッサが、予測文字確率の系列を使用して、潜在的な文候補を計算し、文候補内の予測される単語間の文字系列に空白を自動的に挿入するように更にプログラムされている、態様 86～116 のいずれか 1 つに記載の方法。

118. 対象の試行非発話運動と関連付けられる脳電気信号データを記録することであって、対象が、意図された文の単語の試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、記録することと、

試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する分類モデルを使用して脳電気信号データを分析することと、を更に含む、態様 86～117 のいずれか 1 つに記載のコンピュータ実装方法。

119. 試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、態様 118 に記載の方法。

120. 手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、態様 119 に記載の方法。

121. 試行非発話運動のイベントラベルを脳電気信号データの記録中の時点に割り当てることを更に含む、態様 118～120 のいずれか 1 つに記載のコンピュータ実装方法。

122. 復号の正確度を評価することを更に含む、態様 86～121 のいずれか 1 つに記載の方法。

123. 神経記録デバイスを使用して、対象による試行発話と関連付けられる脳電気信号データを記録することであって、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをコンピューティングデバイスのプロセッサに送信する、記録することと、

プロセッサを使用して、対象による試行発話と関連付けられる記録された脳電気信号データから単語、句、又は文を復号することと、を更に含む、態様 86～122 のいずれか 1 つに記載の方法。

124. 対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データから文を復号するためのコンピュータ実装方法であって、コンピュータが、

a) 対象による意図された文の単語の文字の試行スペリングと関連付けられる記録された脳電気信号データを受信するステップと、

b) 電気信号データの記録中の任意の時点で試行スペリングが発生している確率を算出し、対象による試行スペリング中の文字生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、

c) 対象による試行文字生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測文字確率の系列を算出する文字分類モデルを使用して、脳電気信号データを分析するステップと、

d) 予測文字確率の系列に基づいて潜在的な文候補を計算し、文候補内の予測単語間の文字系列に空白を自動的に挿入するステップであって、文字系列内の復号される単語は、対象によって使用される言語の語彙内の単語のみに制約される、計算し、挿入するステップと、

e) 予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の

10

20

30

40

50

単語の確率を提供する言語モデルを使用して潜在的な文候補を分析し、文内の最も可能性の高い単語系列を決定するステップと、

f) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する、コンピュータ実装方法。

125. 記録された脳電気信号データが、対象による文字の試行スペリングの検出された開始の周りのタイムウィンドウ内でのみ使用される、態様124に記載のコンピュータ実装方法。

126. 対象が意図された文の単語の各文字の試行スペリングをいつ開始すべきかを指示する一連のゴーキューを対象に表示することを更に含む、態様124又は125に記載のコンピュータ実装方法。

127. 各ゴーキューに先行して、ゴーキューの提示へのカウントダウンが表示され、次にスペリングされる文字のカウントダウンが、各ゴーキューの後に自動的に開始される、態様126に記載のコンピュータ実装方法。

128. 一連のゴーキューが、各ゴーキュー間の設定時間間隔を提供される、態様126又は127に記載のコンピュータ実装方法。

129. 対象が、各ゴーキュー間の設定時間間隔を制御することができる、態様128に記載のコンピュータ実装方法。

130. ゴーキューに後続するタイムウィンドウ内で記録された脳電気信号データが文字分類に使用される、態様122～127のいずれか1つに記載のコンピュータ実装方法。

131. 対象の試行非発話運動と関連付けられる記録された脳電気信号データを受信することであって、対象が、意図された文の単語の試行スペリングの開始若しくは終了を示すか、又は外部デバイスを制御するために、試行非発話運動を実施する、受信することと、

試行非発話運動と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、対象が非発話運動を試行した確率を算出する分類モデルを使用して脳電気信号データを分析することと、を更に含む、態様124～130のいずれか1つに記載のコンピュータ実装方法。

132. 試行非発話運動が、頭部、腕、手、足、又は脚の試行運動を含む、態様131に記載の方法。

133. 手の試行運動が、想像上の手のジェスチャ又は想像上の手の握りを含む、態様132に記載の方法。

134. 機械学習アルゴリズムが、試行スペリング若しくは試行非発話運動の検出又は文字分類に使用される、態様124～133のいずれか1つに記載のコンピュータ実装方法。

135. 言語モデルに従って、より頻繁に出現する単語に、それほど頻繁に出現しない単語よりも大きい重みを割り当てることを更に含む、態様124～134のいずれか1つに記載のコンピュータ実装方法。

136. 対象による試行スペリング中の文字生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンに関する情報を含む対象のユーザプロファイルを格納することを更に含む、態様124～135のいずれか1つに記載のコンピュータ実装方法。

137. 電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、態様124～136のいずれか1つに記載のコンピュータ実装方法。

138. 電気信号データが、70Hz～150Hzの高ガンマ周波数範囲、及び0.3Hz～100Hzの低周波数範囲内の神経振動を含む、態様137に記載のコンピュータ実装方法。

139. 復号の正確度を評価することを更に含む、態様124～138のいずれか1つに記載のコンピュータ実装方法。

140. 対象による試行発話と関連付けられる記録された脳電気信号データから文を復号することを更に含む、コンピュータが、

10

20

30

40

50

a) 対象による試行発話と関連付けられる記録された脳電気信号データを受信するステップと、

b) 任意の時点で試行発話が発生している確率を算出し、対象による試行発話中の単語生成の開始及び終了を検出するために、発話検出モデルを使用して、記録された脳電気信号データを分析するステップと、

c) 対象による試行単語生成と関連付けられる記録された脳電気信号データ内の電気信号のパターンを識別し、予測単語確率を算出する単語分類モデルを使用して、脳電気信号データを分析するステップと、

d) 単語分類モデルからの算出された単語確率を、予測単語系列確率を算出するために単語系列内の前の単語又は句を所与として次の単語の確率を提供する言語モデルを使用した文内の予測単語系列確率と組み合わせて使用することによって、文復号を実施し、単語分類モデル及び言語モデルを使用して決定された予測単語確率に基づいて、文内の最も可能性の高い単語系列を決定するステップと、

e) 記録された脳電気信号データから復号された文を表示するステップと、を含むステップを実施する、態様 124~139 のいずれか 1 つに記載のコンピュータ実装方法。

141. 機械学習アルゴリズムが、発話検出及び単語分類、並びに文復号に使用される、態様 140 に記載のコンピュータ実装方法。

142. 人工ニューラルネットワーク (ANN) モデルが、発話検出及び単語分類に使用され、隠れマルコフモデル (HMM)、ビタビ復号モデル、又は自然言語処理技法が、文復号に使用される、態様 141 に記載のコンピュータ実装方法。

143. コンピュータ内のプロセッサによって実行されるとき、プロセッサに、態様 124~142 のいずれか 1 つに記載の方法を実行させるプログラム命令を含む、非一時的コンピュータ可読媒体。

144. 態様 143 に記載の非一時的コンピュータ可読媒体と、対象による意図された文の単語の文字の試行スペリングと関連付けられる脳電気信号データを復号するための命令と、を備える、キット。

145. 対象のコミュニケーションを支援するためのシステムであって、システムが、対象による試行発話、意図された文の単語の文字の試行スペリング、若しくは試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録するために、対象の脳の感覚運動皮質領域内のロケーションに位置するように適合されている電極を備える神経記録デバイスと、

態様 124~142 のいずれか 1 つに記載のコンピュータ実装方法に従って、記録された脳電気信号データから文を復号するようにプログラムされているプロセッサと、

コンピューティングデバイスと通信するインターフェースであって、当該インターフェースが、対象の頭部上のロケーションに位置するように適合されており、インターフェースが、神経記録デバイスから脳電気信号データを受信し、脳電気信号データをプロセッサに送信する、インターフェースと、

記録された脳電気信号データから復号された文を表示するためのディスプレイコンポーネントと、を備える、システム。

146. 対象が、構語障害、脳卒中、外傷性脳損傷、脳腫瘍、又は筋萎縮性側索硬化症のために当該コミュニケーションが困難である、態様 145 に記載のシステム。

147. 神経記録デバイスのロケーションが、腹側感覚運動皮質内にある、態様 145 又は 146 に記載のシステム。

148. 電極が、感覚運動皮質領域の表面上又は感覚運動皮質領域内に位置するように適合されている、態様 145~147 のいずれか 1 つに記載のシステム。

149. 電極が、硬膜下腔内の脳の感覚運動皮質領域の表面上に位置するように適合されている、態様 148 に記載のシステム。

150. 神経記録デバイスが、脳侵入電極アレイを含む、態様 145~149 のいずれか 1 つに記載のシステム。

151. 神経記録デバイスが、皮質脳波記録 (ECOG) 電極アレイを含む、態様 14

10

20

30

40

50

5～150のいずれか1つに記載のシステム。

152．電極が、深部電極又は表面電極である、態様145～151のいずれか1つに記載のシステム。

153．電気信号データが、高ガンマ周波数成分特徴及び低周波数成分特徴を含む、態様145～152のいずれか1つに記載のシステム。

154．電気信号データが、70Hz～150Hzの高ガンマ周波数範囲、及び0.3Hz～100Hzの低周波数範囲内の神経振動を含む、態様153に記載のシステム。

155．インターフェースが、対象の頭蓋に取り付けられた経皮台座コネクタを備える、態様145～154のいずれか1つに記載のシステム。

156．インターフェースが、経皮台座コネクタに接続可能であるヘッドステージを更に備える、態様155に記載のシステム。

157．プロセッサが、コンピュータ又はハンドヘルドデバイスによって提供される、態様145～156のいずれか1つに記載のシステム。

158．ハンドヘルドデバイスが、携帯電話又はタブレットである、態様157に記載のシステム。

159．態様145～158のいずれか1つに記載のシステムと、対象による試行発話、単語の試行スペリング、若しくは試行試行非発話運動、又はこれらの組み合わせと関連付けられる脳電気信号データを記録し、復号するために、システムを使用するための命令と、を備える、キット。

【実施例】

【0179】

上記に提供される開示から理解され得るように、本開示は、多種多様な用途を有する。したがって、以下の実施例は、当業者に、本発明の作製及び使用方法の完全な開示及び説明を提供するために提示されるものであり、発明者が発明とみなす範囲を限定することを意図せず、また、以下の実験が行われる全て又は唯一の実験であることを表すことを意図するものではない。使用される数字（例えば、量、寸法など）に対する正確性を確保する努力がなされているが、ある程度の実験誤差及び偏差が考慮されるべきである。当業者は、本質的に同様の結果を得るために変更又は修正することができる様々な重要でないパラメータを容易に認識するであろう。

【0180】

実施例1：重度の麻痺を患うヒトの単語を復号するための発話神経プロテーゼ序論

構語障害とは、発話能力の喪失である。構語障害は、脳卒中、外傷性脳損傷、及び筋萎縮性側索硬化症を含む様々な状態に起因し得る[1]。重度の運動障害を有する麻痺した個人の場合、構語障害は、家族、友人、及び介護者とのコミュニケーションを妨げ、自己申告の生活の質を低下させる[2]。

【0181】

障害者がカーソル制御を使用して意図されたメッセージをスペルアウトことを可能にするタイピングベースの脳コンピュータインターフェースが進歩している[3～7]。しかしながら、神経信号記録によって駆動される、文字毎に選択するインターフェースは、比較的遅く、面倒であり得る。より効率的で自然なアプローチは、発話を制御する脳領域からの単語全体を直接復号することであり得る。過去10年間で、発話運動皮質が声道の急速な関節運動をどのように調整しているかについての理解が広がっている[8～13]。並行して、技術的努力が、これらの発見を活用して、発話障害のない人々の脳活動から発話を復号することができることを実証している[14～17]。

【0182】

しかしながら、発話復号アプローチが、発話することができない麻痺した個人で機能するか否かは不明である。神経活動は、発話出力がないため意図された発話と正確に整合させることができず、計算モデルを訓練するための障害となっている[18]。加えて、発話制御の基礎となる神経信号が、何年も、又は何十年も発話していない個人において、依

10

20

30

40

50

然として正常であるか否かは不明である。以前の研究では、閉じ込め症候群を患っているヒトが、埋め込まれた2チャンネルの微小電極デバイスを使用して、視聴覚インターフェースを通じて母音及び音素を生成した〔19、20〕。構語障害のヒトの神経活動から完全な単語を確実に復号することが可能か否かは不明なままである。

【0183】

本研究では、遠隔脳幹卒中に起因する重度の麻痺及び構語障害を患っているヒトの神経活動からのリアルタイムの単語及び文の復号を実証する（図1）。本発明者らの調査結果は、直接発話脳コンピュータインターフェースを通じた長期的なコミュニケーション回復のための概念実証を表している。

【0184】

方法

試験概説

本研究は、長期的なコミュニケーション及び運動回復のための皮質脳波記録（ECoG：脳の表面から直接神経活動を記録する方法）及びカスタム復号技法の可能性を評価することを目的とした単一施設臨床試験であるBRAVO試験（BCI Restoration of Arm and Voice function、clinicaltrials.gov登録番号NCT03698149）の一環として実施された。本研究で使用されたECoGデバイスは、米国食品医薬品局による治験用機器の適用免除の承認を受けた。執筆時点で、1人の臨床試験参加者（「Bravo-1」、本研究の参加者）のみがECoGデバイスを埋め込まれている。

【0185】

参加者

参加者は、研究開始時に36歳であった右利きの男性である。20歳で、彼は、重度の痙性四肢麻痺及び構語障害（言語聴覚士及び神経学者によって診断された、図5）をもたらした、右椎骨動脈解離に関連する広範囲の両側性ポンティニュー脳卒中を患った。彼は認知的に正常である（ミニメンタルステート試験で評価した）。彼はうなり声及びうめき声を発声することはできるが、理解可能な発話を生成することはできない。彼は通常、頭部の残留運動によって制御される補助的なコンピュータベースのタイピングインターフェースを使用してコミュニケーションし、タイピング速度は、1分当たり約5つの正しい単語又は18の正しい文字であった（補足方法S1）。

【0186】

インプラントデバイス

参加者から脳信号を取得するために使用される神経インプラントは、高密度ECoG電極アレイ（PMT Corporation、MN、USA）と、台座コネクタ（Blackrock Microsystems、UT、USA）とのカスタマイズされたハイブリッドである。ECoGアレイは、4mmの中心間隔を有する128の平坦なディスク形状の電極で構成されている。外科的移植の間、発話感覚運動皮質が、開頭術を介して露出され、アレイは、硬膜下腔内の脳の表面上に配置された。硬膜を縫合して閉じ、頭蓋骨弁を交換した。経皮台座コネクタを別個の部位に配置し、小型チタンネジで頭蓋骨に固定した。この台座コネクタは、脳信号を取得し、取り外し可能なデジタルコネクタ及びケーブル（図1）を介してコンピュータに送信することができる、外部からアクセス可能なプラットフォームである。参加者は、2019年初頭にデバイスの外科的埋め込みを受けた。手術は成功し、無事に回復した。電極被覆は、左中心前回、中心後回、後中前頭回、及び後下前頭回の一部を含む、発話処理に関与している複数の皮質領域からのサンプリングを可能にした〔8、10～12〕。

【0187】

神経データ取得及びリアルタイム処理

デジタル信号処理ユニット及び周辺ハードウェア（NeuroPortシステム、Blackrock Microsystems）を使用して、インプラントデバイスの全128チャンネルからの信号を取得し、リアルタイム分析のためにカスタムソフトウェアを実

10

20

30

40

50

行する別のコンピュータに送信した（補足方法 S 2、図 6 及び図 7）[16、21]。このコンピュータ上で、各チャンネルの高ガンマ活性（70～150 Hz の周波数範囲内の神経振動）を測定し、その後の全ての分析及びリアルタイム復号中に使用した。

【0188】

タスク設計

参加者は、孤立単語タスク及び文タスク（補足方法 S 3）の 2 つのタスクにエンゲージした。各タスクの各試験において、参加者は、テキスト標的を視覚的に提示され、次いで、その標的の生成（声に出しての発語）を試行した。

【0189】

孤立単語タスクにおいて、参加者は 50 の英単語からなるセットからの個々の単語の生成を試行した。この単語セットには、介護に関連する単語及び参加者が要求する単語を含む、様々な文を作成するために使用できる一般的な英単語が含まれていた。各試験において、参加者は、これらの 50 の単語のうちの 1 つを提示され、短い遅延の後、視覚ゴーキユーが提示されたときにその単語の生成を試行した。

【0190】

孤立単語タスクにおいて、参加者は、50 単語セットからの単語のみから構成される 50 の英文からなるセットからの単語系列の生成を試行した（補足方法 S 4 及び S 5）。各試験において、参加者は、標的文を提示され、その文の単語を（順番に）自分が快適にできる最速の速度で生成することを試行した。試験全体を通して、神経活動から復号された単語系列をリアルタイムで更新し、参加者にフィードバックとして表示した。

【0191】

モデリング

タスク中に収集された神経活動を使用して、カスタムモデルを訓練、最適化、及び評価した（補足方法 S 6 及び S 7、図 8、補表 S 1）。具体的には、両方とも、深層学習技法を活用して神経活動から予測を行う発話検出及び単語分類モデルを作成した。文タスク中にリアルタイムで参加者の神経活動から文を復号するために、これらの 2 つのモデル、言語モデル、及びビタビデコーダ（図 1）を含む復号パイプラインを使用した。

【0192】

発話検出器は、タスク中の神経活動の各時点进行处理し、試行単語生成イベントの開始及び終了をリアルタイムで検出した（補足方法 S 8、図 9）。孤立単語タスクからの神経データ及びタスクタイミング情報のみを使用して、このモデルを適合させた。

【0193】

検出されたイベントごとに、単語分類器は、検出された開始の 1 秒前から 3 秒後までに及ぶ神経活動进行处理することによって、単語確率のセットを予測した（補足方法 S 9、図 10）。50 単語セット内の各単語と関連付けられる予測された確率は、検出されたイベント中に参加者がその単語の発語を試行した可能性がどの程度であったかを定量化した。孤立単語タスクからの神経データを使用して、このモデルを適合させた。

【0194】

英語では、特定の単語系列は他の単語系列よりも可能性が高くなる。本発明者らは、系列の前の単語を所与として次の単語の確率をもたらす言語モデルを使用することによって、この基礎となる構造を活用した[22、23]（補足方法 S 10）。このモデルは、クラウドソーシングプラットフォーム上でカスタムタスクを使用して得られた 50 単語セットからの単語のみから構成される文のコレクションで訓練した（補足方法 S 4）。

【0195】

本発明者らは、復号パイプラインの最終コンポーネントとしてカスタムビタビデコーダを使用した。これは、単語分類器からの予測単語確率及び言語モデルからの単語系列確率を所与として最も可能性の高い単語系列を決定するタイプのモデルである[24]（補足方法 S 11、図 11）。言語モデルを組み込むことによって、ビタビデコーダは、単語分類器からの予測単語を単に連結することから生じるものよりも妥当な文を復号することができた。

10

20

30

40

50

【0196】

評価

本発明の復号パイプラインのパフォーマンスを評価するために、単語誤り率及び毎分単語数の2つのメトリックを使用して、リアルタイムで復号された文を分析した（補足方法S12）。復号された文の単語誤り率は、編集距離（その文の単語誤り数）を標的文の単語数で除算した値として定義される。毎分単語メトリックは、神経データの1分あたりに復号された単語の数を測定する。また、リアルタイム復号中の本発明のシステムの待ち時間も測定した。

【0197】

参加者の神経活動からの単語生成試行の検出及び分類を更に特徴付けるために、本発明者らは、オフライン分析において、発話検出器及び単語分類器を用いて孤立単語データを処理した（補足方法S13を参照されたい）。パフォーマンスが訓練データの量によってどのように影響されたかを評価するために、本発明者らは、訓練中に使用される試験の数を変化させながら、単語分類器からの予測単語確率を使用して分類正確度を測定した。ここで、分類正確度は、単語分類器が標的単語に最も高い確率を正しく割り当てた予測の割合に等しい。また、神経活動の各チャンネルがモデルの予測に及ぼした影響を測定することによって、各電極が検出及び分類に及ぼした寄与を測定した[17、25]。

【0198】

長期適用のための本発明のアプローチの臨床的生存可能性を調査するために、本発明者らは、孤立単語データを使用して、取得されたECOG信号の経時的な安定性を評価した（補足方法S14）。最初に、単語生成試行中に収集された神経応答の大きさが、81週研究期間の過程にわたって変化したか否かを決定した。また、4つの異なる日付範囲（「早期」、「中間」、「後期」、及び「最後期」）からサンプリングされた神経データを使用してモデルを訓練及び試験し、得られた分類正確度及び電極寄与を比較することによって、検出及び分類パフォーマンスが研究期間全体を通して安定していたか否かを評価した。

【0199】

統計分析

本研究で使用した統計的検定は、対応する有意性主張とともに記載されており、検定の詳細な説明は補足方法S15に記載されている。簡潔には、ウィルコクソン符号順位検定を使用して、復号パフォーマンスを偶然と比較し、言語モデルのパフォーマンスへの影響を評価し（単語誤り率メトリックを用いて）、線形混合効果モデリングを使用して信号安定性を評価し、フィッシャー正確確率検定及び正確マクネマー検定を使用して、異なる日付範囲にわたる分類正確度を比較し、ウィルコクソン符号順位検定を使用して、異なる日付範囲にわたる電極の寄与を比較した。全ての検定について、0.01のアルファレベルを使用した。同じタイプの個々の統計検定において使用される神経データが互いに独立していない場合、複数の比較を考慮するためにホルム・ボンフェローニ補正を使用した。

【0200】

結果

文復号

リアルタイム文復号中、文ブロック（各ブロックが10回の試験を含む）にわたる復号単語誤り率の中央値は、言語モデリングを用いない場合は60.5%、言語モデリングを用いる場合は25.6%であった（図2A）。単一の検定ブロックについて観察された最低単語誤り率は、6.98%であった（言語モデリングを用いて）。単語誤り率は偶然よりも有意に良好であり、言語モデルを組み込むと有意に減少した（ $P < 0.001$ 、片側ウィルコクソン符号順位検定、3方向ホルム・ボンフェローニ補正）。150回の全ての試験にわたって、復号速度の中央値は、全ての復号された単語を含む場合、15.2単語毎分であり、正しく復号された単語のみを含む場合、12.5単語毎分であった（図2B）。92.0%の試験において、検出された単語の数は、標的文の単語の数に等しかった（図2C）。検出された文の長さは、試験の2.67%では少なくとも1つの単語が短す

ぎ、試験の 5.33%では少なくとも 1つの単語が長すぎた。15の全ての文ブロックにわたって、5つの発話イベントが、ブロック内の最初の試験の前に誤って検出され、リアルタイムの復号及び分析から除外された（他の全ての検出された発話イベントは含まれた）。ほぼ全ての標的文について、言語モデルを使用したときに平均編集距離が減少した（図 2D）。更に、文の半分以上が誤りなく復号された（言語モデリングを用いて 150回の試験のうち 80回、ゼロの編集距離によって示された）。復号中の言語モデルの使用は、文法的及び意味的に妥当でない単語予測を補正することによってパフォーマンスを改善した（図 2E）。リアルタイム単語予測と関連付けられる平均待ち時間は、4.0秒であると推定された（標準偏差は 0.91秒であった）。

【0201】

単語検出及び分類

皮質活動の検出されたタイムウィンドウを使用した孤立単語生成試行のオフライン分析中、分類正確度は、訓練データの量が増大するにつれて増大した（利用可能な全てのデータを使用した場合は最大 47.1%、図 3A）。パフォーマンスは、訓練データの最初の 4時間はより急速に改善し、その後の 5時間はそれほど急速ではなかったが、プラトーではなかった。孤立単語データにおける 9000語の生成試行のうち、98%が首尾よく検出され（191回の試行は、検出されたイベントに関連付けられなかった）、及び 968回の検出されたイベントがスプリアスであった（試行と関連付けられなかった。追加の孤立単語分析結果については、図 12及び図 13を参照されたい）。単語分類パフォーマンスに寄与する電極は、主に、腹側感覚運動皮質（vSMC）の最も腹側の態様に局在化され、vSMCの背側の態様の電極は、発話検出及び単語分類の両方のパフォーマンスに寄与した（図 3B）。全体的に、電極寄与は、単語分類よりも発話検出の方に多く分布し、総寄与のうちの 50%以上が、単語分類器の上位 37個の電極、及び発話検出器の上位 50個の電極からのものであった。単語混同分析は、単語標的の大部分にわたって一貫した分類正確度を明らかにした（図 3C、行正規化された混同行列の対角線に沿った分類正確度の平均 47.1%及び標準偏差 14.5%）。

【0202】

長期信号安定性

81週間の研究期間全体を通して、単語生成試行中に比較的安定した単一試験神経活動パターンを観察した（図 4A）。全ての電極及び孤立単語試行にわたって、発話試行中の神経応答の大きさに対する、埋め込み以来の時間のわずかに全体的な否定的影響があった（勾配 = -0.00021、SE = 0.000011、 $P < 0.001$ 、線形混合効果モデリング、129方向ホルム・ボンフェローニ補正、図 14）。しかしながら、個々の電極モデリングは、128個の電極のうち 4個のみにおいて有意な効果を明らかにした（1個は肯定的、3個は否定的、 $P < 0.01$ 、線形混合効果モデリング、129方向ホルム・ボンフェローニ補正）。

【0203】

別個の日付範囲からの孤立単語データのサブセットに関して発話検出器及び単語分類器を訓練及びテストすることによって、分類正確度が最も早いサブセットについて最も低く、残りのサブセットにわたって相対的に一貫していることが分かった（「早期」対「後期」比較について $P = 0.0015$ 、他の全ての比較について $P > 0.01$ 、両側フィッシャー正確確率検定、10方向ホルム・ボンフェローニ補正、図 4B）。2つの最新のサブセットにおけるデータを評価するとき、分類正確度は、他のサブセットからのデータとは対照的に、同じサブセット内からのデータに関してモデルを訓練するとき有意に高かった（「後期」及び「最後期」サブセットについて $P < 0.001$ 、他のサブセットについて $P > 0.01$ 、両側正確マクネマー検定、10方向ホルム・ボンフェローニ補正）。4つのサブセットにわたって電極寄与に有意な変化はなかった（全て $P > 0.32$ 、両側ウィルコクソン符号順位検定、未修正）。

【0204】

考察

重度の麻痺者の皮質活動の高分解能記録を使用して、完全な単語及び文をリアルタイムで復号することができることを実証した。本発明者らの深層学習モデルは、神経活動からの単語生成試行を検出して分類することができ、これらのモデルを言語モデリング技術とともに使用して、様々な意味のある文を復号することができた。神経インターフェースから記録された信号は、研究期間全体を通して安定性を呈し、外科的埋め込み後最大90週間であっても復号に成功することを可能にした。まとめると、これらの結果は、発話神経機能代替技術の恩恵を受ける可能性のある麻痺者に即座に実用的な意味をもたらす。

【0205】

神経活動からの単語及び文の復号の以前の実証は、正常な発話を保持し、コミュニケーションに支援技術を必要としない参加者で行われた[14~17]。発話することができないヒトとの発話を復号するとき、意図された発話と神経活動との間の正確な時間的整合性の欠如は、モデル訓練中に重大な課題を提起する。ここでは、この時間整合問題を検出技法[16、26、27]、並びに、モデルアンサンブル及びデータ増大（補足方法S9で説明）などの機械学習の進歩を活用して、小さい時間変動に対する耐性を高める分類器によって管理した[28、29]。追加的に、本発明の復号モデルは、腹側感覚運動皮質の神経活動パターンを活用しており、これは、この領域が正常な発話生成に関与していることを示唆する以前の研究と一致している[8、11、12]。この結果は、運動機能低下の数年後の四肢麻痺者における四肢関連皮質運動表現の以前の所見に類似した、構語障害の15年以上後の機能性皮質発話表現の持続性を実証している[30]。

【0206】

不完全な単語分類パフォーマンスにもかかわらず、言語モデリング技法の組み込みは、文試験の半分超で完璧な復号を可能にした。この改善は、単語分類器からの追加の確率的情報（検出された単語生成試行ごとの最も可能性の高い単語識別子を超えた）を活用し、デコーダが新しい入力を所与として以前の誤りを補正することを可能にすることによって促進された。これらの結果は、神経記録から発話を復号するときに言語情報を統合する利点を実証している。発話復号アプローチは、一般的に、30%未満の単語誤り率で使用可能になり[31]、これは本発明のアプローチが臨床環境で即座に適用できることを示唆している。

【0207】

長期脳コンピュータインターフェース（BCI）を設計する際の基本的な考慮事項は、神経記録モダリティ（例えば、侵襲的対非侵襲的）の選択、及びこの選択が取得される神経信号の分解能、空間カバレッジ、及び安定性に及ぼす影響である。以前の運動制御BCI研究において、皮質脳波記録（ECOG、本研究で使用されている記録モダリティ）は、他の記録モダリティと比較して、長い評価期間にわたって比較的高い信号安定性を有することが実証されている[4、32-34]。しかし、これらの復号の取り組みは、限られたチャンネルカウント及び空間カバレッジによって制約されていた。本発明の高密度ECOGデバイスによって、広い空間カバレッジ及び高い空間分解能を活用して、研究全体を通して比較的安定した皮質活動を観察しながら、単語が確実に復号された（3つの電極のみが時間の経過とともに神経応答の大きさを大幅に減少させた）。オフライン分類パフォーマンスは改善され、その後、研究の最初の数週間後にほとんど安定した。これは、埋め込み後の早期の治癒中の脳組織の落ち着きによって説明される可能性があり得る[35、36]。このインプラントデバイス及び研究参加者を対象とした最近のカーソル制御研究と一致して[37]、本発明の結果は、ECOGベースのBCIが、時折モデルの再較正を行うことで、数ヶ月間にわたって一貫した発話復号パフォーマンスを維持することができることが示されている。全体として、本発明者らの知見は、てんかんに対する応答性神経刺激[35、36]及び長期BCI対照[34、37]に対するECOGベースのインターフェースの長期生存性、安全性、及び信号安定性の実証に加わって、これらの属性を高密度ECOGを有する発話BCIを含むように拡張した。

【0208】

発話は典型的には、健康なヒトにとって最も速く、最も自然で、最も効率的なコミュニ

ケーション方法である〔38〕。本発明の現在の復号速度は、130語毎分を超えることが多い自然発話速度よりもはるかに遅いが〔38, 39〕、これらの結果は、構語障害を患う麻痺者の皮質信号からの直接発話復号の早期の実現可能性を実証している。この原理の証明から、本発明によれば、新規のデコーダを開発して評価して、より大きい語彙によってより幅広い種類の文の生成を可能にすることができる。最終的には、復号正確度、柔軟性、及び速度を向上させるための今後の研究を通じて、本発明では、重度のコミュニケーション障害を患う人々のために、発話ベースの神経機能代替義肢のコミュニケーションの可能性を最大限に実現することを目標とする。

【0209】

参考文献：

1. Beukelman DR, Fager S, Ball L, Dietz A. A AC for adults with acquired neurological conditions: A review. *Augmentative and Alternative Communication* 2007; 23 (3) : 230-42.
2. Felgoise SH, Zaccheo V, Duff J, Simmons Z. Verbal communication impacts quality of life in patients with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 2016; 17 (3-4) : 179-83.
3. Sellers EW, Ryan DB, Hauser CK. Noninvasive brain-computer interface enables communication after brainstem stroke. *Science translational medicine* 2014; 6 (257) : 257 re7.
4. Vansteensel MJ, Pels EGM, Bleichner MG, et al. Fully Implanted Brain-Computer Interface in a Locked-In Patient with ALS. *New England Journal of Medicine* 2016; 375 (21) : 2060-6.
5. Pandarinath C, Nuyujukian P, Blabe CH, et al. High performance communication by people with paralysis using an intracortical brain-computer interface. *ELife* 2017; 6 : 1-27.
6. Brumberg JS, Pitt KM, Mantie-Kozłowski A, Burnison JD. Brain-Computer Interfaces for Augmentative and Alternative Communication: A Tutorial. *Am J Speech Lang Pathol* 2018; 27 (1) : 1-12.
7. Linse K, Aust E, Joos M, Hermann A, Oliver DJ. Communication Matters-Pitfalls and Promise of Hightech Communication Devices in Palliative Care of Severely Physically Disabled Patients With Amyotrophic Lateral Sclerosis. 2018; 9 (July) : 1-18.
8. Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. *Nature* 2013; 495 (7441) : 327-32.

10

20

30

40

50

9. Lotte F, Brumberg JS, Brunner P, et al. Electrographic representations of segmental features in continuous speech. *Frontiers in Human Neuroscience* 2015;09 (February):1–13.

10. Guenther FH, Hickok G. Neural Models of Motor Speech Control. In: *Neurobiology of Language*. Elsevier; 2016. p. 725–40.

11. Mugler EM, Tate MC, Livescu K, Templer JW, Goldrick MA, Slutzky MW. Differential Representation of Articulatory Gestures and Phonemes in Precentral and Inferior Frontal Gyri. *The Journal of Neuroscience* 2018;4653:1206–18.

12. Chartier J, Anumanchipalli GK, Johnson K, Chang EF. Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. *Neuron* 2018;98(5):1042–1054.e4.

13. Salarì E, Freudenburg ZV, Branco MP, Arnoutse EJ, Vansteensel MJ, Ramsey NF. Classification of Articulator Movements and Movement Direction from Sensorimotor Cortex Activity. *Sci Rep* 2019;9(1):14165.

14. Herff C, Heger D, de Pesters A, et al. Brain-to-text: decoding spoken phrases from phone representations in the brain. *Frontiers in Neuroscience* 2015;9(June):1–11.

15. Anumanchipalli GK, Chartier J, Chang EF. Speech synthesis from neural decoding of spoken sentences. *Nature* 2019;568(7753):493–8.

16. Moses DA, Leonard MK, Makin JG, Chang EF. Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nat Commun* 2019;10(1):3096.

17. Makin JG, Moses DA, Chang EF. Machine translation of cortical activity to text with an encoder-decoder framework. *Nat Neurosci* 2020;23(4):575–82.

18. Martin S, Iturrate I, Millan J del R, Knight RT, Pasley BN. Decoding Inner Speech Using Electrographic: Progress and Challenges Toward a Speech Prosthesis. *Front Neurosci* 2018;12:422.

19. Guenther FH, Brumberg JS, Wright EJ, et al. A Wireless Brain-Machine Interface for Real-Time Speech Synthesis. *PLoS ONE* 2009;4(12):e8218.

20. Brumberg JS, Wright EJ, Andreasen DS, G

10

20

30

40

50

uenther FH, Kennedy PR. Classification of intended phoneme production from chronic intracortical microelectrode recordings in speech-motor cortex. *Front Neurosci* 2011;5:65.

21. Moses DA, Leonard MK, Chang EF. Real-time classification of auditory sentences using evoked cortical activity in humans. *J Neural Eng* 2018;15(3):036005.

22. Kneser R, Ney H. Improved backing-off for M-gram language modeling. In: 1995 International Conference on Acoustics, Speech, and Signal Processing. Detroit, MI, USA: IEEE; 1995. p. 181-4.

23. Chen SF, Goodman J. An empirical study of smoothing techniques for language modeling. *Computer Speech & Language* 1999;13(4):359-93.

24. Viterbi AJ. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm. *IEEE Transactions on Information Theory* 1967;13(2):260-9.

25. Simonyan K, Vedaldi A, Zisserman A. Deep Inside Convolutional Networks: Visualizing Image Classification Models and Saliency Maps. In: Bengio Y, LeCun Y, editors. Workshop at the International Conference on Learning Representations. Banff, Canada: 2014.

26. Kanas VG, Mporas I, Benz HL, Sgarbas KN, Bezzerianos A, Crone NE. Real-time voice activity detection for ECoG-based speech brain machine interfaces. In: 19th International Conference on Digital Signal Processing. 2014. p. 862-5.

27. Dash D, Ferrari P, Dutta S, Wang J. NeuroVAD: Real-Time Voice Activity Detection from Non-Invasive Neuromagnetic Signals. *Sensors* 2020;20(8):2248.

28. Sollich P, Krogh A. Learning with ensembles: How overfitting can be useful. In: Touretzky DS, Mozer MC, Hasselmo ME, editors. *Advances in Neural Information Processing Systems* 8. MIT Press; 1996. p. 190-196.

29. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burgess CJC, Bottou L, Weinberger KQ, editors. *Advances in Neural Information Processing Systems* 25. Curran Associates, Inc.; 2012. p

10

20

30

40

50

. 1097-1105.

30. Shoham S, Halgren E, Maynard EM, Norman RA. Motor-cortical activity in tetraplegics. *Nature* 2001; 413 (6858) : 793-793.

31. Watanabe S, Delcroix M, Metze F, Hershey JR. New era for robust speech recognition: exploiting deep learning. Berlin, Germany: Springer-Verlag; 2017.

32. Chao ZC, Nagasaka Y, Fujii N. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkey. *Front Neuroeng* 2010; 3:3.

33. Freudenbourg ZV, Branco MP, Leinders S, et al. Sensorimotor ECoG Signal Features for BCI Control: A Comparison Between People With Locked-In Syndrome and Able-Bodied Controls. *Front Neurosci* 2019; 13:1058.

34. Pels EGM, Aarnoutse EJ, Leinders S, et al. Stability of a chronic implanted brain-computer interface in late-stage amyotrophic lateral sclerosis. *Clinical Neurophysiology* 2019; 130 (10) : 1798-803.

35. Rao VR, Leonard MK, Kleen JK, Lucas BA, Mirro EA, Chang EF. Chronic ambulatory electrocorticography from human speech cortex. *NeuroImage* 2017; 153: 273-82.

36. Sun FT, Arcot Desai S, Tcheng TK, Morrell MJ. Changes in the electrocorticogram after implantation of intracranial electrodes in humans: The implant effect. *Clinical Neurophysiology* 2018; 129 (3) : 676-86.

37. Silversmith DB, Abiri R, Hardy NF, et al. Plug-and-play control of a brain-computer interface through neural map stabilization. *Nat Biotechnol* 2020.

38. Hauptmann AG, Rudnicky AI. A comparison of speech and typed input. In: *Proceedings of the workshop on Speech and Natural Language-HLT '90*. Hidden Valley, Pennsylvania: Association for Computational Linguistics; 1990. p. 219-24.

39. Waller A. Telling tales: unlocking the potential of AAC technologies. *International Journal of Language & Communication Disorders* 2019; 54 (2) : 159-69.

[0210]

実施例2：単語復号のための補足方法

方法S1．参加者の補助タイピングデバイス

補助タイピングデバイスの説明

参加者は、多くの場合、市販のタッチスクリーンタイピングインターフェース（To b

10

20

30

40

50

i i Dynavox) を使用して他のヒトとコミュニケーションし、頭及び首の残留運動を使用して野球帽に取り付けられた長い(約18インチ)プラスチック製のスタイラスで制御する。デバイスは、参加者がスタイラスによって選択することができる文字、単語、及び他のオプション(句読点など)を表示し、参加者がテキスト文字列を構築することを可能にする。所望のテキスト文字列を作成した後、参加者は自分のスタイラスを使用して、テキスト文字列を合成して可聴発話波形にするアイコンを押すことができる。所望のメッセージをスペルアウトし、デバイスにそれを合成させるこのプロセスは、介護者及び訪問者との参加者の典型的なコミュニケーション方法である。

【0211】

タイピング速度評価タスク設計

本発明のシステムで達成された神経ベースの復号速度と比較するために、参加者がカスタムタスクでタイピングインターフェースを使用している間の参加者のタイピング速度を測定した。この課題の各試験において、画面上に単語又は文を提示し、参加者はタイピングインターフェースを使用してその単語又は文を打ち込んだ。参加者に、自身のインターフェースで単語サジェスト又は自動補完オプションを一切使用しないように指示したが、補正機能(バックスペースや元に戻すオプションなど)の使用は許可されていた。標的単語又は文が最初に画面に表示されてから、参加者が標的の最後の文字を入力するまでの時間を測定した。次いで、この持続時間及び標的単語又は発言を使用して、各試験の1分当たりの単語数及び1分当たりの正しい文字の数を測定した。

【0212】

合計35回の試験(25語及び10文)を使用した。参加者に提示されたときに句読点が含まれていたが、参加者はタスク中に句読点を入力しないように指示された。標的単語及び文は次のとおりである。

1. Th i r s t y
2. I
3. T i r e d
4. A r e
5. U p
6. H o w
7. O u t s i d e
8. Y o u
9. B a d
10. C l e a n
11. H a v e
12. T e l l
13. H e l l o
14. G o i n g
15. R i g h t
16. C l o s e r
17. W h a t
18. S u c c e s s
19. I t
20. F a m i l y
21. T h a t
22. H e l p
23. D o
24. A m
25. O k a y
26. I t i s g o o d .
27. I a m t h i r s t y .

28. They are coming here.
29. Are you going outside?
30. I am outside.
31. Faith is good.
32. My family is here.
33. Please tell my family.
34. My glasses are comfortable.
35. They are coming outside.

【0213】

タイピング速度結果及び考察

このタイピングタスクの全ての試験にわたって、参加者のタイピング速度の平均±標準偏差は、1分当たり5.03±3.24個の正しい単語又は1分当たり17.9±3.47個の正しい文字であった。

これらのタイピング速度は、本発明のアプローチのリアルタイム復号速度よりも遅いが、タイピングインターフェースの無制限の語彙サイズは、本発明のアプローチにまさる重要な利点である。参加者がタイピングインターフェースを用いて達成できる1分当たりの正しい文字数を所与として、インターフェース内の文字をこのタスクからの50語に置き換えると、本発明者らのアプローチで達成されたものよりも高い復号速度及び正確度が得られる可能性がある。しかしながら、このタイピングインターフェースはあまり自然ではなく、発話試行よりも多くの身体的労力を必要とすると考えられ、タイピングインターフェースが本発明者らのアプローチよりも疲れやすい可能性があることを示唆している。

【0214】

方法S2. 神経データ取得及びリアルタイム処理

初期データ取得及び前処理ステップ

埋め込み皮質脳波記録 (ECoG) アレイ (PMT Corporation) は、4 mmの中心間隔を有する16×8格子構成に配置構成された電極を含む。矩形ECoGアレイは、長さ6.7 cm、幅3.5 cm、及び厚さ0.51 mmであり、電極接点は、2 mmの接点直径を有するディスク形状である。神経データを処理及び記録するために、信号をECoGアレイから取得し、複数のハードウェアデバイスを伴ういくつかのステップで処理した (図6及び図7)。まず、経皮台座コネクタ (ブラックロックマイクロシステムズ) に接続されたヘッドステージ (取り外し可能デジタルリンク、Blackrock Microsystems) が、埋め込み電極アレイから電位を取得した。台座はオスコネクタ、ヘッドステージはメスコネクタである。このヘッドステージは、0.3 Hz〜7.5 kHzのハードウェアベースのバターワースフィルタを使用して信号に対するバンドパスフィルタリングを実施した。次いで、デジタル化された信号 (16ビット、ビット解像度当たり250 nV) をHDMIケーブルを通じてデジタルハブ (Blackrock Microsystems) に送信し、デジタルハブは、光ファイバケーブルを通じてNeuroportシステム (Blackrock Microsystems) にデータを送信した。早期の記録セッションにおいて、デジタルヘッドステージがヒト研究に使用することが承認される前に、ヒト患者ケーブル (Blackrock Microsystems) を使用して、台座をフロントエンド増幅器 (Blackrock Microsystems) に接続した。増幅器は、光ファイバを通じてNeuroportシステムに送信する前に、号を増幅してデジタル化する。このNeuroportシステムは、30 kHzでECoGデータの128チャンネル全てをサンプリングし、ソフトウェアベースの回線ノイズキャンセルを適用し、500 Hzでアンチエイリアシングローパスフィルタリングを実施し、次いで、処理された信号を1 kHzで別のリアルタイム処理マシン (Colfax International) にストリーミングした。Neuroportシステムはまた、30 kHz (リアルタイム処理コンピュータからのマイクロフォン入力及びスピーカ出力) での関連する音響の同期録音を取得、ストリーミング、及び保存した。

【0215】

更なる前処理及び特徴抽出

Linuxマシン(64ビットUbuntu 18.04、48 Intel Xeon Gold 6146 3.20GHzプロセッサ、500GBのRAM)であるリアルタイム処理コンピュータが、リアルタイム神経発話認識(rtNSR) [1、2]と呼ばれるカスタムソフトウェアパッケージを使用して、入来する神経データを分析及び処理し、タスクを実行し、リアルタイム復号を実施し、タスクデータ及びメタデータをディスクに格納した。このソフトウェアを使用して、取得した全ての神経信号に対して次の前処理ステップをリアルタイムで実施した。

【0216】

(全ての電極にわたって)取得したECOGデータの各時間サンプルに共通平均基準を適用した。これは、マルチチャネルデータの共有ノイズを低減するための標準的な技法である [3、4]。

【0217】

高ガンマ帯域(72.0、79.5、87.8、96.9、107.0、118.1、130.4、及び144.0Hz、小数点以下第1位に四捨五入)で対数的に増大する中心周波数を有する8つのバンドパス有限インパルス応答(FIR)フィルタを適用した。これらの390次フィルタの各々は、Parkes-McClellanアルゴリズムを使用して設計された [5]。

【0218】

ヒルベルト変換を近似するためにParkes-McClellanアルゴリズムによって設計された170次FIRフィルタを使用して、各バンド及びチャネルについて分析振幅値を計算した。各バンド及びチャネルについて、元の信号(フィルタ次数の半分である85サンプルの遅延)を実成分として、元の信号のヒルベルト変換(このFIRフィルタによって近似)を虚成分として使用して、分析信号を推定した [6]。その後、これらの分析信号の各々の大きさを計算することによって分析振幅値を得た。この分析振幅算出は、バンドパスされた信号の4つおきのサンプルにのみ適用され、分析振幅は200Hzに減少した。

【0219】

8つのバンドにわたる分析振幅値を平均化することによって、各チャネルについて単一の高ガンマ分析振幅測度を計算した。

【0220】

30秒のスライドウィンドウによるWelchの方法を使用して、各チャネルについて高ガンマ分析振幅値をzスコア化した [7]。

【0221】

これらの高ガンマ分析振幅zスコア時系列(200Hzでサンプリングされた)を、全ての分析において、及びオンライン復号中に使用した。

【0222】

ハードウェアインフラストラクチャの可搬性及びコスト

本研究において、使用されたハードウェアは相当に大きかったが、依然として可搬性であり、ほとんどのハードウェアコンポーネントは、長さ及び幅が各々約76cmのモバイルラック上にあった。参加者の寝室又は参加者の住居の近くの小規模事務室で、全てのデータ収集及びオンライン復号タスクを実施した。臨床試験全体を通してハードウェアの全ての使用を監督したが、記録を開始するために必要なハードウェア及びソフトウェアセットアップ手順は簡単であった。介護者は、数時間の訓練後、適切な規制当局の承認を得て、当方の直接の監督なしに参加者が使用できるように本発明のシステムを準備することが可能である。システムを使用のためにセットアップするには、介護者は次のステップを実施する。

1. 経皮コネクタキャップを取り外して清掃する。これは、システムが使用されていない間、経皮コネクタの外部電気接点を保護する

10

20

30

40

50

2. 経皮コネクタ、デジタルリンク、及び経皮コネクタの周りの頭皮領域を清掃する
 3. デジタルリンクを経皮コネクタに接続する
 4. コンピュータの電源を入れ、ソフトウェアを起動する
 5. 画面が使用のために参加者の正面に適切に位置付けられていることを確認する
- その後、システムからディスエンゲージするために、介護者は次のステップを実施する。

1. ソフトウェアを閉じ、コンピュータの電源を切る
2. 経皮コネクタからデジタルリンクを外す
3. 経皮コネクタ、デジタルリンク、及び経皮コネクタの周りの頭皮領域を清掃する
4. 経皮コネクタキャップを経皮コネクタに戻す

【0223】

ハードウェアインフラストラクチャ全体は、主に新しいNeuroportシステムの比較的高いコスト（本研究で使用された他のハードウェアデバイスのコストと比較して）のために、相当に高価であった。しかしながら、最近の研究において、比較的安価で可搬性の脳コンピュータインターフェースシステムが、（本研究で使用されているシステムなどのBlackrock Microsystemデバイスを含む典型的なシステムと比較して）システムパフォーマンスの大幅な低下なしに展開できることが実証されている[8]。本研究の実証は、復号パフォーマンスを犠牲にすることなく、本発明のハードウェアインフラストラクチャの将来の反復をより安価で、より可搬性に行うことができることを示唆している。

【0224】

計算モデリングインフラストラクチャ

リアルタイム処理コンピュータから収集したデータを、本発明者らの研究室の計算及び格納サーバインフラストラクチャにアップロードした。ここでは、複数のNVIDIA V100 GPUを使用して計算時間を短縮し、復号モデルを適合させて最適化した。次いで、最終化されたモデルを、オンライン復号のためにリアルタイム処理コンピュータにダウンロードした。

【0225】

方法S3. タスク設計

全てのデータは、一連の「ブロック」として収集された。各ブロックは約5分又は6分間続き、複数の試験からなる。タスクには、孤立単語タスク及び文タスクの2種類があった。

【0226】

孤立単語タスク

孤立単語タスクにおいて、参加者は、50単語セットからの個々の単語の生成を試行し、同時に、当方はオフライン処理のために参加者の皮質活動を記録した。この単語セットは、次の基準に基づいて選ばれた。

1. それらの単語を使用して様々な文を容易に作成できること。
2. それらの単語を使用して、基本的な介護ニーズを容易に伝えることができること。
3. それらの単語を含めることに、参加者が関心を有すること。本発明者らは、参加者が市販のコミュニケーション支援技術を通じて本発明者らに提供したフィードバックを使用して、50単語セットのいくつかのバージョンを繰り返した。
4. 意味のある様々な文を作成するのに十分に大きい、十分な神経ベースの分類パフォーマンスを可能にするのに十分に小さい単語数を含めることが所望されること。この後者の基準は、デバイスの埋め込み後（本研究で分析されたデータのいずれかを収集する前）の参加者による探索的な予備的評価によって通知された。この50単語セットに含まれる単語のリストは、このセクションの最後に記載されている。

【0227】

タスクブロックの持続時間を短く保つために、この単語セットを3つの互いに素なサブセットに任意に分割する。2つのサブセットは各々20個の単語を含み、3つ目のサブセットには残りの10個の単語が含まれる。このタスクの各ブロック中、参加者は、これら

10

20

30

40

50

のサブセットのうちの1つに含まれる各単語の生成を2回試行し、結果として、(単語サブセットのサイズに応じて)ブロック当たり合計40個又は20個のいずれかの単語の生成が試行された。第3のより小さいサブセットの3つのブロックにおいて、参加者はそのサブセットの10個の単語の生成を各々4回(通常の2つではなく)試行した。

【0228】

このタスクのブロックの各試験は、黒色背景のブランク画面から開始した。1秒後(又は、非常に少数のブロックでは1.5秒)、現在の単語サブセットの単語の1つが、両側を4つのピリオド文字で囲まれた白色テキストで画面に表示された(例えば、現在の単語が「Hello」であった場合、テキスト「...Hello...」が表示される)。次の2秒間、両側で外側のピリオド(表示されるテキスト文字列の最初及び最後の文字)が500ミリ秒ごとに消えていき、視覚的にカウントダウンを表す。単語の両側の最後のピリオドが消えると、テキストが緑色に変わり、4秒間画面に表示されたままになる。白色から緑色へのこの色の遷移は、各試験のゴーキューを表し、参加者はテキストが緑色になったらすぐに単語の生成を試行するように指示された。その後、タスクは引き続き次の試験に進んだ。単語提示順序は、各タスクブロック内でランダム化された。参加者は、術前インタビュー中に当方が提示した一連の可能なパラダイムオプションから、一貫したカウントダウンタイミグを使用して、各試験において自身の生成試行をゴーキューとより良好に整合させることができたことを主張する、このカウントダウンスタイルのタスクパラダイムを選んだ。

【0229】

文タスク

文タスクにおいて、参加者は、50文セットから文の生成を試行し、同時に、当方は、参加者の神経活動を処理し、テキストに復号した。これらの文は、50単語セットの単語のみから構成されていた。これらの50個の文は、可能な文のコーパスから半ランダムに選択された(方法S5を参照されたい)。この50文セットに含まれる文のリストは、このセクションの最後に記載されている。タスクブロックの持続時間を短く保つために、この文セットを、各々が10個の文を含む5つの互いに素なサブセットに任意に分割する。このタスクの各ブロック中、参加者は、これらのサブセットのうちの1つに含まれる各文の生成を1回試行し、結果として、ブロック当たり合計10個の文の生成が試行された。

【0230】

このタスクのブロックの各試験は、両方とも黒色背景を有する上半分及び下半分に水平方向に分割されたブランク画面から開始した。2秒後、現在の文サブセット内の文のうちの1つが、白色テキストで画面の上半分に示された。参加者は、テキストが画面に表示されるとすぐに、自分が快適にできる最速の速度で文内の単語の生成を試行するように指示された。標的文が参加者に表示されていた間、参加者の皮質活動は、発話検出モデルによってリアルタイムで処理された。取得された神経信号から試行単語生成が検出されるたびに、サイクル楕円のセット(1、2、及び3つのピリオド文字の間で毎秒サイクルするテキスト文字列)が、発話イベントが検出されたことを示すフィードバックとして画面の下半分に追加された。次いで、単語分類、言語、及びビタビ復号モデルを使用して、対応する神経活動及び現在の試験内の任意の前の検出されたイベントからの復号された情報を所与として現在の検出されている発話イベントと関連付けられる最も可能性の高い単語を復号した。新しい単語が復号されるたびに、その単語は、スクリーンの下半分の関連するサイクル楕円テキスト文字列に置き換わり、参加者に更なるフィードバックを提供した。観察された神経活動を所与とした試験において最も可能性の高い単語系列を維持したビタビ復号モデルは、しばしば、新しい発話イベントを所与として以前の発話イベントの予測を更新し、新しい情報が利用可能になったときに、フィードバックテキスト文字列内の以前に復号された単語を変化させた。直近の発話イベントの検出された開始から所定の時間が経過した後、文標的テキストは白色から青色に変わり、試験の復号部分が終了し、復号された文がその試験について最終化されたことを示した。この所定の時間は、ブロックタイプに応じて9秒又は11秒であった(次の段落を参照)。3秒後、タスクは引き続き次の

試験に進んだ。

【0231】

文タスクの2つのタイプのブロック、すなわち、最適化ブロック及びテストブロックを収集した。これら2つのタイプのブロックの差は次のとおりである。

1. 最適化ブロックは、ハイパーパラメータ最適化を実行するために使用され、テストブロックは、復号システムのパフォーマンスを評価するために使用された。

2. 最適化ブロックを収集するときには中間（最適化されていない）モデルを使用し、テストブロックを収集するときには、最終化された（最適化された）モデルを使用した。

3. 検出された発話試行及び復号された単語系列は、このタスク中にフィードバックとして参加者に常に提供されたが、最適化ブロックの収集中に、参加者は、発話イベントが見逃された場合は単語を繰り返さないように、又は、フィードバックを使用してどの単語の生成を試行したかを変更するように指示された。本発明では、ハイパーパラメータ最適化手順で使用するデータの完全性を保護するために、これらの指示を含めた（参加者が不完全な発話検出のために自身データの挙動を変更した場合、プロンプトされた単語系列と参加者が実際に試行した単語系列との間の不一致が、最適化手順を妨げる可能性がある）。しかしながら、テストブロック中に、標的文の作成を試行するときに、参加者にフィードバックを考慮に入れるよう奨励した。例えば、試行単語生成が検出されなかった場合、参加者は、次の単語に進む前に、生成試行を繰り返すことができる。

4. 最適化ブロック中に、各試験における復号された単語系列が最終化されたときに制御された所定の時間（前の段落を参照）を9秒に設定した。テストブロック中に、このタスクパラメータを11秒に設定して、参加者に復号パイプラインから提供されたフィードバックを組み込むための余分な時間を与えた。

【0232】

また、参加者が50単語から質問へのカスタム応答を生成できる、よりオープンエンドな環境で復号アプローチを使用することができることを実証するために、文タスクの会話的変形例も収集した。タスクのこの変形例において、反復を試行する標的文で促される代わりに、参加者は、会話パートナーを模倣する質問又は陳述で促され、プロンプトへの応答の生成を試行するように指示された。会話プロンプト及び参加者へのタスク指示のこの変更以外では、タスクのこの変形例は通常のバージョンと同一であった。本発明では、文タスクのこの変形例から収集されたデータを用いたいかなる分析も実行しなかった。それは実証目的でのみ使用された。タスクのこの変形例は、本文の図1に示される。

【0233】

単語及び文リスト

本研究で使用された50単語セットは次のとおりである。

1. A m
2. A r e
3. B a d
4. B r i n g
5. C l e a n
6. C l o s e r
7. C o m f o r t a b l e
8. C o m i n g
9. C o m p u t e r
10. D o
11. F a i t h
12. F a m i l y
13. F e e l
14. G l a s s e s
15. G o i n g
16. G o o d

17. Goodbye
 18. Have
 19. Hello
 20. Help
 21. Here
 22. Hope
 23. How
 24. Hungry
 25. I
 26. Is
 27. It
 28. Like
 29. Music
 30. My
 31. Need
 32. No
 33. Not
 34. Nurse
 35. Okay
 36. Outside
 37. Please
 38. Right
 39. Success
 40. Tell
 41. That
 42. They
 43. Thirsty
 44. Tired
 45. Up
 46. Very
 47. What
 48. Where
 49. Yes
 50. You

10

20

30

【0234】

本研究で使用された50文セットは次のとおりである。

1. Are you going outside?
 2. Are you tired?
 3. Bring my glasses here.
 4. Bring my glasses please.
 5. Do not feel bad.
 6. Do you feel comfortable?
 7. Faith is good.
 8. Hello how are you?
 9. Here is my computer.
 10. How do you feel?
 11. How do you like my music?
 12. I am going outside.
 13. I am not going.
 14. I am not hungry.

40

50

15. I am not okay.
16. I am okay.
17. I am outside.
18. I am thirsty.
19. I do not feel comfortable.
20. I feel very comfortable.
21. I feel very hungry.
22. I hope it is clean.
23. I like my nurse.
24. I need my glasses.
25. I need you.
26. It is comfortable.
27. It is good.
28. It is okay.
29. It is right here.
30. My computer is clean.
31. My family is here.
32. My family is outside.
33. My family is very comfortable.
34. My glasses are clean.
35. My glasses are comfortable.
36. My nurse is outside.
37. My nurse is right outside.
38. 番号
39. Please bring my glasses here.
40. Please clean it.
41. Please tell my family.
42. That is very clean.
43. They are coming here.
44. They are coming outside.
45. They are going outside.
46. They have faith.
47. What do you do?
48. Where is it?
49. Yes.
50. You are not right.

10

20

30

40

50

【0235】

方法S4. 文コーパスの収集

文タスクのドメイン固有の言語モデルを訓練する（及びこのタスクの標的文のセットを取得する）ために、Amazon Mechanical Turkタスクを使用して、50単語セットの単語のみを含む自然な英語文の偏りのないコーパスをクラウドソーシングした。ウェブベースのインターフェースは、50語を表示するように設計され、Mechanical Turk作業員（「ターカー」と称される）は、次の基準を満たす文を作成するように指示された。

各文は、50単語セットからの単語のみから構成されなければならない。

各個々のターカーの文応答には、重複が存在してはならない。

各文は文法的に有効でなければならない。

各文の長さは8単語以下でなければならない。

【0236】

追加的に、ターカーは異なる文に異なる単語を使用することを奨励された（ただし、単

語は常に50単語セットに制限されている)。このタスクにおいて、収集された文の方言の影響を制限するために、米国のターカーのみが許可された。偽の提出及びスパム業者を排除した後、コーパスには187人のターカーからの3415個の文(1207個の一意の文)が含まれていた。

【0237】

方法S5. 文標的セットの作成

Amazon Mechanical Turk コーパスから文タスクの標的として使用される50の文のセットを抽出するために(このコーパスに関する詳細についてはMethod S4を参照)、最初にこの選択プロセスを、コーパスに複数回登場した文のみを考慮するように制限した。本発明者らは、標的セットの特異的な文の選択を妨げるために、この包含基準を課した。その後、残りの文からランダムにサンプリングし、文法的な間違い又は望ましくない内容(「Family is bad(家族は悪いです)」など)が含まれている場合は、いくつかのサンプルを破棄した。50文サンプルのセットが作成された後、50単語セット内の単語の少なくとも90%がこの文セット内に少なくとも1回出現することを確実にするためにチェックが実施された。このチェックが失敗した場合、チェックが合格し、文タスクに設定された標的文がもたらされるまで文のサンプリングプロセスを再度実行した。

【0238】

本研究で使用される50文セットを最終的に得た文サンプリング手順中、次の22文が破棄された。

1. Good family is success (良い家族は成功です)
2. Tell success (成功を伝えてください)
3. Bring computer (コンピュータを持ってくる)
4. Tell that family (その家族に伝えてください)
5. I going outside (外出している私)
6. You are hungry (あなたはお腹が空いています)
7. I feel very bad (私はとても気分が悪いです)
8. I need glasses (私には眼鏡が必要です)
9. I need computer (私にはコンピュータが必要です)
10. You need my help (あなたには私の助けが必要です)
11. You are coming closer (あなたは近づいています)
12. Tell you right (あなたに正しく伝えてください)
13. I am closer (私はもっと近くにいます)
14. It is bad outside (外は悪いです)
15. Success is not coming (成功は近づいていません)
16. I like nurse (私は看護師が好きです)
17. Family is bad (家族は悪いです)
18. I tell you (言っておきます)
19. That nurse is thirsty (その看護師は喉が渇いています)
20. Need help (助けが必要です)
21. They are very thirsty (彼らは非常に喉が渇いています)
22. Where is computer (コンピュータはどこにありますか)

【0239】

標的文セットには、可能な50語のうち45語が含まれていた。次の5つの単語は、標的文セットに現れなかった。

1. Closer
2. Goodbye
3. Help
4. Success
5. Up

【0240】

しかしながら、単語分類器は、50単語セット内の各単語を生成するための孤立試行、及び推論中に50単語全てにわたって計算された確率について訓練されたため、これらの5つの単語は、参加者の神経活動から復号された文に依然として現れる可能性がある。

【0241】

方法S6．データ編成

孤立単語データ：サブセット作成

合計で、48日間の記録にわたる291のタスクブロックで22時間30分の孤立単語タスクを収集し、単語当たり196回の試験（試行生成）を行った（合計9800回の試験）。これらのブロックを、単一の最適化サブセット及び10個の交差検証サブセットの、11個の互いに素なサブセットに分割した。最適化サブセットは、単語当たり合計16回の試験を含んでおり、各交差検証サブセットは単語当たり18回の試験を含んでいた。

【0242】

時間にわたって同様に分散されたサブセットを作成するために、最初にブロックを時系列で並べ替えた。次に、この順序付きリスト内の等間隔のインデックスで発生したブロック（最古のブロックから最新のブロックまで）を最適化サブセットに割り当てた。次いで、交差検証サブセットラベルを循環しながらブロックを反復することによって、残りのブロックを交差検証サブセットに割り当てた。このアプローチからわずかに逸脱したのは、各サブセットに単語当たりの所望の試験回数が含まれていることを確認するためだけであった。これは、任意の単一のサブセットが特定の時間期間からのデータの過剰表現を有することを防止したが、本発明者らの不規則な記録スケジュールは、サブセットが時間的に等間隔であったブロックを含むことを防止した（図8を参照）。

【0243】

ハイパーパラメータ最適化中に最適化サブセット内のデータに関してモデルを評価した（方法S7を参照されたい）。別途明記しない限り、このプロセス中に見つかったハイパーパラメータ値を全ての孤立単語分析に使用した。

【0244】

このプロセス中に見つかったハイパーパラメータ値を使用して、10個の交差検証サブセットで10フォールド交差検証を実施し、9つのサブセットに対してモデルを適合させ、各フォールドにおいて保持サブセットを評価した。別途明記しない限り、最適化サブセット内の試験は、孤立単語評価中に直接使用されなかった。

【0245】

孤立単語データ：学習曲線スキーム

訓練データの量がパフォーマンスにどのように影響するかを評価するために、10個の交差検証サブセットを使用して学習曲線スキームを生成した。このスキームにおいて、発話検出器及び単語分類器を、9つの異なる量の訓練データとの交差検証を使用して評価した。具体的には、 $N \in [1, 9]$ の各整数値について、各フォールドにおいてN個のランダムに選択されたサブセットのみに対して訓練しながら、孤立単語データによる10フォールド交差検証評価を実施した。このアプローチを通じて、訓練データの量が異なるにもかかわらず、利用可能な全ての試験をNの各値について評価し、いずれの個々の評価においても訓練データと試験データとの間に重複はなかった。この学習曲線スキームの分析の最終セット（ $N=9$ ）は、利用可能な全てのデータを用いた完全な10フォールド交差検証分析と同等であり、学習曲線結果を除いて、この分析セットを使用して、報告された全ての孤立単語結果（本文の図3に示す電極寄与及び混同行列を含む）を計算した。各サブセット内の単語当たり18回の試行生成で、この学習曲線スキームを使用した9セットの分析は、訓練中の単語当たり18、36、54、72、90、108、126、144、及び162回の試行をその順序で含んでいた。キュレーションされている検出されたイベントを使用して単語分類器が適合されたため、全ての試験が各分析セットで評価されたわけではない（詳細については、方法S13及び方法S8を参照されたい）。

【0246】

孤立単語データ：安定性サブセット

単語検出及び分類を駆動する信号が研究期間全体を通してどれほど安定していたかを評価するために、孤立単語データを使用して、異なる日付範囲中に収集されたデータを含む4つの日付範囲サブセットを定義した。「早期」、「中間」、「後期」、及び「最後期」と名付けられたこれらの日付範囲サブセットは、それぞれ、埋め込み後9～18週、18～30週、33～41週、及び88～90週に収集されたデータを含んでいた。正確な18週間のマークの日に収集されたデータは、「早期」サブセットの一部であり、「中間」サブセットではないと考えられた。これらのサブセットの各々は、対応する日付範囲内の利用可能なデータからランダムに（置換なしで）抽出された各単語について20回の試験を含んでいた。試験は、孤立単語交差検証サブセットからのみサンプリングされた（最適化サブセットからではない）。本文の図4において、これらのサブセットの日付範囲は、本研究のデータ収集の開始に対して表現されている（デバイスの埋め込み日に対して表現されているのではなく）。これらのサブセットの各々の中で、本発明者らは更に、各単語の2回の試験を含む10の互いに素なサブセット（これらのサブセットを、4つの日付範囲サブセットから明確に区別するために、このセクションでは「ピース」と呼ばれる）にデータを分割する。これらの4つの日付範囲サブセットを使用して、サブセット内スキーム、サブセット間スキーム、及び累積サブセットスキームの3つの評価スキームを定義した。

10

【0247】

サブセット内スキームは、各日付範囲サブセット内の10個のピースを使用して10フォールド交差検証を実施することを伴った。具体的には、日付範囲サブセットの各ピースは、その日付範囲サブセットの残りのピースからの全てのデータを適合させるモデルを使用して評価された。サブセット内スキームを使用して、単語分類器が訓練及びテスト中に使用する全ての発話イベントを検出した（各日付範囲サブセット及び各評価スキームについて）。各日付範囲サブセットの各個々の交差検証フォールド内で使用された訓練データは、常に単語当たり18回の試験から構成されていた。

20

【0248】

サブセット間スキームは、他の日付範囲サブセットからのデータを適合させるモデルを使用して、日付範囲サブセット内のデータを評価することを伴った。このスキームにおいて、日付範囲サブセットの各ピースが、他の日付範囲サブセットの各々からランダムにサンプリングされた（置換されていない）単語当たり6回の試験を適合させるモデルを使用して評価されたことを除いて、サブセット内スキームが再現された。各日付範囲サブセットの各個々の交差検証フォールド内で使用された訓練データは、常に単語当たり18回の試験から構成されていた。

30

【0249】

累積サブセットスキームは、様々な量のデータを適合させるモデルを使用して、「最後期」サブセットからのデータを評価することを伴った。このスキームにおいて、（日付範囲サブセットごとに定義された10個のピースを使用して）4つの交差検証評価を実施した。第1の評価において、「最後期」サブセットからのデータを、10フォールド交差検証を使用する単語分類器によって分析した（これは、「最後期」サブセット内評価と同一であった）。第2の評価において、「後期」サブセットからの全てのデータが、各交差検証フォールドの訓練データセットに追加されたことを除いて、第1の評価からの交差検証分析を繰り返した。第3の評価は、「中間」及び「後期」サブセットからの全てのデータも訓練中に含まれたことを除いて同様であり、第4の評価においては、「早期」、「中間」、及び「後期」サブセットからの全てのデータが訓練中に含まれた。

40

【0250】

これらのスキームが信号安定性を分析するためにどのように使用されたかの説明については、方法S14を参照されたい。

【0251】

文データ

50

合計で、7日間の記録にわたる25のタスクブロックで2時間4分の文タスクを収集し、文当たり5回の試験（試行生成）を行った（合計250回の試験）。これらのブロックを、2つの互いに素なサブセット、すなわち、文最適化サブセット及び文テストサブセットに分割した。各文の2回の試験を含む文最適化サブセットを使用して、オンラインテストの前に本発明の文復号パイプラインを最適化した。これらのブロックを収集するときは、最適化されていないモデルを使用した。その後、これらのブロックからのデータを使用して、オンラインテストのための本発明のモデルを最適化した（方法S7に記載されているハイパーパラメータ最適化手順を参照）。これらのブロックは、最適化のためにのみ使用され、更なる文復号分析には含まれなかった。

【0252】

各文の3回の試験を含むテストサブセットに含まれるブロックの結果を使用して、復号パフォーマンスを評価した。これらのブロックは、最適化されたモデルを使用して収集された。

【0253】

（いずれかのサブセットから）文タスク中に収集された神経データに対して、いかなるモデルも直接適合させなかった。

【0254】

方法S7．ハイパーパラメータ最適化

パフォーマンス評価中に使用されるモデルハイパーパラメータの最適値を求めるために、ハイパーパラメータ最適化手順を使用して、モデルパフォーマンスを測定するように設計した目的関数を用いて、カスタム探索空間からサンプリングされたハイパーパラメータ値の多くの可能な組み合わせを評価した。各ハイパーパラメータ最適化手順中に、所望の数の組み合わせを試験し、全ての組み合わせにわたって最も低い（最良の）目的関数値と関連付けられる組み合わせを、そのモデル及び評価タイプの最適なハイパーパラメータ値組み合わせとして選択した。関連する目的関数値を測定するために使用されたデータは、最適なハイパーパラメータ値が評価するために使用されるデータとは別個のものであった（試験セットの評価中に使用されるハイパーパラメータ値は、決して、その試験セット内のデータに対して最適化することによっては選択されなかった）。合計9つのハイパーパラメータを最適化するために、3つのタイプのハイパーパラメータ最適化手順を使用した（ハイパーパラメータ及びそれらの最適値については表S1を参照）。

【0255】

孤立単語データによる発話検出最適化

孤立単語データによって発話検出器を最適化するために、最適化手順中に確率的にハイパーパラメータ値組み合わせをサンプリングする `hyperopt Python` パッケージ [9] を使用した。この手順を使用して、平滑化サイズ、確率閾値、及び時間閾値持続時間ハイパーパラメータを最適化した（方法S8に記載）。これらの閾値化ハイパーパラメータは、発話確率が予測された後にのみ適用されたため、これらのハイパーパラメータは、発話検出器を駆動する人工ニューラルネットワークモデルの訓練又は評価に影響を与えなかった。最適化手順の各反復において、現在のハイパーパラメータ値組み合わせを使用して、既存の発話確率から検出された発話イベントを生成した。式S5に与えられる目的関数を使用して、各ハイパーパラメータ値組み合わせによるモデルパフォーマンスを測定した。各検出ハイパーパラメータ最適化手順において、停止する前に1000のハイパーパラメータ値組み合わせを評価した。

【0256】

方法S6に記載されるように、他の9つの交差検証サブセットからのデータに関して訓練された発話検出モデルを使用して、10の交差検証データサブセットの各々における孤立単語ブロックの発話確率を計算した。最適化サブセット内のブロックの発話確率を計算するために、10の交差検証サブセット全てからのデータに関して訓練された発話検出モデルを使用した。その後、最適化サブセットのブロックによるハイパーパラメータ最適化を実施し、それによって、10個の交差検証サブセット（学習曲線及び安定性分析を含む

10

20

30

40

50

）内のデータの評価中に使用された最適なハイパーパラメータ値組み合わせを得た。

【0257】

最適化サブセット内のブロック（単語分類器のハイパーパラメータ最適化中に使用された）の検出されたイベントを生成するために、10個のクロス検証サブセットからのデータのサブセットを用いて別個のハイパーパラメータ最適化を実施した。50個の単語の各々の16回の試験を含むこのサブセットは、10個の交差検証サブセットからブロックをランダムに選択することによって作成された。次いで、これらのブロックについてすでに計算されている予測発話確率を使用して、この新しいサブセットを用いてハイパーパラメータ最適化を実施した（前の段落で説明されているように）。その後、得られた最適なハイパーパラメータ値組み合わせを使用して、最適化サブセット内のブロックの発話イベントを検出した。

10

【0258】

孤立単語データによる単語分類の最適化

孤立単語データによって単語分類器を最適化するために、事前定義された探索空間からのランダムにサンプリングされたハイパーパラメータ値組み合わせによって、並列化されたハイパーパラメータ最適化を実施するRay Pythonパッケージ[10]を使用した。このハイパーパラメータ最適化アプローチは、「非同期逐次半減アルゴリズム」（ASHA）[11]に基づいたスケジューラを使用し、これは、積極的な早期停止を実施して、パフォーマンスの低いハイパーパラメータ値組み合わせを完全に評価する前に破棄する。単一のハイパーパラメータ値組み合わせの評価に関連する計算複雑度が高く、多数のハイパーパラメータ組み合わせが評価される場合、このアプローチはベイズハイパーパラメータ最適化アプローチをパフォーマンスにおいて上回ることが示されている[10]。各単語分類器を含む深層人工ニューラルネットワークモデルのアンサンブルを訓練するためには長い計算時間が必要になるため、単語分類ハイパーパラメータを最適化するために、このアプローチを使用した。本発明者らの拡張データセットを使用して、NVIDIA V100 GPU上で単一のネットワークを訓練するには、エポック当たり約28秒が必要であった。各ネットワークは、平均して約25エポックの訓練を必要とした（ただし、各エポックの持続時間は早期停止のために異なる場合がある）。この近似は、単一のネットワークが訓練に700秒を必要としたことを示している。ハイパーパラメータ最適化中に4つのネットワークからなるアンサンブルを使用したため、単一のハイパーパラメータ値組み合わせについて単語分類器を訓練するために約46分40秒の合計GPU時間が必要であった（評価及びリアルタイム予測中に使用される単語分類器については、これは各々10個のネットワークからなるアンサンブルを含み、分類器ごとのおおよその訓練時間は1時間56分40秒であった）。これらの訓練時間を所与として多数のハイパーパラメータ値組み合わせを評価するために、計算効率の高いハイパーパラメータ最適化アルゴリズム（ここで使用されているASHAアルゴリズムなど）を使用することが有益であった。

20

30

【0259】

単語分類器について2つの異なるハイパーパラメータ最適化を実施し、両方とも最適化中の目的関数として保持された試行のセットの交差エントロピー損失を使用した（方法S9の式S6を参照）。各最適化は、ハイパーパラメータ値の300個の異なる組み合わせを評価した。第1の最適化について、10個全ての交差検証サブセットからのデータに関して訓練しながら、最適化サブセットを保持セットとして使用した。得られたハイパーパラメータ値組み合わせを、孤立単語分析に使用した。第2の最適化について、3週間のオンライン文復号テストブロック以内に収集されたブロックから各単語の4回の試験をランダムに（置換なしに）選択することによって、保持セットを作成した。この最適化のための訓練セットは、この保持セット内の試験を除いて、（交差検証及び最適化サブセットからの）孤立単語データの全てを含んでいた。得られた最適なハイパーパラメータ値組み合わせを、文復号に関連する他のハイパーパラメータのオフライン最適化中及びオンライン文復号中に使用した。

40

50

【0260】

文データによる最適化

文最適化サブセットを使用して、閾値検出ハイパーパラメータ（方法S8を参照）、初期単語平滑化値（言語モデルについては方法S10を参照）、及び言語モデルスケール係数（ビタビデコードについては方法S11を参照）のハイパーパラメータ最適化を実施した。この手順では、最初に発話検出器（孤立単語最適化サブセットを含む、全ての孤立単語データに関して訓練された）を使用して、全ての文最適化ブロックの発話確率を予測した。次いで、これらの予測発話確率を使用して、文復号中に使用するために孤立単語データに関して訓練及び最適化された単語分類器、並びに言語モデル及びビタビデコードを使用して、全ての最適化文ブロックにわたってハイパーパラメータ最適化を実行した（方法S6を参照）。ハイパーパラメータ最適化中の目的関数として、試験全体の平均復号単語誤り率（単語分類器、言語モデル、及びビタビデコードを用いて各試験における検出されたイベントを評価することによって計算される）を使用した。Hyperopt Pythonパッケージ[9]を使用して、最適化中に100個のハイパーパラメータ値組み合わせを評価した。得られた最適なハイパーパラメータ値組み合わせを、オンライン復号を用いた文テストブロックの収集中に使用した。

10

【0261】

方法S8．発話検出モデル

オフライン訓練及び評価のためのデータ準備

孤立単語データを用いた発話検出器の教師あり訓練及び評価のために、発話イベントラベルを神経時点に割り当てた。これらのブロック中のタスクタイミング情報を使用して、各神経時点のラベルを決定した。準備、発話、及び休止の3つのタイプの発話イベントラベルを使用した。

20

【0262】

各孤立単語試験内で、標的発言がカウントダウンアニメーションで画面に表示され、2秒後に発言が緑色に変わって、ゴーキューを示した。この2秒ウィンドウ（ゴーキューに対して[-2, 0]秒）中に収集された全ての神経時間点を準備としてラベル付けした。ゴーキューに関連して、[0.5, 2]秒の間に収集された神経時点を発話として、[3, 4]の間に収集された時点休止としてラベル付けした。参加者の応答時間のばらつきが訓練に及ぼす影響を減らすために、ゴーキューに対して[0, 0.5]秒及び[2, 3]秒の時間期間（発話時間期間を取り巻く時間期間）を訓練データセットから除外した。評価中、これらの時間期間は、それぞれ、準備及び休止としてラベル付けされた。

30

【0263】

検出器が発話準備から試行発話生成を神経的に明確に区別することを可能にするために、準備ラベルを含めた。これは、試行発話生成に関連する神経活動が、発話準備に関連する活動よりも単語分類器によってより容易に区別可能であるという仮定によって動機付けられた。

【0264】

発話検出モデルアーキテクチャ及び訓練

PyTorch 1.6.0 Pythonパッケージを使用して、発話検出モデルを作成し、訓練した[12]。

40

【0265】

発話検出アーキテクチャは、3つの長・短期記憶（LSTM）層のスタックであり、潜在次元サイズ（150、100、及び50）が減少しており、各層に0.5のドロップアウトが適用された。再帰層は、入力データの新しい個々の時間サンプルによって更新することができる時間を通して内部状態を維持することができ、これによって、それらは時間的に動的なプロセスを用いたリアルタイム推論に適したものになる[13]。LSTMは、元の再帰層と比較して長期依存性をモデリングするのに適しているため、本発明においてはこれを特に使用する。LSTMには、最後の潜在次元を3つのクラス（休止、発話、及び準備）にわたる確率に投影するための全結合層が後続する。最近の研究[14]では

50

、同様のモデルが明示的な発話を検出するために使用されているが、本発明者らのアーキテクチャは独立して設計されている。このアーキテクチャの概略図が、図 9 に示される。

【0266】

y は一連の神経データウィンドウを表し、 l はそれらのウィンドウの対応する一連のラベルを表し、 y_n はデータ系列内のインデックス n におけるデータウィンドウ、 l_n はラベル系列内のインデックス n における対応するラベルである。発話検出モデルは、状態ラベル $L = \{\text{休止、準備、発話}\}$ のセットから l_n の 3 つの可能な値にわたる確率の分布 $Q(l_n | y_n)$ を出力する。予測分布 Q は、暗黙的にモデルパラメータに依存する。次の式で表されるデータ及びラベルの系列を使用して、真の分布に関するこの分布の交差エントロピー損失を最小限に抑えるように、発話検出モデルを訓練した。

10

【数 1】

$$H_{P,Q}(l | y) \approx -\frac{1}{N} \sum_n^N w_{fp,n} \log Q(l_n | y_n), \quad (S1)$$

定義は以下のとおりである。

P : 割り当てられた状態ラベル l によって決定される状態の真の分布。

N : サンプルの数。

$H_{P,Q}(l | y)$: l の真の分布に対する予測分布の交差エントロピー。

$\log$: 自然対数。

20

【0267】

ここで、 N 個のサンプルによって観察されたデータの下で、サンプル平均による真の分布の期待値を近似する。

【0268】

訓練中、発話ラベルが誤って予測された任意のフレームに 0.75 の偽陽性重み付けを適用した。この修正により、式 S1 からの交差エントロピー損失は、次のように再定義される。

【数 2】

$$\begin{aligned} H_{P,Q}(l | y) &= E_P[-\log Q(l | y)] \\ &\approx -\frac{1}{N} \sum_{n=1}^N \log Q(l_n | y_n), \end{aligned} \quad (S2)$$

30

式中、 $w_{fp,n}$ は、サンプル n の偽陽性重みであり、以下のように定義される。

【数 3】

$$w_{fp,n} := \begin{cases} 0.75 & \text{if}(l_n \neq \text{発話かつ}) \\ 1 & \text{そうでない場合} \end{cases} \quad \left(\operatorname{argmax}_{l \in L} [Q(l | y_n)] = \text{発話} \right) \quad (S3)$$

40

【0269】

この重み付けの結果、発話生成試行中に発生したと誤って分類されたサンプルに関連する損失は、他のサンプルと比べて 75% しか重み付けされなかった。この重み付けは、孤立単語データを評価するために使用された発話検出モデルの訓練中にのみ適用された。この重み付けを適用して、完全な発話イベントを検出することを選択するようにモデルを促した。これは、生成試行が検出されるのを妨げる可能性のある発話生成試行中の発話確率の変動を抑制した。これは、単語分類器の訓練及び評価中に関連付けられる検出された発話イベントを有する孤立単語試行の数を効果的に増大させた。

50

【0270】

典型的には、LSTMモデルは、処理の各時間ステップを通して逆伝播を展開する通時的逆伝播（BPTT）によって訓練される[15]。本発明者らの孤立単語タスク構造の周期性のために、BPTTにのみ依存することによって、モデルがこの構造を学習し、発話イベントの神経徴候を学習しようとするのではなく、全てのゴーキューにおいてイベントを予測する可能性がある。これを防ぐために、勾配が逆伝播することができる時間の長さを制限するアプローチである打ち切り型BPTTを使用した[16、17]。訓練データにおいて500ミリ秒スライドウィンドウを定義することによって、これを手動で実装した。これらのウィンドウは重複が非常に大きく、ウィンドウ間で1つの神経サンプル（5ms）のみがシフトしていた。訓練中にこれらのウィンドウを y_n 値として使用した。ここで、 l_n は、ウィンドウ内の最終時点に割り当てられたラベルに等しい。これにより、ウィンドウ内で訓練データを処理することによって、勾配は一度に500ミリ秒しか逆伝播しないようにされた。これは、タスクの周期性を学習するのに十分な長さではなかった（各試験のゴーキューの間の時間は、典型的には7秒であった）。オンライン及びオフライン推論中、データはウィンドウ内で処理されず、代わりに時点ごとに処理された。

10

【0271】

訓練中、Adamオプティマイザを使用して、学習率は0.001とし、残りのAdam最適化パラメータについてはデフォルト値として、式S2[18]を所与として交差エントロピーを最小化した。孤立単語データに関して発話検出器を評価するとき、方法S6に記載されている10フォールド交差検証スキームを使用した。文データに対してオフライン及びオンライン推論を実行するとき、10個の交差検証サブセットの孤立単語データの全てに関して訓練された発話検出器のバージョンを使用した。訓練中、訓練セットを、訓練セット及び検証セットに更に分割し、検証セットを使用して早期停止を実施した。連続して5エポックにわたってモデルパフォーマンスが改善されなくなり（検証セットの交差エントロピー損失が前のエポックで計算された最低値+損失許容値を下回らない場合）、少なくとも10エポックが完了するまでモデルを訓練し、その時点でモデル訓練を停止し、最低損失と関連付けられるモデルパラメータを保存した。損失許容値は0.001に設定されたが、モデル訓練には有意な影響を与えなかったようである。

20

【0272】

発話イベント検出

テスト中、ニューラルネットワークは、ブロックからの入力神経データを所与として各クラス（休止、準備、発話）の確率を予測した。試行発話イベントを検出するために、予測発話確率に閾値化を適用した。この閾値化アプローチは、本発明者らの以前の研究で使

用したアプローチと同一である[2]。まず、スライドウィンドウ平均を使用して確率を平滑化した。次に、平滑化確率に閾値を適用して、各フレームを二値化した（発話の場合は1、それ以外の場合は0の値）。その後、時間閾値を適用することによって、これらの二値化された値を「デバウンス」した。このデバウンスステップは、検出器が実際の変化とみなす前に、発話の存否（二値化された値によって示されるものとしての）の変化が最小限の持続時間にわたって維持されることを必要とした。具体的には、発話開始は、二値化された値が0から1に変化し、所定数の時点（又はそれ以上）にわたって1のままであった場合にのみ検出された。同様に、発話終了は、二値化された値が1から0に変化し、同じ所定数の時点（又はそれ以上）にわたって0のままであった場合にのみ検出された。予測確率から発話イベントを取得するこのプロセスは、平滑化ウィンドウのサイズ、確率閾値、及び時間閾値持続時間の3つの検出閾値化ハイパーパラメータによってパラメータ化された。ハイパーパラメータ最適化を使用して、これらのパラメータの値を決定した（以下のセクション及び方法S7を参照されたい）。

30

40

【0273】

検出スコア及びハイパーパラメータ最適化

孤立単語データによる検出閾値化ハイパーパラメータのハイパーパラメータ最適化中、以前の研究で使

50

た〔2〕。検出スコアは、各ブロックのフレームレベル及びイベントレベル正確度の加重平均である。

【0274】

フレームレベル正確度は、発話中に神経時点が発生したか否かを予測する発話検出器の能力を測定する。理想的には、発話検出器は、（例えば、各実際の発話イベントの小さいサブセットを検出するのとは対照的に）実際の試行発話イベントの持続時間に及んだイベントを検出する。フレームレベル正確度 α_{frame} は次のように定義した。

【数4】

$$\alpha_{\text{frame}} := \frac{w_p F_{\text{TP}} + (1 - w_p) F_{\text{TN}}}{w_p F_{\text{P}} + (1 - w_p) F_{\text{N}}} \quad 10$$

可変定義は以下のとおりである。

w_p ：陽性重み分数。陽性フレーム（発話試行中に発生した神経時点を正しく識別する）を、陰性フレーム（発話試行中に発生しなかった神経時点を正しく識別する）に対して正しく検出することの重要性を制御するために使用した。

F_P ：実際の陽性フレームの数（データ準備中に発話ラベルが割り当てられた時点の数）。

F_{TP} ：検出された真陽性フレームの数（試行発話イベント中に発生したと正しく識別された時点の数）。

F_N ：実際の陰性フレームの数（データ準備中に準備又は休止としてラベル付けされた時点の数）。

F_{TN} ：検出された真陰性フレームの数（試行発話イベント中に発生していないと正しく識別された時点の数）。

【0275】

本研究では、 $w_p = 0.75$ を使用した。これにより、発話検出器は偽陰性誤りを行うよりも偽陽性誤りを行うことを選好するように促された。

【0276】

イベントレベル正確度は、試行単語生成中に発話イベントを検出する検出器の能力を測定する。イベントレベル正確度 α_{event} を次のように定義した。

【数5】

$$\alpha_{\text{event}} := \max \left(0, \frac{E_{\text{TP}} - E_{\text{FP}} - E_{\text{FN}}}{E_P} \right), \quad (\text{S4})$$

可変定義は以下のとおりである。

E_{TP} ：真陽性検出イベントの数（実際の単語生成試行に対応した検出された発話イベントの数）。

E_{FP} ：偽陽性検出イベントの数（実際の単語生成試行に対応しなかった検出された発話イベントの数）。

E_{FN} ：偽陰性イベントの数（いかなる検出されたイベントにも関連付けられなかった実際の単語生成試行の数）。

E_P ：実際の単語生成試行の数（試験の回数）。

【0277】

検出されたイベントをキュレーションした後、イベントレベル正確度を算出した。これ

10

20

30

40

50

は、各試験と検出されたイベント（又は検出されたイベントがないこと。更なる詳細については、次のセクションを参照されたい）とのマッチングを伴った。イベントレベル正確度は、0～1に及び1の値は、偽陽性又は偽陰性検出イベントがなかったことを示す。

【0278】

これら2つの正確度測定値を使用して、検出スコアを次のように計算する。

$$S_{\text{detection}} = \omega_F \alpha_{\text{frame}} + (1 - \omega_F) \alpha_{\text{event}},$$

式中、 ω_F はフレームレベル正確度重みである。単語分類器は、発話イベントの検出された開始に対する神経活動の固定持続時間タイムウィンドウに依拠しているため、検出された終了を正確に予測することは、参加者が単語の生成を試行するたびにイベントを首尾よく検出することよりも重要性が低かった。これを踏まえて、フレームレベル正確度よりも

10

【0279】

孤立単語データを用いた3つの検出閾値化ハイパーパラメータの最適化中、主な目標は、検出スコアを最大化するハイパーパラメータ値を求めることであった。また、時間閾値持続時間ハイパーパラメータの小さい値を選択する補助目標も含めた。長い時間閾値持続時間は、より短い発言を逃す可能性を高め、持続時間が十分に大きい場合、リアルタイム発話検出に遅延を加えるため、この補助目標を含めた。これらの目標の両方をカプセル化した、このハイパーパラメータ最適化手順中に使用される目的関数は、次のように表現することができる。

20

【数6】

$$c_{hp}(\Theta) = (1 - s_{\text{detection}})^2 + \lambda_{\text{time}} \theta_{\text{time}},$$

(S5)

可変定義は以下のとおりである。

$c_{hp}(\Theta)$ ：ハイパーパラメータ値組み合わせ Θ を使用した目的関数の値。

λ_{time} ：時間閾値持続時間に適用されるペナルティ。

θ_{time} ： Θ に含まれる3つのパラメータのうちの1つである時間閾値の持続時間値。ここでは、 $\lambda_{\text{time}} = 0.00025$ を使用した。

30

【0280】

この目的関数は、孤立単語試験の発話イベントを検出するために使用された検出モデルの最適化中にのみ使用した。文データとともに使用するための検出モデルを準備するときは、異なる目的関数を使用した。ハイパーパラメータ最適化手順の更なる詳細については、方法S7及び表S1を参照されたい。

【0281】

孤立単語データの検出されたイベントキュレーション

孤立単語ブロックの神経データを処理し、発話イベントを検出した後、検出されたイベントをキュレーションして、各イベントを実際の単語生成試行と照合した（そして、対応する検出されたイベントを有しない単語生成試行、及び単語生成試行に対応しない検出されたイベントを識別した）。単語分類器の訓練及び評価中に、このキュレーション手順を使用して、イベントレベル正確度の算出（式S4）中の偽陽性及び偽陰性イベント検出の数を測定し、試験を神経データと照合した。このキュレーション手順は、文データでは使用しなかった。

40

【0282】

検出されたイベントをキュレーションするために、各試験で次の手順を実行した。（ゴーキューに対して）－1.5秒～3.5秒に及ぶタイムウィンドウ内に発生した、検出された開始を全て識別した。このタイムウィンドウ外の検出された開始を有するイベントは、偽陽性イベントと考えられ、EFPの値を計算するときに含まれた。

【0283】

このタイムウィンドウ内に正確に1つの検出された開始があった場合、関連する検出さ

50

れたイベントを試験に割り当てた。

【0284】

他の状態で、このタイムウィンドウ内に検出された開始がなかった場合、検出されたイベントを試験に割り当てなかった（これは偽陰性イベントと考えられ、 E_{FN} の値を計算するときに含まれた）。

【0285】

それ以外の場合、このタイムウィンドウ内に2つ以上の検出された開始があり、次の手順を実施してこれらの検出されたイベントを処理した。

【0286】

これらの検出された開始のうちの正確に1つがゴーキューの後に発生した場合、その検出された開始と関連付けられる検出されたイベントを試験に割り当てた。

【0287】

他の状態で、これらの検出された開始のうちのいずれもがゴーキューの後に発生しなかった場合、最新の検出された開始と関連付けられる検出されたイベントを試験に割り当てた（これは、検出された開始がゴーキューに最も近い検出されたイベントであった）。

【0288】

他の状態で、ゴーキューの後に2つ以上の検出された開始が発生した場合、これらの検出された開始と関連付けられる各検出されたイベントの長さを計算し、最長の検出されたイベントを試験に割り当てた。同点が発生した場合、検出されたイベントには、開始がゴーキューに最も近い検出されたイベントを試験に割り当てた。

【0289】

試験に割り当てられなかったこれらの検出されたイベントの各々は、偽陽性イベントと考えられ、 E_{FP} の値を計算するときに含まれた。

【0290】

偽陰性は、いくつかの試験を検出されたイベントと関連付けられないようにするため、実際に分析ステップに使用される試験の数は、報告された試験の数よりも少ない場合がある。例えば、各単語のN回の試験が分析ステップに使用されたと述べる場合、そのステップにおいて単語分類器によって分析された試験の実際の回数は、偽陰性検出があった回数に応じて、1つ以上の単語についてN未満であり得る。

【0291】

方法S9．単語分類モデル

オフライン訓練及び評価のためのデータ準備

孤立単語データを用いた単語分類器の訓練及び評価中に、各試験について、（利用可能な場合、方法S8に記載されている検出キュレーション手順によって決定される）検出された開始の時間を取得した。各試験での評価中、単語分類器は、検出された開始に対して-1〜3秒に及ぶ高ガンマ活性のタイムウィンドウを所与として、50単語の各々が、参加者が生成を試行していた標的単語である確率を予測した。

【0292】

訓練サンプルの数を増大し、学習された特徴マッピングの神経入力の小さい時間的変動に対するロバスト性を向上させるために、モデル適合中に、開始時間をジッタリングすることによって、訓練データセットを追加の試験コピーによって拡張した。これは、教師あり画像分類のためにニューラルネットワークを訓練するために使用されるデータ拡張技法の十分に確立された使用と同様である[19]。具体的には、各試験について、各 $\epsilon \in \{-1, -0.667, -0.333, 0, 0.333, 0.667, 1\}$ について検出された開始に対して $(-1 + \alpha) \sim (3 + \alpha)$ 秒に及ぶ神経タイムウィンドウを得た。これらのタイムウィンドウの各々は、訓練サンプルとして含まれ、試験からの関連する標的単語をラベルとして割り当てられた。

【0293】

オフライン及びオンライン訓練及び評価中に、各タイムウィンドウ内の高ガンマ活性を、活性を単語分類器に渡す前にダウンサンプリングした。これは、以前の研究で人工ニュー

10

20

30

40

50

ーラルネットワーク（ANN）による発話復号を改善することが示されている〔20〕。
Scipy Pythonパッケージ内のデシメーション関数を使用して、各電極の高ガンマ活性を6分の1に（200Hzから33.3Hzに）デシメーションした〔21〕。
この関数は、信号をデシメーションする前に、8次チェビシェフI型アンチエイリアシングフィルタを適用する。デシメーション後、全ての電極にわたるユークリッドノルムが1に等しくなるように、神経活動の各時間サンプルを正規化した。

〔0294〕

単語分類モデルアーキテクチャ及び訓練

TensorFlow 1.14 Pythonパッケージを使用して、単語分類モデルを作成し、訓練した〔22〕。

〔0295〕

単語分類ANNアーキテクチャ内で、神経データは、2サンプルストライド及び2サンプルカーネルサイズによる時間的畳み込みによって処理され、これは、データのより高次元の表現を作成しながら、時間的に神経活動を更にダウンサンプリングした。時間畳み込みは、時系列データからロバストな特徴を抽出するための一般的なアプローチである〔23〕。次いで、この表現は、時系列データの非線形分類に使用されることが多い、2つの双方向ゲート付き再帰ユニット（GRU）層からなるスタックによって処理された〔24〕。その後、softmax活性化を有する全結合（高密度）層が、最終的なGRU層から50語にわたる確率値に潜在次元を投影する。ドロップアウト層は、正規化のために各中間表現の間で使用される。このアーキテクチャの概略図が、図10に示される。

〔0296〕

y は一連の高ガンマタイムウィンドウを表し、 w はそれらのウィンドウの対応する標的単語ラベルを表し、 y_n はデータ系列内のインデックス n におけるタイムウィンドウ、 w_n はラベル系列内のインデックス n における対応するラベルである。単語分類器は、50単語セット W からの w_n の50個の可能な値にわたる確率の分布 $Q(w_n | y_n)$ を出力する。予測分布 Q は、暗黙的にモデルパラメータに依存する。次の式で表されるデータ及びラベルの系列を使用して、真の分布に関するこの分布の交差エントロピー損失を最小限に抑えるように、単語分類器を訓練した。

〔数7〕

$$\begin{aligned} H_{P,Q}(w | y) &= \mathbb{E}_P[-\log Q(w | y)] \\ &\approx -\frac{1}{N} \sum_{n=1}^N \log Q(w_n | y_n), \end{aligned} \quad (S6)$$

定義は以下のとおりである。

P ：割り当てられた単語ラベル w によって決定されるラベルの真の分布。

N ：サンプルの数。

$H_P, Q(w | y)$ ： w の真の分布に対する予測分布の交差エントロピー。

$\log$ ：自然対数。

〔0297〕

ここで、 N 個のサンプルによって観察されたデータの下で、サンプル平均による真の分布の期待値を近似する。

〔0298〕

訓練中、Adamオプティマイザを使用して、学習率は0.001とし、残りのAdam最適化パラメータについてはデフォルト値として、式S6〔18〕を所与として交差エントロピーを最小化した。各訓練セットを、訓練セット及び検証セットに更に分割し、検証セットを使用して早期停止を実施した。連続して5エポックにわたってモデルパフォーマンスが改善されなくなる（検証セットの交差エントロピー損失が前のエポックで計算された最低値を下回らない場合）までモデルを訓練し、その時点でモデル訓練を停止し、最

低損失と関連付けられるモデルパラメータを保存した。訓練は典型的には、20～30エポック続いた。各エポック後にモデルパラメータに勾配更新を適用するとき、全てのパラメータ更新値にわたる勾配のユークリッドノルムが1よりも大きい場合（これらの値を学習率でスケールする前）、次いで、勾配爆発を防ぐために、そのユークリッドノルムが1に等しくなるように勾配が正規化された〔25〕。

【0299】

訓練データの過剰適合を低減するために、各単語分類器は、各々が同一のアーキテクチャ及びハイパーパラメータ値を有するが、異なるパラメータ値（重み）を有する10個のANNモデルのアンサンブルを含んだ〔26〕。訓練中、各ANNは、ランダムモデルパラメータ値で初期化され、同じ訓練サンプルを使用して個々に適合されたが、各ANNは、確率的勾配更新中に異なる順序でサンプルを処理した。このプロセスにより、10個の異なるモデルパラメータセットが得られた。評価の間、アンサンブルされたANNの10個全てが、各入力神経タイムウィンドウを処理し、本発明者らは、各ANNの予測分布 $Q(w_n | y_n)$ を平均して、神経タイムウィンドウ y_n を所与として w_n の50個の可能な値の各々の全体的な予測単語確率を計算した。

【0300】

ハイパーパラメータ最適化手順を使用して、訓練中に直接学習されなかったモデルパラメータの値を選択した。本発明者らは、1つはオフライン孤立単語分析のためのものであり、もう1つはオンライン文復号のためのものである、2つの異なるハイパーパラメータ値組み合わせを計算した。ハイパーパラメータ探索を高速化するために、10個のセット全部ではなく、ハイパーパラメータを探索するときに4つのANNモデルのアンサンブルを使用した。詳細については、方法S7及び表S1を参照されたい。

【0301】

文タスクの修正

オンライン文復号について、全ての孤立単語データに関して単語分類器の修正バージョンを訓練した。このバージョンの単語分類器のハイパーパラメータ最適化中、保持セットには、研究期間の終わり近くに収集されたブロックからランダムにサンプリングされた各単語の4回の試験が含まれていた（詳細については、方法S7を参照）。ハイパーパラメータ最適化後、次いで、各単語の4回の試行からなるこの保持セットを検証セット（早期停止を実行するために使用）として使用し、残りの孤立単語データの全てを訓練セットとして使用することによって、選択されたハイパーパラメータを用いて単語分類器を訓練した。この訓練手順中、訓練中に使用される損失関数に単一の修正を追加した。各訓練サンプルを、50単語セットの単語のみからなるコーパス内の標的単語ラベルの発生頻度によって重み付けされた。より頻繁に発生した単語には、より大きい重みが割り当てられた。単語の出現頻度を計算するために使用されるコーパスは、Amazon Mechanical Turkからクラウドソーシングされ、言語モデルの訓練に使用されたものと同じコーパスである（方法S4を参照）。単語分類器に、低頻度単語（「glasses」など）の分類パフォーマンスを犠牲にして、高頻度単語（「I」など）の試行生成中に検出された神経タイムウィンドウを正しく分類することに焦点を当てるように促すために、この修正を含めた。

【0302】

この修正により、式S6からの損失関数を次のように修正することができる。

【数8】

$$H'_{p,Q}(w|y) \approx -\frac{1}{N} \sum_{n=1}^N \xi(w_n) \log Q(w_n | y_n), \quad (S7)$$

可変定義は以下のとおりである。

$H'_{p,Q}(w|y)$ ：修正された交差エントロピー損失関数。

$\xi(w_n)$ ：単語出現頻度重み付け関数

【 0 3 0 3 】

単語出現頻度重み付け関数は、次のように定義される。

【数 9】

$$\xi(w_n) := \frac{\kappa_{w_n}}{\sum_{\omega \in W} \kappa_{\omega}} \quad (S8)$$

式中、

【数 1 0】

$$\kappa_{w_n} \quad 10$$

は、標的単語ラベル w_n が参照コーパス内で出現した回数であり、

【数 1 1】

$$\sum_{\omega \in W} \kappa_{\omega}$$

は、参照コーパスの総単語数であり、 W は 50 単語セットである。

【 0 3 0 4 】

【数 1 2】

$$\bar{\xi}$$

20

を、次のように定義する。

【数 1 3】

$$\bar{\xi} := \frac{1}{|W|} \sum_{\omega_i \in W} \frac{\kappa_{\omega_i}}{\sum_{\omega \in W} \kappa_{\omega}} \quad (S9)$$

30

式中、 W は、50 単語セットの濃度（50 に等しい）を示す。したがって、

【数 1 4】

$$\bar{\xi}$$

は、平均単語出現頻度が 1 であるように、式 S 8 の各単語頻度をスケールリングするように作用し、これは、損失値が式 S 6 から生じる損失値と同程度であるように、目的関数をスケールリングする。

【 0 3 0 5 】

方法 S 1 0 . 言語モデリング

40

モデル適合及び単語系列確率

文復号中に使用するために言語モデルを適合させるために、最初に Amazon Mechanical Turk タスクを使用して訓練コーパスをクラウドソーシングした（詳細については、方法 S 4 を参照されたい）。このコーパスには、50 単語セットからの単語のみから構成された 3415 個の文が含まれていた。最も一般的な文に対する言語モデルの過剰適合を抑制するために、これらの応答から作成された訓練コーパスには、各固有の文の最大 15 個のインスタンスのみを含めた。

【 0 3 0 6 】

次に、訓練コーパス内の各文から $n \in \{1, 2, 3, 4, 5\}$ の全ての N グラムを抽出した。ここで、N グラムは、 n 語の長さを有する単語系列である [27]。例えば、この

50

アプローチで「I hope my family is coming（私の家族が来ることを願っています）」という文から抽出されるNグラム（タプルで表される）は、次のようになる。

1. (I)
2. (Hope)
3. (My)
4. (Family)
5. (Is)
6. (Coming)
7. (I, Hope)
8. (Hope, My)
9. (My, Family)
10. (Family, Is)
11. (Is, Coming)
12. (I, Hope, My)
13. (Hope, My, Family)
14. (My, Family, Is)
15. (Family, Is, Coming)
16. (I, Hope, My, Family)
17. (Hope, My, Family, Is)
18. (My, Family, Is, Coming)
19. (I, Hope, My, Family, Is)
20. (Hope, My, Family, Is, Coming)

10

20

30

40

50

【0307】

このようにして訓練コーパス内の全ての文から抽出されたNグラムを使用して、nltk Pythonパッケージを用いて、5次補間Kneser-Ney Nグラム言語モデルを適合させた[28、29]。このモデルには、nltk内で指定されたデフォルト値である0.1の割引係数が使用された。この言語モデルアーキテクチャの詳細は、様々なコーパスモデリングタスクに対するより単純なNグラムアーキテクチャのパフォーマンスを上回る能力の特徴付けとともに、既存の文献[27、28]に見出される。

【0308】

（抽出されたNグラムによって指定されるような）訓練コーパス内の特定の単語系列の出現頻度を使用して、言語モデルは、 $(n-1)$ 又はそれに先行するより少ない単語の系列である、その単語の文脈を所与として出現する任意の単語の条件付き確率をもたらすように訓練された。これらの確率は、 $p(w_i | c_{i,n})$ として表すことができ、ここで、 w_i は、いくつかの単語系列における位置*i*にある単語であり、 $c_{i,n}$ は、それがNグラムの一部であると仮定するその単語の文脈であり（このNグラムは、*n*個の単語を含む単語系列であり、 w_i がその系列の最後の単語である）、 $n \in \{1, 2, 3, 4, 5\}$ である。単語 w_i の文脈は、以下のタプルとして定義される。

【数15】

$$c_{i,n} := (w_{i-(n-1)}, \dots, w_{i-1}). \quad (S10)$$

【0309】

$n=1$ の場合、文脈は、空のタプルである $()$ である。 $n=2$ の場合、単語 w_i の文脈は (w_{i-1}) であり、 w_i の前の単語を含む単一要素タプルである。本研究で使用された言語モデルによれば、このパターンは $n=5$ まで続き、単語 w_i の文脈は $(w_{i-4}, w_{i-3}, w_{i-2}, w_{i-1})$ であり、系列内の、 w_i の前の4つの単語を（順番に）含むタプルである。各 w_i について、 $w_i \in W$ である必要であった。ここで、 W は50単語セットである。この要件には、文脈 $c_{i,n}$ に含まれる単語が含まれていた。

【0310】

文独立性

文タスク中、各文は、タスクブロック内の他の文とは独立して復号された。言語モデルによる推論中に使用した文脈 c_i 、 n は、先行する単語のみを含み得たが、 w_i と同じ文内でもあった（文脈が2つ以上の文にまたがることは決してなかった）。推論中に使用した文脈における値 i と n との関係は、次のように表現することができる。

【数16】

$$n = \text{分}(i+1, m), \quad (\text{S11})$$

10

式中、 m は、モデルの次数（このモデルの場合、 $m=5$ ）であり、 $i=0$ は、文内の最初の単語のインデックスを指定する。 n のこの定義を、式 S10 で指定された c_i 、 n の定義に置き換えると、以下ようになる。

【数17】

$$c_i := (w_{i-\min(i, m-1)}, \dots, w_{i-1}) \quad (\text{S12})$$

式中、 c_i は、文試験内の単語 w_i の文脈である。この置換は、言語モデルから得られる単語確率の形式を $p(w_i | c_i)$ に単純化する。

20

【0311】

最初の単語の確率

このタスクでは、文は常に独立して復号されていたため、空のタプルは、文の最初の単語である w_0 の推論を実行するときの文脈としてのみ使用された。推論中に言語モデルによって得られた $p(w_0 | c_0)$ の値を使用する代わりに、コーパスからの直接の単語カウント及び2つの異なるタイプの平滑化を使用した。まず、次の確率を計算した。

【数18】

$$\phi(w_0 | c_0) = \frac{k_{w_0} + \delta}{N + \delta |W|}, \quad (\text{S13})$$

30

式中、 k_{w_0} は、単語 w_0 が訓練コーパス内の文の最初の単語として出現した回数であり、 N は、訓練コーパス内の文の総数であり、 δ は、加算平滑化係数である。ここで、加算平滑化係数は、正規化前の全てのカウント k_{w_0} に加算される値であり、これは、確率分布を平滑化する（確率分布の分散を低減する）[27]。本研究では、 $N=3415$ 、 $\delta=3$ 、 $W=50$ である。

【0312】

次いで、これらの $\phi(w_0 | c_0)$ 値を平滑化して、最初の単語の確率にわたる確率分布がどれほど平坦であるかを更に制御した。これは、言語モデルによる最初の単語の確率の予測がどれほど「信頼できる」かを制御することとして解釈することができる（より平坦な確率分布は信頼性がより低いことを示す）。この平滑化の程度を制御するためにハイパーパラメータを使用し、ハイパーパラメータ最適化手順が、テスト中にどの程度の平滑化が最適であったかを決定することを可能にした（ハイパーパラメータ最適化手順の説明については、方法 S7 及び表 S1 を参照されたい）。次の式を使用して、この平滑化を実施した。

40

【数19】

$$p(w_0 | c_0) = \frac{\phi(w_0 | c_0)^\psi}{\sum_{w_j \in W} \phi(w_j | c_0)^\psi}, \quad (\text{S14})$$

50

式中、 Ψ は、最初の単語の平滑化ハイパーパラメータ値である。 $\Psi > 1$ のとき、最初の単語の確率の分散が増大し、平滑性がより低くなる。 $\Psi < 1$ のとき、最初の単語の確率の分散が低減し、平滑性がより高くなる。 $\Psi = 1$ のとき、 $p(\omega_0 | c_0) = \phi(\omega_0 | c_0)$ となる。式S14の分母は、平滑化された確率を再正規化して、それらの合計が1になるように使用されることに留意されたい。

【0313】

本研究で使用したビタビ復号モデルは、文復号アプローチ中に $p(\omega_i | c_i)$ 値を再スケーリングする別個のハイパーパラメータである言語モデルスケーリング係数(LMSF)を含んでいた(詳細については、方法S11を参照)。このハイパーパラメータが全ての言語モデル確率に及ぼした影響は、 ψ が最初の単語の確率に及ぼした影響に類似する。これにより、ハイパーパラメータ最適化手順は、言語モデルの確率を最適にスケーリングしたLMSF値、及びその後適用されたスケーリングと比較して最初の単語の確率を最適に平滑化した値を求めるように促されたはずである。

【0314】

リアルタイム実施態様

リアルタイム復号中の迅速な推論を確実にするために、言語モデル及び ω_i と c_i との全ての可能な組み合わせの平滑化ハイパーパラメータ値を用いて $p(\omega_i | c_i)$ 値を予め計算して、これらの値をhdf5ファイルに格納した[30]。このファイルは、リアルタイム復号中にルックアップテーブルとして機能した。値はファイル内の多次元配列に格納され、h5py Pythonパッケージを使用したリアルタイム復号中にテーブルへの効率的なルックアップクエリが実行された[31]。より大きい語彙サイズを必要とするこの復号アプローチの将来の反復では、kenlm言語モデル[32]などのリアルタイム推論に十分な計算効率もある、より洗練された言語モデルを使用することがより適切であり得る。

【0315】

方法S11. ビタビ復号

文復号手順の隠れマルコフモデル表現

文試験中、参加者が生成を試行した単語系列と、発話検出器によって提供される神経活動タイムウィンドウの系列との間の関係は、隠れマルコフモデル(HMM)として表現することができる。このHMMにおいて、各観察された状態 y_i は、任意の特定の試験の検出されたタイムウィンドウの系列内のインデックス i における神経活動のタイムウィンドウであり、各隠れた状態 q_i は、系列内の最初の単語からインデックス i におけるまでの、参加者が生成を試行した単語を含むNグラムである(図11)。ここで、 $q_i = \{\omega_i, c_i\}$ であり、 ω_i は、配列内のインデックス i における単語であり、 c_i は、その単語の文脈である(式S12で定義される。方法S10を参照されたい)。

【0316】

このHMMの放出確率は、 $p(y_i | q_i)$ であり、これは、Nグラム q_i を所与として神経タイムウィンドウ y_i を観測する尤度を指定する。ここで、

【数20】

$$w_i(y_i \perp w_j | w_i \forall j \neq i), p(y_i | q_i)$$

が $(y_i | \omega_i)$ に単純化されていることを所与として、 ω_i の試行生成と関連付けられる神経活動のタイムウィンドウは、他の全ての試行単語生成から条件的に独立していると仮定する。単語分類器は、確率 $p(\omega_i | y_i)$ を提供した。これはベイズの定理を適用し、平坦な事前確率分布を仮定することによって $p(y_i | \omega_i)$ の値として直接使用された。

【0317】

このHMMの遷移確率は、 $p(q_i | q_{i-1})$ であり、これは、インデックス $(i -$

1)におけるNグラムが q_{i-1} であることを所与として、 q_i がインデックス i におけるNグラム(参加者が生成を試行した最終単語として w_i を含む、最大 n 個の単語からなる系列)である確率を指定する。ここで、 q_{-1} は空集合として定義することができ、 q_0 が系列内の最初の単語であることを示す。 C_i 内の任意の要素は q_{i-1} に含まれ、 w_i は q_{i-1} に含まれていない q_i 内の唯一の単語であるため、 $p(q_i | q_{i-1})$ は、言語モデルによって提供される単語系列事前確率である $p(w_i | c_i)$ に単純化される。この単純化には、 q_i が q_{i-1} と互換性がない場合(例えば、 c_i 内の最後の単語が q_{i-1} 内の最後から2番目の単語に等しくない場合)、 $p(q_i | q_{i-1}) = 0$ という主張が暗示されている。

【0318】

ビタビ復号実施態様

参加者が文タスク中に生成を試行した単語を予測するために、この基礎となるHMM構造を有するビタビ復号アルゴリズムを実装した。ビタビ復号アルゴリズムは、動的プログラミングを使用して、隠れ状態事前遷移確率及び実測状態放出尤度を所与として隠れ状態の最も可能性の高い系列を計算する[33、34]。最も可能性の高い隠れ状態系列を決定するために、このアルゴリズムは、隠れ状態系列空間(q_i 値の様々な組み合わせ)を通る様々な「経路」の確率を反復的に計算する。ここで、これらのビタビ経路の各々は、隠れた状態(特定の単語系列)を通る特定の経路及び神経活動を所与としてその経路と関連付けられる確率によってパラメータ化された。新しい単語生成試行が検出されるたびに、このアルゴリズムは、既存のビタビ経路ごとに、検出された神経活動のタイムウィンドウ及び関連付けられる既存のビタビ経路内の先行する単語を所与として各有効な新しい単語に遷移する確率を計算することによって、新しいビタビ経路のセットを作成した。既存の経路からの新しいビタビ経路の作成は、次の再帰式を使用して表現することができる。

【数21】

$$V_i = \left\{ v_{i-1} + \log p(y_i | q_{i,v_{i-1}}) + L \log p(q_{i,v_{i-1}} | q_{i-1,v_{i-1}}) \mid v_{i-1} \in V_{i-1} \wedge w_i \in W \right\}, \quad (S15)$$

可変定義は以下のとおりである。

V_j : 文試験内のインデックス j における単語生成試行後に作成された全てのビタビ経路からなるセット。

v_j : V_j 内のビタビ経路。これらのビタビ経路のそれぞれは、Nグラム($q_0, \dots, q_j$) (又は、同等に、単語($w_0, \dots, w_j$))及び神経活動を所与として出現しているその単語系列の対数確率によってパラメータ化されたが、これらの方程式は、対数確率値の再帰的計算のみを記述する(各ビタビ経路と関連付けられる単語の追跡は暗黙的に仮定される)。

q_j, v_k : 単語 w_j 及びその単語の文脈を含むNグラム q_j 。この文脈は、ビタビ経路 v_k の隠れ状態系列内の最新の単語から決定される。

$p(y_j | q_j, v_{j-1})$: Nグラム q_j, v_{j-1} を所与として実測神経活動 y_j の尤度を指定する放出確率。

$p(q_i, v_{i-1} | q_{i-1}, v_{i-1})$: Nグラム q_{i-1}, v_{i-1} からNグラム q_i, v_{i-1} に遷移する事前確率を指定する遷移確率。

L : 単語分類器からの放出確率に対して、言語モデルからの遷移確率の重みを制御するために使用したハイパーパラメータである、言語モデルスケール係数(ハイパーパラメータ最適化手順の記述については、方法S7及び表S1を参照)。

W : 50単語セット。

$\log$: 自然対数。

【0319】

前のセクションで説明された単純化を使用して、式S15は、以下の式に単純化され得る。

10

20

30

40

50

【数 2 2】

$$V_i = \left\{ v_{i-1} + \log p(w_i | y_i) + L \log p(w_i | c_{i,v_{i-1}}) \mid v_{i-1} \in V_{i-1} \wedge w_i \in W \right\}, \quad (S16)$$

式中、 c_j 、 v_k は、ビタビ経路 v_k から決定される単語 w_j の文脈であり、 $p(w_i | y_i)$ は、（単語分類器から取得される）放出確率であり、 $p(w_i | c_i, v_{i-1})$ は、（言語モデルから取得される）遷移確率である。各文試験の開始時に、インデックス i をゼロにリセットし（各試験の最初の単語を w_0 とした）、以前の試験からの任意の既存のビタビ経路を破棄した。再帰を初期化するために、 V_{-1} を、その隠れ状態系列としての空集合及びゼロの関連する対数確率を有する単一のビタビ経路を含むシングルトン集合として定義した。数値的安定性及び計算効率のために、実際に対数確率を使用した。

10

【0 3 2 0】

ビーム探索を介したビタビ経路プルーニング

式 S 1 6 で指定されるように、新しい放出確率 $p(w_i | y_i)$ が単語分類器から取得されたとき、本発明者らのビタビデコーダは、 V_{i-1} 内の各既存の経路を各可能な次の N グラム q_i に遷移させることによって作成された経路から構成される、ビタビ経路の新しいセット V_i を計算した。その結果、インデックス i に対して作成された新しいビタビ経路の数は、 $|V_{i-1}| \times W$ （インデックス $i-1$ における既存のビタビ経路の数に 50 を乗算した値）に等しくなった。介入なしでは、インデックスが増大するにつれてビタビ経路の数は指数関数的に増大する（ $|V_i| = |W|^{(i+1)}$ ）。

20

【0 3 2 1】

指数関数的な増大を防ぐために、ビーム幅が β であるビーム探索を、各新しいビタビ経路セット V_i に、作成直後に適用した。このビーム探索は、各新しいビタビ経路セットに β の最大サイズを強制し、 β 個の最も可能性の高い経路（関連付けられる対数確率が最大の経路）を保持し、残りをプルーニング（破棄）した。 $|V_i| \leq \beta$ の場合、全ての経路を保持した。ビーム探索手順を含むように方程式 S 1 6 を拡大すると、文復号中に実際に使用したビタビ復号更新式の最終セットが得られる。

【数 2 3】

$$V'_i = \left\{ v_{i-1} + \log p(w_i | y_i) + L \log p(w_i | c_{i,v_{i-1}}) \mid v_{i-1} \in V_{i-1} \wedge w_i \in W \right\} \quad (S17)$$

30

$$V_i = \left\{ v_{i,j} \mid j \in \left\{ 0, \dots, \text{分}(\beta, |V'_i|) - 1 \right\} \right\}, \quad (S18)$$

式中、 V'_i 、文試験内（プルーニング前）のインデックス i における単語生成試行後に作成された全てのビタビ経路の集合であり、 $v_{i,j}$ は、 V'_i 内のビタビ経路を対数確率の降順でソートすることによって作成されたベクトルのインデックス j における要素である（ソート中、紐づけは任意に破棄される）。

40

【0 3 2 2】

方法 S 1 2 . 文復号評価

文タスクブロック中に行われたオンライン予測を使用して、本発明の復号パイプライン（発話検出器、単語分類器、言語モデル、及びビタビデコーダ）のパフォーマンスを評価した（テストサブセット内。方法 S 6 を参照されたい）。具体的には、各試験のアクティブフェーズ（参加者がプロンプトされた文標的の生成を試行するように指示された各試験の部分）中に参加者の神経活動からリアルタイムで復号された文を分析した。オフラインで、（リアルタイム復号中に無視された）非アクティブタスクフェーズ中に誤って検出された偽陽性発話イベントの数をカウントした。これらの偽陽性イベントは、ブロック内の最初の試験の前にのみ発生し、このカウントは本文の結果セクションに報告される。

50

【0323】

単語誤り率及び編集距離

復号結果の品質を測定するために、各試験における標的文と復号された文との間の単語誤り率（WER）を計算した。WERは、参照（標的）と復号された文との間の編集（レーベンシュタイン）距離を算出し、次いで編集距離を参照文内の単語数で除算することによって算出される、予測単語系列の質を測定するために一般的に使用されるメトリックである。ここで、編集距離測定値は、復号された文の単語誤りの数として解釈することができる（本文の図2では、編集距離は「単語誤りの数」又は「誤りカウント」と称される）。これは、復号された文を参照文に変換するために必要な挿入、欠失、及び置換の最小数として計算される。下記において、例示的な復号された文（各矢印の左側）を標的文「I am good（私は元気です）」に変換するために使用することができる各タイプの編集操作を実証する。いずれの場合も、例示的な復号された文は、標的文に対して1の編集距離を有する。

挿入：I good → I am good

欠失：I am very good → I am good

置換：I am going → I am good

【0324】

編集距離及びWERが小さいほど、パフォーマンスが向上する。言語モデル及びビタビデコーダがある状態及びない状態で行われた予測を使用して、編集距離及びWERを計算した。

【0325】

本文内の図2Aに示されるブロックレベルWERを計算するために、まず、（本文内の図2Dに示される）各文試験についての編集距離を計算した。次いで、ブロックレベルWERを、テストブロック内の全ての試験にわたる編集距離の合計を、全ての試験にわたる標的文単語長の合計で除算した値として計算した。ブロックレベルのWERを測定するこのアプローチは、短い文を長い文と比較して過大評価しないため、単に試験レベルWER値を平均化するよりも選好された。例えば、ブロックレベルWERを計算するために試験レベルWERを単純に平均した場合、標的文「I am thirsty」を有する試験の1つの誤りは、標的文「My family is very comfortable」を有する試験の1つの誤りよりも大きい影響を引き起こす。これは、ブロックレベルWER測定の望ましい側面ではなかった。

【0326】

文タスクによる本発明の復号アプローチの偶然パフォーマンスを評価するために、言語モデル及びビタビデコーダ（いかなる神経データにも依存しない）からランダムに生成された文を使用してWERを測定した。これらの文を生成するために、各試験で次の手順を実行した。

ステップ1：空の単語系列から開始する。

ステップ2：現在の単語系列を文脈として使用して、言語モデルから単語確率を取得する。

ステップ3：ステップ2の単語確率をサンプリングの重みとして使用して、50単語セットから単語をランダムにサンプリングする。

ステップ4：ステップ3からの単語を現在の単語系列に追加する。

ステップ5：現在の単語系列の長さが試験の標的文の長さと同しくなるまで、ステップ2～4を繰り返す。

【0327】

各試験のランダムに生成された文を用いて、前のパラグラフに記載された方法を使用してブロックレベルWERを計算することにより、偶然パフォーマンスを測定した。偶然パフォーマンスを測定する方法は、言語モデル、及び各試験の標的文と同じ文長を使用するため、真の偶然パフォーマンスを過大評価することに留意されたい（これは、発話検出モデルが各試験で常に正しい単語数を検出したと仮定することと同等である）。

【0328】

毎分単語数及び復号された単語の正確性

復号速度を測定するために、毎分単語数（WPM）メトリックを使用した。各試験について、試験内の検出された単語の数をカウントし、そのカウントを検出された試験持続時間で除算することによって、WPM値を計算した。各検出された試験期間を、参加者のモニタ上に文プロンプトが現れた時間（ゴーキュー）と、試験内で発話検出器から単語分類器に渡された最後の神経時間サンプルの時間との間の経過時間として算出した。

【0329】

単語が正確に復号された速度を測定するために、正しく復号された単語のみをカウントしながらWPMも計算した。各試験でいずれの単語が正しく復号されたかを決定するために、次のステップを実施した。

ステップ1： $n = 1$ 、 $\omega = 0$ で開始する。

ステップ2：復号された文の最初の n 個の単語と標的文の最初の n 個の単語との間のWERを計算する。

ステップ3：このWERが ω 以下であり、かつ $\omega \neq 1$ の場合、復号された文のインデックス n の単語は正しいとみなされる（ $n = 1$ は文の最初の単語のインデックスである）。それ以外の場合、インデックス n の単語は正しくないとみなされる。

ステップ4： ω がこのWER値に等しいものとし、 n を1だけ増分する。

ステップ5：復号された文の各単語が正しいか又は正しくないとみなされるまで、ステップ2～4を繰り返す。

【0330】

システム待ち時間算出

リアルタイム文復号中の復号パイプラインの待ち時間を推定するために、まず、待ち時間を計算するために使用する文テストブロックのうちの1つをランダムに選択した。インフラストラクチャ及びモデルパラメータは、文テストブロック間で同一であったため、任意のブロックからの待ち時間の分布が、全てのブロックにわたる待ち時間の分布を表すはずであると仮定した。これは、（本発明の観点から、及び参加者の観点から）全ての文テストブロックにわたって待ち時間に顕著な差がないことによって更に裏付けされた。文テストブロックをランダムに選択した後、ブロックのビデオ記録を使用して、各復号された単語が画面に表示された時間を識別した。次いで、各リアルタイム単語予測の待ち時間を、単語出現時間（ビデオから）と、その単語と関連付けられる神経活動の検出ウィンドウに含まれる最終的な神経データ点の時間（ブロックと関連付けられる結果ファイルから取得される、その単語生成試行の確率を予測するために単語分類器によって使用される神経データの最終時点）との差として計算した。これらの差を使用することによって、計算された待ち時間は、その予測を行うために必要とされる関連付けられる神経データの全てを取得した後に、系列内の次の単語を予測するためにシステムが必要とする時間量を表した。ビデオと結果ファイルタイムスタンプとの間のタイミングは、全てのブロックの開始時に再生される短いビーブ音を使用して同期された（各ブロック中にスピーカ出力も取得され、結果ファイルに格納された。方法S2を参照）。全ての試行にわたって、このブロックには42個の復号された単語があった。

【0331】

このアプローチを使用して、本発明者らは、リアルタイム単語予測に関連する平均待ち時間が4.0秒であることを見出した（標準偏差0.91秒）。

【0332】

方法S13．孤立単語評価

分類正確度、交差エントロピー、及び検出誤り

孤立単語データのオフライン交差検証評価中（方法S6を参照）、単語分類器を使用して、各試験における単語生成試行と関連付けられる神経データから単語確率を予測した。発話検出器からのキュレーションされた検出イベントと関連付けられる神経活動のタイムウィンドウを使用して、これらの単語確率を計算した（方法S8を参照されたい）。これ

らの予測単語確率から、分類正確度を、標的単語が最も高い予測確率を有する単語と等しかった試験の割合として計算した。また、これらの予測確率を使用して、交差エントロピーを計算しており、これは、予測確率から標的単語のアイデンティティを決定するために必要な追加情報の量を測定する。交差エントロピーを計算するために、まず各試験において標的単語の予測確率を取得した。次いで、交差エントロピー（ビット単位）を、これらの確率の全てにわたる負の対数（底 2）の平均として算出した。キュレーションされた検出イベントを使用してこれらのメトリックを計算することに加えて、それらを使用して行われた検出誤りの数も測定した。具体的には、偽陰性の数（検出されたイベントと関連付けられない試験の数）及び偽陽性の数（検出されたイベントの数は試験と関連付けられなかった）の 2 つのタイプの検出誤りを測定した。これらの検出誤りを別個に報告した（分類正確度及び交差エントロピーは、正しく検出された試験によってのみ計算され、検出誤りに対してペナルティは科されなかった）。

10

【0333】

発話検出器及び単語分類器の両方を適合させるために使用されるデータの量を変化させる学習曲線スキームを使用してこれらの分析を行った（方法 S 6 で詳述）。この学習曲線スキームの分析の最終セットは、利用可能な全てのデータを使用することと同等であった。学習曲線スキーム内の分析の全てのセットについて、発話検出器は、キュレーションされた検出發話イベントを提供した。単語分類器を適合させ、単語確率を予測するために、これらのキュレーションされた検出イベントの開始に整合した神経データを使用した。

20

【0334】

学習曲線スキームのための訓練データ量の測定

発話検出モデルと単語分類モデルとは異なる訓練手順を使用しているため、学習曲線スキームの分析の各セットに対して、各タイプのモデルによって使用される神経データの量を別個に測定した。各単語分類器について、モデルを適合させるために使用される検出イベントの数に 4 秒（分類器が使用する神経タイムウィンドウのサイズ）を乗算した。学習曲線スキームの各分析セットは 10 フォールド交差検証を使用したため、この結果として、各分析セットに使用される訓練データの量の 10 個の測度が得られた。10 フォールドにわたって平均を計算することによって、分析の各セットについて単語分類器を適合させるために使用されるデータの平均量の単一の測度を得た。

30

【0335】

各発話検出モデルは、神経活動の個々の時点を予測するためのスライドウィンドウによって適合され、試験よりもタスクブロック当たりの多くの訓練サンプルが得られた。ここで、各訓練サンプルは、タスクブロック内の個々の時点に対応するスライドウィンドウ訓練手順からの単一のウィンドウであった。過剰適合を防ぐために早期停止を使用したため、実際には各発話検出器はモデル適合中に利用可能な全てのデータを使用することはなかった。しかしながら、利用可能なデータの量を増大させることは、（例えば、長い期間にわたって収集されたブロックからのデータを有することによって）訓練データの多様性を増大させることができ、これはまた、検出器が訓練されるエポックの数及び訓練された検出モデルのロバスト性に影響を与えることができる。訓練中に各発話検出器に利用可能なデータの量を測定するために、単純に、利用可能な訓練サンプルの数をサンプリングレート（200 Hz）で除算した。訓練中に各発話検出器によって実際に使用されたデータの量を測定するために、使用された訓練サンプルの数をサンプリングレートで除算した。10 フォールドにわたって平均を計算することによって、利用可能なデータの平均量、及び分析の各セットについて単語分類器を適合させるために実際に使用された平均量を測定した。

40

【0336】

電極寄与（サリエンス）

各電極が検出及び分類パフォーマンスにどの程度寄与したかを測定するために、それぞれ発話検出及び単語分類モデルを駆動する人工ニューラルネットワーク（ANN）を用いて電極寄与（サリエンス）を計算した。画像分類に最も有用な画像領域の識別中に畳み込

50

みANNによって実証されたサリエンス算出方法を使用した〔35〕。また、この方法を本発明者らの以前の研究で使用して、再帰ANN及び畳み込みANNを用いた発話復号に最も有用な電極を測定している〔20〕。

【0337】

各タイプのANNの電極サリエンスを算出するために、まず、入力特徴に関するANNの損失関数の勾配を算出した。入力特徴は、発話検出器のブロック全体にわたる、又は単語分類器の検出タイムウィンドウにわたる、高ガンマ活性の個々の時間サンプルであった。各入力特徴について、ANNを介して入力層に勾配を逆伝播させた。次いで、各電極と関連付けられる、得られた勾配値の経時的な（各ブロック又は試験内の）ユークリッドノルムを計算した。ここでは、勾配のノルムを使用して、各入力に対する損失関数の感度の大きさを測定した（感度の方向は度外視した）。次に、ユークリッドノルム値のブロック又は試験にわたる平均を計算し、各電極について単一のサリエンス値をもたらした。最後に、電極サリエンスの各セットを正規化して、合計が1になるようにした。

【0338】

発話検出器及び単語分類器の10フォールド交差検証評価を使用して、学習曲線スキームの分析の最終セット中にこれらのサリエンスを計算した。各フォールドのテストセットにおいて評価されたブロック及び試験を使用して、勾配を計算した。また、信号安定性分析中にもサリエンスを計算した（方法S14を参照されたい）。

【0339】

情報転送速度

システムが単位時間あたりに通信する情報の量を測定する情報転送速度（ITR）メトリックは、脳コンピュータインターフェースを評価するために一般的に使用される〔36〕。既存の文献〔2、36、37〕に記載されている定式化と同様に、本研究では以下の式を使用してITRを計算した。

【数24】

$$ITR = \frac{1}{T} \left[\log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1} \right) \right], \quad S19$$

式中、Nは、一意の標的の数であり、Pは、予測正確度であり、Tは、各予測の平均持続時間である。本研究では、N=50（単語セットのサイズ）及びT=4秒（分類器が単語確率を計算するために使用する神経タイムウィンドウのサイズ）とした。（学習曲線スキームの分析の最終セットからの）孤立単語データによる完全な交差検証分析の平均分類正確度に等しいPを設定した。この式では、次のように仮定する。

【0340】

平均して、全ての可能な単語標的の、任意の試験において実際の単語標的である事前確率（すなわち、神経データに依存しない確率）は同じであった。これは、単語標的ごとに同数の孤立単語試験が収集されたため、合理的である。

【0341】

Pに使用される分類正確度は、（訓練データの量を所与として）単語分類器の全体的な正確度を表し、試験間で一貫していた。本発明者らの交差検証分析により、収集された全ての試験のパフォーマンスを評価することができたため、これは妥当な仮定であるはずである。

【0342】

平均して、正しくない各単語標的の、任意の試験において最も高い確率値を割り当てられる確率は同じであった。これは、本発明の結果について実際には正確には当てはまらないが（図3に示す混同行列で明らかのように、一部の単語は平均して他の単語よりもわずかに頻繁に予測されている）、これは典型的には、この式を使用している他の研究では正確には当てはまらず、一般的に、許容できる単純化仮定とみなされる。

【0343】

式 S 1 9 を使用して、I T R を計算し、その結果を図 1 2 のキャプションに報告した。
【 0 3 4 4 】

I T R は、単語分類器（発話検出器からの検出神経ウィンドウを使用した）からの孤立単語予測についてのみ計算された。言語モデルからの単語系列確率が、上記のリストからの仮定（1）及び（3）に違反するため、文データに関する復号パイプライン全体（言語モデルを含む）の I T R の算出は、著しく複雑になる [3 8]。いくつかの復号された文の単語長が対応する標的文と異なるという事実もまた、I T R 計算をより困難にする。単純にするために、単語分類器出力を使用した I T R のみを報告することとした。この I T R 測定は、他の脳コンピュータインターフェースアプリケーション（本発明の特定の言語モデリングアプローチとは無関係）で報告されている弁別モデルのパフォーマンスと比較しても、より容易に行うことができる。

10

【 0 3 4 5 】

潜在的な音響汚染の調査

最近の研究で、R o u s s e l l は、音響信号が電気生理学的記録を直接「汚染」する可能性があり、それによって、電気生理学的記録方法によって記録された信号の時間スペクトル内容が、同時に発生する音響波形と強く関連させられることを実証している [3 9]。本発明者らの神経記録に音響汚染が存在するか否かを評価するために、[3 9] に記載されている汚染識別方法を本発明者らのデータセットに適用した（下記に記載する、いくつかのわずかな手順上の逸脱がある）。

【 0 3 4 6 】

まず、この分析で考慮するために、2 4 個の孤立単語タスクブロック（8 1 週間の研究期間にわたって時系列的に分布していた）からなるセットをランダムに選択した。各ブロックから、1 k H z で記録された神経活動（共通平均に対する再参照又は高ガンマ特徴抽出を使用しては処理されなかった）と及び 3 0 k H z で記録されたマイクロフォン信号を取得した。これらのマイクロフォン信号は、（方法 S 2 で説明されているように）すでに神経信号に同期されていた。次いで、マイク信号を 1 0 0 0 H z にダウンサンプリングして、神経データに一致させた。次に、[3 9] で実行されたように、本発明者らは、各時点での信号から前の 1 秒間にわたるその平均値を差し引くことによって、マイクロフォン信号を「センタリング」した。

20

【 0 3 4 7 】

次いで、各電極チャンネルから記録された神経活動及び記録されたマイクロフォン信号のスペクトログラムを計算した。スペクトログラムを、短時間フーリエ変換の絶対値として計算した。計算効率のために、本発明のアプローチにおいては、[3 9] からわずかに逸脱して、2 の累乗を使用した。[3 9] で使用されている 2 0 0 ミリ秒のウィンドウとは対照的に、2 5 6 サンプル（各ウィンドウには 2 5 6 m s のデータが含まれている）のスライドウィンドウ内でフーリエ変換を計算し、0 ~ 5 0 0 H z の間で均等に離間した中心周波数を有する 1 2 9 個の周波数帯域を得た。各スライドウィンドウは、3 2 時間サンプルだけ離間され、[3 9] で使用された 5 0 H z のレートとは対照的に、約 3 1 H z においてスペクトログラムサンプルを得た。大量の「沈黙」タスクセグメント（参加者が発話を試行していなかったセグメント）を含めると、音響汚染を見つけることに対して分析が偏ってしまうため、試験間沈黙に対応する時間期間をスペクトログラムから切り取った。具体的には、各試験でゴーキューの 0 . 5 秒前から 3 . 5 秒後までの間に発生したデータから計算されたスペクトログラムのみを保持した。これらの時間期間には、参加者が沈黙している間に記録されたサンプルがまだ含まれていたが、このアプローチは、考慮されたデータにおける沈黙の全体的な割合を劇的に減少させた。

30

40

【 0 3 4 8 】

次いで、各マイクロフォンスペクトログラムと各電極の対応するスペクトログラムとの間の（個々の周波数帯域内の）経時相関を測定した。神経チャンネルとマイクロフォン信号との間の小さい相関は、音響汚染の決定的な証拠ではなく、共有された電気ノイズの存在、及び試行発話生成中に誘発される純粋に生理学的な神経応答の特性を含む、相関に影響

50

を与える可能性のある多くの要因がある。狭い周波数帯域内の相関を計算することによって、結果として生じる相関は、音響汚染を示す可能性がより高く（しかし、保証はない）、例えば、音響信号中の300Hzでのスペクトルパワーは、電気生理学的信号中のその周波数での神経振動と強く相関することは期待されない。スペクトログラムにわたって相関行列を集約し、電極及び周波数帯域ごとに1つの要素を含む、考慮される全てのデータにわたる全体的な相関行列を取得した。この手順は、各ブロックからの（切り取られた）神経スペクトログラムと音響スペクトログラムとを連結し、次いで、全てのデータにわたって単一の相関行列を計算することと同等であった。

【0349】

任意の潜在的な音響汚染を更に特徴付けるために、本発明者らは、周波数の関数としての神経スペクトログラムと音響スペクトログラムとの間の相関を、マイクロフォンのパワースペクトル密度（PSD）と比較した。本研究の中心的な仮説は、埋め込まれた電極から記録される神経活動が試行発話生成に因果的に関連しているということであるため、相関はゼロではないと予測した。しかしながら、このPSDとともに増大及び減少する神経スペクトログラムと音響スペクトログラムとの間の強い相関は、音響汚染の強い証拠となる。ここでは、マイクロフォンPSDを、全てのスペクトログラムサンプル及びブロックにわたる（周波数次元に沿った）マイクロフォンスペクトログラムの平均として計算した（周波数帯域ごとに単一の値をもたらす）。

【0350】

方法S14．安定性評価

単語生成試行中に記録された神経信号の安定性を評価するために、モデルを訓練及び試験するために使用されるデータがサンプリングされた日付範囲を変化させながら、発話検出器及び単語分類器を用いて分類正確度及び電極寄与（サリエンス）を計算した。4つの日付範囲サブセット（「早期」、「中間」、「後期」、及び「最後期」）、及び方法S6で定義された3つの評価スキーム（サブセット内、サブセット間、及び累積サブセット）を使用してこれらの分析を実施した。

【0351】

まず、各サブセットについてキュレーションされた検出時間をもたらすために、発話検出モデルは、サブセット内訓練スキームを使用した。結果として、サブセットについての全てのキュレーションされた検出イベントが、同じサブセットからのデータのみによって適合された発話検出モデルから取得された。検出されたイベントキュレーション手順中に検出されたイベントと関連付けられなかったため、各サブセットにおける更なる分析から除外された試験の割合は、「早期」、「中間」、「後期」、及び「最後期」サブセットについて、それぞれ2.3%、3.8%、0.8%、及び1.5%であった。単語分類器は、これらのキュレーションされた検出イベントの開始に整合した神経データを使用して訓練及びテストされた。

【0352】

各日付範囲中に記録された神経信号が同様の量の弁別的情報を含んでいるか否かを判定するために（及び経時的な全体的記録品質の低下の尤度を評価するために）、サブセット内評価スキームを使用して計算された異なる日付範囲サブセットからの分類正確度を比較した。分類モデルによって学習された空間マップの安定性を評価するために、サブセット内評価スキームを使用して各日付範囲サブセットについて電極サリエンス（寄与）も計算した。

【0353】

訓練及びテストデータの時間的近接性が分類パフォーマンスに影響を及ぼしたか否かを決定するために（及びサブセット内正確度が全て類似していても、日付範囲サブセット間で基礎となる神経活動に有意な変化があったか否かを評価する）、サブセット内及びサブセット間の分類正確度をサブセットごとに個別に比較した。サブセット内及びサブセット間の比較を図14に示す。

【0354】

数ヶ月の記録にわたって収集された皮質活動が、頻繁な再較正なしにモデルパフォーマンスを改善するために蓄積され得るか否かを評価するために、累積サブセット評価スキームを使用して訓練データの量を変化させながら、「最後期」サブセットに対して分類正確度を計算した（本文の図 4 に示されている）。この評価スキームの訓練データ量を測定するために、学習曲線分析における単語分類器の訓練データ量を測定するために、方法 S 13 に記載されているものと同じ方法を使用した。

【0355】

方法 S 15 . 統計テスト

単語誤り率信頼区間

単語誤り率（W E R）の 95% 信頼区間を計算するために、結果の各セット（偶然、言語モデルなし、及び言語モデルあり）に対して次のステップを実行した。

10

【0356】

1. ブロックレベル W E R を単一の配列にコンパイルする（各ブロックに 1 つずつ 15 個の要素がある）。

【0357】

2. この配列から 15 個の W E R 値をランダムにサンプリングし（置換あり）、これらの値から中央値 W E R を計算して格納する。

【0358】

3. 100 万個の中央 W E R 値が計算されるまで、ステップ 2 を繰り返す。

【0359】

4. 信頼区間を、ステップ 3 からの中央 W E R 値の集合の 2.5 及び 97.5 パーセントイルとして計算する。

20

【0360】

分類正確度信頼区間

信号安定性分析中に得られる分類正確度の 95% 信頼区間を計算するために、各日付範囲サブセット（「早期」、「中間」、「後期」、及び「最後期」）及び各評価スキーム（サブセット内、サブセット間、及び累積サブセット）について以下のステップを実行した。

【0361】

1. 各交差検証フォールドからの分類正確度を単一の配列にコンパイルする（各フォールドに 1 つずつ 10 個の要素がある）。

30

【0362】

2. この配列から 10 個の分類正確度をランダムにサンプリングし（置換あり）、これらの値から平均分類正確度を計算して格納する。

【0363】

3. 100 万個の平均分類正確度が計算されるまで、ステップ 2 を繰り返す。

【0364】

4. 信頼区間を、ステップ 3 からの平均分類正確度の集合の 2.5 及び 97.5 パーセントイルとして計算する。

【0365】

40

【表 1】

補足表 S 1. ハイパーパラメータ定義及び値

モデル	ハイパーパラメータの探索空間タイプ 説明		値範囲	最適値 ¹
発話 検出器	平滑化サイズ	均一(整数)	[1,80]	(8,5,22)
	確率閾値	均一	[0.1,0.9]	(0.297,0.319,0.592)
	時間閾値持続時間	均一(整数)	[25,150]	(79,82,93)
単語 分類器	GRU 層の数	均一(整数)	[1,3]	(2,2)
	GRU 層当たりのノー ド	均一(整数)	[64,512]	(434,420)
	ドロップアウト分数	均一	[0.5,0.95]	(0.704,0.646)
	畳み込みカーネルのサ イズ及びスキップ	均一(整数)	[1,2]	(2,2)
言語 モデル	最初の単語	対数的	[0.001,1000]	0.576
	平滑化(ψ)	均一		
ビタビ デコーダ	言語モデル スケール係数(L)	対数的 均一	[0.1,10]	0.913

¹発話検出ハイパーパラメータについては、3つの値がリストされている。1つ目は、孤立単語最適化サブセットに対して検出器を最適化するとき求められた最適値(単語分類器による評価のための交差検証サブセット内の単語生成試行を検出するために使用)であり、2つ目は、プールされた交差検証サブセットのサブセットに対して検出器を最適化するとき求められた最適値(単語分類器のハイパーパラメータ最適化中に使用するために、孤立単語最適化サブセット内の単語生成試行を検出するために使用)であり、3つ目は、文最適化サブセットを用いた復号パイプラインのハイパーパラメータ最適化中に求められた最適値(オンライン文復号中に使用される値)である。単語分類ハイパーパラメータについては、2つの値がリストされている。1つ目は、孤立単語最適化サブセットに対して分類器を最適化するとき求められた最適値(全ての孤立単語評価に使用される値)であり、2つ目は、研究期間の終わり近くに孤立単語試験の小規模なサブセットに対して分類器を最適化するとき求められた最適値(オフライン文最適化及びオンライン文復号に使用される値)である。言語モデリング及びビタビ復号ハイパーパラメータについては、文最適化サブセットを用いて復号パイプラインを最適化するとき、リストされた最適値が求められた(オンライン文復号に使用される値)。

【0366】

補足参考文献

1. Moses DA, Leonard MK, and Chang EF. Real-time classification of auditory sentences using evoked cortical activity in humans. *Journal of Neural Engineering* 2018; 15: 036005.
2. Moses DA, Leonard MK, Makin JG, and Chang EF. Real-time decoding of question-and-answer speech dialogue using human corti

cal activity. Nature Communications 2019; 10.

3. Ludwig KA, Miriani RM, Langhals NB, Joseph MD, Anderson DJ, and Kipke DR. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. Journal of neurophysiology 2009; 101:1679–89.

4. Williams AJ, Trumpis M, Bent B, Chiang CH, and Viventi J. A Novel iECOG Electrode Interface for Comparison of Local and Common Averaged Referenced Signals. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI: IEEE, 2018: 5057–60.

5. Parks TW and McClellan JH. Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase. IEEE Transactions on Circuit Theory 1972; 19:189–94.

6. Romero DET and Jovanovic G. Digital FIR Hilbert Transformers: Fundamentals and Efficient Design Methods. In: MATLAB—A Fundamental Tool for Scientific Computing and Engineering Applications—Volume 1. 2012: 445–82.

7. Welford BP. Note on a Method for Calculating Corrected Sums of Squares and Products. Technometrics 1962; 4:419–9.

8. Weiss JM, Gaunt RA, Franklin R, Boninger ML, and Collinger JL. Demonstration of a portable intracortical brain–computer interface. Brain–Computer Interfaces 2019; 6: 106–17.

9. Bergstra J, Yamins DLK, and Cox DD. Making a Science of Model Search: Hyper-parameter Optimization in Hundreds of Dimensions for Vision Architectures. Icml 2013: 115–23.

10. Liaw R, Liang E, Nishihara R, Moritz P, Gonzalez JE, and Stoica I. Tune: A Research Platform for Distributed Model Selection and Training. arXiv:1807.05118 2018.

11. Li L, Jamieson K, Rostamizadeh A, et al. A System for Massively Parallel Hyperparameter Tuning. arXiv:1810.05934 2020.

12. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32. Ed

10

20

30

40

50

. by Wallach H, Larochelle H, Beygelzimer A, d'Alch'e-Buc F, Fox E, and Garnett R. Curran Associates, Inc., 2019:8024-35.

13. Hochreiter S and Schmidhuber J. Long Short-Term Memory. *Neural Computation* 1997;9:1735-80.

14. Dash D, Ferrari P, Dutta S, and Wang J. NeuroVAD: Real-Time Voice Activity Detection from Non-Invasive Neuromagnetic Signals. *Sensors* 2020;20:2248.

15. Werbos P. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE* 1990;78:1550-60.

16. Elman JL. Finding Structure in Time. *Cognitive Science* 1990;14:179-211.

17. Williams RJ and Peng J. An Efficient Gradient-Based Algorithm for On-Line Training of Recurrent Network Trajectories. *Neural Computation* 1990;2:490-501.

18. Kingma DP and Ba J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980* 2017.

19. Krizhevsky A, Sutskever I, and Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems* 25. Ed. by Pereira F, Burges CJC, Bottou L, and Weinberger KQ. Curran Associates, Inc., 2012:1097-105.

20. Makin JG, Moses DA, and Chang EF. Machine translation of cortical activity to text with an encoder-decoder framework. *Nature Neuroscience* 2020;23:575-82.

21. Virtanen P, Gommers R, Oliphant TE, et al. *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*. *Nature Methods* 2020;17:261-72.

22. Martin Abadi, Ashish Agarwal, Paul Barham, et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

23. Zhang Y, Chan W, and Jaitly N. Very Deep Convolutional Networks for End-to-End Speech Recognition. *arXiv:1610.03022* 2016.

24. Cho K, Merriënboer B van, Gulcehre C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv:1406.1078* 2014.

25. Pascanu R, Mikolov T, and Bengio Y. On the difficulty of training recurrent neu

10

20

30

40

50

ral networks. In: Proceedings of the 30th International Conference on Machine Learning. Ed. by Dasgupta S and McAllester D. Vol. 28. Proceedings of Machine Learning Research. Atlanta, Georgia, USA: PMLR, 2013: 1310–8.

26. Sollich P and Krogh A. Learning with ensembles: How overfitting can be useful. In: Advances in Neural Information Processing Systems 8. Ed. by Touretzky DS, Mozer MC, and Hasselmo ME. MIT Press, 1996: 190–6.

27. Chen SF and Goodman J. An empirical study of smoothing techniques for language modeling. Computer Speech & Language 1999; 13: 359–93.

28. Kneser R and Ney H. Improved backing-off for M-gram language modeling. In: 1995 International Conference on Acoustics, Speech, and Signal Processing. Vol. 1. Detroit, MI, USA: IEEE, 1995: 181–4.

29. Bird S, Klein E, and Loper E. Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc., 2009.

30. Group TH. Hierarchical Data Format. 1997.

31. Collette A. Python and HDF5: unlocking scientific data. "O'Reilly Media, Inc.", 2013.

32. Heafield K. KenLM: Faster and Smaller Language Model Queries. In: Proceedings of the Sixth Workshop on Statistical Machine Translation. WMT '11. Association for Computational Linguistics, 2011: 187–97.

33. Viterbi AJ. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm. IEEE Transactions on Information Theory 1967; 13: 260–9.

34. Jurafsky D and Martin JH. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. 2nd. Upper Saddle River, New Jersey: Pearson Education, Inc., 2009.

35. Simonyan K, Vedaldi A, and Zisserman A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. In: Workshop at the International Conference on Learning Representations. Ed. by Bengio Y and LeCun Y. Banff, C

10

20

30

40

50

anada, 2014.

36. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, and Vaughan TM. Brain-computer interfaces for communication and control. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology* 2002;113:767-91.

37. Mugler EM, Patton JL, Flint RD, et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. *Journal of neural engineering* 2014;11:35015-15.

38. Speier W, Arnold C, and Pouratian N. Evaluating True BCI Communication Rate through Mutual Information and Language Models. *PLoS ONE* 2013;8. Ed. by Wennekers T:e78432.

39. Roussel P, Godais GL, Bocquelet F, et al. Observation and assessment of acoustic contamination of electrophysiological brain signals during speech production and sound perception. *Journal of Neural Engineering* 2020;17:056028.

【0367】

実施例3：麻痺者における発話神経プロテーゼを使用した一般化可能なスペリング
序論

脳卒中及び筋萎縮性側索硬化症などの壊滅的な神経学的状態は、発話を通じたコミュニケーション能力の喪失である構語障害をもたらす可能性がある¹。構語障害患者は、正常な言語技能及び認知を有することができるが、麻痺は、補助デバイス进行操作する能力を阻害し、家族、友人、及び介護者とのコミュニケーションを著しく制限し、自己報告の生活の質を低下させる可能性がある²。

【0368】

脳コンピュータインターフェース（BCI）は、神経活動を復号して意図されたメッセージにすることによって、そのような患者へのコミュニケーションを回復させる可能性を有する^{3, 4}。既存のコミュニケーションBCIは、典型的には、想像上の腕及び手運動を文字に復号して、意図された文のスペリングを可能にすることに依拠する^{5, 6}。このアプローチの実装は有望な結果を呈しているが、発話又はテキストにする直接的な会話の自然な試行を復号することは、コミュニケーションBCIにまさる、より速く、より自然な制御を提供し得る。実際、将来のBCIユーザを対象とした最近の調査では、多くの患者が、腕及び手で駆動する神経機能代替よりも発話駆動の神経機能代替を選好することが示唆されている⁷。追加的に、脳が発話を生成するために声道運動を表現する方法の理解⁸⁻¹¹及び健常な発話者の脳活動からのテキスト復号の実証¹²⁻¹⁵においていくつかの最近の進歩があり、脳活動からの試行発話の復号がコミュニケーション回復のための実行可能なアプローチであり得ることを示唆している。

【0369】

これを評価するために、本発明者らは最近、構語障害及び麻痺を患っているヒトの、発話を試行したときの皮質活動から完全な単語をリアルタイムで直接復号するための発話神経プロテーゼを開発した¹⁶。このアプローチは、有望な復号正確度及び速度を呈したが、初期の研究では予備的な50語の語彙に焦点を当てていた。限られた語彙での直接的な

単語復号は即座の実用的な利点があるが、少なくとも1,000語のより大きい語彙へとアクセスを拡張すると、自然英語文の内容の85%超がカバーされ¹⁷、コミュニケーション支援技術の日常的な効果的な使用が可能になる¹⁸。したがって、強力な補完技術は、現在の発話復号アプローチを拡張して、頻繁かつ一般的に使用される単語を表現するために、依然として高速で直接的な単語復号を可能にしながら、ユーザが大量の一般化可能な語彙から意図されたメッセージをスペルアウトことを可能にし得る。これとは別に、この先行研究では、参加者は声に出しての発話を試行することによって神経プロテゼを制御していたため、音声出力をまったく生成できない潜在的なユーザにとってこのアプローチが実行可能か否かは不明であった。

【0370】

ここでは、NATOフォネティックアルファベットからの26個のアルファベットコードワードの発語の沈黙試行のリアルタイム復号が、麻痺及び構語障害を患う参加者において非常に正確で迅速なスペリングを可能にし得ることを実証する。訓練セッション中に、参加者に、個々のコードワード及び手運動の生成を試行するように指示し、埋め込まれた128チャンネル皮質脳波記録（ECOG）アレイからの同時記録皮質活動を使用して、分類及び検出モデルを訓練した。訓練後、参加者は、対応するアルファベットコードワードを黙って発語する試行を使用して、1,152語の語彙によってリアルタイムで文をスペルアウトするスペリングタスクを実行した。ビーム探索アルゴリズムが、分類モデルからの予測コードワード確率を使用して、復号された単語の間に空白を自動的に挿入しながら、神経活動を所与として最も可能性の高い文を見つけた。スペリングを開始するために、参加者は黙って発話を試行し、発話検出モデルはECOG活動から直接この開始信号を識別した。意図した文をスペルアウトした後、参加者はスベラからディスエンゲージするために手運動を試行した。分類モデルがECOG活動からこの手動コマンドを識別したとき、大規模なニューラルネットワークベースの言語モデルが、ビーム探索から潜在的な文候補を再スコアリングし、文を確定した。事後シミュレーションでは、本発明者らのシステムは9,000超の単語の大きい語彙によく一般化された。

【0371】

結果

リアルタイムスペリングパイプラインの概要

構語障害及び麻痺を患う参加者が、自身の感覚運動皮質上に埋め込まれた高密度皮質脳波記録（ECOG）アレイから得られた信号を使用してメッセージを黙ってスペルアウトことを可能にする文スペリングパイプラインを設計した（図15）。コピータイピング及び会話タスク条件下でスペリング系をテストした。コピータイピングタスク条件の各試験において、参加者は画面上に標的文を提示され、次いでその文の複製を試行した。会話タスク条件において、参加者が自身に提示された質問に対して自発的に選択した回答をスペルアウトした試験、及び参加者が任意のプロンプトされていない文をスペルアウトした試験の、2つのタイプの試験があった。リアルタイムテストの前は、1日の再較正は発生しなかった。モデルパラメータ及びハイパーパラメータは、以前のセッションからのデータのみを使用して適合された。

【0372】

参加者が文のスペリングを始める準備ができたとき、参加者は黙って任意の単語の発話を試行した（図15A）。本発明では、沈黙発話試行を、発声せずに発話を明確に発することを試行する意志的試行として定義する。一方、参加者の神経活性を各電極から記録し、高ガンマ活性（HGA、70～150Hz間）及び低周波信号（LFS、0.3～100Hz、図15B）を同時に抽出するように処理した。スペリングを開始するために、発話検出モデルは、結合特徴ストリーム（HGA+LFS特徴を含む、図15C）内のデータの各時点処理して、この最初の沈黙発話試行を検出した。

【0373】

発話試行が検出されると、ペーシングされたスペリング手順が開始された（図15D）。この手順では、下線に続いて3つの点が白色テキストで画面に表示された。点は1つずつ

つ消えていき、カウントダウンを表した。最後の点が消えた後、下線は緑色に変わって、ゴーキューを示す。その時点において、参加者は、文の最初の文字に対応する N A T O コードワードを黙って発語することを試行した。ゴーキューの直後の 2.5 秒の間隔中に得られた結合特徴ストリームからの神経特徴のタイムウィンドウは、神経分類器に渡された（図 15 E）。ゴーキューの直後、次の文字のカウントダウンが自動的に開始された。次いで、この手順を、参加者がそれから自発的にディスエンゲージするまで繰り返した（このセクションの後半で説明する）。

【0374】

神経分類器は、神経特徴の各タイムウィンドウを処理して、26 個のアルファベットコードワードにわたる確率を予測した（図 15 F）。ビーム探索アルゴリズムが、予測文字確率の系列を使用して潜在的な文候補を計算し、適切な場合に自動的に文字系列に空白を挿入し、言語モデルを使用して言語学的に妥当な文に優先順位を付ける。リアルタイム文スペリング中、ビーム探索は、コミュニケーション支援アプリケーションに関連する一般的な単語を含む、事前定義された 1,152 語の語彙から構成される文のみを考慮した。タスクの任意の時点で最も可能性の高い文は、参加者に常に見えていた（図 15 D）。ビーム探索は、より多くの予測を受け取った後に間違いを修正することができるため、表示された文に間違いがあってもスペリングを継続するように参加者に指示した。

【0375】

文全体を黙ってスペルアウトすることを試行した後、参加者は、スペリング手順からディスエンゲージするために右手を握ることを試行するように指示された（図 15 H）。神経分類器は、神経特徴の各 2.5 秒のウィンドウからこの試行手運動の確率を予測し、この確率が 80% を超える場合、スペリング手順を停止し、復号された文を最終化した（図 15 I）。文を最終化するために、不完全な単語を有する文を最初に潜在的な候補のリストから削除し、次いで残りの文を別個の言語モデルによって再スコアリングした。次いで、最も可能性の高い文を参加者の画面上で更新した（図 15 G）。短い遅延の後、画面がクリアされ、タスクは引き続き次の試験に進んだ。

【0376】

リアルタイムテストの前に検出及び分類モデルを訓練するために、参加者が孤立標的タスクを実行したときにデータを収集した。このタスクの各試験では、N A T O コードワードが画面に表示され、参加者は対応するゴーキューにおいてコードワードを黙って発語することを試行するように指示された。いくつかの試験では、コードワードの代わりに手動コマンドを表すインジケータが提示され、参加者は、それらの試験のためにゴーキューにおいて右手を握ることを想像するように指示された。

【0377】

復号パフォーマンス

スペリングシステムのパフォーマンスを評価するために、コピータイピングタスク中に参加者が 150 個の文（コミュニケーション支援コーパスから選択された 75 個の一意の文の各々の 2 回の繰り返し、表 S 1 を参照）のスペリングを試行したときに、参加者の神経活動から文をリアルタイムで復号した。単語誤り率（W E R）、文字誤り率（C E R）、毎分単語数（W P M）、及び毎分文字数（C P M）メトリックを使用して、復号された文を評価した（図 16）。文字及び単語の場合、誤り率は、予測文を参加者に表示された標的文に変換するために必要な文字又は単語の削除、挿入、及び置換の最小数である編集距離を、それぞれ標的文の文字又は単語の総数で除算した値として定義される。これらのメトリックは、一般的に、自動発話認識システム¹⁹及び脳コンピュータインターフェースアプリケーション^{6, 16}の復号パフォーマンスを評価するために使用される。

【0378】

リアルタイムテストブロックにわたって、中央値 C E R が 6.13%、中央値 W E R が 10.53%（99% 信頼区間（C I）[2.25, 11.6] 及び [5.76, 24.8]）であることを観察した（各ブロックは、複数の文スペリング試験を含んでいた、図 16 A、図 16 B）。150 文にわたって、105 文（70%）が誤りなく復号され、7

10

20

30

40

50

5文のうち69文(92%)が、それらが試行された2回のうちの少なくとも1回完璧に復号された。追加的に、150文にわたって、139文(92.7%)が、試行された手の握りの高い分類正確度によって可能にされた、正しい文字数で復号された(図16E)。また、テストブロックにわたって中央値CPMが29.41、中央値WPMが6.86(99%CI[29.1, 29.6]及び[6.54, 7.12])であり、個々のブロックのスプリング速度は30.79CPM及び8.60WPMと高かった(図16C、16D)。これらの速度は、参加者が市販のTobii Dynavoxタイピング支援デバイスを使用した際に観察された17.37CPM及び4.16WPM(99%CI[16.1, 19.3]及び[3.33, 5.05])の中央値速度よりも高い(本発明者らの以前の研究で測定されたものとして¹⁶)。

10

【0379】

分類器、ビーム探索、及び言語モデルの復号パフォーマンスへの個々の寄与を理解するために、これらのリアルタイムコピータイピングタスクブロック中に収集されたデータを使用してオフライン分析を実行した(図16A、図16B)。システムの偶然パフォーマンスを調べるために、ビーム探索及び言語モデルを引き続き使用しながら、モデルの予測をランダムに生成された値に置換した。これは、リアルタイム結果よりも有意に悪いCER及びWERをもたらした($z = 7.09$ 、 $P = 8.08 \times 10^{-12}$ 及び $z = 7.09$ 、 $P = 8.08 \times 10^{-12}$)。これは、神経信号の分類がシステムパフォーマンスにとって重要であり、システムパフォーマンスが単に制約された語彙及び言語モデリング技術に依拠していなかったことを実証している。

20

【0380】

神経分類器のみが試行文をどの程度良好に復号することができるかを評価するために、神経分類器のみを使用して、神経活動の各個々の2.5秒のウィンドウについて最も可能性の高い文字から構成された文字系列を、対応する標的文字系列と比較した。この比較中、全ての空白文字は無視された(リアルタイム復号中、これらの文字はビーム探索によって自動的に挿入された)。これは、35.1%(99%CI[30.6, 38.5])の中央値CERをもたらし、これは、偶然($z = 7.09$ 、 $P = 8.08 \times 10^{-12}$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)よりも有意に低く、沈黙コードワード生成試行中の神経活動のタイムウィンドウが弁別可能であったことを示している。これは、64.9%の分類器正確度率に対応する。この条件の中央値WERは100%(99%CI[100.0, 100.0])であり、言語モデリング又は空白文字の自動挿入がなければ、予測文字系列が対応する標的文字系列と一致することはなかった。

30

【0381】

ビーム探索によってどれだけ復号が改善されたかを測定するために、言語モデリングを一切組み込むことなく、神経分類器の予測をビーム探索に渡し、語彙内の単語のみから構成されるように文字系列を制約した。これは、各タイムステップで最も可能性の高い文字のみを使用するよりも、CER及びWERを有意に改善した(それぞれ $z = 4.51$ 、 $P = 6.37 \times 10^{-6}$ 及び $z = 6.61$ 、 $P = 1.19 \times 10^{-10}$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。単語系列の可能性を組み込んだ言語モデリングを使用しない結果として、システムは、「Do not do that again(二度とそれをしないでください)」の代わりに「Do not do that again」などの無意味な文を予測することがある(図16F)。したがって、完全なリアルタイムスプリングパイプラインを完成させるために言語モデリングを含むことによって、一切の言語モデリングなしでシステムを使用するよりも、中央値CERを有意に6.13%、中央値WERを10.53%に改善し(それぞれ $z = 5.53$ 、 $P = 6.34 \times 10^{-8}$ 及び $z = 6.11$ 、 $P = 2.01 \times 10^{-9}$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)、復号中に英語の自然な構造を組み込むことの利点を示している。

40

【0382】

50

高ガンマ活性及び低周波信号における弁別的な内容

脳活動から発話を復号するための以前の取り組みは、典型的には、復号中に高ガンマ周波数範囲（70～170 Hz、しかし正確な境界は異なる）の内容に依拠している^{12, 13, 24}。しかしながら、最近の研究は、低周波数成分（0～40 Hz）もまた、発話及び想像上の発話の復号に使用され得ることを実証しており^{14, 15, 25-27}、ただし、各周波数範囲に含まれる弁別的情報の差は、あまり理解されていないままである。

【0383】

脳活動から発話を復号するための以前の取り組みは、典型的には高ガンマ活性（HGA）のみを使用した^{12, 13, 15}、本発明のスペリングシステムはまた、復号中に低周波信号（LFS）も使用した。分類器への入力は、分類前に（アンチエイリアシングフィルタを用いて）33.33 Hzにダウンサンプリングされたため、分類中に使用されたLFSは、0.3～16.67 Hzの信号成分のみを含んでいた。孤立単語タスクの最新の9, 132回の試験（これらの試験の各々で、参加者はコードワードを黙って発話することを試行した）を使用して、HGAのみ、LFSのみ、及び両方の特徴タイプを使用して10フォールド交差検証モデルを訓練した。LFSのみを使用するモデルは、HGAのみを使用するモデルよりも高いコードワード分類正確度を実証し、両方の特徴タイプ（HGA + LFS）を使用するモデルのパフォーマンスは、他の2つのモデル（全ての比較について $P < 0.001$ 、3方向ホルム・ボンフェローニ補正を伴う両側マン・ホイットニーのU検定、図17A、図24）を上回り、56.4%の中央値分類正確度を達成した（図25）。

【0384】

次いで、HGA、LFS、及びHGA + LFSを使用して訓練された神経分類モデルに対する各電極及び特徴タイプの相対的寄与を調査した。各モデルについて、まず、電極の値へのわずかな変化がモデルの予測に及ぼした影響を測定することによって、分類への各電極の寄与を計算した²⁸。HGAモデルの電極寄与は、主に、腹側感覚運動皮質（vSMC）、弁蓋部、及び三角部に対応するグリッドの腹側部分に局在化された（図17B）。LFSモデルの寄与は、はるかにより広範であり、中心前回及び中心後回におけるvSMCの背側側面に対応するグリッドのより多くの背側及び後方部分をカバーした（図17D）。HGAモデル及びLFSモデルからの寄与は、0.501のスピアマン順位相関と適度に相関した（特徴タイプ当たりの $n = 128$ の電極寄与、 $P < 0.01$ ）。HGA + LFSモデルにおけるHGA及びLFSからの別個の寄与は、それぞれ、HGAのみ及びLFSのみのモデルの寄与と高度に相関していた（特徴タイプ当たり $n = 128$ の電極寄与、（それぞれ0.922及び < 0.963 の両方のスピアマン順位相関について $P < 0.01$ 、図17C、図17E）。これらの知見は、復号中に最も有用であった2つの特徴タイプに含まれる情報が冗長ではなく、比較的明確に区別できる皮質領域から記録されたことを示している。

【0385】

HGA及びLFS特徴を更に特徴付けるために、LFSが特徴又は時間次元を増大させたか否かを調査した。これは、復号正確度の増大に寄与する可能性がある。まず、HGA、LFS、HGA + LFS特徴セットの特徴次元に対して主成分分析（PCA）を実施した。得られた主成分（PC）は、HGA及びLFS特徴セットについて（電極チャンネルにわたる）空間変動性を捕捉し、HGA + LFS特徴セットについて（それぞれ、電極チャンネル及び特徴タイプにわたる）空間及びスペクトル変動性を捕捉した。次いで、分散の80%超を説明するために必要な主成分（PC）の最小数を算出した。分散の80%超を説明するために、LFSは、HGAよりも有意に多くの特徴PCを必要とした（ $z = 12.2$ 、 $P = 7.57 \times 10^{-34}$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、図17F）。結合HGA + LFS特徴セットは、個々のHGA又はLFS特徴よりも有意に多くの特徴PCを必要とし（それぞれ $P = 6.20 \times 10^{-38}$ 及び $P = 1.60 \times 10^{-33}$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定、図17F）、LFSが単純に各電極でHGAを複製したのではなく、代

わりに固有の特徴分散を追加したことを示唆している。

【0386】

特徴の時間的内容を評価するために、まず、同様のPCAアプローチを使用して時間次元を測定した。LFS特徴が、HGA及びHGA+LFS特徴セットの両方よりも有意に多くの時間的PCを必要とすることを観察した（それぞれ、 $P = 2.72 \times 10^{-39}$ 及び $P = 1.37 \times 10^{-38}$ 、図17G、3方向ホルム・ボンフェローニ補正を伴う両側マン・ホイットニーのU検定）。LFS特徴が、80%超の分散を説明するために、HGA及びHGA+LFS特徴セットの両方よりも有意に多くの時間的PCを必要とすることを観察した（それぞれ、 $z = 12.2$ 、 $P = 7.57 \times 10^{-34}$ 及び $z = 12.2$ 、 $P = 7.57 \times 10^{-34}$ 、図17G、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定）。各特徴タイプの固有の時間次元は、HGA+LFS特徴セット内で同じままであったため、HGA+LFS特徴のこの大きい分散を説明するために必要な時間PCの数は、個々の特徴タイプの対応する数の間にあった。次いで、各特徴タイプの時間分解能が復号パフォーマンスにどのように影響するかを評価するために、幅が異なるガウスフィルタを用いて各特徴時系列を時間的に平滑化した。より広いガウスフィルタは、より多くの時間平滑化を引き起こし、実効的に信号を時間的にぼかし、したがって時間分解能を低下させる。

10

【0387】

LFS特徴の時間的な平滑化は、HGA又はHGA+LFS特徴を平滑化するよりも有意に分類正確度を低下させた（ウィルコクソン符号順位統計量 $= 737.0$ 、 $P = 4.57 \times 10^{-5}$ 及び統計量 $= 391.0$ 、 $P = 1.13 \times 10^{-8}$ 、3方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン符号順位検定、図17H）。（ウィルコクソン符号順位統計量 $= 1460.0$ 、 $P = 0.443$ ）。これは、時間的PCA比較の結果とほぼ一致する。まとめると、これらの結果は、LFSの時間的内容が、HGAよりも高い変動性を有し、より多くの発話関連弁別情報を含んでいたことを示す。

20

【0388】

NATOコードワードと文字との間の神経的弁別性の差

本発明のシステムの制御中、参加者は、単純に文字自体を発語するのではなく、各文字を表すためにNATOコードワード（「a」の代わりに「alpha」、「b」の代わりに「beta」など）を黙って発語することを試行した。コードワードを生成する試行に関連する神経活動は、音声変動性の増大及び発言長の延長に起因して、文字よりも弁別可能であると仮定した。これをテストするために、まず、参加者が26個の英語文字の各々を、それらを表すNATOコードワードの代わりに発語することを試行した孤立標的タスクの修正バージョンを使用してデータを収集した。その後、10フォールド交差検証分析において各コードワード及び各文字を黙って発語する最新の29回の試行からのHGA+LFS特徴を使用して分類モデルを訓練し、テストした。実際、コードワードは、文字よりも有意に高い正確度で分類された（ $z = 3.78$ 、 $P = 1.57 \times 10^{-4}$ 、両側ウィルコクソン順位和検定、図18A）。

30

【0389】

各タイプの発言（コードワード又は文字のいずれか）の神経的弁別可能性の間のモデルに依存しない比較を実施するために、HGA+LFS特徴セットを使用して各発言の最近クラス距離を計算した。ここで、各発言は単一のクラスを表し、距離は同じタイプの発言間でのみ計算された。コードワード又は文字についてのより大きい最近クラス距離は、その発言を生成する沈黙試行と関連付けられる神経活性化パターンが、それぞれ、他のコードワード又は文字からより明確に区別されるため、その発言が、神経特徴空間においてより弁別可能であることを示す。コードワードの最近クラス距離は、文字よりも全体的に有意に高かった（ $z = 2.98$ 、 $P = 2.85 \times 10^{-3}$ 、両側ウィルコクソン順位和検定、図18B）が、文字の代わりにコードワードを使用する場合に全ての文字がより大きい最近クラス距離を有するわけではなかった（図18C）。

40

【0390】

50

沈黙発話試行と公然発話試行との間の誘発神経活動の区別

スプリングシステムは、沈黙発話試行によって制御されており、同じ参加者が同様の発話復号システムを制御するために公然発話試行（声に出して発話する試行）を使用した本発明者らの以前の研究とは異なる¹⁶。2つのタイプの発話試行間の神経活動及び復号パフォーマンスの差を評価するために、参加者がコードワードを声に出して（黙ってではなく公然と）発話することを試行するように指示された孤立標的タスクのバージョンを収集した。公然発話試行と沈黙発話試行との間の差を可視化するために、異なるコードワード及び電極の誘発HGAを比較した。2つのタイプの発話試行についての誘発神経活動の空間パターンは、類似性を呈しており、2つの電極についての誘発HGAの検査は、いくつかの神経集団が各発話タイプについて同様に応答する一方で、他の集団は同様には応答しないことを示唆する（図19B、図19C、図26）。沈黙発話試行と公然発話試行との間の識別的な神経コンテンツを比較するために、発話試行と関連付けられるHGA+LFS特徴を使用して10フォールド交差検証分類分析を実施した（図19D）。まず、各発話タイプ（沈黙又は公然）について、その発話タイプで収集されたデータを使用して分類モデルを訓練した。分類モデルが各発話タイプと関連付けられる神経表現の類似性を活用してパフォーマンスを向上させることができるか否かを決定するために、また、1つの発話タイプに関して事前に訓練し、次いで、もう1つの発話タイプを微調整することによってモデルを作成した。次いで、各発話タイプと関連付けられる保持データに対して各分類モデルをテストし、結果のペアの28の組み合わせ全てを比較した。沈黙データのみに対して訓練されたが、公然データに対して試験されたモデル、及びその逆のモデルは、偶然を上回る分類正確度をもたらした（それぞれ、36.3%、99%CI[35.0, 37.5]及び33.5%、99%CI[31.0, 35.0]の中央値正確度、偶然正確度は3.85%である）。しかしながら、両方の発話タイプについて、同じタイプに対する訓練及びテストは、有意により高いパフォーマンスをもたらした（ $P < 0.01$ 、両側ウィルコクソン順位和検定、28方向ホルム・ボンフェローニ補正）。他の発話タイプを使用した事前訓練モデルは、分類正確度の増大をもたらしたが、その増大はより控えめであり、公然発話タイプでは有意ではなかった（公然について中央値正確度は2.33%増大し、 $z = 2.65$ 、 $P = 0.033$ 、沈黙について中央値正確度は10.4%増大し、 $z = 3.78$ 、 $P = 4.40 \times 10^{-3}$ 、両側ウィルコクソン順位和検定、28方向ホルム・ボンフェローニ補正）。まとめると、これらの結果は、沈黙発話試行及び公然発話試行中に誘発された神経活性化パターンがいくつかの類似点を共有していたが、同一ではなかったことを示唆している。

10

20

30

【0391】

より大きい語彙及び代替タスクへの一般化可能性

1,152語の語彙は、多種多様な一般的な文のコミュニケーションを可能にしたが、本発明のアプローチがより大きい語彙サイズにどれだけ良好にスケーリングすることができるかも評価した。具体的には、3,303語、5,249語、9,170語のサイズを有する、大規模英語コーパスにおけるそれらの単語の頻度に基づいて選択された3つの大きい語彙を使用した、コピータイピングスプリング結果をシミュレートした。各語彙について、ビーム探索中に使用された言語モデルを再訓練して、新しい単語を組み込んだ。文を最終化するとき使用される大規模言語モデルは、任意の英語テキストに一般化するように設計されているため、これらの分析のために変更されなかった。

40

【0392】

新たな語彙の各々によって高いパフォーマンスが維持されており、3,303語、5,249語、及び9,170語の語彙について、中央値文字誤り率（CER）はそれぞれ7.18%（99%CI[2.25, 11.6]）、7.93%（99%CI[1.75, 12.1]）、及び8.23%（99%CI[2.25, 13.5]）であった（図20A、中央値リアルタイムCERは、1,152語を含む元の語彙では6.13%（99%CI[2.25, 11.6]）であった）。中央値単語誤り率（WER）は、それぞれ12.4%（99%CI[8.01, 22.7]）、11.1%（99%CI[8.01,

50

23.1])、及び13.3% (99%CI [7.69, 28.3])であった(図20B、元の語彙について、WERは10.53% (99%CI [5.76, 24.8])であった)。全体的に、任意の2つの語彙によるCER又はWERの間に有意差は見られなかった(全ての比較で $P > 0.01$ 、6方向ホルム・ボンフェローニ補正を伴う両側ウィルコクソン順位和検定)。これは、流暢なコミュニケーションを可能にするより大きい語彙サイズへの本発明のスペリングアプローチの一般化可能性を示している。

【0393】

最後に、コピータイピングタスク構造を超えた行動文脈に対するスペリングアプローチの一般化可能性を評価するために、参加者が会話型タスク条件にエンゲージするときのパフォーマンスを測定した。この条件の各試験において、参加者に質問(画面上のテキストとしての)が提示されたか、又はいかなる刺激も提示されなかった。次いで、参加者は、提示された質問に対する自発的に選択された回答、又は、刺激が提示されなかった場合は任意の文をスペルアウトすることを試行した。復号された各文の正確性を測定するために、その文が、自身が意図した文と正確に一致したか否かを示すために、参加者に頷くように頼んだ。文が完璧に復号されなかった場合、参加者は市販のコミュニケーション支援デバイスを使用して、自身の意図したメッセージをスペルアウトした。このリアルタイム会話型タスク条件の28回の試験にわたって、中央値CERは14.8% (99%CI [0.00, 29.7])であり、中央値WERは16.7% (99%CI [0.00, 44.4])であった(図20C、図20D)。コピータイピングタスクと比較して、復号誤り率がわずかに増大していることが観察された。これは、参加者が言語モデルではうまく表現されない不完全な文(「going out (外出)」及び「summer time (夏時間)」など)を使用して応答したことに起因する可能性がある。それにもかかわらず、これらの結果は、本発明のスペリングアプローチが、ユーザが質問への応答、及びプロンプトなしの、自発的に選択されたメッセージを生成することを可能にすることを実証している。

【0394】

考察

ここで、構語障害を患う麻痺者が、黙って発話する試行を使用して、意図されたメッセージをリアルタイムでスペルアウトするために神経プロテーゼを制御することができることを実証した。個々の文字を表す音声的に豊富なコードワード、及び文末コマンドを示すための試行手運動によって、本発明者らは深層学習及び言語モデリング技法を使用して、皮質脳波記録(ECoG)信号から文を復号した。これらの結果は、完全な沈黙制御を可能にし、非発話運動コマンドを含む高周波及び低周波ECoG特徴の両方を活用して文を最終化し、スペリングを通じて大規模語彙文の復号を容易にし、デバイスの埋め込みから128週間を超えて関連する皮質活動の継続的な安定性を実証することによって、同じ参加者での以前の単語復号所見を大幅に拡大する²⁰。

【0395】

スペリング脳コンピュータインターフェース(BCI)の以前の実施態様は、ユーザが、画面上の文字に視覚的に注意を向けることによって^{29, 30}、又は2次元コンピュータカーソルを制御するために運動想像を使用することによって^{4, 5}、又は文字を手書きしようと試行することによって⁶、意図されたメッセージをタイプすることができることを実証している。運動皮質内の侵入微小電極アレイを使用したBCIパフォーマンスは、過去20年間で着実に改善されており³¹⁻³³、最近では、この参加者は正常に発話することができたが、1人の参加者で1分当たり90文字までのスペリング速度を達成した⁶。本発明の結果は、スペリングBCIアプリケーションのための即座に実用的で臨床的に実行可能な制御モダリティのリストを拡張して、埋め込みECoGアレイを使用した沈黙試行発話を含むようにし、これは、発話の相対的な自然さに起因して一部の患者による日常使用に好ましい場合があり、より広い皮質カバレッジを有する侵襲性の低い非侵入性電極アレイの使用を通じて、患者にわたってより長期的に口バストであり得る⁷。

【0396】

事後分析において、より多くの言語情報がスペリングパイプラインに組み込まれるにつれて、復号パフォーマンスが向上したことを示した。この情報は、1, 152語の語彙でリアルタイム復号を容易にするのを助け、可能な出力としての多種多様な一般的及び臨床的に関連する文を可能にした。更に、オフラインシミュレーションを通じて、基本的な流暢さのための推定語彙サイズ閾値を超え、一般的なコミュニケーションを可能にする9,000超の一般的な英語単語を含む語彙で、このスペリングアプローチを検証した^{34, 35}。これらの結果は、言語モデリングが神経ベースの発話復号を大幅に改善することができるという一貫した知見に加わり^{12, 15, 20}、汎用コミュニケーション支援システムのための発話ベースのスペリングアプローチの即時実行可能性を実証している。

【0397】

本研究では、構語障害のヒトによる沈黙発話試行中に記録された神経信号を、発話神経プロテゼを駆動するために効果的に使用することができることを示した。これらの信号が、公然発話試行中に記録された信号と同様の発話運動表現を含んでいたという仮説を裏付けて、公然発話試行を分類するためだけに訓練されたモデルが、沈黙発話試行の偶然を超える分類を達成することができ、逆もまた同様であることを示した。追加的に、分類パフォーマンスに最も寄与する電極の空間的位置付けは、公然発話及び沈黙発話の両方で同様であり、これらの電極の多くは、構音発話運動処理に大きく関与する脳領域である腹側感覚運動皮質に位置していた^{8-10, 36}。

【0398】

全体的に、これらの結果は、想像上の発話に対する効果的な代替行動戦略としての沈黙試行発話を更に検証し、同じ参加者による公然発話試行の復号を伴う本発明者らの以前の研究からの知見を拡大し²⁰、発話試行中の残留発声の生成が発話神経プロテゼを制御するために必須ではないことを示している²⁰。これらの知見は、完全な声帯麻痺（ロッキン症候群など）を患う個人ののための試行発話制御の実行可能性を示しているが、これらの個人を対象とした将来の研究は、公然発話試行、沈黙発話試行、及び純粋に想像上の発話の間の神経的差異、並びに、特定の医学的状態がこれらの差異にどのように影響するかについての理解を深めるために必要である。記録方法、タスク設計、及びモデリング技術を含む、ここで説明するアプローチは、発話関連の神経科学的調査と、発話運動皮質がまだ損なわれておらず、発話を試行する精神的能力があると仮定した、声帯麻痺の重症度にかかわらず患者によるBCI開発の両方に適切であると期待される。

【0399】

外側発話運動皮質脳領域にわたる空間カバレッジを可能にすることに加えて、埋め込みECOGアレイはまた、実行又は試行される手運動の間に典型的に関与する手運動（「手のノブ」）皮質領域内の神経集団への同時アクセスを提供した³⁷。本発明のアプローチは、BCIを制御するために2つの皮質領域を組み合わせる最初のものである。これにより、最終的に本発明の参加者は手運動を試行することができた。これは信頼性の高い検出が可能であり、98.43%の分類正確度（99%CI [95.31, 99.22]）で沈黙発話試行から高度に弁別可能であり、任意の特定の文のスペリングがいつ終了したかを示すことができた。これは、事前に指定された時間間隔が経過した後、又は外部的に文が完成したときに⁵、若しくは文⁶を終端させるために頭部の動きを必要としたときに、文のスペリングを終了する以前のスペリングBCI実施態様と比較して、好ましい停止機構であり得る。また、沈黙発話試行がスペリングを開始することを可能にすることにより、システムは、参加者によって自発的に参加され、ディスエンゲージされ得、これは、実用的なコミュニケーションBCIのための重要な設計上の特徴である。試行手運動は、マルチモーダルコミュニケーションBCIのこの最初のデモンストレーションでは単一の目的のためにのみ使用されたが、同じ参加者による別個の研究は、非発話運動の想像を使用していくつかの明確に異なるコマンドを示すことができることを示唆している³⁸。

【0400】

将来のコミュニケーション神経プロテゼでは、限られた、頻繁に使用される語彙からの完全な単語又は句の迅速な復号²⁰、及び語彙外の項目のより低速の、一般化可能なス

10

20

30

40

50

ペリングを可能にする組み合わせアプローチを使用することが可能であり得る。以前の発話復号研究¹³で検証されたように、複数のタスク及び語彙にわたって集約されたデータを使用して、異なる目的の発話モデルを交差訓練するために、転移学習方法が使用され得る。埋め込み経皮コネクタに関する臨床及び規制ガイドラインが、参加者が現在のスペリングシステムを独立して使用することを可能にするのを妨げたが、完全に埋め込み可能なECoGアレイ及び復号パイプラインをオペレーティングシステムのアクセシビリティ機能と統合するソフトウェアアプリケーションの開発は、自律的な使用を可能にする可能性がある。深層学習技法、言語モデリング、並びにECoG記録によって提供される信号の安定性及び空間カバレッジによって促進されて、将来のコミュニケーション神経プロテーゼは、重度の麻痺及び構語障害を患うユーザが、意図されたメッセージを生成する自然的な沈黙発話試行と、高レベルの対話的コマンドを発行する非発話運動を使用して、支援技術及び個人デバイスを制御することを可能にする可能性がある。

10

【0401】

方法

臨床試験概要

本研究は、BCI Restoration of Arm and Voice (BRAVO) 臨床試験 (ClinicalTrials.gov 登録番号 NCT03698149) の一環として実施した。この単一施設臨床試験の目標は、ECoG 及びカスタム復号方法が、支援ニューロテクノロジーがコミュニケーション及び移動性を回復させることを可能にすることができるかを決定することである。米国食品医薬品局は、本研究で使用された神経インプラントについて治験用機器の適用免除を承認した。研究プロトコルは、カリフォルニア大学サンフランシスコ校のヒト研究に関する審議会によって承認された。データ安全監視委員会が、試験の完了前に原稿の結果を公開することに合意した。参加者は、神経インプラント、実験プロトコル、及び医療上のリスクに関する詳細が徹底的に説明された後、本研究に参加するためのインフォームドコンセントを与えた。

20

【0402】

参加者

研究開始時に36歳であった参加者は、広範囲のポンティアヌ脳卒中を経験した後、神経科医及び言語聴覚士によって重度の痙性四肢麻痺及び構語障害と診断された。参加者は完全に認知的に正常である。彼はうなり声及びうめき声を発声することはできるまでであるが、理解可能な発話を生成することはできず、(自己報告の説明による)彼の状態に起因して、声に出して発話する彼の試行は、異常な努力を要する。彼は通常、コミュニケーションするために頭部の残留運動によって彼が制御する補助的なコンピュータベースのインターフェースに依拠する。この参加者は、この臨床試験の一部として以前の研究に参加しているが^{16, 20}、これらの研究からの神経データは本研究では使用されていない。

30

【0403】

神経インプラント

神経インプラントデバイスは、高密度皮質脳波記録 (ECoG) アレイ (PMT) 及び経皮コネクタ (Blackrock Microsystems) から構成されていた。ECoGアレイは、4mmの中心間隔を有する格子構成に配置構成された128個のディスク形状の電極を含んでいた。このアレイは、下前頭回の背側後部側面、中前頭回の後部側面、中心前回、及び中心後回の前方側面を含む、発話生成と関連付けられる皮質領域にわたって、脳の左半球の軟膜表面に外科的に埋め込まれた^{8, 10, 32}。経皮コネクタを頭蓋骨に埋め込んで、ECoGアレイから取り外し可能デジタルヘッドステージ及びケーブル (Neuroplex E; Blackrock Microsystems) に電気信号を伝導し、取得した脳活動を最小限に処理及びデジタル化し、データをコンピュータに送信した。このデバイスは、2019年2月に一切の外科的合併症なしで埋め込まれた。デバイス及び外科的処置のより多くの詳細は、同じデバイス及び参加者を用いた本発明者らの以前の研究に見出され得る¹⁶。

40

【0404】

50

データ取得及び前処理

いくつかのハードウェアコンポーネント及び処理ステップを含むパイプラインを使用して、埋め込みECOGアレイから神経特徴を取得した（図22を参照されたい）。ヘッドステージ（取り外し可能デジタルコネクタ、NeuroPlex E、Blackrock Microsystems）を経皮台座コネクタに接続した。経皮台座コネクタは、ECOGアレイからの神経信号をデジタル化し、HDMI接続を介してデジタルハブ（Blackrock Microsystems）に送信した。次いで、デジタルハブは、光ファイバケーブルを介してNeuroportシステム（Blackrock Microsystems）にデジタル化信号を送信した。Neuroportシステムは、ノイズキャンセル及びアンチエイリアシングフィルタを信号に適用した後、1kHzでイーサネット接続を通じて別個のリアルタイムコンピュータ（Colfax International）に信号をストリーミングした。

10

【0405】

リアルタイム処理コンピュータ上で、カスタムPythonソフトウェアパッケージ（rtNSR）を使用して、ECOG信号を処理及び分析し、リアルタイムタスクを実行し、リアルタイム復号を実行し、データ及びタスクメタデータを格納した^{16, 33, 34}。このソフトウェアパッケージを使用して、まず、ECOGデータの各時間サンプルに（全ての電極チャンネルにわたって）共通平均基準を適用した。共通平均参照は、共有ノイズを低減するために、一般的にマルチチャンネルデータセットに適用される^{35, 36}。次いで、これらの再参照された信号が2つの並列処理ストリームで処理されて、Parksclellanアルゴリズム³⁷を使用して設計されたデジタル有限インパルス応答（FIR）フィルタを使用して、高ガンマ活性（HGA）及び低周波数信号（LFS）特徴が抽出された（図22参照）。簡潔に言えば、これらのFIRフィルタを使用して、高ガンマ周波数帯域（70～150Hz）の信号の分析振幅及び信号のアンチエイリアスバージョン（100Hzのカットオフ周波数を有する）を計算した。高ガンマ分析振幅及びダウンサンプリング信号からの時間同期値を200Hzの単一の特徴ストリームに結合した。次に、30秒のスライドウィンドウを使用して各チャンネル及び各特徴タイプの値をzスコアリングして、実行中の統計情報を計算した。最後に、アーチファクト除去アプローチを実装した。このアプローチは、zスコアの大きさが10を超える少なくとも32の特徴を含む神経時点を識別し、これらの時点のそれぞれを前の時点からのzスコア値に置き換え、実行中のzスコア統計を更新するときにこれらの時点を見捨てる。リアルタイム復号中及びオフライン分析において、HGA特徴としてzスコア高ガンマ分析振幅を使用し、LFS特徴としてのzスコアダウンサンプリング信号（並びにHGA+LFS特徴セットとしての2つの組み合わせ）を使用した。神経分類器は、（16.67Hzのカットオフ周波数を有するアンチエイリアスフィルタを使用して）推論のためにそれらを使用する前に、これらの特徴ストリームを更に6分の1にダウンサンプリングしたが、発話検出器はそうしなかった。

20

30

【0406】

参加者の住居の近くの小規模事務室で、全てのデータ収集及びリアルタイム復号タスクを実施した。本発明者らの研究室のサーバインフラストラクチャにデータをアップロードし、このインフラストラクチャでホストされているNVIDIA V100 GPUを使用して復号モデルを訓練した。記録ハードウェア、参加者とのタスク設定手順、及び臨床試験プロトコルに関する追加情報は、本発明者らの以前の研究¹⁶で提供されている。

40

【0407】

タスク設計

孤立標的タスク及び文スペリングタスクの2つの一般的なタイプのタスク中に参加者による神経データを記録した（図21）。孤立標的タスクの各試験では、テキスト標的が両側の4つの点とともに画面に表示された。両側の点、点がなくなるまで1つずつ消え、その時点でテキスト標的が緑色に変わり、ゴーキューを表す。このゴーキューにおいて、参加者は、標的がNATOコードワード又は英語文字のいずれかであるか否かを標的に（

50

現在のタスク指示に応じて、黙って又は声に出して）発語することを試行した。標的が「R i g h t」という単語を含むテキスト文字列及び右を指す矢印である場合、参加者は代わりに右手を握ることを試行した。孤立標的タスク中に収集された神経データを使用して、検出及び分類モデルを訓練及び最適化し、分類器パフォーマンスを評価した（方法 S 1 を参照）。

【0408】

文スペリングタスクは、結果セクションの冒頭及び図 1 5 に説明されている。簡潔に言えば、参加者は、全スペリングパイプライン（以下のサブセクションで説明されている）を使用して、コピータイピングタスク条件で標的として提示された文をスペリングするか、又は、会話型タスク条件で任意の文をスペリングした。参加者が予測された文を遡及的に変更することを可能にする機能は実装しなかったが、言語モデルは、追加の文字予測を受信した後、文中の以前に予測された単語を変更することができる。文スペリングタスク中に収集されたデータを使用して、ビーム探索ハイパーパラメータを最適化し、全スペリングパイプラインを評価した。

【0409】

モデリング

参加者がコードワード及び手動コマンドの生成を試行したときに、孤立標的タスク中に収集されたデータを使用して検出及び分類モデルを適合させた。これらのモデルをオフラインで適合させた後、訓練されたモデルをリアルタイムテスト中に使用するためにリアルタイムコンピュータに保存した。これらの 2 つのモデルに加えて、文スペリングを可能にするために言語モデルも使用した。モデルハイパーパラメータの値を選択するために、保持検証データセットに対してハイパーパラメータ最適化手順を使用した（表 S 2 を参照されたい）。

【0410】

発話検出

参加者がいつスペリングシステムへのエンゲージを試行していたかを判断するために、リアルタイム沈黙発話検出モデルを開発した。以前の実施態様と同様に、このモデルは、再帰ニューラルネットワーク層の一種である長・短期記憶層を使用して、リアルタイムで神経活動を処理し、黙って発話する試行を検出した¹⁶。このモデルは、L F S 及び H G A の両方の特徴（合計 2 5 6 個の個々の特徴）を 2 0 0 H z で使用した。

【0411】

発話検出モデルは、教師あり学習及び打ち切り型通時的逆伝播を使用して訓練された。訓練のために、神経データの各時点、その時点でのタスクの現在の状態に応じて、「休止」、「発話準備」、「運動」、及び「発話」の 4 つのクラスのうち 1 つとしてラベル付けした。発話確率のみが、スペリングシステムにエンゲージするためにリアルタイム評価中に使用されたが、他のラベルは、検出モデルが発話する試行を他の行動から明確に区別するのを助けるために訓練中に含まれた。発話検出モデルについての更なる詳細については、方法 S 2 及び図 2 3 を参照されたい。

【0412】

分類

孤立標的試験又は 2 . 5 秒の文字復号サイクル i と関連付けられる神経活動 x_i のタイムウィンドウからの試行されたコードワード又は手動コマンド y_i を分類するように、人工ニューラルネットワーク (ANN) を訓練した。訓練手順は、最大尤度推定の形式であり、 θ によってパラメータ化され、神経活動 x_i に対して条件付けられた ANN 分類器を所与として、モデル適合中の本発明者らの目標は、訓練ラベルの確率を最大化するパラメータ θ^* を見つけることであった。これは、次の最適化問題として書くことができる。

【数 2 5】

$$\theta^* = \operatorname{argmax}_{\theta} \prod_i p_{\theta}(y_i|x_i) = \operatorname{argmax}_{\theta} \sum_i \log p_{\theta}(y_i|x_i) = \operatorname{argmin}_{\theta} - \sum_i \log p_{\theta}(y_i|x_i)$$

10

20

30

40

50

【0413】

確率的勾配降下及びAdamオプティマイザを使用して、最適なパラメータ θ^* を近似した³⁸。

【0414】

神経時系列データの時間的動態をモデリングするために、入力層、及び後続する2層の双方向ゲート付き再帰ユニット（GRU）³⁹の、合計3層に対して1次元時間量み込みを伴うANNを使用した。最後のGRU層の最終出力に出力行列を乗算した後、softmax関数を適用して、 x_i を所与として

【数26】

y_i

10

27個のラベルの各々の推定確率を得た。神経分類器を適合させるために使用されるデータ増強、ハイパーパラメータ最適化、及び訓練手順に関する更なる詳細については、方法S3を参照されたい。

【0415】

文スペリングのための分類器アンサンブル：文スペリング中に、ランダムパラメータ初期化によって引き起こされる過剰適合及び望ましくないモデリング分散を低減することによって分類パフォーマンスを改善するためにモデルアンサンブルを使用した⁴⁰。具体的には、同じ訓練データセット及びモデルアーキテクチャを使用して、ただし異なるランダムパラメータ初期化を用いて、10個の別個の分類モデルを訓練した。次いで、神経活動 x_i の各タイムウィンドウについて、これらの10個の異なるモデルからの予測をともに平均化して、最終予測を生成した

【数27】

y_i

20

。

【0416】

文スペリングのための増分的分類器再校正

文スペリングパフォーマンスを改善するために、（孤立標的タスクからのデータに加えて）前のセッションからの文のスペリングタスク中に記録されたデータに関して、文スペリング中に使用される分類器を訓練した。これらの分類器を訓練する際に高品質の文スペリングデータのみを含める取り組みにおいて、文字誤り率が0で復号された文からのデータのみを使用した。

【0417】

ビーム探索

文スペリング中、本発明者らの目標は、神経データ X を所与として最も可能性の高い文テキスト s^* を計算することであった。Hannun他¹⁹の定式化を使用して、神経データからのその尤度及び調整済み言語モデル事前確率の下での尤度を所与として s^* を求めた。これにより、単語系列確率を神経分類器からの予測に組み込むことができた。これは、式的に次のように表現することができる。

$$s^* = \arg \max_s p_{nc}(s | X) p_{lm}(s)^a | s |^b$$

【0418】

ここで、 $p_{nc}(s | X)$ は、神経活動の各ウィンドウを所与とした神経分類器の下での s の確率であり、これは、神経活動 x_i の各ウィンドウについて神経分類器によって与えられる s 内の各文字の確率の積に等しい。 p_{lm} は、言語モデル事前確率の下での文 s の確率である。ここでは、Nグラム言語モデルを使用して p_{lm} を近似した。 $N=3$ である本発明者らのNグラム言語モデルは、文中の先行する2つの単語を所与とした各単語の確率を提供する。このとき、文の言語モデルの下での確率は、それに先行する2つの

30

40

50

単語を所与とした各単語の確率の積として受け取られる（方法 S 5 を参照）。

【0419】

Hannun 他¹⁹のように、N グラム言語モデル事前確率が強すぎると仮定し、ハイパーパラメータ α を用いてそれを下方修正した。本発明者らはまた、言語モデルに、より多くの単語を含む文を選好するよう促すために単語挿入ボーナス β を含め、その下にある文の確率 $p_{lm}(s)$ を、 s 内の単語の数が増大するにつれて減少させる言語モデルの暗黙の結果を相殺した。 $|s|$ は、 s の濃度を表し、 s 内の単語数に等しい。文 s が部分的に完成していた場合、 $p_{lm}(s)$ 及び $|s|$ を計算するとき、 s 内の最後の空白文字の前の単語のみが考慮された。

【0420】

次いで、Hannun 他¹⁹のような反復ビーム探索アルゴリズムを使用して、各時点 $t = \tau$ で s^* を近似した。 $t = \tau - 1$ からの B 個の最も可能性の高い文のリスト（又は、 $t = 1$ の場合、単一の空文字列要素を含むリスト）を候補プレフィックスのセットとして使用し、ここで、 B はビーム幅である。次いで、候補プレフィックス l 及び $p_{nc}(c | x_\tau) > 0.001$ を有する各英語文字 c について、 l 及び後続する c を考慮して新しい候補文を構築した。追加的に、各候補プレフィックス l 、並びに、 $p_{nc}(c^+ | x_\tau) > 0.001$ である、英語文字及び後続する空白文字から構成される各テキスト文字列 c^+ について、 l 及び後続する c^+ を考慮することによって、より新しい候補文を構築した。ここでは、ビーム探索全体を通して、各 c 及び対応する c^+ について $p_{nc}(c^+ | x_\tau) = p_{nc}(c | x_\tau)$ を考慮した。次に、制約された語彙を所与として、有効ではなかった単語又は部分的に完成した単語を含む任意の結果の候補文を破棄した。次いで、各残りの候補文

【数28】

$\tilde{l}$

を

【数29】

$$p(\tilde{l}) = p_{nc}(\tilde{l} | X_{1:\tau}) p_{lm}(\tilde{l})^\alpha |\tilde{l}|^\beta$$

によって最スコアリングした。次いで、最も可能性の高い候補文 s^* をフィードバックとして参加者に表示した。

【0421】

ハイパーパラメータ最適化を使用して、 α 、 β 、及び B の値を選択した（更なる詳細については、方法 S 4 を参照）。

【0422】

任意の時点 t において、試行手運動コマンド（文最終化コマンド）の確率が 80 % を超える場合、ビーム探索の前の反復からの B 個の最も可能性の高い文を処理して、不完全な単語又は語彙外の単語を有する任意の文を削除した。次いで、各残りの文の確率

【数30】

$\hat{l}$

が、次のように再計算された。

【数31】

$$p(\hat{l}) = p_{nc}(\hat{l}|X_{1:t-1})p_{lm}(\hat{l})^{\alpha}|\hat{l}|^{\beta}p_{gpt2}(\hat{l})^{\alpha_{gpt2}}$$

【0423】

ここで、

【数32】

$$p_{gpt2}(\hat{l})$$

10

は、GPT-2の低パラメータ変形例であるDistilGPT-2言語モデルの下での

【数33】

$$\hat{l}$$

の確率を示し（更なる詳細については方法S5を参照）、 α_{gpt2} は、ハイパーパラメータ最適化を通じて設定されたスケールングハイパーパラメータを表す。次いで、この定式化を所与として最も可能性の高い文

【数34】

$$\hat{l}$$

20

を参加者に表示し、最終的な文として保存した。

【0424】

ビーム探索アルゴリズムの更なる詳細については、方法S4を参照されたい。

【0425】

パフォーマンス評価

文字誤り率（CER）及び単語誤り率（WER）：

CER及びWERは、以前の研究^{6, 16}にあるように短文によって過度に影響されるため、CERとWERを、文スペリングブロック内の予測された各文と標的文との間の文字又は単語編集距離を合計し、次いで、この数をブロック内の全ての標的文にわたる文字又は単語の総数で除算した値として報告した。各ブロックには、2回～5回の文試験が含まれていた。

30

【0426】

会話型タスク条件中のパフォーマンスの評価：

文スペリングタスクの会話型条件のCER及びWERを算出するためのグランドトゥールース文を得るために、各ブロックを完了した後、参加者にそのブロックからの質問及び復号された文を想起させ、その後、復号された文ごとに、参加者は、復号された文が正しいことを確認するか、又は、市販のコミュニケーション支援デバイスを使用して意図された文をタイプした。評価に使用した各ブロックには、2回～4回の文試験が含まれていた。

40

【0427】

毎分文字数及び単語数：

各文スペリング（コピータイピング）ブロックの毎分文字数及び毎分単語数を次のように算出した。

【数35】

50

$$\text{速度} = \frac{\sum_i N_i}{\sum_i D_i}$$

10

【0428】

ここで、 i は各試験をインデックス付けし、 N_i は試験 i のために復号された単語又は文字（空白文字を含む）の数を示し、 D_i は試行 i の持続時間を示している（分単位。試験 i の最後のコードワードに対応する神経活動のウィンドウが終了した時間と試験 i の最初のコードワードのゴーキューの時間との間の差として計算される）。

20

【0429】

電極寄与

孤立標的タスク中に記録されたデータを使用して電極寄与を計算するために、Simon⁴¹のように、時間の経過とともに入力特徴に関して分類器の損失関数の導関数を計算し、予測モデル出力が、各時点での各電極及び特徴タイプ（HGA又はLFS）の入力特徴値へのわずかな変化によってどれだけ影響を受けたかの測度をもたらした。次いで、これらの値のL2ノルムを経時的に算出し、全ての孤立標的試験にわたって結果の値を平均し、その分類器の各電極及び特徴タイプについて単一の寄与値をもたらした。

30

【0430】

相互検証

各フォールドについて、孤立標的タスクの層別交差検証フォールドを使用した。各フォールドを、データの90%を含む訓練セットと、残りの10%を含む保持テストセットとに分割した。次いで、訓練データセットの10%を検証セットとして選択した。

【0431】

神経特徴主成分の分析

HGA及びLFSの神経特徴を特徴付けるために、ブートストラップされた主成分分析を使用した。まず、各NATOコードワードについて、そのコードワードについての最初の318回の沈黙試行孤立標的試験から、神経活動（ゴーキューからゴーキューの2.5秒後までに及ぶ）のキュー整合タイムウィンドウをランダムにサンプリングした（置換あり）。分類のための各特徴ストリームの役割を明確に理解するために、分類器によって使用される信号を得るために、信号を6分の1にダウンサンプリングした。次いで、各コードワードのデータを平均化し、各電極及び特徴セット（HGA、LFS、及びHGA+LFS）の時間にわたる26の試験平均を得た。次いで、これを次元 $N \times T \times C$ の行列に配置構成した。ここで、 N は特徴の数（HGA及びLFSについては128、HGA+LFSについては256）、 T は各2.5秒ウィンドウ内の時点の数であり、 C は各特徴の試験平均活動を連結することによる、NATOコードワードの数（26）である。次いで、この行列の特徴次元に沿って主成分分析を実施した。追加的に、各コードワードの試験平均データを、次元 $T \times N \times C$ の行列に配置構成した。次いで、時間次元に沿って主成分分析を

40

50

実施した。各分析について、測定手順を100回実施して、分散の80%超を説明するために必要な主成分の最小数の代表的な分布を得た。

【0432】

最近クラス距離比較

コードワード及び文字の最近クラス距離を比較するために、まず、各コードワード及び文字について47回の沈黙試行孤立標的試験にわたる、結合HGA+LFS特徴セットの1,000回のブートストラップ反復の平均を算出した。次いで、各対組み合わせ間の差のフロベニウスノルムを計算した。各コードワードについて、そのコードワードと任意の他のコードワードとの間の計算された最小の距離を最近クラス距離として使用した。次いで、文字についてこのプロセスを繰り返した。

10

【0433】

より大きい語彙への一般化可能性

リアルタイム文スペリング中、参加者は、一般的な単語及び臨床介護に関連する単語を含む1,152語の語彙から構成される文を作成した。本発明のシステムの一般化可能性を評価するために、3つのより大きい語彙を使用したオフラインシミュレーションにおいて文スペリングアプローチをテストした。これらの語彙の最初のものは、「Oxford 3000」単語リストに基づいており、これは、Oxford英語コーパスにおけるそれらの頻度及び英語話者との関連性に基づいて選択された3,000の中核単語から構成されている⁴²。2つ目は「Oxford 5000」単語リストに基づいており、これは「Oxford 3000」リストを、追加の2,000個の頻繁で関連性の高い単語で増補したものである。3つ目は、1兆語を超えるテキストのコーパスであるGoogleのTrillion Word Corpus内の最も頻度の高い10,000語に基づいた語彙であった⁴³。このリストに含まれていた非単語（「f」、「gp」、及び「ooo」など）を排除するために、3文字以下から構成される単語を、「Oxford 5000」リストに表示されていない場合に除外した。これらの3つの語彙の各々を、まだ含まれていない元の1,152語の語彙からの単語で補完した後、3つの最終的な語彙は、3,303語、5,249語、及び9,170語を含んでいた（これらのサイズは、語彙が紹介されたのと同じ順序で与えられている）。

20

【0434】

各語彙について、ビーム探索手順中に使用されたNグラム言語モデルを、新しい語彙の下で有効であったNグラムで再訓練し（方法S5を参照）、ビーム探索中により大きい語彙を使用した。次いで、リアルタイムテスト中に使用されたのと同じハイパーパラメータを使用して、文スペリング実験をオフラインでシミュレートした。

30

【0435】

試験拒絶

文スペリングタスクのコピータイピング条件中、参加者は、フィードバックとして表示された復号文がどれほど正確であったかにかかわらず、各意図された文を黙ってスペリングすることを試行するように指示された。しかしながら、少数の試験中に、参加者は間違いを犯した（例えば、間違ったコードワードを使用すること、又は、文の中で自分の場所を忘れることによって）ことを自己報告し、時として自身の試行を停止した。これは主に、彼がまだインターフェースに慣れようとしている間の最初の文スペリングセッション中に発生した。参加者のパフォーマンスではなく、本発明のシステムのパフォーマンスを評価することに焦点を当てるために、これらの試験（合計163回の試験のうち13回の試験）をパフォーマンス評価分析から除外し、パフォーマンス評価中に所望の量の試験（75の固有の文の各々について2回の試験）を維持するために、参加者に後続のセッションにおいてこれらの試験において再び文をスペリングすることを試行させた。パフォーマンスメトリックを評価するときに、これらの拒絶された文を含めると、リアルタイムのスペリングブロック中に観察された中央値CER及びWERは、それぞれ8.52%（99%CI [3.20, 15.1]）及び13.75%（99%CI [8.71, 29.9]）にわずかに増大した。

40

50

【0436】

文のスペリングタスクの会話型条件中に、参加者が間違いを犯したと自己報告した場合（コピータイピング条件のように）、又は意図された単語が1, 152語の語彙外にあった場合、試験は拒絶された。いくつかのブロックでは、参加者は、ブロックが終了した後に意図した回答を報告するように求められたとき、意図した回答の1つを忘れたことを示した。この会話型タスク条件にはグラントゥールスがなかったため、この試験を分析に使用することはできなかった。39回の元の会話型文スペリング試験のうち、参加者は2つの試行で迷子になり、6回の試験中に語彙外の単語を使用しようとし、3回の試験中にグラントゥールス文を忘れた（パフォーマンス評価のために28回の試験を残した）。参加者が語彙外の意図された単語を使用したブロックを組み込むと、CER及びWERはそれぞれ15.7%（99%CI [6.25, 30.4]）及び17.6%（99%CI [12.5, 45.5]）の中央値にわずかに上昇した。

10

【0437】

統計テスト

本研究で使用された統計検定は全て、図のキャプション及び本文に記載されている。簡潔に言えば、両側ウィルコクソン順位和検定を使用して、任意の2つの観察群を比較した。観察がペアリングされたとき、代わりに両側ウィルコクソン符号順位検定を使用した。基礎となる神経データが互いに独立していない比較に対して、ホルム・ボンフェローニ補正を使用した。0.01未満のP値を有意であると考えた。並べ替え検定を使用して、スピアマン順位相関のP値を計算した。各並べ替えについて、1つの観察群をランダムにシャッフルし、次いで相関を決定した。シャッフルされていない観察について計算されたスピアマン順位相関よりも大きい相関値を有する並べ替えの割合として、p値を計算した。報告されたメトリックの周りの任意の信頼区間について、ブートストラップアプローチを使用して99%信頼区間を推定した。各反復（合計2000回の反復）において、データ（交差検証フォールド当たりの正確度など）を置換してランダムにサンプリングし、希望のメトリック（中央値など）を算出した。次いで、信頼区間を、ブートストラップされたメトリックのこの分布に関して計算した。

20

【0438】

文献目録

1. Beukelman, D. R., Fager, S., Ball, L. & Dietz, A. AAC for adults with acquired neurological conditions: A review. *Augment. Altern. Commun.* 23, 230–242 (2007).
2. Felgoise, S. H., Zaccardo, V., Duff, J. & Simons, Z. Verbal communication impacts quality of life in patients with amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Front. Degener. Amyotroph. Lateral Scler. Front. Degener.* 17, 179–183 (2016).
3. Brumberg, J. S., Pitt, K. M., Mantie-Kozlowski, A. & Burnison, J. D. Brain-Computer Interfaces for Augmentative and Alternative Communication: A Tutorial. *Am. J. Speech Lang. Pathol.* 27, 1–12 (2018).
4. Vansteensel, M. J. et al. Fully Implanted Brain-Computer Interface in a Locked-In Patient with ALS. *N. Engl. J. Med.* 375, 2060–2066 (2016).
5. Pandarinath, C. et al. High performance communication by people with paralysis u

30

40

50

sing an intracortical brain-computer interface. *eLife* 6, 1–27 (2017).

6. Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M. & Shenoy, K. V. High-performance brain-to-text communication via handwriting. *Nature* 593, 249–254 (2021).

7. Branco, M. P. et al. Brain-Computer Interfaces for Communication: Preferences of Individuals With Locked-in Syndrome. *Neurorehabil. Neural Repair* 35, 267–279 (2021).

8. Bouchard, K. E., Mesgarani, N., Johnson, K. & Chang, E. F. Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495, 327–332 (2013).

9. Carey, D., Krishnan, S., Callaghan, M. F., Sereno, M. I. & Dick, F. Functional and Quantitative MRI Mapping of Somatomotor Representations of Human Supralaryngeal Vocal Tract. *Cereb. Cortex* 27, 265–278 (2017).

10. Chartier, J., Anumanchipalli, G. K., Johnson, K. & Chang, E. F. Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. *Neuron* 98, 1042–1054.e4 (2018).

11. Lotte, F. et al. Electrocorticographic representations of segmental features in continuous speech. *Front. Hum. Neurosci.* 09, 1–13 (2015).

12. Herff, C. et al. Brain-to-text: decoding spoken phrases from phone representations in the brain. *Front. Neurosci.* 9, 1–11 (2015).

13. Makin, J. G., Moses, D. A. & Chang, E. F. Machine translation of cortical activity to text with an encoder-decoder framework. *Nat. Neurosci.* 23, 575–582 (2020).

14. Mugler, E. M. et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. *J. Neural Eng.* 11, 035015–035015 (2014).

15. Sun, P., Anumanchipalli, G. K. & Chang, E. F. Brain2Char: a deep architecture for decoding text from brain recordings. *J. Neural Eng.* 17, 066015 (2020).

16. Moses, D. A. et al. Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria. *N. Engl. J. Med.* 385, 217–227 (2021).

17. Adolphs, S. & Schmitt. Lexical Coverage of Spoken Discourse. *Appl. Linguist.* 24, 425

10

20

30

40

50

– 438 (2003) .

18. van Tilborg, A. & Deckers, S. R. J. M. Vocabulary Selection in AAC: Application of Core Vocabulary in Atypical Populations. *Perspect. ASHA Spec. Interest Groups* 1, 125–138 (2016) .

19. Hannun, A. Y. , Maas, A. L. , Jurafsky, D. & Ng, A. Y. First-Pass Large Vocabulary Continuous Speech Recognition using Bi-Directional Recurrent DNNs. *ArXiv14082873 Cs* (2014) .

20. Silversmith, D. B. et al. Plug-and-play control of a brain-computer interface through neural map stabilization. *Nat. Biotechnol.* 39, 326–335 (2020) .

21. Rezeika, A. et al. Brain-Computer Interface Spellers: A Review. *Brain Sci.* 8, 57 (2018) .

22. Sellers, E. W. , Ryan, D. B. & Hauser, C. K. Noninvasive brain-computer interface enables communication after brainstem stroke. *Sci. Transl. Med.* 6, 257re7–257re7 (2014) .

23. Gilja, V. et al. A high-performance neural prosthesis enabled by control algorithm design. *Nat. Neurosci.* 15, 1752–1757 (2012) .

24. Kawala-Sterniuk, A. et al. Summary of over Fifty Years with Brain-Computer Interfaces—A Review. *Brain Sci.* 11, 43 (2021) .

25. Serruya, M. D. , Hatsopoulos, N. G. , Paninski, L. , Fellows, M. R. & Donoghue, J. P. Instant neural control of a movement signal. *Nature* 416, 141–142 (2002) .

26. Wolpaw, J. R. , McFarland, D. J. , Neat, G. W. & Forneris, C. A. An EEG-based brain-computer interface for cursor control. *Electroencephalogr. Clin. Neurophysiol.* 78, 252–259 (1991) .

27. Laufer, B. What percentage of text-lexis is essential for comprehension. *Spec. Lang. Hum. Think. Think. Mach.* 316323, (1989) .

28. Webb, S. & Rodgers, M. P. H. Vocabulary Demands of Television Programs. *Lang. Learn.* 59, 335–366 (2009) .

29. Nourski, K. V. et al. Sound identification in human auditory cortex: Differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. *Brain Lang.* 148, 37–50 (2015) .

10

20

30

40

50

30. Conant, D. F. , Bouchard, K. E. , Leonard, M. K. & Chang, E. F. Human sensorimotor cortex control of directly-measured vocal tract movements during vowel production. *J. Neurosci.* 38, 2382–17 (2018) .

31. Gerardin, E. et al. Partially Overlapping Neural Networks for Real and Imagined Hand Movements. *Cereb. Cortex* 10, 1093–1104 (2000) .

32. Guenther, F. H. & Hickok, G. Neural Models of Motor Speech Control. in *Neurobiology of Language* 725–740 (Elsevier, 2016) .

33. Moses, D. A. , Leonard, M. K. & Chang, E. F. Real-time classification of auditory sentences using evoked cortical activity in humans. *J. Neural Eng.* 15, (2018) .

34. Moses, D. A. , Leonard, M. K. , Makin, J. G. & Chang, E. F. Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nat. Commun.* 10, 3096 (2019) .

35. Ludwig, K. A. et al. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *J. Neurophysiol.* 101, 1679–89 (2009) .

36. Williams, A. J. , Trumpis, M. , Bent, B. , Chiang, C. -H. & Viventi, J. A Novel μ ECoG Electrode Interface for Comparison of Local and Common Averaged Referenced Signals. in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 5057–5060 (IEEE, 2018) . doi:10.1109/EMBC.2018.8513432.

37. Parks, T. W. & McClellan, J. H. Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase. *IEEE Trans. Circuit Theory* 19, 189–194 (1972) .

38. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *ArXiv1412.6980 Cs* (2017) .

39. Cho, K. et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. in *1724–1734* (2014) . doi:http://dx.doi.org/10.3115/v1/D14-1179.

40. Fort, S. , Hu, H. & Lakshminarayanan, B. Deep Ensembles: A Loss Landscape Perspective. *ArXiv1912.02757 Cs Stat* (2020) .

41. Simonyan, K. , Vedaldi, A. & Zisserman, A. Deep Inside Convolutional Networks: Visual

10

20

30

40

50

Listing Image Classification Models and saliency Maps. ArXiv1312.6034 Cs (2014).

42. About the Oxford 3000 and 5000 word lists at Oxford Learner's Dictionaries. <https://www.oxfordlearnersdictionaries.com/us/about/wordlists/oxford3000-5000>.

43. Brants, Thorsten & Franz, Alex. Web 1T 5-gram Version 1. 20971520 KB (2006) doi: 10.35111/CQPA-A498.

【0439】

10

実施例4：公然発話試行対沈黙発話試行に関する参加者調査

参加者に、沈黙発話試行又は公然発話試行のいずれかを使用したスペリングシステムの制御について、次の質問をした。各質問の後に参加者の回答が提供される。

【0440】

1. 公然発話試行によるコミュニケーションに、スペリングシステムをどのくらいの間快適に使用できると思いますか？(How long do you think you could comfortably use the spelling system for communication with overt-speech attempts?) 回答：15分(15 minutes)

2. 沈黙試行によるコミュニケーションに、スペリングシステムをどのくらいの間快適に使用できると思いますか？(How long do you think you could comfortably use the spelling system for communication with silent-speech attempts?) 回答：30分(30 minutes)

20

3. 公然発話試行でスペリングシステムを使用する快適さを1～10の尺度でランク付けしてください。(Can you please rank your comfort using the spelling system with overt-speech attempts on a scale from 1-10?) 回答：5

4. 沈黙発話試行でスペリングシステムを使用する快適さを1～10の尺度でランク付けしてください。(Can you please rank your comfort using the spelling system with silent-speech attempts on a scale from 1-10?) 回答：8

30

5. ゴーキューが出てから公然発話試行でスペリングシステムを使用するまでの間に最低限必要な時間はどのくらいですか？(What is the minimum amount of time you need between go cues to use the spelling system with overt-speech attempts?) 回答：4秒(4 seconds)

6. ゴーキューが出てから沈黙発話試行でスペリングシステムを使用するまでの間に最低限必要な時間はどのくらいですか？(What is the minimum amount of time you need between go cues to use the spelling system with silent-speech attempts?) 回答：2.5秒(2.5 seconds)

40

7. スペラデバイスを制御するために沈黙発話試行を使用することは、公然発話試行を使用することと比べてどうですか？(How does using silent-speech attempts compare to using overt speech attempts to control the speller device?)

(a) 沈黙は公然よりもはるかに簡単です(Silent is much easier than overt)

(b) 沈黙は公然よりも簡単です(Silent is easier than o

50

vert)

(c) 沈黙は公然と同じです (Silent is the same as overt)

(d) 沈黙は公然よりも難しいです (Silent is harder than overt)

(e) 沈黙は公然よりもはるかに難しいです (Silent is much harder than overt)

回答: (a) 沈黙は公然よりもはるかに簡単です

【0441】

参加者の回答を下記に要約する。全体的に、参加者は、スペリング神経プロテゼを制御するために、沈黙発話試行を非常に選好する。

【表2】

質問	公然	沈黙
どのくらいの間快適にデバイスを使用できますか?	15 分	30 分
コミュニケーションのためにデバイスを使用する快適さはどのくらいですか(1~10)?	5	8
ゴーキューの間に最低限必要な時間はどのくらいですか?	4 秒(4 seconds)	2.5 秒(2.5 seconds)
沈黙発話試行は公然よりもどの程度簡単ですか?	(n/a)	はるかに簡単

【0442】

実施例5: データ再正規化

記録セッションにわたって神経特徴一貫性を促進するために、全ての神経特徴について実行中の30秒のzスコアを使用した(図22を参照されたい)。しかしながら、参加者の右手を握る試行中に記録された神経活動は、典型的には、沈黙発話試行中に記録された活動と比較して、信号の大きさが異なる。結果として、実行中のzスコアを使用する場合、発話内容(文字及びNATOコードワード試験)のみ又は試行手運動試験のみを有するいくつかの孤立標的タスクブロックは、発話試験及び手運動試験の両方を有する孤立標的ブロックとは異なる神経特徴ベースラインを有した。

【0443】

これを軽減するために、同じ日に記録された文字及びNATOコードワード孤立標的ブロックと、試行手運動孤立標的ブロックとを共同的に再正規化した。各記録日について、かつ各発話タイプ(沈黙又は公然)について独立して、その日に記録された全ての試行発話試験と手運動試験を、これらの試験と関連付けられる神経特徴(zスコア正規化のない高ガンマ活性及び低周波信号)のタイムウィンドウを(時間次元に沿って)連結することによって組み合わせた。これらの神経特徴タイムウィンドウは、各試験のゴーキューの2秒前から3.5秒後までの範囲であった。これらの正規化されていない信号における潜在的な信号アーチファクトの影響を低減するために、各特徴(各特徴タイプの各電極チャンネル)の信号大きさを、その特徴について記録された信号大きさの1パーセンタイル及び99パーセンタイル以内にクリップした。次いで、その日に記録された各試験の神経特徴を、特徴量平均を差し引き、連結されたデータ行列の特徴量標準偏差で除算して再正規化した。(同じ日に両方のタイプのデータが記録されていない場合)試行発話のみ又は試行手運動のみを含むタスクブロックの中には、このように再正規化されなかったものがあることに留意されたい。追加的に、いくつかの試行手運動ブロックは、公然及び沈黙の両方の試行NATOコードワード孤立標的も記録された日に記録されたため、これは、再正規化されなかったブロック(これらのブロックは、試行発話のみを含むブロックと同じ日に記録されなかった)、公然発話試行のみを含むブロックとともに再正規化されたブロック、及び沈黙試行発話のみを含むブロックとともに再正規化されたブロックの、3つの可能なタイプの試行手運動ブロックがあることを意味した。再正規化されなかったタスクブロックからのデータは、図22に記載される実行中の30秒のzスコア正規化手順及び自動ア

ーチファクト拒絶を使用した。

【0444】

実施例6：スペリング復号のための補足情報

セクションS1．孤立標的タスク

参加者が黙って（又は時には公然と）促された発言を発語すること、又は、孤立標的タスク中に促された運動を実施することを試行したときの参加者の神経活動を記録した。本文の方法セクションで説明されているように、孤立標的タスクの各試験は、テキストの両側に4つの点がある参加者の画面上の単一の発話又は運動標的のテキスト提示から開始した。これらの点は、一度に1つずつ（テキストの両側で同時に）一定の速度で消え、参加者にタスクタイミングを提供した。最後の点が消えると、テキスト標的は緑色に変わり、ゴーキューを表した。このゴーキューにおいて、参加者は標的を生成しようとするように指示された。テキスト標的は、画面がクリアされ、次の試験が開始される前に、参加者の画面上に短い間隔にわたって残った。

10

【0445】

発話検出及び神経分類モデルを訓練するための孤立標的パラダイムを用いて、以下の4つ発言セットを収集した。

- 1．26個の英語文字
- 2．26個のNATOコードワード
- 3．26個のNATOコードワード及び試行された手の握り
- 4．試行された手の握り及び3つの他の試行運動

20

【0446】

孤立標的タスクの各ブロック内で、カウントダウン点が消えた速度 τ_p 及びゴーキューの後に標的テキストが画面に残った持続時間 τ_t は、単一のブロック内では試験にわたって同一であった。しかしながら、これら2つのタスク間隔パラメータは、ブロック間で変化した。試行運動ブロックについては、点当たり $\tau_p \in [0.35, 0.5]$ 秒及び $\tau_t = 4.0$ 秒を使用した。他の全ての孤立標的ブロックについて、点当たり $\tau_p \in [0.45, 1.5]$ 秒及び $\tau_t \in [0.45, 6.0]$ 秒を使用した。

【0447】

セクションS2．発話検出モデル

沈黙試行発話イベントがいつ発生したかを識別するために、リアルタイムで神経特徴を分析するように発話検出モデルを設計した。この発話検出器を使用して、リアルタイム文スペリング中にスペリングシステムの自発的なエンゲージを可能にした。発話検出器を訓練及び評価するために使用される全てのデータは、試行された手の握り又は沈黙試行発話のいずれかの試験であった（公然試行発話データは使用されなかった）。

30

【0448】

データ準備

26個のNATOコードワードの試行を含む孤立標的タスクブロック、26個のNATOコードワードの試行及び試行された右手の握りを含むブロック、及び試行された手の握りを含む様々な試行運動を含むブロック（そこから本発明者らは、試行された手の握りのみを使用した）からのデータを使用して、発話検出器を訓練した。発話検出器を訓練するために、神経特徴データの各時点をラベル付けするために「発話準備」、「発話」、「運動」、及び「休止」の4つのカテゴリを使用した。標的NATOコードワードの参加者の画面上への出現と、関連するゴーキューとの間の時点は、発話準備としてラベル付けされた。NATOコードワード試行のゴーキューとそのゴーキューの1秒後との間の時点は、発話としてラベル付けされた。試行された手の握りのゴーキューとそのゴーキューの2秒後との間の時点は、運動としてラベル付けされた。試行の割り当てられた時間期間の終了（発話のゴーキューの1秒後、又は手の握りの2秒後）とその試行の終了（試行間の間隔にわたって画面がクリアされたとき）との間の時点は、訓練されなかった。発話検出器のための訓練データは、試行運動孤立標的タスクのブロックを含んだ。試行運動のみを含むブロックの場合、試行運動試験中の、試行された手の握りの試行ではない時点は無視され

40

50

た。他の全ての時点は、休止としてラベル付けされた。

【0449】

発話検出器は、200Hzにおいて低周波信号（LFS）及び高ガンマ活性（HGA）の両方を特徴として使用した。これは、分類器とは異なることに留意されたい。分類器もこれらの特徴を使用した。これらの特徴を33.3Hzに更にダウンサンプリングした。

【0450】

モデルアーキテクチャ及び訓練

Python 3.6.6及びPyTorch 1.6.0を使用して、発話検出器を作成し、訓練した[1]。発話検出器は、それぞれ100、50、及び50ノードを有する3つの長・短期記憶（LSTM）層のスタックを含んでいた。LSTM層に後続して、潜在次元を4つのクラス（発話準備、発話、休止、及び運動）にわたる確率に投影する単一の全結合層があった。モデルは、特徴ストリームから各時点を連続的に処理し、確率の連続ストリーム（200Hzでの神経特徴時点当たり1つの予測確率ベクトル）を出力した。モデルの概略図が、図23に示される。

【0451】

発話検出モデルは、修正交差エントロピー損失を最小限に抑えるように訓練された。交差エントロピー損失は、もともと次のように定義されている。

【数36】

$$H_{P,Q}(\ell | y) = \mathbb{E}_P[-\log Q(\ell | y)]$$

$$\approx -\frac{1}{N} \sum_{n=1}^N \log Q(\ell_n | y_n), \quad (S1)$$

式中、

- P：割り当てられたクラスラベル ℓ によって決定される、クラスの真の分布。
- N：サンプルの数。
- $H_{P,Q}(\ell | y)$ ： ℓ の真の分布に対する予測分布の交差エントロピー。
- $\log$ ：自然対数。

【0452】

この損失を修正して、運動としてラベル付けされたが、発話として予測された時点、発話としてラベル付けされたが、運動として予測された時点、及び休止としてラベル付けされたが、発話として予測された時点の3つのタイプの誤った予測に追加のペナルティを追加した。実際には、 w_n を1.1と定義した。これらの修正により、式S1で定義された交差エントロピー損失は、次のように再定義される。

【数37】

$$H_{P,Q}(\ell | y) \approx -\frac{1}{N} \sum_n^N w_n \log Q(\ell_n | y_n), \quad (S2)$$

式中、 w_n は、サンプル n のペナルティ重みであり、次のように定義される。

【数38】

$$w_n := \begin{cases} 1.1 & (l_n = \text{運動}) \text{かつ} \left(\arg \max_{l \in L} [Q(l | y_n)] = \text{発話} \right) \text{である場合} \\ 1.1 & (l_n = \text{発話}) \text{かつ} \left(\arg \max_{l \in L} [Q(l | y_n)] = \text{運動} \right) \text{である場合} \\ 1.1 & (l_n = \text{休止}) \text{かつ} \left(\arg \max_{l \in L} [Q(l | y_n)] = \text{発話} \right) \text{である場合} \\ 1 & \text{そうでない場合} \end{cases} \quad (\text{S3})$$

10

【0453】

このペナルティ修正を使用して、発話検出器が偽陽性の間違いを犯す尤度を低減した（参加者が実際には手を握ることを試行していたときに、試行発話イベントを誤って検出するなど）。

【0454】

〔2〕で以前に説明したように、本発明者らは、発話検出器を訓練するために、打ち切り型通時的逆伝播（BPTT）を使用した。簡潔に言えば、モデルが予測を行うためにタスク周期性に依拠しないようにするために、発話検出モデルを一度に500ミリ秒だけ逆伝播させることによって、手動でBPTTを実装した。Adamオプティマイザを使用して、学習率は0.001とし、残り最適化パラメータについてはデフォルト値として、式S2〔3〕を所与として交差エントロピー損失を最小化した。過剰適合を防ぐために、保持検証セットに対して早期停止を使用し、最終層を除く各LSTM層に対して0.5のドロップアウトを使用した。全ての訓練ステップについて、4つの可能なクラスの間で（同じ数の訓練例を含む）クラスを平衡させた。

20

【0455】

イベント検出

リアルタイム文スペリング中、発話検出器は、LFS及びHGAの時点を連続的に処理し、沈黙発話確率のストリームを生成した。〔2〕の補足セクションS8に記載されているのと同じアプローチを使用して、この確率ストリームから沈黙発話イベントを特定した。簡潔に言えば、発話確率は、まず、移動ウィンドウ平均を使用して時間的に平滑化された。次いで、確率閾値を使用して平滑化確率を二値化した。最後に、バイナリ状態の変化（発話の不在から発話の存在へ、又はその逆）が、変化が発話の開始又は終了とみなされる前に特定の持続時間よりも長く持続しなければならないことを要求することによって、これらの二値化された値を「デバウンス」した。これらの3つのパラメータ値は、ハイパーパラメータ最適化を介して選択され、表S2に列挙されている。

30

【0456】

ハイパーパラメータ最適化

ハイパーパラメータ最適化プロセスは、本発明者らの以前の研究と同一である〔2〕。簡潔に言えば、Hyperopt Pythonパッケージ〔4〕を使用して、検出スコアに基づいてコスト関数を最小化することにより、3つの検出ハイパーパラメータを最適化した。〔2〕の補足セクションS8で定義されるように、検出スコアは、個々の時点が発話又は非発話としてどの程度正確に予測されたか、及び検出器が一般的に発話試行イベントをどの程度正確に識別したかの両方を包含する測度である。ハイパーパラメータを最適化するために使用されるコスト関数は、時間閾値パラメータを最小化しながら検出スコアを最大化することを模索する（沈黙発話試行を検出するのに必要な時間を最小化することが所望されたため）。コスト関数は、次のように定義された。

40

【数39】

$$c_{hp}(\Theta) = (1 - s_{\text{detection}})^2 + \lambda_{\text{time}} \theta_{\text{time}}, \quad (\text{S4})$$

50

式中、

- $c_{hp}(\theta)$: ハイパーパラメータ値組み合わせ θ を使用した目的関数の値。
- λ_{time} : 時間閾値持続時間に適用されるペナルティ。
- θ_{time} : θ に含まれる 3 つのパラメータのうちの 1 つである時間閾値の持続時間値。

ここでは、 $\lambda_{time} = 0.00025$ を使用した。

【0457】

発話確率に適用された検出パラメータのみを最適化したため、訓練されたモデルからのタスクブロックのセットにわたって発話確率を計算し、これらのブロックからの発話確率を使用してハイパーパラメータ組み合わせを評価することができた。孤立標的ブロックに関してモデルを訓練した後、そのモデルを使用して、NATOコードワード沈黙発話試行及び試行された手の握りを含む孤立標的タスクの 12 個の保持ブロックの発話確率を予測した。リアルタイム文スペリングタスクには沈黙発話試行及び手の握りの試行の両方のタイプの試行が含まれているため、沈黙発話試行及び手の握りの試行の両方を含むブロックに対して最適化することを選択した。1000 回の最適化反復の後、最小のコスト値を有する最適化実行から最終的なハイパーパラメータを選択した。

10

【0458】

セクション S3. 分類モデル

データ準備

26 個の NATOコードワードの試行を含む孤立標的タスクブロック、26 個の NATOコードワードの試行及び試行された右手の握りを含むブロック、及び試行された手の握りを含む様々な試行運動を含むブロック（そこから本発明者らは、試行された手の握りのみを使用した）からのデータを使用して、分類器を訓練した。特徴タイプ、発話タイプ、及び発言セットの比較中に使用される分類器については、孤立標的タスクブロックからのデータのみが使用された。

20

【0459】

リアルタイム文スペリング（及び関連するオフライン分析）のための分類器の訓練中に、復号された文が 0.0 文字誤り率（CER）を有する文スペリング（コピータイピング）試験も含めた。これらの文スペリング試験は、公然発話試行のデータの 3.06% を構成し（公然発話試行を伴う予備的な文スペリング試験は収集されたが、評価中には使用されなかった）、沈黙発話試行のデータの 22.7% を構成した。これらの分類器について、転移学習アプローチを使用して、公然発話試行を事前に訓練し、次いで、沈黙発話試行を微調整した（特に明記されている場合を除く。更なる詳細については、このセクションの後半で説明する）。テスト中に使用された任意の試験と同じセッション（又は関連するオフライン分析については、その試験の進行中のセッション）中に記録された分類器訓練中の文スペリング試験は含めなかった。分類器は評価セッション中に再校正又は更新されなかった。特定の評価のための特定のデータセットの使用法は、下記の表に記載されている。

30

40

50

【表 3】

評価及び本文の関連する図	孤立標的(公然)	孤立標的(沈黙)	文スペリング(公然)	文スペリング(沈黙)
リアルタイム文スペリングパフォーマンス評価及び関連するオフライン分析(ビーム探索、言語モデル、語彙セット、及びタスク条件評価)[図 16 及び図 20]	事前訓練	微調整	事前訓練	微調整及びテスト
特徴タイプ(HGA 対 LFS)比較[図 17]		訓練及びテスト		
オフライン発話タイプ(沈黙対公然)比較[図 18]	訓練*及びテスト	訓練*及びテスト		
オフライン発言セット(文字対コードワード)[図 19]		訓練*及びテスト		

*＝該当する場合、事前訓練及び微調整に使用される（詳細は図 18 を参照）

【0460】

評価に使用されるデータとハイパーパラメータ最適化に使用されるデータとの間に重複はなかった。

【0461】

各孤立標的試験について、神経特徴（200 Hz における高ガンマ活性（HGA）及び低周波信号（LFS）特徴）の関連するタイムウィンドウを、ゴーキューの 2 秒前から 4 秒後までと定義した。この神経特徴ウィンドウは、訓練及びテストに実際に使用されているウィンドウ（「アーキテクチャ及び訓練」サブセクションにおいて下記に詳述）よりも大きかった。これは、このより大きい試験関連ウィンドウからより小さいウィンドウが引き出される時間ジッタリングデータ増強を利用したためである。次いで、デシメーション前は 16.67 Hz のアンチエイリアシングフィルタを適用して、神経活動を 33.33 Hz まで 6 分の 1 にデシメーションした。各サンプルを正規化して、全ての神経特徴にわたって（各電極チャンネルにわたって、また、HGA 及び LFS 特徴タイプについては別個に）1 の L2 ノルムを有するようにした。リアルタイム推論及びオフライン評価のために、神経活動の関連するタイムウィンドウの間に、組み合わせられた（連結された）HGA + LFS 特徴を使用した。したがって、各訓練例について、形状（T, C）の神経活動 x_i の行列を有しており、T は時間ステップの数であり、C は 256 個の特徴（128 個の電極の各々からの 2 つの特徴）を指す。1 つの特徴ストリームのみが特定の分析に使用されていた場合、C は 128 に等しい。

【0462】

モデリングアーキテクチャ及び訓練

沈黙発話試行中の参加者の神経活動の時間的及び空間的動態をモデリングするために、関連するゴーキューの後の神経特徴の 2.5 秒ウィンドウを所与として参加者がどの NATO コードワード（又は想像上の手の握り）を生成したかを分類するように、人工ニューラルネットワークを訓練した。系列タスク [7] に対して他の再帰アーキテクチャ（長・短期記憶ネットワークなど）[6] を上回るパフォーマンスを示したゲート付き再帰ユニット（GRU）層 [5] を使用した。

【0463】

分類器において、神経特徴は、まず、重み W 及びバイアス項 b によってパラメータ化された 1 次元畳み込み層によって処理された。この結果として、次のように定義された出力表現 h_n （隠れ層 n の出力）が得られる。

【数 40】

$$h_{1,j} = b_j + \sum_{k=0}^{C-1} W[j,k] * x_i[:,k], \quad (S5)$$

式中、 $h_{1,j}$ は、隠れ層 1 の出力の要素 j であり、 $*$ は、有効な相互相関演算子を示し、 C は、入力行列 x_i における神経特徴の数を指す。

【0464】

次いで、この表現を n 個の GRU 層からなるスタックに渡した。各ユニットは、それぞれ入力状態及び隠れ状態に作用する重み及びバイアスである W_i 、 b_i 、 W_h 、及び b_h によってパラメータ化された。各行列の一部は、リセットゲート r_t 、更新ゲート z_t 、及び新規ゲート n_t に専用であった。

【0465】

各時点 t で、GRU は、以下を計算した。

$$r_t = \sigma(W_{ir} x_t + b_{ir} + W_{hr} h_{(t-1)} + b_{hr})$$

$$z_t = \sigma(W_{iz} x_t + b_{iz} + W_{hz} h_{(t-1)} + b_{hz})$$

$$n_t = \tanh(W_{in} x_t + b_{in} + r_t * (W_{hn} h_{(t-1)} + b_{hn}))$$

$$h_t = (1 - z_t) * n_t + z_t * h_{(t-1)}$$

式中、 $*$ はアダマール積を示し、 σ はシグモイド関数を示し、 h_t はこの層の各時点 t での出力である。基本的に、GRU は、 z_t を使用して、(リセット機能を組み込んで) 新しい活動を所与として以前の値から隠し状態をどの程度更新するかを各時点で決定した。各層の出力 h_n は、次の層への入力として使用される。訓練中に、過剰適合を最小限に抑えるために、ドロップアウト [8] を使用して、ハイパーパラメータ最適化によって決定した確率 $p_{dropout}$ で h_n の要素を 0.0 にランダムに設定した。

【0466】

正確度を向上させるために、双方向 GRU 層を使用した。これは、各 GRU において、入力がコピーされ、後方に反転され、ネットワークへの入力として使用されたことを意味する。これにより、前方及び後方の表現を学習し、クラス確率を予測する際の文脈として使用することができた。

【0467】

26 個の NATO コードワードにわたる予測確率分布を計算し、最終 GRU 層の最終時点をもととして想像上の手の握りを計算するために、これに行列 W_{out} を乗算して、 W_{out} が形状 $(N_{hn}, 27)$ を有するバイアス項 b を加算した。 N_{hn} は、最終 GRU 層内の隠れユニットの数に対応する。次いで、これらの活性化に softmax 関数を適用し、各ウィンドウ i 及び各要素 (クラス) k の出力ベクトル

【数 4 1】

$\hat{y}$

の値を次のようにした。

【数 4 2】

$$\hat{y}_{i,k} = \frac{\exp((W_{out} h_n)_k)}{\sum_j \exp((W_{out} h_n)_j)}, \quad (S6)$$

ここで、

【数 4 3】

10

20

30

40

50

y_i

は、サンプル x_i 及び本発明のニューラルネットワークモデル θ のパラメータを所与とした可能な出力クラスにわたる多項分布と考えることができる。

【0468】

訓練中の目標は、神経活動及び θ を所与として、本発明のラベル付けされた訓練データの尤度を最大化することであった。これは、最適化問題として記述することができる。

【数44】

$$\begin{aligned}\theta^* &= \arg \max_{\theta} \prod_i p_{\theta}(y_i | x_i) \\ &= \arg \max_{\theta} \sum_i \log(p_{\theta}(y_i | x_i)).\end{aligned}\tag{S7}$$

【0469】

同等の最適化問題を解くために、ミニバッチ確率的勾配降下を使用してこの問題の解を近似した。

【数45】

$$\theta^* = \arg \min_{\theta} \sum_i -\log(p_{\theta}(y_i | x_i)).\tag{S8}$$

【0470】

具体的には、収束率を向上させるために、勾配の平均及び非中心分散の適応推定を組み込んだAdamオブティマイザ[9]を使用する。PyTorch 1.6.0[10]を使用して、ニューラルネットワークモデル及び最適化手順を実装した。検証セット正確度の改善なしに5エポック後にモデルを早期に停止し、最高の検証セット正確度に対応するモデルパラメータを使用した。

【0471】

リアルタイム推論のために、[2]のように、パフォーマンスを向上させるために10個のモデル予測を平均化することによってモデルをアンサンブルした。

【0472】

2.69秒の神経特徴ウィンドウを使用して訓練され、次いで、2.5秒のウィンドウを使用してテストされたモデルを使用した。この矛盾は、文のスペリング評価ブロックの収集前にタスクタイミングが変更されたことによって引き起こされた。具体的には、当初、文のスペリング中に2.69秒の文字復号サイクルを使用し、それに応じて分類器を訓練することを計画していたが、最終的には2.5秒の文字復号サイクルを使用してペーシングを高速化することを決定した。分類器は、柔軟なウィンドウ長を有する入力に対して推論を実施するように設計されたため、2.5秒のウィンドウをシームレスに、パフォーマンスを著しく低下させることも一切なく、評価することができた。

【0473】

拡大

分類器パフォーマンスを強化するために、一般化を改善し、画像[11、12]及び神経活動[13、14]の両方の過剰適合を低減することが示されているデータ拡大を使用した。以下の拡大は、関連するラベル y_i を変更することなく、訓練（テストではない）中の神経活動 x_i の各試験に順次適用された。

1. 時間ジッタリング：以下になるように、時間シフト τ によって神経特徴をシフトする。

$$x_i(t) = x_i(t - \tau) \cup (-j, j)$$

式中、 j はハイパーパラメータである。

2. 時間マスキング：以下になるように、神経特徴のいくつかの時点をも 0 に設定する。

$$x_i[t_0 : t_1] = (1 - \delta_p), \quad t_1 = t_0 + s \cup (0, b)$$

式中、 t_0 は、 x_i 内のランダムに引き出された時点であり、 p は、 δ_p が 1 である確率であり、時点は 0 に設定される。 b 及び p は両方とも、ハイパーパラメータである。

3. スケーリング：次のようになるように、神経特徴の大きさをスケーリングする。

$$x_i = \alpha x_i, \quad \alpha \cup [\alpha_{min}, \alpha_{max}],$$

式中、 α_{min} 及び α_{max} はハイパーパラメータである。

4. 加法ノイズ：次のようになるように、神経特徴 x_i にランダムガウスノイズの行列を追加する。

【数 4 6】

$$x_i = x_i + N(0, \sigma_n^2),$$

式中、 σ_n はハイパーパラメータである。

5. チャネル毎ノイズ：次のようになるように、神経特徴を、ガウス分布から各チャネル c へとランダムにサンプリングされる値だけオフセットする。

【数 4 7】

$$x_i[:, c] = x_i + N(0, \sigma_{ch}^2),$$

式中、 σ_{ch} はハイパーパラメータであり、全ての特徴にわたって共有される。

【0 4 7 4】

モデルの事前訓練及び微調整

リアルタイムのコピータイピング結果に対するビーム探索、言語モデル、及び異なる語彙サイズの影響を評価するためにその後オフライン分析中にも使用された、リアルタイム文スペリングに使用される分類器のアンサンブルを訓練するとき、まず、公然発話試行に関してモデルを事前に訓練し、次いで沈黙発話試行に対してそれらを微調整した。具体的には、 10^{-3} の学習率で、公然発話試行を含む初期データセットに関して分類器を訓練した。この初期データセットを訓練セットと検証セットとに分割し、検証セットに関する正確度が 5 エポック連続で改善されなかった後、モデルを早期に停止し、モデルパラメータを最も高い検証正確度に対応するものにリセットした。次いで、これらのパラメータから開始して、沈黙発話試行を含む 2 番目のデータセットに関してモデルを微調整した。これには、同じ早期停止プロセスを用いるが、学習率を 10^{-4} とより小さくして、新しいデータセットに関して、事前に訓練されたモデルを訓練することが含まれる。

【0 4 7 5】

ハイパーパラメータ最適化

分類器について、層の数、各層内の隠れノードの数、カーネルサイズ、ストライド、ドロップアウト率、及び拡張ハイパーパラメータを、非同期ハイパーバンド（ASH）方法〔15〕を使用してRayソフトウェアパッケージで最適化した。Hyperoptソフトウェアパッケージを使用して、各評価実行後の次のハイパーパラメータセットを提案した〔16〕。探索空間及び最終値はS2に詳述されており、300個の可能なハイパーパラメータセットを検索した。

〔0476〕

任意の文スペリングタスクブロックを収集する前に記録された孤立標的ブロックからの公然及び沈黙試行試験からの全ての神経データを、ハイパーパラメータ最適化中に保持検証データセットとして使用した。このプロセス中に、残りの孤立標的試験を訓練データとして使用した。ハイパーパラメータ探索の各評価実行中に、アルゴリズムによって決定されたハイパーパラメータのセットを使用して新しいモデルを初期化し、次いで、モデルの訓練を開始した。微調整の前にモデル事前訓練を実施したため、まず、公然発話試行中に記録されたデータに関してモデルを訓練して評価した。訓練の各エポックの後、現在のハイパーパラメータセットを用いて、これらの公然試行試験に関するモデル正確度を評価した。ASHは各ステップにおいて正確度を使用して、パフォーマンスの低いハイパーパラメータ組み合わせを早期に終了させるため、この事前訓練プロセス中に正確度を0.1によってスケールリングして、微調整が開始されると正確度が低下した場合にプロセスが早期に終了することを防止した。

〔0477〕

通常どおりモデルを早期に停止し、最高正確度に対応するパラメータを復帰させた。次いで、それらのパラメータから開始して、 10^{-3} の学習率でデータセットの沈黙試行部分に関してモデルを微調整（及び評価）した。ここでは、ハイパーパラメータ組み合わせをより迅速に評価するために、最終訓練手順で使用されたもの（ 10^{-4} ）よりも高い学習率を意図的に使用した。ASHは、微調整プロセス中にスケールリングされていない正確度値を監視した。ハイパーパラメータ最適化データセットに関する正確度が5エポック連続で改善しなかった後、ハイパーパラメータ最適化反復を終了し、最良の正確度をそのハイパーパラメータセットのスコアとして維持した。

〔0478〕

得られた最適な神経分類器ハイパーパラメータをリアルタイム文スペリングブロック及び分析の全てに使用し、ハイパーパラメータ最適化に使用されたブロックは、全ての分析で評価ブロックとして使用されることから除外された。

〔0479〕

各リアルタイム文スペリング評価セッションの前に、復号された文のCERが0.0であったコピータイピング文スペリング試験からの任意の以前に記録されたデータを含む、その日以前に利用可能な全てのデータに関して10個の神経分類器モデルを訓練した。本発明の記録セッションは連続した日ではなかったため、新しい分類器の訓練に使用できる最新のデータは、常に特定のセッションの少なくとも3日前であった（例えば、次の録画セッションが4日目の場合、最新のデータは1日目からのものであり、2日目と3日目には記録はない）。モデルがセッションの途中で更新されることはなく、1日のモデルの再較正なしで、全てのリアルタイム文スペリング評価を実行した。

〔0480〕

セクションS4. 適合ビーム探索

本文の方法セクションで説明したように、〔17〕のように適合したプレフィックスビーム探索を使用して、可能性のある転写しのセットにわたって、以下を最大化する文字列（空白文字を含む）を含む転写し*を見つけた。

$p_{nc}(l|X)p_{lm}(l)$, (S9)

ここで、 X は神経活動のウィンドウのセット $x_1, \dots, x_T$ であり、 $p_{nc}(l|X)$ は、 X を所与とした l の神経分類器の下での確率であり、 $p_{lm}(l)$ は、言語モデ

ル事前確率の下での転写 l の確率である。[17]のように、Nグラム言語モデルからの言語モデル事前確率は制約が強すぎると仮定したため、重み付けパラメータ(α)を使用してそれを重要視するのを止め、単語数が増大するにつれて文 l の確率が暗黙のうちに減少することを補うために単語挿入ボーナス β を追加し、ビーム探索が以下に最大化しようとする式を修正した。

$$P_{nc}(l|X)p_{lm}(l)^{\alpha}|l|\beta, (S10)$$

式中、 $|l|$ は、転写 l から得られる単語系列の濃度である。 α 及び β の両方は、保持文スペリングデータに対するハイパーパラメータ最適化によって求められたハイパーパラメータであった。 $P_{lm}(l)$ を近似するためにNグラム言語モデルを使用した。完全なアルゴリズムは、アルゴリズム1に詳述されている。

【0481】

文最終化

試行手運動(文最終化コマンド)の確率が80%を超える場合、予測された文が最終化された。具体的には、候補文の現在のリストを(ビーム探索から)プルーニングして、不完全な単語又は語彙外の単語を含む文を削除した。次いで、残りの各候補文 l の確率を次のように更新した。

【数48】

$$p_{finalized}(l) = p(l)p_{gpt2}(l)^{\alpha_{gpt2}}, (S11)$$

式中、 $p_{finalized}(l)$ は、文 l の最終化された確率であり、 $p(l)$ は、式S10の下での文 l の確率であり、 $p_{gpt2}(l)$ は、 $Distil-GPT2$ [18]を使用した l の確率であり、 α_{gpt2} は、ハイパーパラメータ最適化を通じて見出されるスケールパラメータである。次いで、最も可能性の高い文 l を最終文として使用した。

【0482】

ハイパーパラメータ最適化

最適なハイパーパラメータ α 、 β 、 α_{gpt2} 、及び B を見つけるために、3つのセッションにわたって記録されたコピータイピング文スペリングデータを含む最適化データセットを収集して、スペリングシステムのパフォーマンス評価の前にこれらのパラメータを調整した。これらの3つのセッション中、参加者は75個のコピータイピング文のうち35個のスペリングを試行した。これらの35個の文のうち、参加者が10回試行した15個のランダムに選択された文、参加者が9回試行した5個の文、及び参加者が1回試行した15個の文があった。残りの40個の文は、リアルタイム評価の前に参加者には見えなかった。次いで、これらの文をオフラインで使用して、 α 、 β 、 α_{gpt2} 、及び B を最適化した。

【0483】

アルゴリズム1制約ビーム探索神経活動の T 個のウィンドウ及び $p(c|x_1:T)$ (c は文字である)を所与として、このアルゴリズムは、制約された語彙 V 内の単語から構成される最も可能性の高い文 l^* を見つける。 l^+ を与えるために l に文字を追加した後、 l^+ 内の最後の単語が、全ての可能な単語及び部分的な単語 $\in V$ から構成される $V_{partial}$ であることを確認する。関数 w_{final} は、最後の空間の後の全ての文字を抽出する。自動的に空白を挿入するために、語彙は A^+ 内の全てのテキスト列を考慮し、ここで、 $A^+ = A \cup A_{space}$ であり、 A は単一の英語文字を含むテキスト文字列のセットであり(“a”, “b”, “c”, . . . , “z”)、 A_{space} は A と同じセットであるが、各文字の後に空白文字が付加されている(“a”, “b”, “c”, . . . , “z”)。 $P(c|x_i)$ に等しい空白を有する文字 c の確率(空白を含まないその文字の確率)を設定する。ここで、関数 $W(l)$ が、各空白で文字 l の系列をセグメント化し、最後の空箱に

後続する任意の文字を切り捨て、 l 内の完成した単語のリストを生成するものとする。 $p_{lm}(W(l^+) | W(l))$ が、 $n-1$ 個の選好する単語を所与とした l^+ の最後の単語の確率を与え、 N グラム言語モデルの使用を可能にするものとする。ビーム探索で考慮される文字の確率閾値は、 10^{-3} に設定された。 B は、ビーム幅（ビーム探索で使用されるビームの数）である。

【数 4 9】

```

sents =  $\{(\emptyset, 0)\}$ 
for  $i = 1, \dots, T$  do
  new_sents =  $\{\}$ 
  for  $\ell, \log p(\ell) \in \text{sents}$  do
    for  $c \in (A^+)$  do
      if  $p(c|x_i) < 0.001$  then
        次の文字に進む
      end if
       $\ell^+ \leftarrow c$  を  $l$  に追加する
      if  $w_{\text{final}}(\ell^+) \in \mathcal{V}_{\text{partial}}$  then
        if  $c \in A_{\text{space}}$  then
           $\log p(\ell^+) \leftarrow \log p(\ell) + \alpha \log p_{lm}(W(\ell^+) | W(\ell)) + \log p(c | x_i) +$ 
             $\beta \log |W(\ell^+)| - \beta \log(\max(1, |W(\ell)|))$ 
        else
           $\log p(\ell^+) \leftarrow \log p(\ell) + \log p(c | x_i)$ 
        end if
         $(l^+, \log p(l^+))$  を new_sents に加える
      end if
    end for
  end for
  sents  $\leftarrow$  new_sents 内の  $B$  個の最も可能性のあるプレフィックス
end for
sents 内の最も可能性のあるプレフィックスを返す

```

【0 4 8 4】

分類器と同様に、非同期ハイパーバンド法 [15] を Ray パッケージ [16] で使用し、Hyperopt を使用して各反復後の次のハイパーパラメータセットを提案した。500 個のハイパーパラメータセットを探索し、リアルタイム文スペリング評価の最初の日に使用するのに最適な単語誤り率を生成するセットを選択した。評価の最初の日の後、その日の間に収集されたデータのみを使用してハイパーパラメータ最適化手順を再実行した。進行中の全てのリアルタイム文スペリング評価セッション中に、この 2 回目の最適化実行中に見つかったハイパーパラメータ値を使用した。

【0 4 8 5】

ビームエッジのない事例

リアルタイム評価セッション中に記録されたコピータイピング文スペリング試行のうちの 3 つについて、ビーム探索は有効な文を使い果たした。これは、参加者が、有効な文候補を作成することができる文字系列がビーム探索による考慮のための閾値を超えないような間違いを犯した場合に発生した。

【0 4 8 6】

リアルタイム評価セッションの最初の日に、これが発生した場合、神経分類器から取得された最も可能性の高い文字を単純に出力する（空白なし）。リアルタイム評価の 2 日目の前に、その時点で（ビーム探索に有効な文候補を含まなくなる直前に）最も可能性の高

い文候補を出力するようにビーム探索アルゴリズムを修正し、その後、試験の残りの部分で神経分類器から取得した最も可能性の高い文字を出力した。追加的に、リアルタイム評価セッションの最初の日について、ビーム探索で考慮される文字の確率閾値（アルゴリズム 1 を参照）を 10^{-3} に設定した。リアルタイム評価の 2 日目について、閾値を同じままにしたが、ビーム探索アルゴリズムを修正して、3 つ未満の文字（及び空白付きの対応物）が確率 $> 10^{-3}$ を有する場合、有効なビームがなくなるのを避けるために、13 個の最も可能性の高い文字（及び空白付きの対応物）を考慮した。

【0487】

セクション S 5 . 言語モデリング N グラムモデリング

ビーム探索プロセス中、各ビームを新しい文字で更新していたときに、信頼性が高いと同時に、大規模なニューラルネットワークベースの言語モデルよりも迅速に予測を生成することも可能であるため、トリグラム言語モデルを使用した。

【0488】

基本的な N グラム式は、次のように、位置 k における単語 w_k の確率を有すると定義される。

【数 5 0】

$$p(w_k \mid w_{k-1}, \dots, w_{k-n+1}) = \frac{C(w_k, \dots, w_{k-n+1})}{C(w_{k-1}, \dots, w_{k-n+1})}, \quad (\text{S12})$$

式中、 C は、各 N グラムがコーパス内で発生する回数をカウントする関数である。

【0489】

改善された N グラムモデリングは、バックオフ及び割引によって達成することができる【19】。高次の N グラムは疎であり得るため、バックオフは、より高次の N グラムの確率を推定するために、より低次の N グラムモデルを使用することを指す。N グラム確率

【数 5 1】

$$p(w_i \mid w_{i-n+1}^{i-1})$$

は、式 S 1 3 に示されるように、より低次の N グラム

【数 5 2】

$$p(w_i \mid w_{i-n+2}^{i-1})$$

に直接依存する（すなわち、トリグラム確率は、バイグラム及びユニグラム確率に依存する）。割引は、N グラム確率を計算する前に定数が各 N グラムのカウントから除去され、

このように除去された確率質量が重み付けされた、より低次 N グラムモデルを通じて再分配される N グラム確率分布の正規化の一形態である。更なる詳細については、[20]を参照されたい。

【0490】

以下の定式化を使用して、割引によるバックオフを実装した。

【数53】

$$p(w_i | w_{i-n+1}^{i-1}) = \frac{\max(C(w_{i-n+1}^i) - \delta, 0)}{\sum_{w_i} C(w_{i-n+1}^i)} + \alpha(w_{i-n+1}^{i-1})p(w_i | w_{i-n+2}^{i-1}). \quad (\text{S13})$$

10

【0491】

ここで、 α は割引係数であり、

【数54】

$$\alpha(w_{i-n+1}^{i-1})$$

20

は次のように定義される。

30

【数55】

$$\alpha(w_{i-n+1}^{i-1}) = \frac{\delta N_{1+}(w_{i-n+1}^{i-1})}{\sum_{w_i} C(w_{i-n+1}^i)}, \quad (\text{S14})$$

式中、 N_{1+} は、先行する $n-1$ 個の単語の後に現れる一意の単語の数（式 S13 において、 $\max$ がゼロでないものを選択する回数）を表す。

【数56】

40

$$\sum_{w_i} C(w_{i-n+1}^i) = 0$$

であるときはいつでも、0 除算を避けるために、より低次のモデル確率を直接使用する。

50

【0492】

また、Kneser-Ney平滑化（[21]）を使用して、S13に暗示されているユニグラムモデルを改善し、単語が発生する、明確に異なる文脈タイプの数を表す単語増殖（word fertility）に置き換えた。単語文脈増殖を使用して、次の割合を書くことができる。

【数57】

$$p(w) \propto |\{w' : C(w', w) > 0\}|, \quad (S15)$$

10

式中、 $w \cdot$ は単語増殖であり、 $|\cdot|$ は濃度演算を指す。

【0493】

ここで、ユニグラムモデルを次のように書き換えることができる。

【数58】

$$p(w) = \frac{|\{w' : C(w', w) > 0\}| + \alpha_{kn}}{\sum_{w \in V} |\{w' : C(w', w) > 0\}| + N\alpha_{kn}}, \quad (S16)$$

20

式中、 V は、訓練語彙内の単語の集合であり、 N は、語彙内の単語の総数であり、 α_{kn} は、見えない単語が0の確率を有し、頻度の低い単語が過度に重いペナルティを受けることを防止する平滑化ハイパーパラメータである。実際には、固定割引係数 $\delta = 0.9$ 、固定Kneser-Ney平滑化係数 $\alpha_{kn} = 0.003$ を定義した。

【0494】

言語モデルを訓練するために、nltkのTwitterコーパス[22]及びCornell moviesコーパス[23]の2つのコーパスを使用した。それらの発話内容のカジュアル性及び会話性を考慮して、これら2つのコーパスを選択した。任意の所与の語彙を用いて、その語彙からの単語のみから構成された両方のコーパスからの全てのトリグラムに関してNグラムモデルを訓練した。訓練の前に、推論中の文開始のモデリングを可能にするために、両方のコーパス内の各文の開始前に2つの文開始トークンを挿入した。

30

【0495】

文最終化言語モデル

文スペリング中の最終化後の文をスコアリングするために、OpenAIのGPT-2言語モデル[24]に基づいているが、パラメータがより少ないDistilGPT-2ニューラルネットワークベース言語モデル[18]を使用した。

【0496】

補足参考文献

1. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32. Ed. by Wallach H, Larochelle H, Beygelzimer A, d'Alch'e-Buc F, Fox E, and Garnett R. Curran Associates, Inc., 2019: 8024-35.

40

2. Moses DA, Metzger SL, Liu JR, et al. Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria. New England Journal of Medicine 2021; 385: 217-27.

50

3. Kingma DP and Ba J. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 2017.

4. Bergstra J, Yamins DLK, and Cox DD. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. Icml 2013: 115–23.

5. Cho K, Van Merriënboer B, Bahdanau D, and Bengio Y. On the properties of neural machine translation: Encoder–decoder approaches. arXiv preprint arXiv:1409.1259 2014.

6. Hochreiter S and Schmidhuber J. Long short-term memory. Neural computation 1997; 9:1735–80.

7. Chung J, Gulcehre C, Cho K, and Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 2014.

8. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, and Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580 2012.

9. Kingma DP and Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.

10. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32. Ed. by Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, and Garnett R. Curran Associates, Inc., 2019:8024–35. (papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf)

11. Krizhevsky A, Sutskever I, and Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in neural information processing systems 2012; 25:1097–105.

12. Reed CJ, Metzger S, Srinivas A, Darrell T, and Keutzer K. SelfAugment: Automatic augmentation policies for self-supervised learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:2674–83.

13. Willett FR, Avansino DT, Hochberg LR, Henderson JM, and Shenoy KV. High-performan

10

20

30

40

50

ce brain-to-text communication via handwriting. *Nature* 2021;593:249–54.

14. Moses DA, Metzger SL, Liu JR, et al. Neuroprosthesis for decoding speech in a paralyzed person with anarthria. *New England Journal of Medicine* 2021;385:217–27.

15. Li L, Jamieson K, Rostamizadeh A, et al. Massively parallel hyperparameter tuning. 2018.

16. Moritz P, Nishihara R, Wang S, et al. Ray: A distributed framework for emerging {AI} applications. In: 13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18). 2018:561–77.

17. Hannun AY, Maas AL, Jurafsky D, and Ng AY. First-pass large vocabulary continuous speech recognition using bi-directional recurrent dnns. *arXiv preprint arXiv:1408.2873* 2014.

18. Sanh V, Debut L, Chaumond J, and Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. 2020. *arXiv:1910.01108 [cs.CL]*.

19. Chen SF and Goodman J. An empirical study of smoothing techniques for language modeling. *Computer Speech & Language* 1999;13:359–94.

20. Jurafsky D and Martin JH. Speech and language processing. Vol. 3. US: Prentice Hall 2014.

21. Kneser R and Ney H. Improved backing-off for m-gram language modeling. In: 1995 international conference on acoustics, speech, and signal processing. Vol. 1. IEEE. 1995:181–4.

22. Bird S, Klein E, and Loper E. Natural language processing with Python: analyzing text with the natural language toolkit. "O'Reilly Media, Inc.", 2009.

23. Danescu-Niculescu-Mizil C and Lee L. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. In: Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, ACL 2011. 2011.

24. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I, et al. Language models are unsupervised multitask learners. *OpenAI blog* 2019;1:9.

25. Romero DET and Jovanovic G. Digital FIR Hilbert Transformers: Fundamentals and Efficient Design Methods. In: MATLAB—A Fundamental Tool for Scientific Computing and Engineering Applications—Volume 1. 2012: 445–82. (intechopen. com /books/matlab—a-fundamental-tool-for-scientific-computing-and-engineering-applications-volume-1 /digital-fir-hilbert-transformers-fundamentals-and-efficient-design-methods)

10

26. Welford BP. Note on a Method for Calculating Corrected Sums of Squares and Products. Technometrics 1962; 4: 419–9.

27. Moses DA, Leonard MK, Makin JG, and Chang EF. Real-time decoding of question-and-answer speech dialogue using human cortical activity. Nature Communications 2019; 10: 3096.

20

30

40

50

【表 4 - 1】

補足表 S 1. コピータイピングタスク文

標的文	1 回目の試験で復号された文	2 回目の試験で復号された文	
good morning	good morning	good for legs	
you have got to be kidding	you have got to be kidding a	you have got to be kidding	
what do you mean	what do you mean	what do you mean	
good to see you	i do i leave you	good to see you	
i think this is pretty good	i think this is pretty good	i think they is pretty good	
i will check	i will check	i will the it	
thank you	thank you	thank you	
please sit down	please sit down	please believe	10
we have to stop	we have to stop	we have to stop	
hand that to me please	hand that time please	have that time always	
i know what you mean	i know what you mean	i know what you mean	
what time is it	what time is it	what time is it	
sit over here with me	sit over here with me	sit over here with me	
no thanks	no thanks	not happen	
you never know	you never know	you never know	
great to see you again	great to show my case in	great to stay in town	
forget about it	forget about it	forget about it	
could you repeat what you said	dog lie on repeat what you said	could you repeat what you said	
where do you live	where do you live	where do you live	20
do not be afraid to ask me	do not be afraid to ask me	do not be afraid to ask me questions	
questions	questions		
i cannot believe it	i can not believe it	i can not believe it	
thanks for telling me	thank for reading me	thanks for telling me	
i do not want that	i do not want that	i do not want that	
that is wonderful	that is work from a	that is wonderful	
what do you think about that	what do you think about that	what do you think about that	
thank you very much	though it very much	thank you very much	
i am glad you are here	i am glad you are here	i am glad you are here	
how are you doing	how are you doing	how are you doing	
i agree	i agree	i agree	
i am okay	i am okay	i am okay	30
tell me what you are doing	tell me what your telling	tell me what you are doing	
how long did it take	how long did it take	how long did it take	
is there anything i can do	is there a nothing i can do	is there anything i can do	
how are things going for you	how are things gives for you	how are things going for you	
do you know what he did	do you know on the ice	do you know what he did	
was there something else	was there to be a high else	was there something else	
where are you going	while are you doing	where are you going	
who is that	who is that	why is that	
tell me about your family	tell me about your family	tell me about your family	
i could probably do better	i could probably do better	i could probably do better	
you can say that again	you can say that again	you can say that open	
i am sorry to hear that	i am to get to hear that	i am sorry to hear that	40

will i see you later i am doing well can that wait until another time how much more is there come talk with me that will be fun how often do you do this how much will it cost bring that over here turn it off i remember the last time i did that i was just kidding i will meet you there i do not really remember i feel cold excuse me for interrupting you are not going to believe this do you understand what i mean what are you talking about which one is it would you like to go with me i do not understand of course i do anything is possible do not do that again let me see that what have you been doing i had a great time easy for you to say i want to go how do you feel that is all right i told you i do not know	will i see you later i am doing well can that wait until another time how much more is there come talk with me that will be fun how often do you do this how much will it cost clinic hat for hat turn it off i remember the last time i did that i was mike kidding i will meet you there i do not really remember i feel weird excuse me for interrupt any you plan to go in on a bit love this do you understand what i mean what are you talking about which one edit a all i was like the white me i do not understand of course its anything is possible do not do that again let me see that what have you been doing i had a great time easy for you to say i want to go how do you feel that is all right i told you i do not know	well i keep by later i am doing fine can that wait until another time how much were in there some take with me that will be fun how often do you do this how much will it cost bring that ever here turn it off i remember to plan new me i did that i was just kidding i will meet you to eat ddonoyrballyrlrefbhrh i feel cold excuse me for interrupting ypuaranpdggingloavlinesoeb do you understand what i mean what are you talking about which one is it would you like to go with me i do not understand of course him anything is possible do not do that again let me see that what have you been doing what a great time easy for you to say i want to go how do you feel that is all right i told you i do not know
--	---	--

【表 5】

補足表 S 2. ハイパーパラメータの定義及び値

モデル	ハイパーパラメータの説明	探索空間タイプ ¹	値範囲	最適値 ²
発話	平滑化サイズ	均一(整数)	[1,80]	78
検出器	確率閾値	均一	[0.1,0.9]	0.304
	時間閾値持続時間	均一(整数)	[25,150]	105
単語	GRU 層の数	均一(整数)	[1,4]	2
分類器	GRU 層当たりのノード	均一(整数)	[128,512]	274
	ドロップアウト分数	均一	[0.3,0.8]	0.545
	畳み込みカーネルのサイズ及びスキップ	均一(整数)	[1,10]	4
	ジッタ量(秒)、j	均一	[0.0,2.0]	0.474
	加法ノイズレベル、 σ_n	均一	[0.0,1.0]	0.0027
	スケール最小値、 $\alpha_{\min}$	均一	[0.8,1.0]	0.955
	スケール最大値、 $\alpha_{\max}$	均一	[1.0,1.2]	1.07
	最大時間マスキング長(秒)、b	均一	[0.00,1.35]	0.871
	時間マスキング確率、p	均一	[0.0,0.5]	0.0478
	チャンネルごとのノイズ、 σ_c	均一	[0.0,1.0]	0.0283
ビーム	言語モデルスケール係数、 α	均一	[0.01,1.0]	(0.642,0.744)
探索	単語挿入重み、 β	均一	[0.0,30.0]	(4.03,10.5)
	維持されたビームの数、B	均一(整数)	[0,750]	(457,739)
	Distil-GPT2 スケール係数、 α_{gpt2}	均一	[0.0,100.0]	(1.53,1.13)

¹「均一(整数)」は、ハイパーパラメータ値が整数であることを強制されたことを示す。²言語モデリング及びビーム探索ハイパーパラメータについて、2つの値がリストされている。1つ目は、文スペリング評価の最初の日の前にコピータイピング文スペリング試験で最適化するときに見つかった最適値であり(この最初の日に使用される)、2つ目は、文スペリング評価の最初の日からのコピータイピング文スペリング試験で最適化するときに見つかった最適値である(2日目及びそれ以降の全ての日に使用される)。

【0497】

上記の発明は、明確な理解のために例示及び例によって多少詳しく説明されてきたが、当業者であれば、本発明の教示に照らして、添付の特許請求の範囲の趣旨又は範囲から逸脱することなく、それらの発明に対して特定の変更及び修正が行われ得ることは、容易に明らかである。また、本発明の範囲は、添付の特許請求の範囲によってのみ限定されることになるため、本明細書で使用される用語は、特定の実施形態のみを説明する目的のためのものであり、限定することが意図されるものではないことも理解されるべきである。

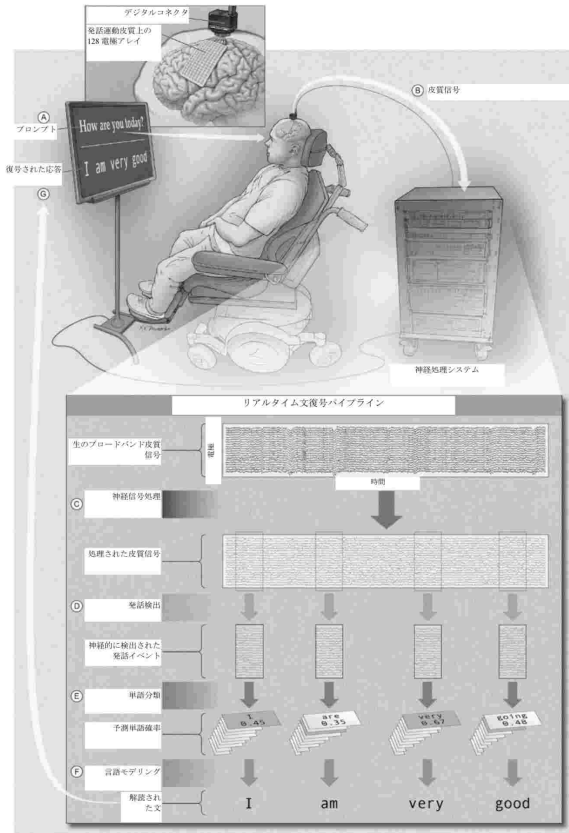
【0498】

したがって、上記は単に本発明の原理を例示するにすぎない。当業者は、本明細書に明示的に記載又は示されていないが、本発明の原理を具現化し、その精神及び範囲内に含まれる様々な配置を考案することができることが理解されるであろう。更に、本明細書に列挙される全ての例及び条件付き言語は、主に、読者が本発明の原理及び当該技術を更に進めるために発明者が寄与する概念を理解するのを助けることを意図しており、そのような具体的に列挙される例及び条件に限定されないと解釈されるべきである。更に、本発明の

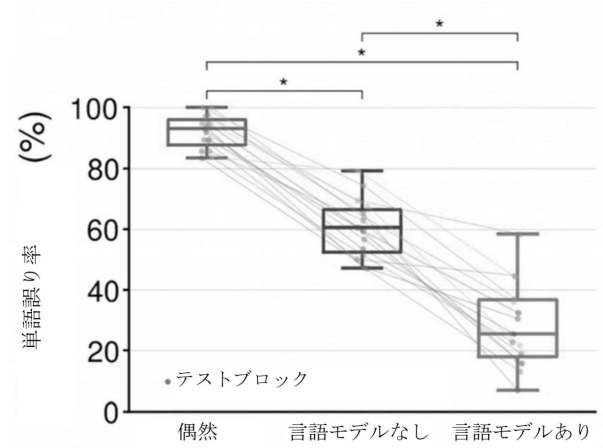
原理、態様、及び実施形態を記載する、本明細書の全ての記述、並びにそれらの具体例は、それらの構造的及び機能的等価物の両方を包含することが意図されている。追加的に、そのような等価物は、構造に関係なく、現在知られている等価物と、将来開発される等価物との両方、すなわち、同じ機能を実行するように開発された任意の要素を含むことが意図される。したがって、本発明の範囲は、本明細書に示され、説明された例示的な実施形態に限定されることを意図されていない。むしろ、本発明の範囲及び精神は、添付の特許請求の範囲によって具現化される。

【図面】

【図 1】



【図 2 A】



10

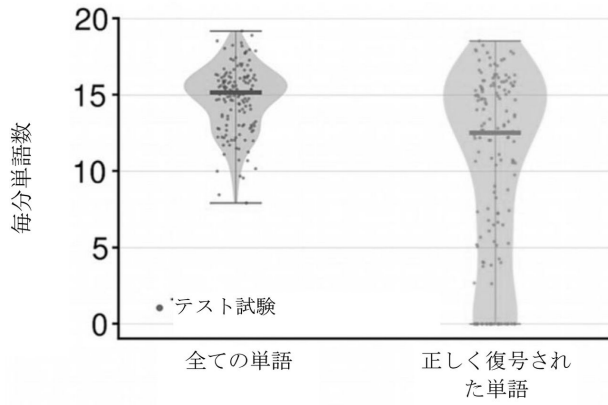
20

30

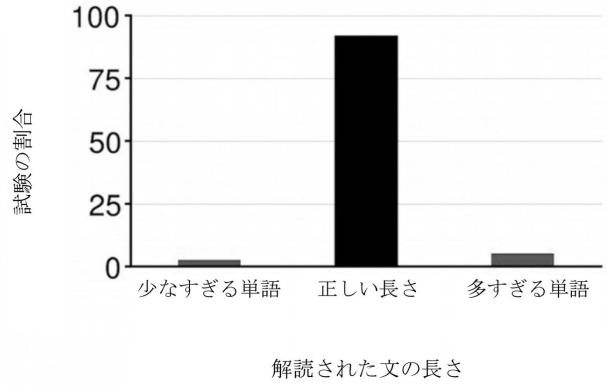
40

50

【図 2 B】



【図 2 C】



10

【図 2 D】



【図 2 E】

標的	言語モデルなし	言語モデルあり
例示的な試験	Hello how are you? I like my nurse. They are going outside. My family is very comfortable. Bring my glasses please. What do you do? How do you like my music?	Hungry how am you I right my nurse They are going outside Glasses family is faith comfortable Please my glasses please What do I you How do you like bad bring

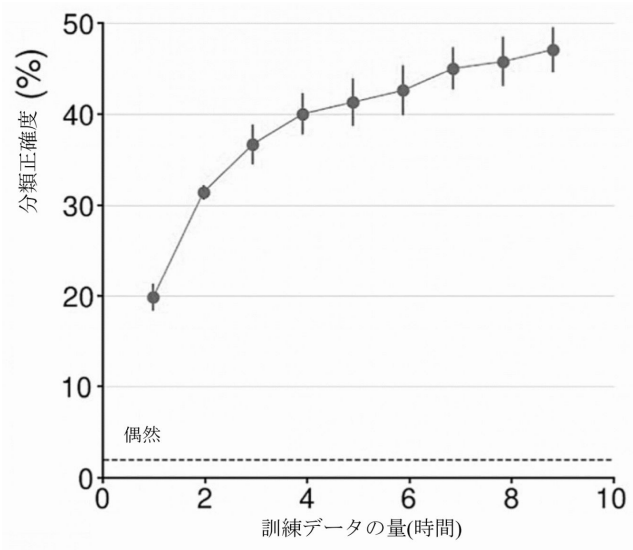
20

30

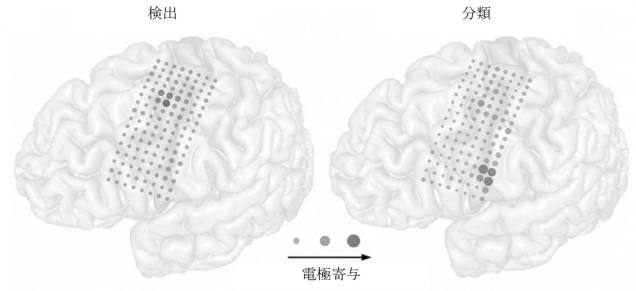
40

50

【図 3 A】

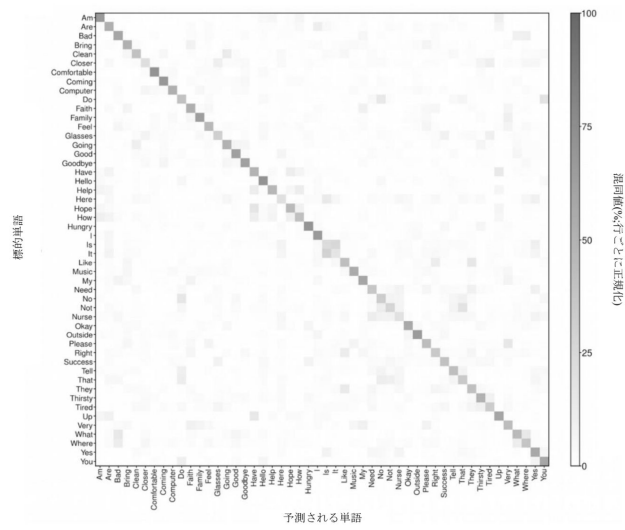


【図 3 B】

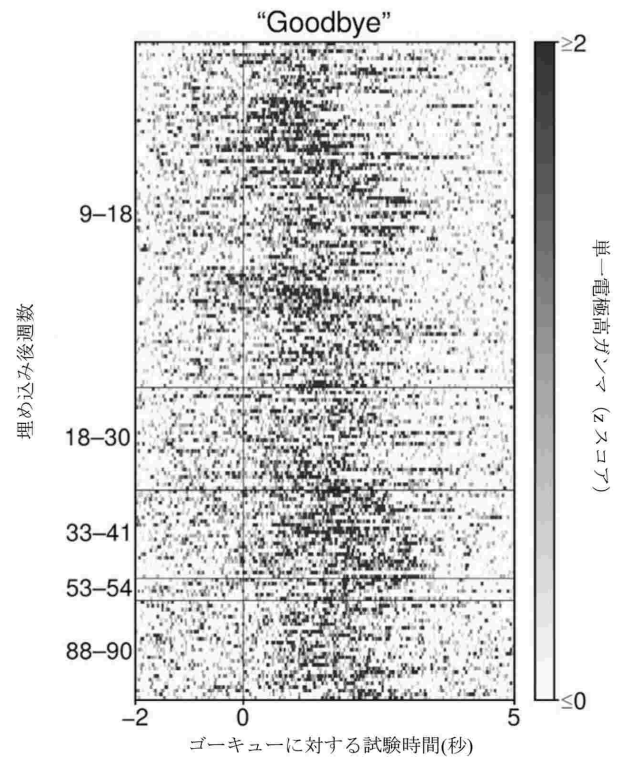


10

【図 3 C】



【図 4 A】



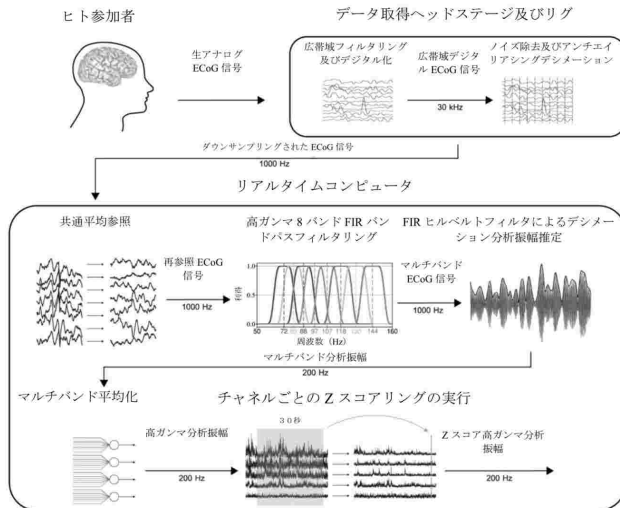
20

30

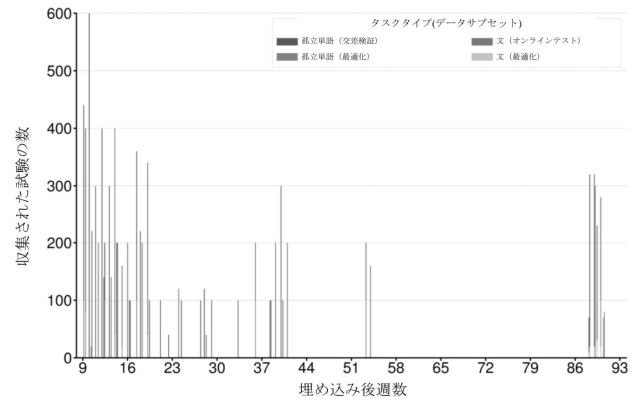
40

50

【図 7】

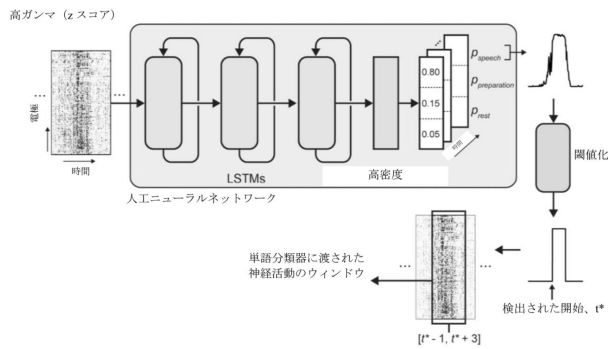


【図 8】

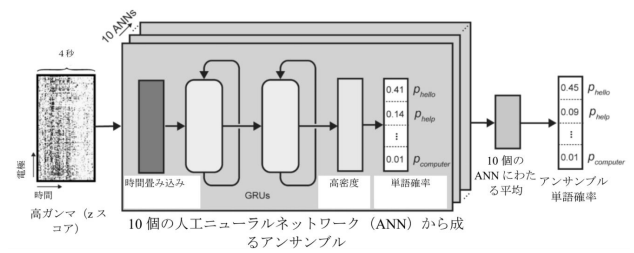


10

【図 9】



【図 10】



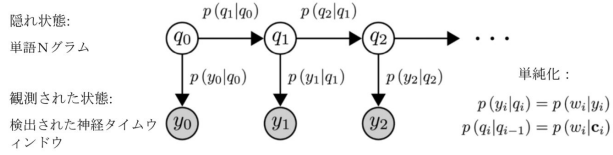
20

30

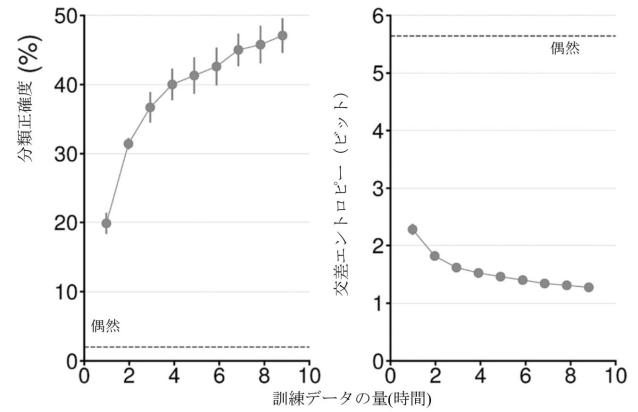
40

50

【図 1 1】

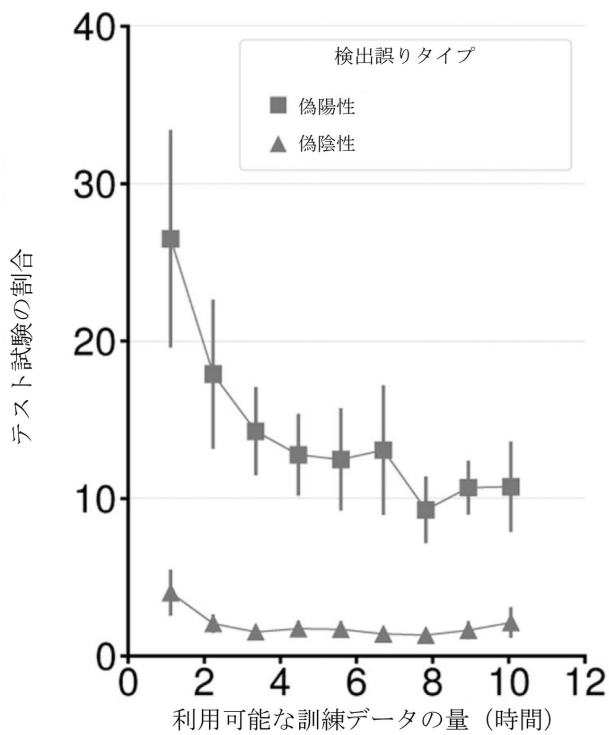


【図 1 2 A】

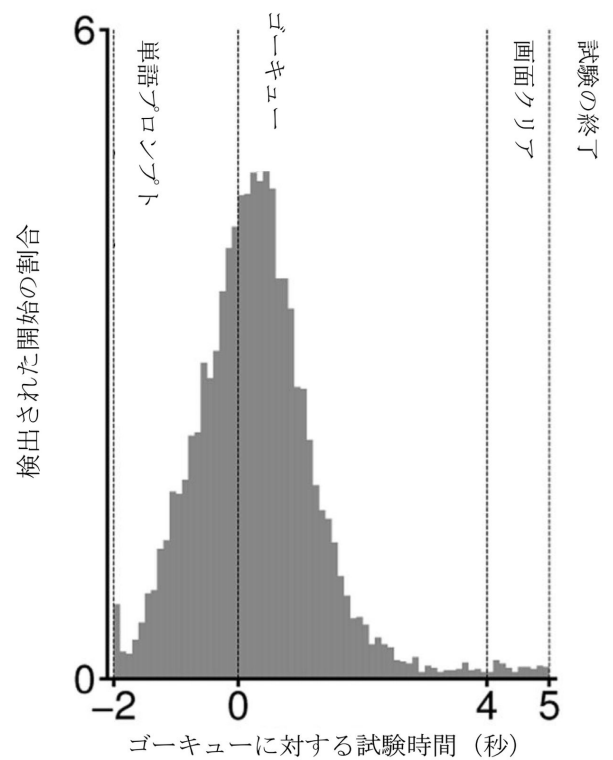


10

【図 1 2 B】



【図 1 2 C】



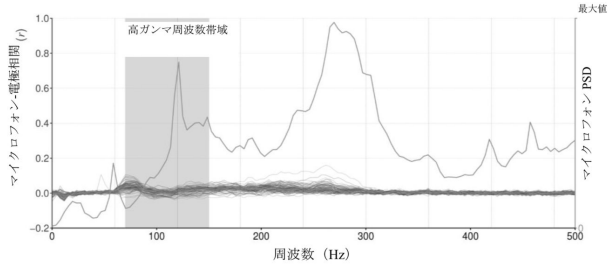
20

30

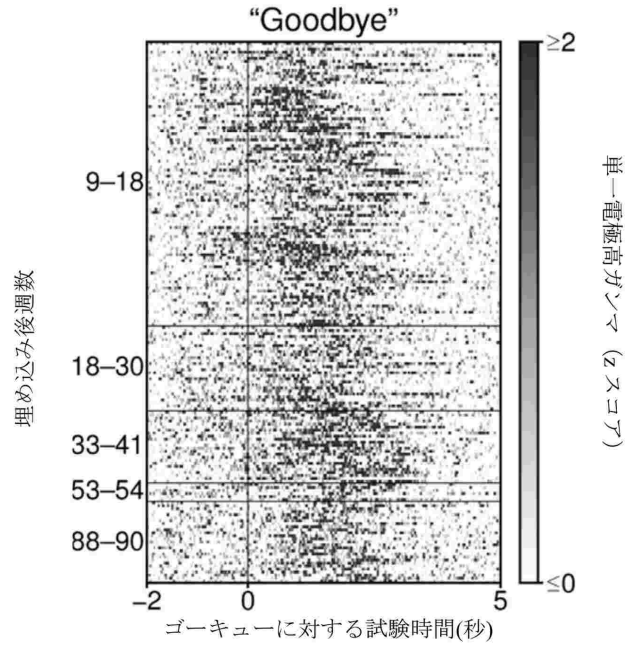
40

50

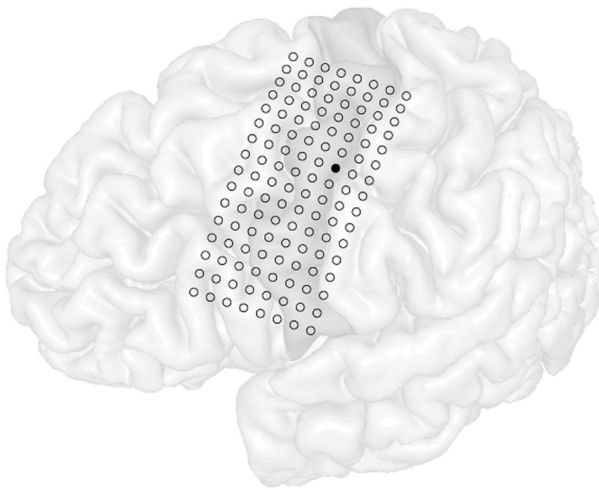
【図 1 3】



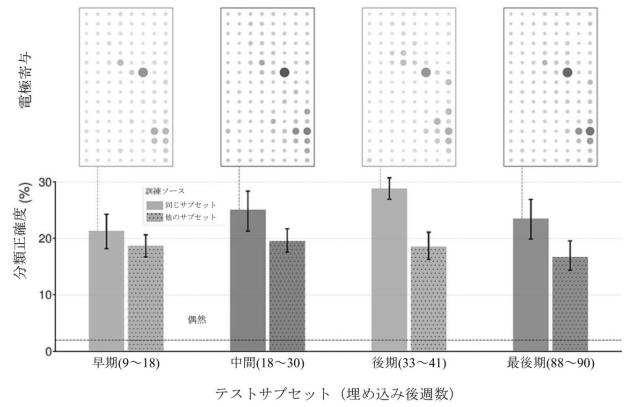
【図 1 4 A】



【図 1 4 B】



【図 1 4 C】



10

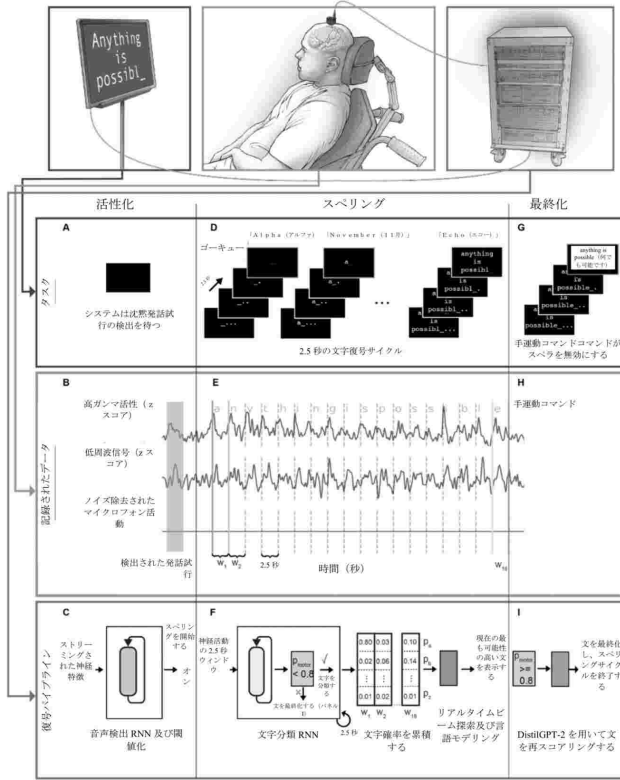
20

30

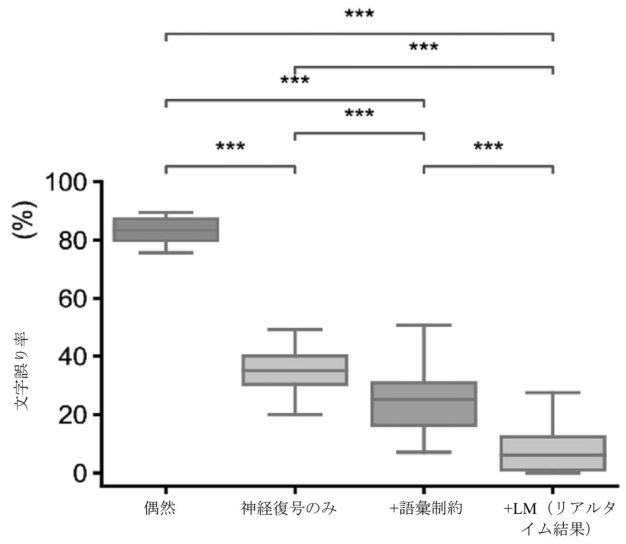
40

50

【図 15】



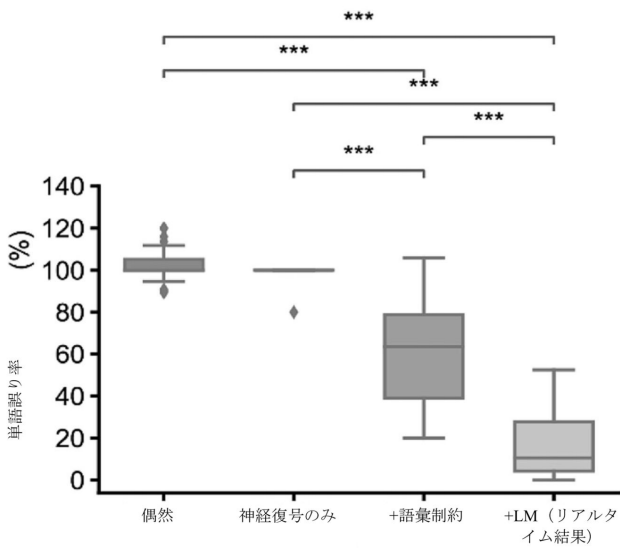
【図 16 A】



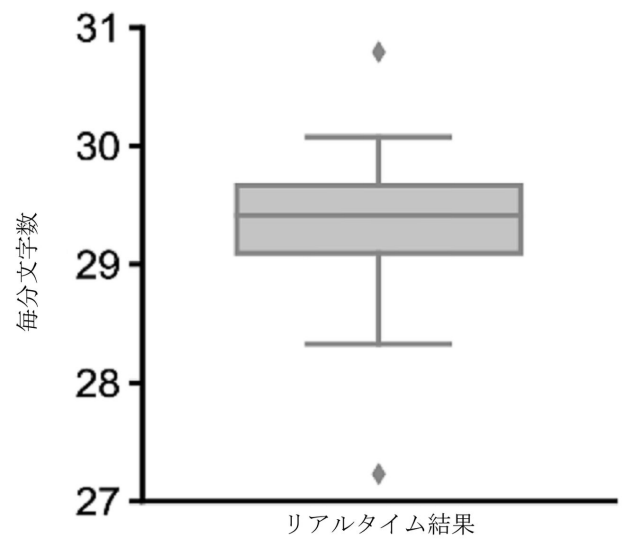
10

20

【図 16 B】



【図 16 C】

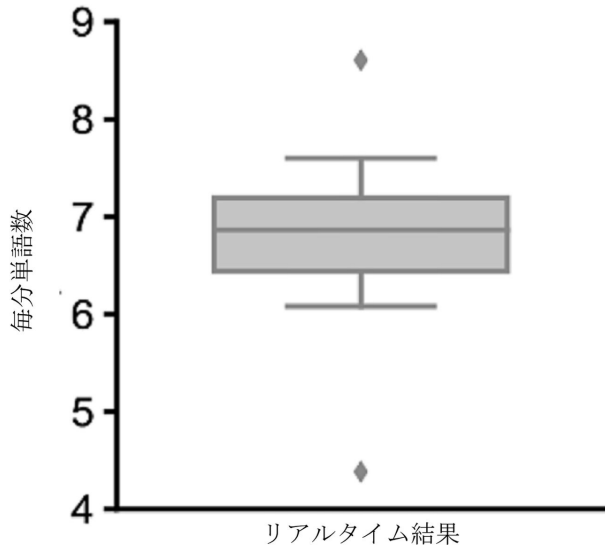


30

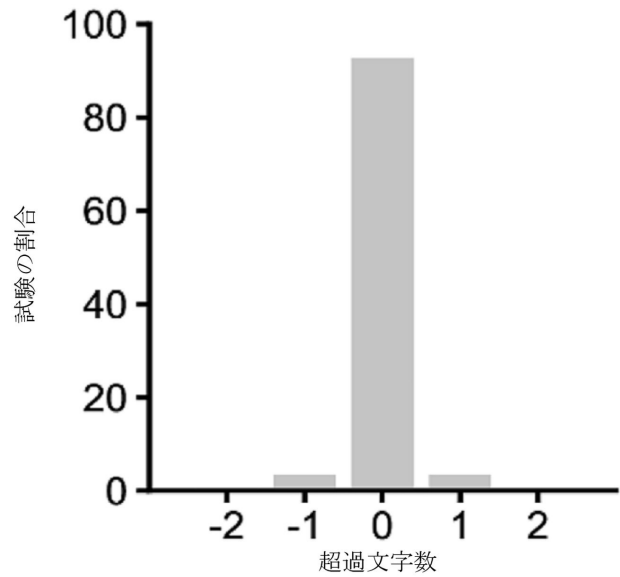
40

50

【図 16 D】



【図 16 E】



10

【図 16 F】

神経復号のみ

Idonotwantthat
Thankyu
Yeucanyaythatagain
Tellmeaboutyourfamily
Fonftoothatagain
Ithinkthiyipretygood
Easyforiouta say
Weavetotop
Youraigtobekvding
Whichoneit
Ithioeanrhengisanx

+語彙制約

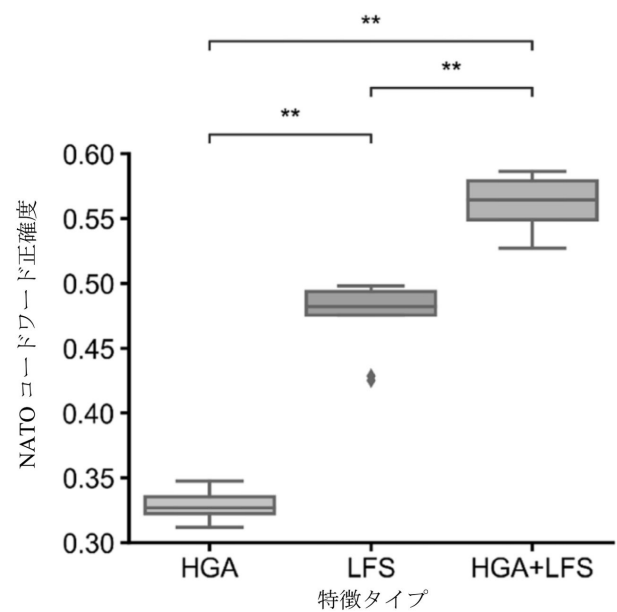
I do not want that
Thank you
You can say that again
Tell me about your family
Do no tooth at again
I think this it pretty good
Easy for i out a say
We have top top
You i a i got to be kidding
Which one it it
It there an a thing is any i

1.正しい文：Which one is it
2.正しい文：Is there anything I can do

+LM (リアルタイム結果)

I do not want that
Thank you
You can say that again
Tell me about your family
Do not do that again
I think this is pretty good
Easy for you to say
We have to stop
You have got to be kidding
Which one edit
Is there a nothing I can do²

【図 17 A】



20

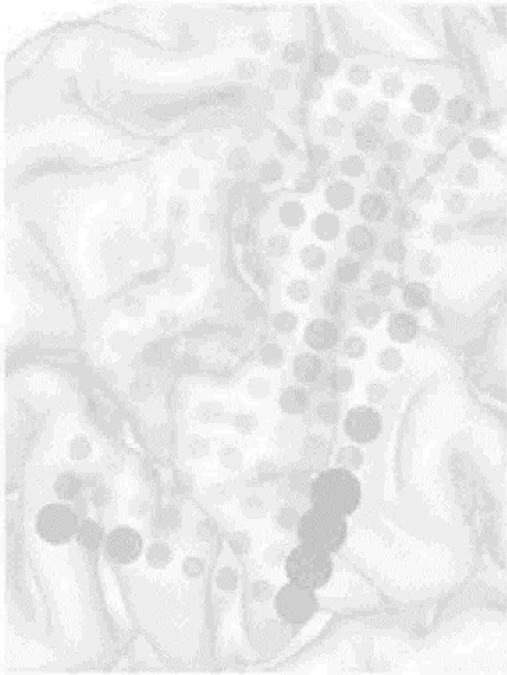
30

40

50

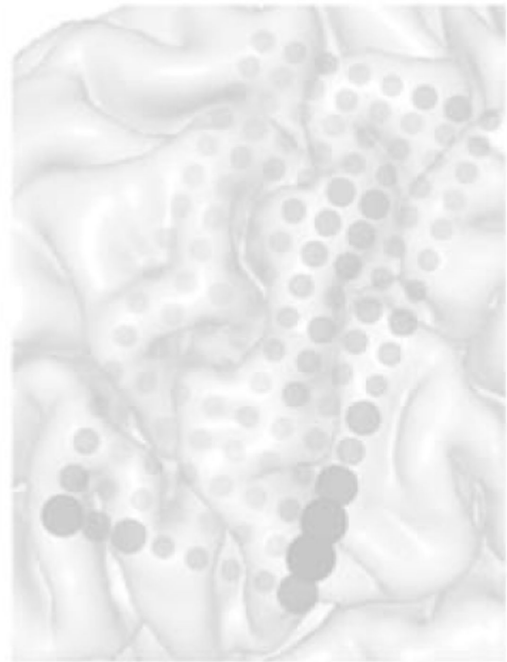
【図 17 B】

HGA 単独



【図 17 C】

HGA+LFS からの HGA

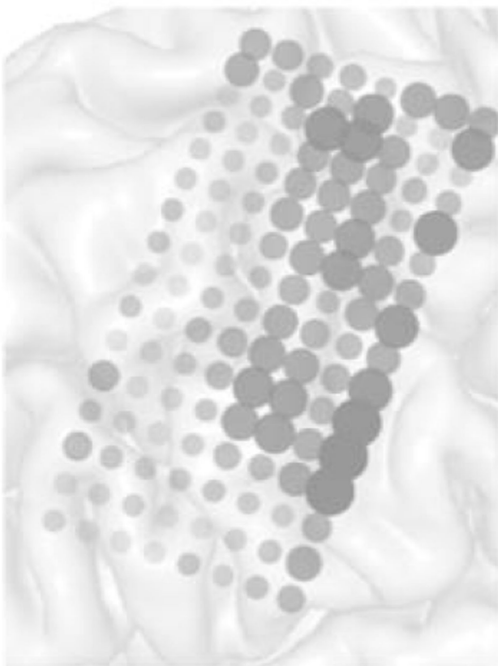


10

20

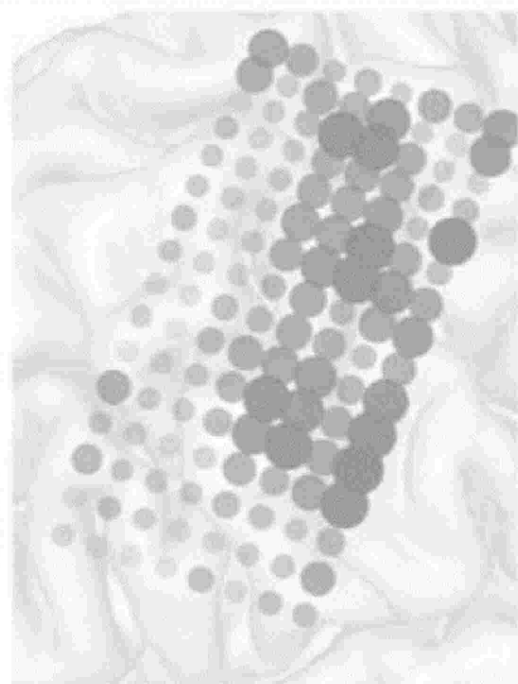
【図 17 D】

LFS 単独



【図 17 E】

HGA+LFS からの LFS

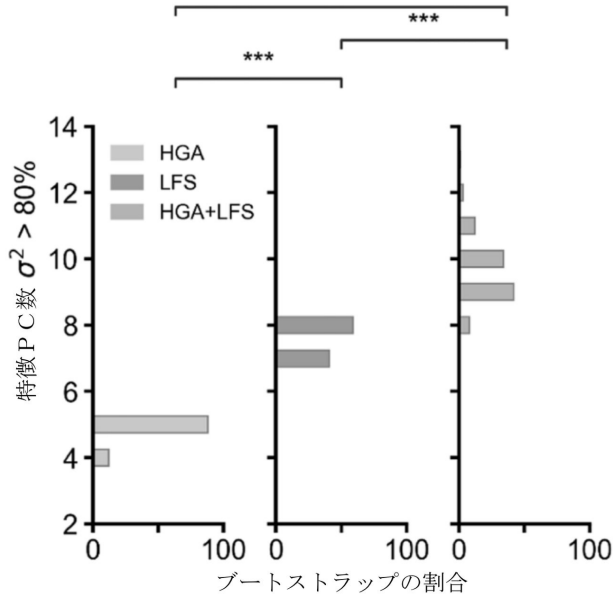


30

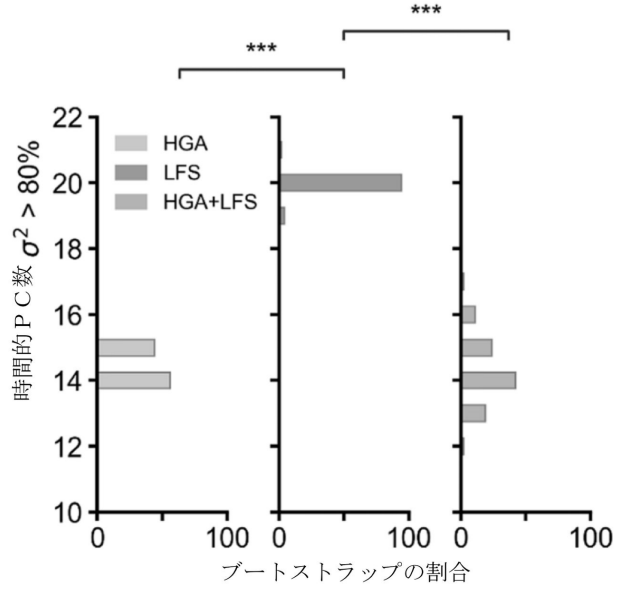
40

50

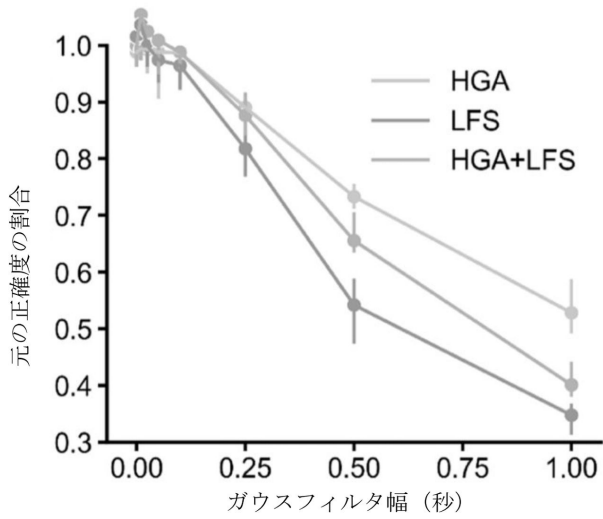
【図 17 F】



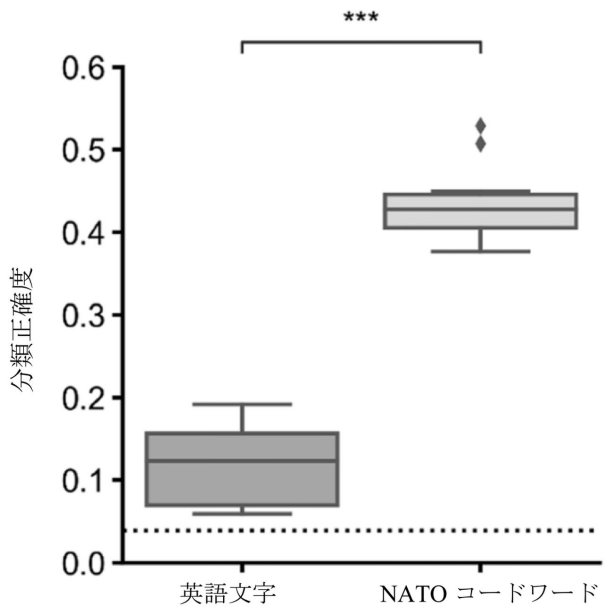
【図 17 G】



【図 17 H】



【図 18 A】



10

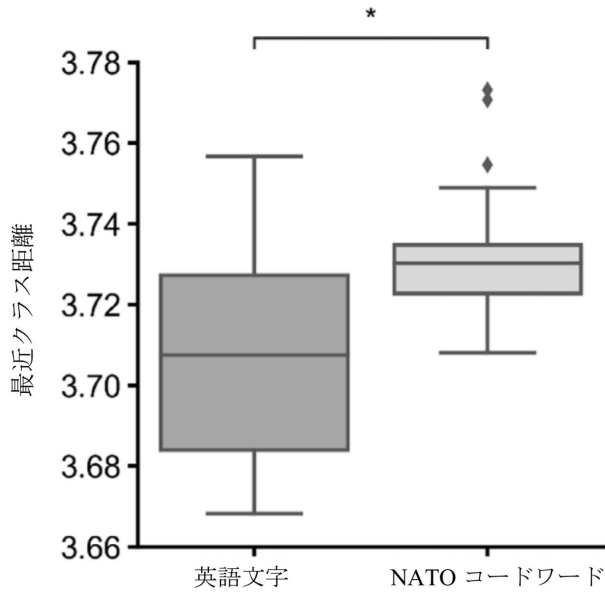
20

30

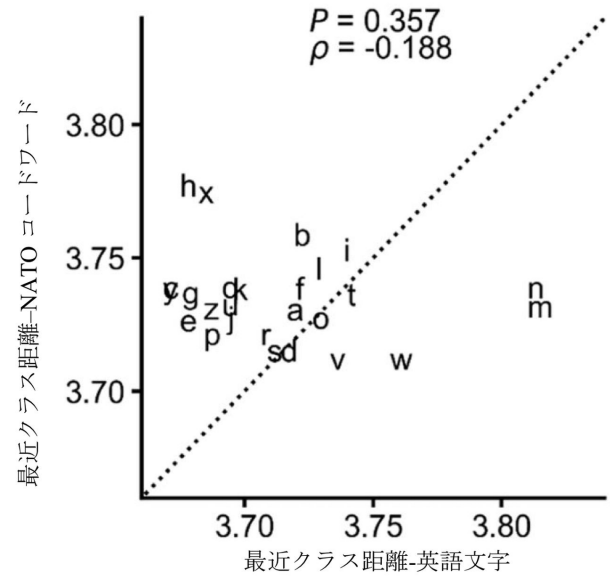
40

50

【図 18 B】

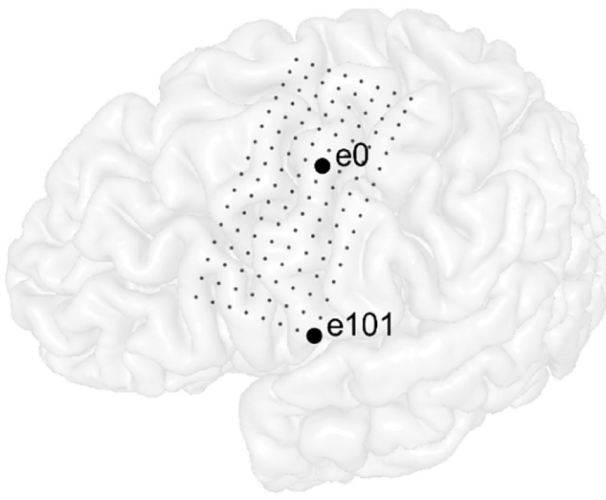


【図 18 C】

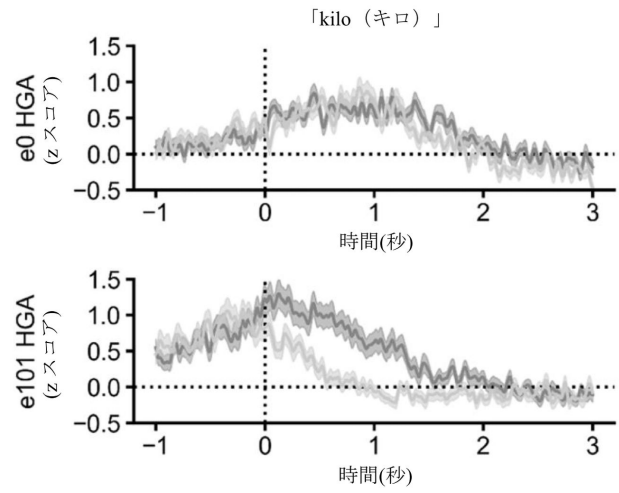


10

【図 19 A】



【図 19 B】



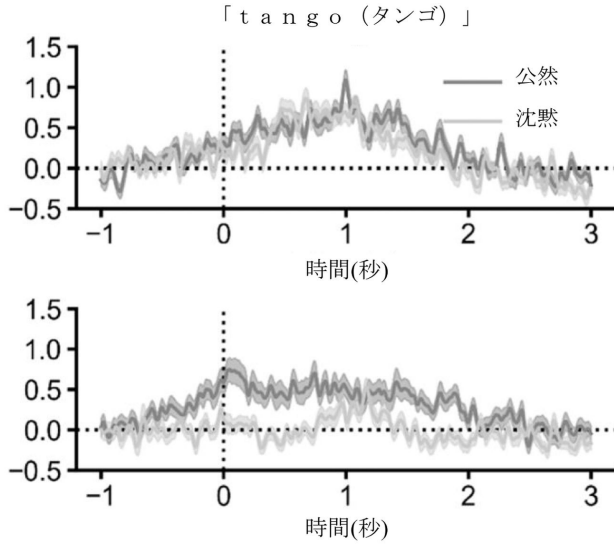
20

30

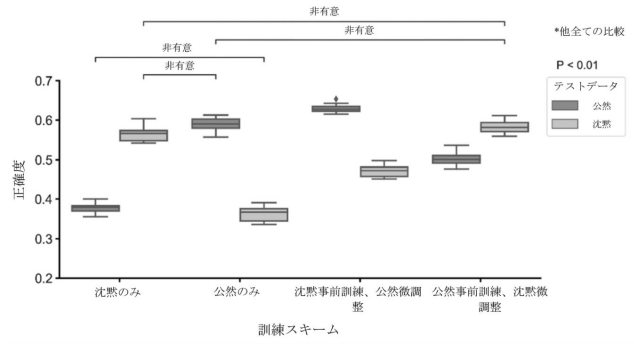
40

50

【図 19 C】

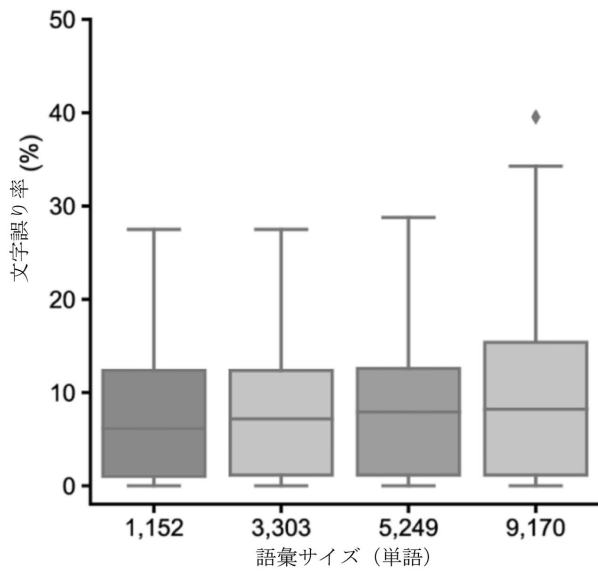


【図 19 D】

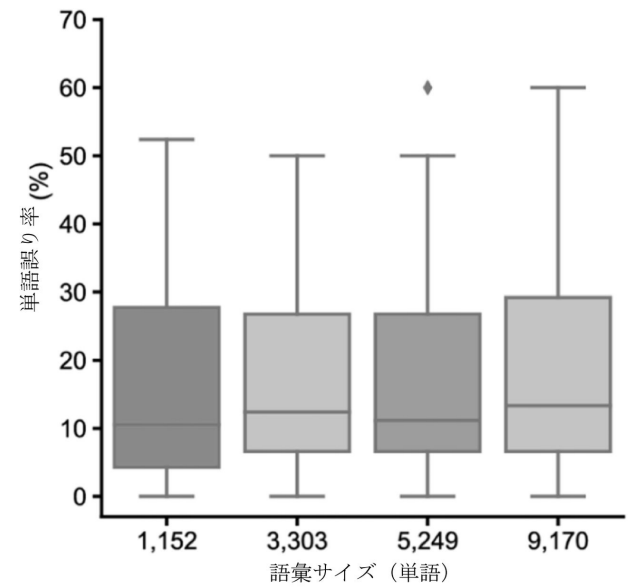


10

【図 20 A】



【図 20 B】



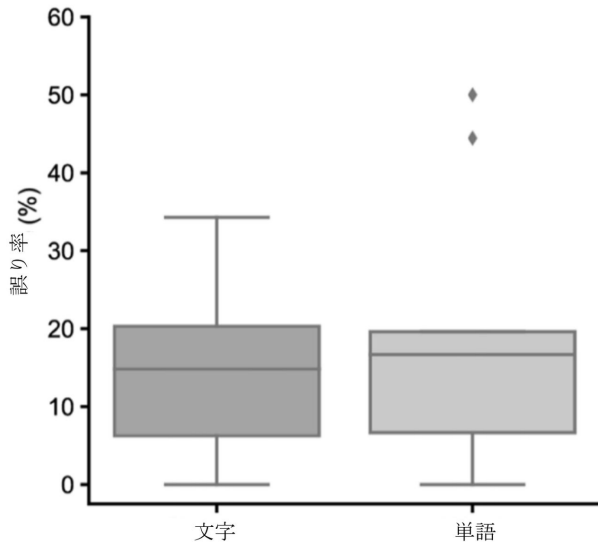
20

30

40

50

【図 20 C】

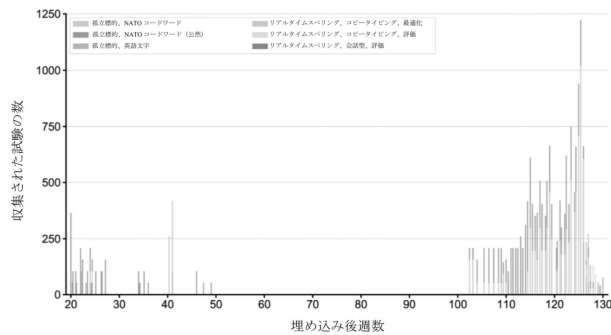


【図 20 D】

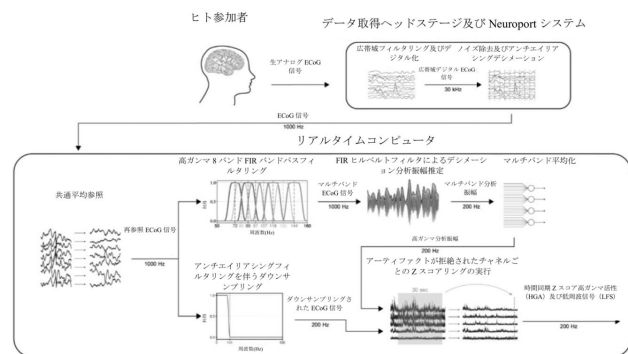
質問	回答 (脳活動から復号)
What is your favorite time of year? (一年の中で好きな時期はいつですか?)	Summer time (夏です)
Where do you live? (どこに住んでいますか?)	I live in California (カリフォルニア州に住んでいます)
What would you change about where you live? (あなたは住んでいる場所について何を変えたいですか?)	The food is terrible (食べ物がいひいですね)
What do you like to watch on TV? (テレビで何を見るのが好きですか?)	I like to watch a movie (映画を見るのが好きです)
(自由回答、参加者に質問は表示されない)	It was nice to see you (お会いできて嬉しかったです)

10

【図 21】



【図 22】



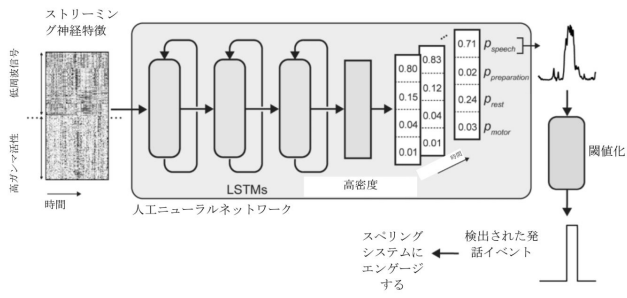
20

30

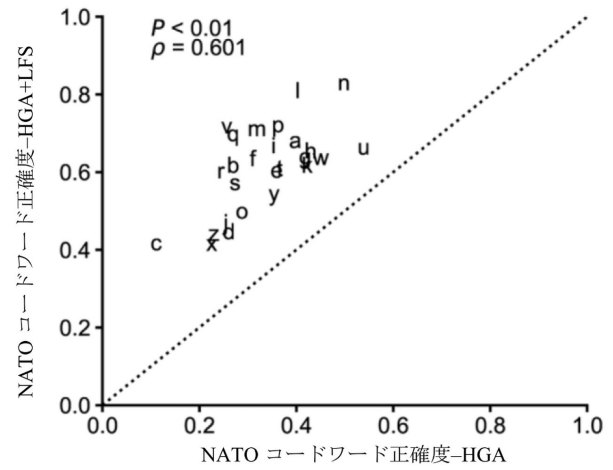
40

50

【図 2 3】

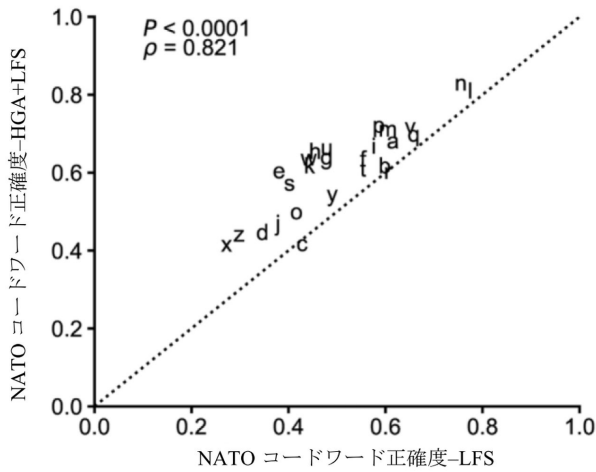


【図 2 4 A】

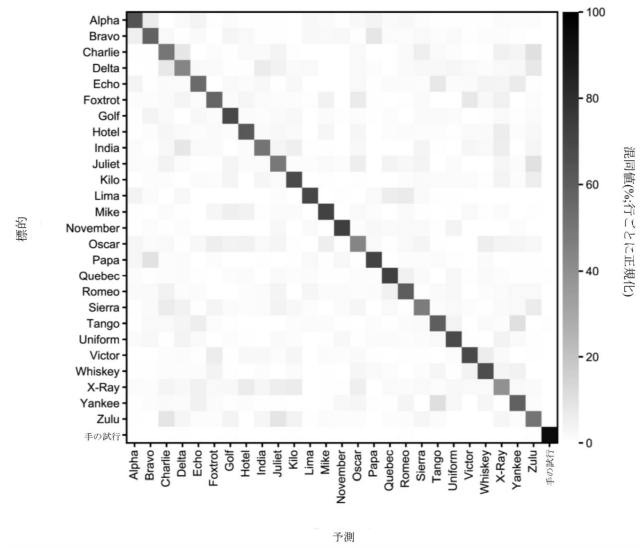


10

【図 2 4 B】



【図 2 5】



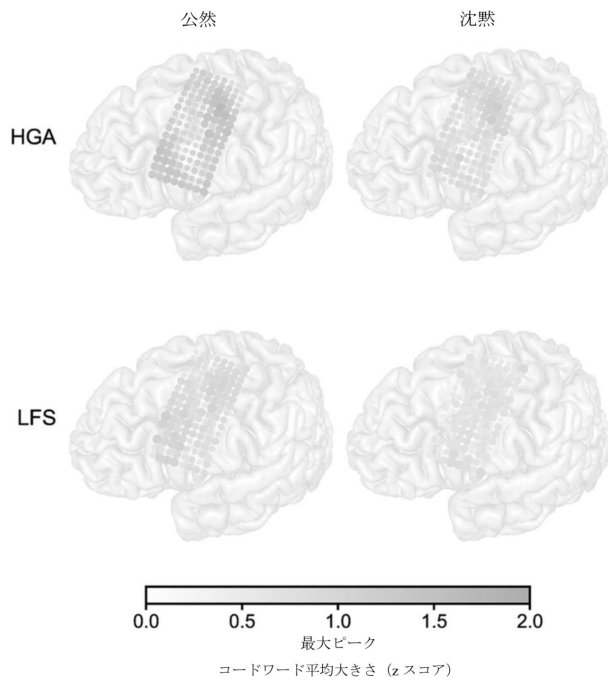
20

30

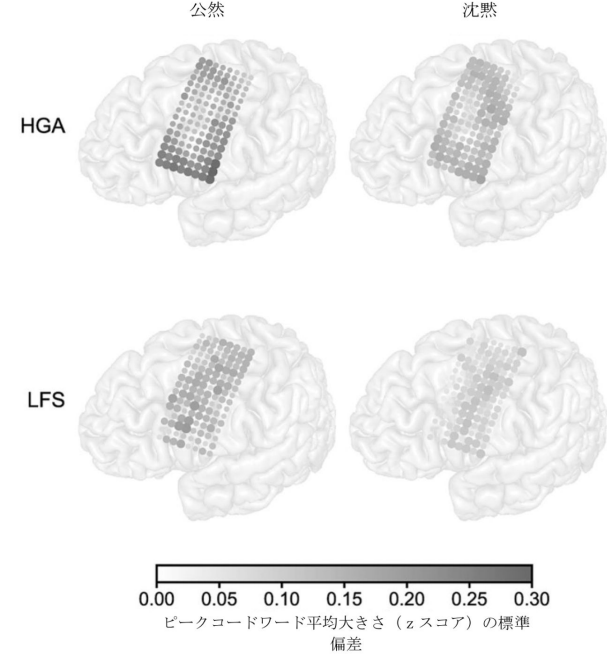
40

50

【図 2 6 A】



【図 2 6 B】



10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US22/31101
A. CLASSIFICATION OF SUBJECT MATTER IPC - INV. A61B 5/372; G10L 15/24 (2022.01) ADD. A61B 5/369; A61B 5/384; G06N 3/02; G06N 3/08; G10L 17/18; G10L 25/78 (2022.01) CPC - INV. A61B 5/372; G10L 15/24 ADD. A61B 5/369; A61B 5/384; G06F 3/015; G06N 3/02; G06N 3/08; G10L 17/18; G10L 25/78 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) See Search History document Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History document Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2015/0338917 A1 (SIA TECHNOLOGY LTD.) 26 November 2015; figures 1, 2, 4A, 4B; paragraphs [0005], [0006], [0137], [0142], [0158], [0160], [0162], [0178], [0196], [0223], [0321], [0322].	1-3, 39-43, 86-88, 124-127
Y	WO 2021/021714 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 04 February 2021; paragraph [00137], [00326], [00471].	1-3, 39-43, 86-88, 124-127
Y	US 2019/0351230 A1 (NEUVANA, LLC) 21 November 2019; paragraph [0387].	127/126/124-127/126/125
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 20 September 2022 (20.09.2022)		Date of mailing of the international search report OCT 05 2022
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300		Authorized officer Shane Thomas Telephone No. PCT Helpdesk: 571-272-4300

Form PCT/ISA/210 (second sheet) (July 2019)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US22/31101

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 4-38, 44-85, 89-123, 128-159
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

フロントページの続き

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. H D M I
2. L i n u x
3. P Y T H O N
4. T E N S O R F L O W
5. T W I T T E R

弁理士 坪倉 道明

(74)代理人 100129713

弁理士 重森 一輝

(74)代理人 100137213

弁理士 安藤 健司

(74)代理人 100143823

弁理士 市川 英彦

(74)代理人 100183519

弁理士 櫻田 芳恵

(74)代理人 100196483

弁理士 川崎 洋祐

(74)代理人 100160749

弁理士 飯野 陽一

(74)代理人 100160255

弁理士 市川 祐輔

(74)代理人 100219265

弁理士 鈴木 崇大

(74)代理人 100203208

弁理士 小笠原 洋平

(74)代理人 100146318

弁理士 岩瀬 吉和

(74)代理人 100127812

弁理士 城山 康文

(72)発明者 モーゼス、デヴィッド

アメリカ合衆国、カリフォルニア・94143、サン・フランシスコ、シックスティーンズ・ストリート・600、スイート・エス272・ジェネンテック・ホール・ボックス・2142、イノベーション・ベンチャーズ

(72)発明者 リュウ、ジェシー

アメリカ合衆国、カリフォルニア・94143、サン・フランシスコ、シックスティーンズ・ストリート・600、スイート・エス272・ジェネンテック・ホール・ボックス・2142、イノベーション・ベンチャーズ

(72)発明者 メッツガー、ショーン

アメリカ合衆国、カリフォルニア・94143、サン・フランシスコ、シックスティーンズ・ストリート・600、スイート・エス272・ジェネンテック・ホール・ボックス・2142、イノベーション・ベンチャーズ

(72)発明者 チャン、エドワード

アメリカ合衆国、カリフォルニア・９４１４３、サン・フランシスコ、シックスティーンズ・スト
リート・６００、スイート・エス２７２・ジェネンテック・ホール・ボックス・２１４２、イノベ
ーション・ベンチャーズ

F ターム (参考) 4C127 AA03 DD01 GG01 GG10 GG11 GG13