

# 日本語初級学習者の筆記テスト時と会話時の脳活動 —fNIRSによる継続的研究(トレンドグラフと統計分析)—<sup>1</sup>

平田 裕\*

## 要旨

本稿は、3年計画のプロジェクトの2年目に行った実験の一部をまとめたものである。この研究の継続的な課題は、「日本語学習者の脳活動が、各種筆記テスト時と日本語での会話時でどのように違うのか明らかにする」というものであるが、今回は、初年度の研究で明らかになった課題を踏まえ、初級学習者2名を対象に同タイプのタスクを繰り返して行うという実験デザインを採用した。これは、個人内データを検証することで脳活動の複雑な個人差に対応しようとしたものである。分析手法の模索・確立も大きな課題の1つであるが、今回、脳活動の傾向把握はトレンドグラフで行い、対象とするタスクと測定チャンネルを絞って統計分析を行った。

トレンドグラフでの傾向把握では、[1] 筆記テスト時と実際の会話時では脳の賦活度という点で大きな相違があると考えられる、[2] 賦活度は[選択/選択パズル] < [訳/会話式] < [会話]の順に大きくなっていると考えられる、[3] 筆記タスクは右脳優位で処理される可能性があることが示唆された。解釈上の制約はあるが、分散分析とBonferroniの方法を用いた多重分析では、「選択パズル vs. 訳」、「訳 vs. 会話式」、「選択パズル vs. 日本語会話」の間で脳の活性化の度合いが同じであると考えられる、すなわち近似性があるという結果になった。

**キーワード:** 筆記テスト形式、脳活動、会話時との比較、fNIRS (functional near-infrared spectroscopy: 近赤外光分光法)

---

\* 立命館大学 大学院 言語教育情報研究科 教授

## 1. はじめに

### 1.1 研究の概要と前年度のまとめ

日頃の日本語教育の現場では、学習成果・教育成果の評価において、筆記テストは大きなウェートを占めるが、どのようなテスト形式であれば、学習者がそれを意識することによって会話力向上につながり、教師側からすると会話力評価につながるのだろうか。本研究は3年計画のプロジェクトの一部であるが、「日本語学習者の脳活動が、各種筆記テスト時と日本語での会話時でどのように質的・量的に違うのか客観的に明らかにする」という全体の研究テーマは、このような普段からの疑問・問題意識がきっかけとなっている。

筆記テストによって学習成果・修得状況を検証することはもちろん必要であり、特に TOEFL、TOEIC、IELTS、日本語能力試験など、大規模な標準化テストの信頼性は極めて高いものになっている。しかし、その一方で、例えば日本語能力試験の N1 取得者であっても、その学習成果が日常生活での会話力に直接つながっていないというのも、学習者自身も教える側である私達も実感するところである。これは、そもそもテストの目的、テストが測ろうとしているもの自体が違うので当然なのであるが、筆記テストを会話力向上やその評価に有効に使うことができれば、教育現場でのメリットは大きいと言える。

外国語教育における筆記テストの妥当性・信頼性の検証は、内容や形式の考察、そしてテスト結果の統計的確認が主であった(Bachman 1990; Bachman and Palmer 1996; 大友・中村 2002; Hughes 2002; 近藤ブラウン 2012 など)。それに対し、本研究では「会話時との脳活動の近似性・相違性」という新しい視点で従来の筆記テストの様々な形式の比較検証を試みている。脳イメージング技術については、近年、多様な手法の確立とそれぞれの技術の著しい進歩、そして、それに伴う普及があり、言語学や応用言語学分野の研究でも利用が次第に増えてきているが、本研究はその中でも近赤外光分光法(fNIRS: Functional Near Infrared Spectroscopy)による測定機器を使って脳活動を調べる。本稿では fNIRS という呼称で測定手法とともに測定機器(島津製作所の FOIRE-3000)も指すものとして使用する。

前年度の研究(平田 2013a, b)では、合計7名の被験者(初級学習者4名、初中級から中級の範囲に入る学習者が3名)を対象に、筆記テスト時と日本語での会話時の脳活動を fNIRS で計測した。タスクは合計5種類で、筆記テストが、[1]「選択穴埋め問題」(例:設問3つに対し、正答3つの選択肢)、[2]「訳問題」、[3]「選択肢問題」(例:設問1つに対し、助詞の選択肢4つ)、[4]「会話方式問題」の4種類、それに [5]「会話」を加えた5種類である。

分析方法としては、脳活動の全体像/傾向の把握を目的として、[1] マッピング解析(FOIRE-3000 提供の分析ツール)、[2] GLM 統計(一般的な GLM での統計分析ではなく、FOIRE-3000 提供の分析ツール)、[3] チャンネル別賦活度順位評価、[4] 総賦活量

比較という4つの分析を行ったが、それぞれのタスクによる脳賦活状態に類似性や特定のパターンを示唆するものもあったが、一目瞭然というような明らかなものではなかった。前年度の研究では研究テーマの複雑さがより具体的に明らかになったと言える。

具体的には、前年度の研究で以下の知見が得られた。[1] 筆記テストの形式としては同形式でも、問題の内容によって脳賦活の部位と賦活の程度は異なる。[2] 同様に、会話タスクにおいても被験者の日本語レベルと話す内容によって、脳の活性化は異なったパターンを示し、賦活総量も著しく違う。[3] 全般的に、中級学習者は筆記テスト時よりも会話時の脳賦活総量が多い。[4] 初級学習者は筆記テストと会話タスクが同程度の脳賦活総量を示すものがある。総括すると、実際の筆記テスト状況、そして会話状況における脳活動は、きわめて複雑、かつ個人差も大きいということが明らかになった。

今回の研究は、前年度にまとめた具体的な課題のうち、主に脳活動の個人差に対応すべく、被験者の人数を絞り、個人内データの比較にフォーカスしたものである。それとともに、タスク種別内差異にも注目し、同種類のタスクの繰り返し回数を増やした。また、測定部位も、ブローカー野だけでなくウィルニッケ野も含むように変更している。今回の研究も、「筆記テスト時と日本語での会話時の脳活動の近似性・相違性」という複雑な研究テーマにアプローチする、1つのステップという位置づけである。前年度の研究については、平田(2013a, b)を参照されたい。なお、本稿で使用するグラフや写真は、カラーページ数の制約もあるので特にカラーである必要がない場合はグレースケールに変換している。

## 2. 先行研究と本研究の研究課題

### 2.1 脳イメージング技術を使った言語学・応用言語学分野での研究

本研究で使用している近赤外光分光法(NIRS: Near Infrared Spectroscopy)は脳イメージング手法の1つであるが、脳機能の測定手法としては他にも脳波測定(EEG: Electroencephalogram)、コンピュータ断層撮影(CT スキャン: Computed Tomographic Scanning)、陽電子断層撮影法(PET: Positron Emission Tomography)、磁気共鳴画像法(MRI: Magnetic Resonance Imaging)、脳磁図(MEG: Magnetoencephalography)などがある。MRIとfMRI(functional MRI)、NIRSとfNIRS(functional NIRS)はほぼ同じ意味で使われることが多いが、前者は画像撮影技術を指すことに対し、後者は特に脳「機能」を観察する場合に短いインターバルで連続撮影(動的撮影)する時に使われている。これらの技術はいわゆる医学・脳科学(神経科学)分野で発展してきたものであり、近年、CT/PET/MRIは脳に限らず医学的研究や治療行為に広く利用されている。しかし、fMRIやfNIRSは測定機器として非常に高額であるため、言語学あるいは応用言語学研究への応用は、かなり進んで来ているとはいえ未だに限られた状況にあると言えるだろう。

脳機能イメージング手法を使った言語処理に関する先駆的研究には、PET を用いた Peterson *et al.* (1988)の実験研究があり、言語と脳機能の研究は特に近年盛んになってきている。これらの研究には主に2種類のタイプがあると言えるだろう。1つは、同一の属性をもつ被験者に対し違うタスクを行い、言語機能と脳賦活部分の対応を特定しようとするタイプの研究である(城生 1997; Scherer *et al.* 2006; 苧阪 2007 など)。もう1つのタイプは、中級学習者と上級学習者、健常者と統語失調症患者、諸条件の違うバイリンガルグループなど、属性の違う被験者グループ間の違いを脳イメージングによって検証しようとするものである(大石 2002; Kubota *et al.* 2005; Taura and Nasu 2012 など)。バイリンガリズムの研究における、個人の言語能力の変化に着目した研究も、被験者の属性の変化(例えば、通時的に見た習得や、言語環境変化に伴う言語能力の喪失など)と脳活動の変化を脳イメージングで捉えようとするものであり、大きな括りでいうと上記の後者のタイプに当てはまると言える(Andrews *et al.* 2013 など)。

言語機能に関係する代表的な部位としては、ブローカー野、ウィルニッケ野、角回、上縁回、聴覚野などが分かっており、右利きの人は左脳側にこれらの言語野がある確率が極めて高い。従来は、「言語産出はブローカー野」、「言語理解はウィルニッケ野」、「文字処理は角回」など、脳の特定の部位に特定の言語機能が分担されていると考えられていた。しかし、近年の研究ではこれらの部位が補完しながら連携して機能しているという見方が有力になってきている(Obler and Gjerlow 1999; Sakai *et al.* 2001; 酒井 2002 など)。言語タスクに対する脳活動の研究で、前頭前野の働きに着目したものとしては、Herrmann *et al.* (2003)、Kubota *et al.* (2005)、Hatta *et al.* (2009)などがある。また、ブローカー野(左脳部位)と前頭前野の働きを示唆する研究としては、大石(2001, 2002)、Gui *et al.* (2004)などがある。ブロードマン脳地図で言語野を示したものを下の図1左にあげる。



(\*左右の図の対比からも分かるように、参照する脳の形状や大きさなどの違いからも、外部からの部位の特定は難しい。)

図1 脳の機能分布(左: 従来のもの vs. 右: DeWitt and Rauschecker の主張)

上述のように、言語機能には脳の様々な部位が関係しており、詳しい機能分布についてはまだ明らかになっていないとされているが(坂井・久光 2011 など)、更に最近の

研究では、DeWitt and Rauschecker (2012) および DeWitt and Rauschecker (2013)が、ウィルニッケ野は従来考えられてきた位置よりも前方にあると主張し、それが担う言語機能についても更に詳細な検証と考察を加えている(上図1の右側参照)。

## 2.2 本研究の研究課題

上で見てきたように、脳イメージング技術を使ったこれまでの研究タイプは言語機能と脳の部位の対応(もしくは対応の度合い)に着目した研究、言い換えれば、人と脳が主たる研究対象である。これに対し、本研究は各種筆記テストおよび会話という、実験で使用するタスク形式自体を主たる研究対象とするものであり、筆者の調べた限りにおいては類する先行研究はなかった。

本研究は、前年度からの継続的な研究として、以下の2つを研究課題とする。

- 1) 前年度に引き続き、最終的な研究目的に合うデータ収集方法、分析方法を模索・検討する。今回採用する方法は、個人内検証、同種類タスクの繰り返し、トレンドグラフによる変化の把握、分析対象チャンネルを絞った形での統計分析である。(ただし、研究デザインの評価のためには実験データを蓄積していく必要があるため、本稿も包括的な検討ではなく、データの蓄積となるところが大きい。)
- 2) 被験者個人別に、各種筆記テスト時と会話時で脳活動の傾向にどのような近似性・相違性があるのかを、トレンドグラフによる視覚的把握と統計での量的分析で検証する。

fNIRSでの測定はデータが大量になり、一回の分析で全ての視点での分析を行えないので、今回は一次分析として、統計処理を行うチャンネルはブローカー野とウィルニッケ野に該当すると考えられるチャンネルに絞る。トレンドグラフでの傾向把握は右脳と左脳の違いにも注目するが、右脳・左脳の違いについての統計分析は行わない。また、前頭前野のデータも収集しているが、今回の分析の対象とはしない。今回の実験のデータを用いて様々な角度での分析が考えられるが、継続的課題とする。

## 3. 研究方法

研究方法の概要としては、前年度同様、タスクとして被験者にタイプの違う筆記テストを数種類受けてもらい、最後に短い会話を数回行うというものである。これらのタスク時の脳活動の状況をfNIRSで測定、様々な角度から分析する。前年度との違いは、[1] 測定部位を変え、ウィルニッケ野を含める、[2] 被験者を少数に絞る(2名)、[3] 同種類のタスクを繰り返す回数を増やし、実験を2日に分けるというものである。

以下、今回採用した研究方法を詳しく見ていく。

### 3.1 fNIRS

前年度同様、今回の脳イメージングには島津製作所の fNIRS 測定システム、FOIRE-3000 を使用した。各種の脳機能の測定手法／機器にはそれぞれメリット・デメリットがあるが、fNIRS に関しては、体の位置や向きに制約が少なく筆記テストを受ける時の実際の状態に近い形で実験/計測ができるという点、そして、測定時に大きな音を発生しないという点で、筆記タスクおよび会話タスクを行う今回の実験に適していると言える。fNIRS の概要については平田(2013)、詳細については島津製作所の HP (2014)も参照されたい。

下の図 2 に前年度使用したプローブホルダーと今回使用したホルダー、およびそれぞれのプローブ配置図を示す。上段が前年度のもので、3×9 個のプローブを長方形に配置したものである。プローブホルダーを装着すると、配置図の中央部分が額をカバーし、紙面左側が右脳側、紙面右側が左脳側となる。下段が今回使用したものになるが、システム上の要件と、前頭前野中央部、ブローカー野近辺、ウィルニッケ野近辺を測定範囲に入れるため、図のような配置を採用した。島津製作所の FOIRE-3000 は仕様の違いにより使用できるプローブの最大本数と測定可能な最大チャンネル数に違いがあり、今回使用した機器では利用可能なプローブ数が赤青合計 32 本と限られていた。

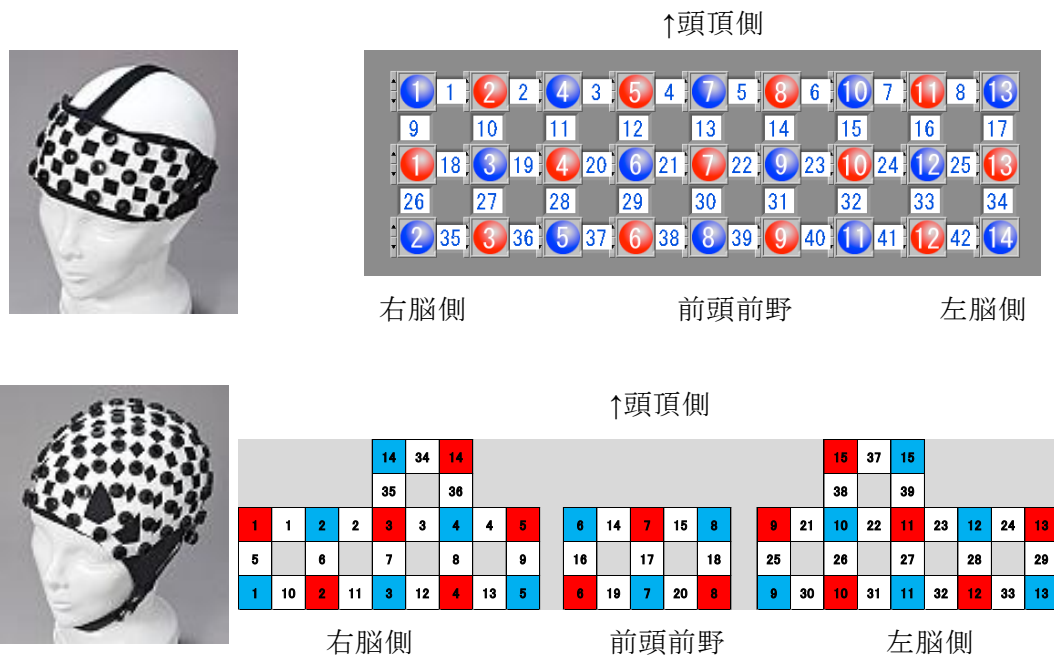


図 2 プローブホルダーとプローブ配置図 (上段前年度、下段今回)<sup>2</sup>

ホルダーに取り付けられた各プローブはそれぞれ3センチずつ離れており、赤色プローブは近赤外光の照射用、青色プローブは、反射して帰ってくる遠赤外光を受光するためのものである。図1のプローブ配置図では赤・青の位置が赤色プローブと青色プローブの位置に対応しており、赤・青の間の白のボックスが実際の計測位置に対応している(チャンネルと呼んでいる)。今回使用したプローブは赤15本、青15本、測定チャンネル数は39チャンネルである。

プローブホルダーの装着にあたっては、鼻根点(ほぼ目と目の間)から頭頂を通る外周(うなじの窪みまで)の長さを測定し、鼻根点からその10分の1の距離のポイントに青7番の受光プローブが来るようにするとともに、プローブ最下列が脳波記録国際10-20法のT3-Fp1-Fz-Fp2-T4のラインにできるだけ一致するように装着する(下の図2参照)。今回のプローブ配置では、右利きで左脳側に主な言語野がある場合、上の図2下段および下の図3右の配置図でチャンネル25、26、30近辺がブローカー野に、チャンネル27近辺がウィルニッケ野に対応すると考えられる。(曲面が強い頭部への装着イメージを紙面で表しているため、チャンネル番号と実際の頭部での位置にはズレがある。また、本研究では機能と特定の部位の対応は研究対象にしていなかったため、ウィルニッケ野についても「近辺」でよいこととする。)

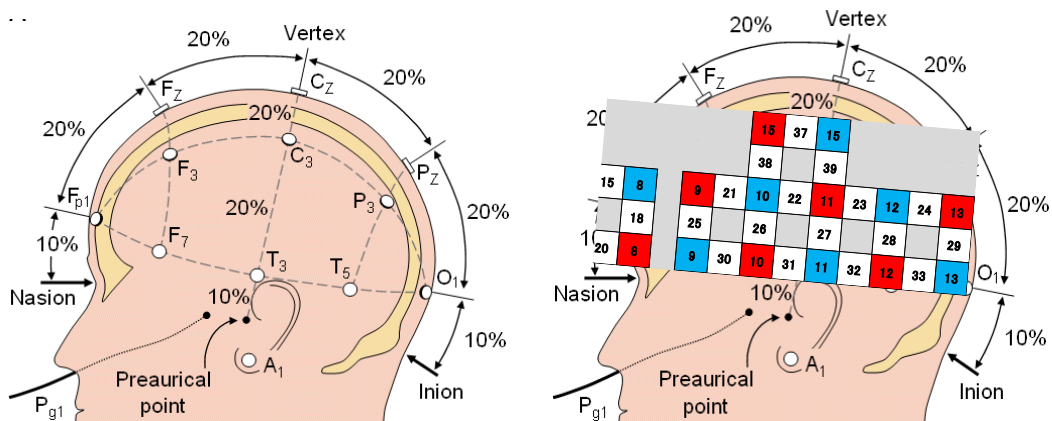


図3 International 10/20 system と今回のプローブ配置イメージ

今回、プローブホルダーの装着およびプローブ配置で特に問題になったのが、全頭をカバーするキャップ型ホルダーの装着性である。右の図4に実際に使用したプローブホルダーに近いものの写真を示すが、プローブを装着するベースは薄い帯状プラスチックで作られており、それを網目状に組む形でフレキシビリティを出している。ホルダー全体の装着性と、頭表とプローブの密着性など、相当の工夫がされているものであ

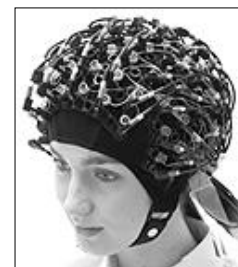


図4 全頭プローブホルダー (島津製作所のHPより)

るが、やはりホルダー自体のサイズの影響と頭部形状の個人差の影響は大きかった。頭頂部を密着させて顎に回したベルトで固定する方法を取ると、個人によってはプローブ最下列が耳よりも下にきたり、頭部が大きい人の場合は耳よりもかなり上部にきたりするのである。今回、その問題に対応するために、頭頂部と顎での固定ではなく、プローブ最下列を T3-Fp1-Fz-Fp2-T4 のラインにできるだけ一致させ、後頭部に紐を回して締め付けるという方法を取った。測定チャンネルを広範囲に配置しようとする、頭表との密着性が充分確保できないプローブが出てきてしまうが、後頭部の方で締め付ける方法では、今回使用した 30 個のプローブの密着性はかなり確保できた。

プローブホルダーのデザイン自体、プローブ間は 3cm という前提で作られており、頭部の個人サイズが反映される脳波記録国際 10-20 法に対応したものではない。プローブ間 3cm という条件から、fNIRS では位置の測定精度は高くないことになるので、上述の「チャンネル 25、26、30 近辺がブローカー野、チャンネル 27 近辺がウィルニッケ野」というのはあくまで目安である。また、今回の実験は 2 日に分けて行っていることもあり、実際の装着状況を目で確認し、対応位置の判定を行う必要がある。プローブホルダーの装着状態は写真に撮っており、後からでもプローブの位置を確認できるようにしている。

実際にfNIRSが測定値として出すものは濃度変化と光路長の積で、単位は $\text{mM} \cdot \text{mm}$ (または $\text{mmol} \cdot \text{mm}$ : ミリモル(L)・ミリメートル)である。データの種類としては、酸素化ヘモグロビン(oxy-hemoglobin: oxy-Hb)、脱酸素化ヘモグロビン(deoxy-hemoglobin: deoxy-Hb)、そして総合ヘモグロビン(total-hemoglobin: total-Hb)の3つについて計測・算出できる。NIRSの原理については、島津製作所のHP (2014)が分かりやすい説明を提供している。

### 3.2 被験者

今回の被験者 2 名の情報を表 1 に示す。彼女達は大学の短期留学プログラムで日本語を学習する留学生であり、日本語の筆記テストには慣れているので、実験タスクのフォーマットに戸惑うことは少ないと考えられる。利き手の調査にはエジンバラ利き手アンケートを使用した(書く、描く、投げる、歯ブラシ、蹴り足などで点数化するもの)。日本語のレベルについては学期当初のプレースメントテストで判定されており、初級教科書の『大地』の 2 冊目(山崎ほか 2009)を使って日本語を学習する初級後半のクラスに所属している。



表 1 被験者一覧

被験者 番号	性別	年齢	国籍	母語	学 日 習 本 歴 語	レ 日 べ 本 ル 語	外 話 以 日 国 せ 外 本 語 る で 語	利 き 手
01	女	19	中国	中国語	2 か月	初級後半	英語、仏語	全て右
02	女	23	アメリカ	英語	2.5 年	初級後半	なし	全て右

中国人被験者の方は、短期留学プログラムの日本語クラスで日本語を学習しているが、英語基準の学部正規留学生であり、英語の運用力は高い。日本語は初修であるが、漫画やアニメなどで日本語に親しんでいることと、漢字知識というアドバンテージで初級後半のクラスに入ったものと考えられる。アメリカ人被験者の方は、2.5 年の学習歴で初級後半クラスに入っているが、欧米で学ぶ日本語学習者としては珍しくないパターンだと言える。

被験者募集の際は、実験参加に対して謝金を払うということで協力者を募集しているが、性別・母語・学習歴・利き手・年齢などの細やかなコントロールは出来ていないのが実情である。今回の募集は、被験者の日本語能力を揃えるという意図で、特定のレベルの日本語クラスでのみ行った。今回の研究自体は個人内比較を目的とするものであり、この 2 名で問題はないと考える。

今回の実験については、立命館大学「人を対象とする研究倫理審査委員会」による審査を受けている(受付番号：衣笠 - 人 -2012)。

### 3.3 タスクデザイン

本研究では、タスクはシンプルで短くという脳実験上の必要性和、できるだけ通常の筆記テストフォーマットを維持するという研究課題上の必要性という 2 つの要件のバランスを取ることが課題となっている。例えば、言語学分野での脳実験は語彙の連想や文法性判断など、いわゆる言語機能に絞った実験タスクが多いのであるが、本研究では「選択問題」や「外国語での会話」など、タスク条件によって変わる脳の使い方全体を研究対象としていると言ってよい。未だ、それをどのように捉えるかを模索している段階である。

前年度の研究では、「実験タスクをより通常の筆記テスト時に近いものにし、そこでの思考過程全体の脳活動を観察する」という目的で、タスク内の問題数、タスク毎の設定時間などを特に統一していなかった。今回のタスクデザインは、同種類のタスクの繰り返し回数を増やすことで、タスク種別ごとの特徴が出るように考えたものである。繰り返し回数を増やすため、そして、タスク毎にまとめて分析できるようにするため、各タスクの設定時間は統一して短く、1 分間に設定した。前年度と同様の種類のタスクをこなそうとすると全体のタスク時間が長くなってしまいうため、実験を 2

日に分けて実施した。

下の表 2 に前年度と今回のタスクデザインを示す。一番左の表が前年度のもの、真ん中と右の 2 つの表が今回のものである。この表には示していないが、タスクの前後には必ず 30 秒の休憩(レスト)を入れている。前年度は 1 回の実験で 5 種類のタスク、延べ 12 タスクを行っており、基本的には同形式の筆記タスクを 2 回続けてから次の種類の筆記タスクに進むようになっている。タスクの種類が変わるたびに次のタスクの説明を入れており、タスク全体での所要時間は 30 分ほどであった。一方、今回のタスクデザインでは 1 回の実験につき延べ 16 タスクであるが、2 つの形式のタスクを交互に繰り返すので(下の表中ではシェードをかけて同タイプタスクを表示)、タスクの種類としては 1 日の実験につき 3 種類となっている(筆記 2 種類と会話)。

表 2 前年度と今回のタスク構成

前年度タスク			今回のタスク 1 日目			今回のタスク 2 日目		
TASK01	選択パズル1*	2分	TASK01	選択1	1分	TASK01	訳1	1分
TASK02	選択パズル2*	3分	TASK02	選択パズル1	1分	TASK02	会話式1	1分
TASK03	訳1 単語	2分	TASK03	選択2	1分	TASK03	訳2	1分
TASK04	訳2 単語	2分	TASK04	選択パズル2	1分	TASK04	会話式2	1分
TASK05	選択1	2分	TASK05	選択3	1分	TASK05	訳3	1分
TASK06	選択2	2分	TASK06	選択パズル3	1分	TASK06	会話式3	1分
TASK07	訳3 表現	2分	TASK07	選択4	1分	TASK07	訳4	1分
TASK08	訳4 従属節/文	3分	TASK08	選択パズル4	1分	TASK08	会話式4	1分
TASK09	会話式 2問	3分	TASK09	選択5	1分	TASK09	訳5	1分
TASK10	会話1	1分	TASK10	選択パズル5	1分	TASK10	会話式5	1分
TASK11	会話2	1分	TASK11	日本語会話1	1分	TASK11	日本語会話1	1分
TASK12	会話3	1分	TASK12	日本語会話2	1分	TASK12	日本語会話2	1分
			TASK13	日本語会話3	1分	TASK13	日本語会話3	1分
			TASK14	母語会話1	1分	TASK14	母語会話1	1分
			TASK15	母語会話2	1分	TASK15	母語会話2	1分
			TASK16	母語会話3	1分	TASK16	母語会話3	1分

\*「選択パズル」と呼ぶ形式は、前年度は「選択穴埋め」と呼んでいたものである。

今回の実験では、1 回の実験内の 2 種類の筆記タスクの説明は最初に 1 回、2 つまとめて行うことで時間短縮を図っている。タスク全体の時間が長すぎると、プローブをつけて動きの制約を受けている被験者の疲労が大きくなるので、今回の実験も 1 日分はだいたい 30 分で全タスクが終了するようにデザインした。タスク所要時間が 30 分であっても、プローブホルダーの装着や被験者情報アンケートの記入なども合わせると、1 人につき約 1 時間かかることになる。

実験全体のデザインとしては、特定の認知/言語行動による脳反応を抽出するため、認知/言語行動の構成要素を厳密に想定した上で「刺激状態(stimulated state)」から「統

制状態(control state)」の差を用いるという方法(いわゆる差分法)が *Petersen et al.* (1988)の研究以降、広く採用されている。例えば、提示した語を単純に口に出して言うというタスクの脳活動量から、提示した語を見て認識するというタスクの脳活動量の差を用いることによって、「口に出す」という運動に対応する脳活動を抽出しようというものである。しかし、本研究の今回の実験も前年度のものと同様、差分法は採用せず、タスクとタスクの間は単純な休憩とした。これは、本研究の目的が特定の認知／言語機能と脳の部位の対応を検証することではないためである。それに加え、「選択肢問題を解く」や「外国語で会話をする」など、本研究が対象とする認知／言語行動が極めて複雑で、差分法を使える条件にはない。現実的な問題／条件としても、今回の実験は30分ほどかかるものであり、実質的な休憩なしに30分間タスクを続けられるものではない。また、差分法の大前提は、全ての認知／言語行動は線形的であり(モジュール的に加算可)、相互作用は存在しないというものであるが、脳活動にはこれがあてはまらないケースがあることも報告されている(Wagner 1999 など)。

検証する筆記テストの種類は前年度と同様である。以下、それぞれのタスクの具体例を見ていく。被験者2名は中国人とアメリカ人であるが、中国人の被験者も日本の大学に英語基準で留学しており、英語は充分堪能である。多言語環境の日本語の授業で英語を媒介語として使うことも珍しくないので、タスクで使う媒介語は英語とし、今回は中国語版を作らなかった。また、本稿ではルビなしであるが、実際の実験に使ったものは漢字に適宜ルビを振っている。

### 3.3.1 選択タスク(1 out of 3) (初日の TASK01, 03, 05, 07, 09)

この形式は1つの設問に対し3つの選択肢の中から正答を選ばせる、いわゆる三択問題である。問題は、様々な品詞およびオノマトペなど、単語レベルの問題である。以下の例は初日の TASK05 である。

#### (1) TASK05

1. あのレストランのステーキは、とても(おもくて ふとくて あつくて )  
びっくりした。
2. 6月からサークルに( 入れて はいられて はいって ) います。
3. 風邪をひいたみたいです。咳が( ある する である ) んです。

### 3.3.2 選択パズルタスク (3 from 3) (初日の TASK02, 04, 06, 08, 10)

これは、3の設問に対し3つの正答選択肢を与え、どれにどれが入るかという組み合わせを考えさせるタイプである。これも、様々な品詞の単語レベルの問題であるが、活用形が問題になる場合はそれも合わせて解答するよう指示をしている。以下の例は初日の TASK06 である。

(2) TASK06 選択肢： 

うまい	じょうず	いい
-----	------	----

1. 旅行のスケジュール、これで ( ) ?
2. 日本語は あまり ( ) ないです。
3. なにか ( ) ものが食べたい。

このタスクだけでも、単語の理解、文章の理解、選択組み合わせ、活用、書くという行為など、様々な認知・行動要素が含まれており、複雑な脳活動になる。

### 3.3.3 会話タスク (初日および2日目の TASK11~16)

会話は、日本語での会話タスクが3つ(TASK11, 12, 13)、母語での会話タスクが3つ(TASK14, 15, 16)である。本研究は日本語の筆記テストと日本語での会話時の脳活動を比較検証することを目的としているが、今回は参考データ採取のため母語での会話タスクも加えた。初日と2日目の両日ともに筆記タスクはタスク番号の奇数グループと偶数グループで同種類のタスクになっているので、まず日本語で、どのように考えて奇数グループのタスクの答えを出したか、どの問題が難しかったかなどを問題例を提示して聞いた(TASK11)。次に、偶数グループのタスクについても同様の質問をし(TASK12)、最後に実験全体についての感想を日本語で説明してもらった(TASK13)。日本語での会話タスクの後、全く同じ内容で被験者の母語で説明してもらった(TASK14, 15, 16)

### 3.3.4 訳タスク (2日目の TASK01, 03, 05, 07, 09)

これは、学習者の母語や媒介語の使用を前提にしたもので、今回の実験では英語から日本語を産出するというものである。問題のレベルは単語からフレーズ、従属節や文全体のものまでである。以下の例は2日目の TASK07 である。

(3) TASK07

1. レポートの日本語を チェックして \_\_\_\_\_ ?  
Could you please (check ...)
2. 電車で 足を \_\_\_\_\_ しまいました。  
being stepped on
3. 学校に \_\_\_\_\_、郵便局に行きました。  
before going

### 3.3.5 会話式タスク (2日目の TASK02, 04, 06, 08, 10)

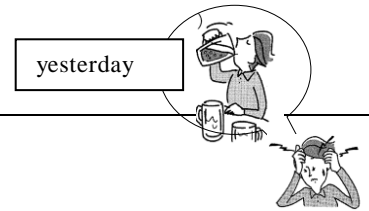
この問題形式のみ、日本語教育でごく普通に行われているタイプではないと思われるが、会話力向上と会話力測定につながる筆記テストの形式を模索する中で、筆者が

実際に授業で使った形式である。筆者は今回の被験者 2 名を直接教えていないが、彼女達が学んでいるクラスの教材の大部分は筆者が作ったものを使用しているため、彼女達もこの問題形式に慣れている。前年度は 1 つのタスクに 2 つの設問というデザインであったが、今回は全てのタスクの時間を最大でも 1 分に設定するため、1 タスクに 1 つの設問にした。以下の例は初日の TASK06 である。

#### (4) TASK06

1. 友達: どうしたの?

自分: \_\_\_\_\_  
【Explain why.】



### 3.4 実験手順

今回の実験も前年度と同様、通常の筆記テストに近いやり方という要件から、会話以外のタスクを紙ベースで被験者に渡し、被験者のペースで解答してもらうという形式をとった。fNIRS でのデータの収集は「連続データ収集方法」ということになる。具体的には、実験全体の説明、それぞれのタスクのやり方の例示、タスク前後の休憩、10 個の筆記タスクと 6 つの会話タスク、これら全てを 35 ページのポケットにまとめ、被験者がページをめくる指示はパワーポイントの画面切り替えでチャイムを鳴らしてコントロールした。タスク全体としては 30 分程度である。

タスクのやり方の説明と休憩はそれぞれ 30 秒、各タスクの時間は最大で 1 分間であるが、解答が早く終わった場合は机上の缶を鉛筆で叩いて合図を送ってもらい、筆者が fNIRS のデータ内にマーキングを入れ、分析時にタスクの早期終了を判別できるようにした<sup>3</sup>。タスクが早く終わった場合も、ポケットのページをめくるのはパワーポイントのチャイムが鳴ってからである。

今回使用した fNIRS 機 FOIRE-3000 はビデオカメラとの同期ができるので、全体のタスク開始から終了まで被験者の様子を録画した。また、会話タスクについてはボイスレコーダーを使って全て録音している。

### 3.5 分析方法

研究のスコープのところでも述べたが、作業量の問題もあり、今回は一次分析として前頭前野部から収集したデータは分析の対象から外す。また、トレンドグラフでの傾向把握は右脳と左脳の違いにも注目するが、右脳・左脳の違いについての統計分析は行わない。統計処理を行うチャンネルはブローカー野とウィルニッケ野に該当すると考えられるチャンネルに絞るものとする。

前年度はマッピング図による賦活パターンの把握を行ったが、今回は 2 つの理由でトレンドグラフによる把握を行う。1 つは、マッピング図は或る時点での賦活状態の

スナップショットであり、タスク期間全体の傾向をみるためにはトレンドグラフの方が向いているからである。マッピング図もムービー機能で時間軸に沿って動かせるが、色変化の動画であるので、論文の分析の中で扱うのは難しい。もう 1 つの理由は FOIRE-3000 の機能上の理由で、今回のような不規則なプローブ配置の場合はマッピング図を頭の位置と対応させて直観的に把握できるようにするのが難しいからである。今回のようにプローブ配置が正方形や長方形でない場合、マッピング図を有効に使うためには入念な配置設定が必要であるが、今回の実験前の設定作業ではそこまで出来なかった。(きれいな矩形でなく、チャンネル配置のギャップが大きい・多い場合、マップ図としての見え方は補正值の影響も大きくなる。)

また、トレンドグラフの配置としては、横 24 コマまで、縦 16 コマまで、かつ、横×縦の積が 64 コマまでという条件がある。今回は左脳・右脳それぞれ横に 9 チャンネルの長さなので単純に並べると画面上は横 18 コマになり、縦は 3 コマ分しか使用できない。測定時は全チャンネルを見るために、前頭部分のコマをパズルのように入れ込んで変則的なレイアウトを取ったが、分析の際は前頭部分を外し、頭部との対応が直感的に分かりやすくする。

トレンドグラフ、および統計処理に使うデータであるが、FOIRE-3000 の機能を使って補正したものを使用する。ベースライン補正は、休憩状態から実験タスクによってどれくらい値が上がったかを見るための基準値補正であるが、ノイズの影響を考慮してタスク開始後 1 秒間の平均値を基準点とした。計測は 0.1 秒毎なので、タスク開始後 10 回計測した値の平均値を引く計算になる。また、今回は FOIRE-3000 のスムージング機能も活用する。これは、平滑化フィルタとしてもよく用いられる Savitzky-Golay 法を使用したもので、FOIRE-3000 に対しスムージング点数とスムージング回数を設定することで自動的に行える。今回はスムージング点数 25 回、とスムージング回数 5 回で補正を行った。下の図 5 に補正なし(左)と補正あり(右)のトレンドグラフを示す。補正の例示だけが目的なので、分析のための配置調整は行わず、右脳側だけに画面を切っている。

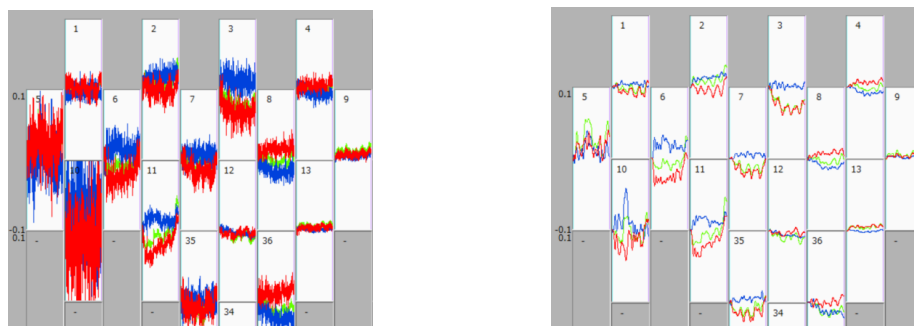


図 5 トrendグラフの補正あり・なし (被験者 01 初日 TASK04 選択パズル)

上の図からも分かるように、fNIRS の測定値はノイズも含み、そのまま見ても傾向が掴みにくいので、トレンドグラフでの分析には補正值を使うこととする。また、本研究のテーマは、「筆記テスト時と日本語での会話時の脳活動の近似性・相違性」であるが、研究の趣旨からすると、fNIRS の実測値ではなく、「実測値が示す“傾向”の近似性・相違性」でよいので、統計処理でも補正後の値を使うものとする。

トレンドグラフを使った分析は、2名の被験者それぞれに対し同種類のタスクのトレンドグラフ画面を並べて一覧を作り、比較した。具体的には、「選択」、「選択パズル」、「初日の日本語会話」、「初日の母語会話」、「訳」、「会話式」、「2日目の日本語会話」、「2日目の母語会話」それぞれについて一覧を作った。例として、下の図6に被験者01の選択タスク5つのトレンドグラフ画面を並べたものを示す。(1つの画面には、配置したチャンネルの数だけのトレンドグラフが並んでいることになる。)

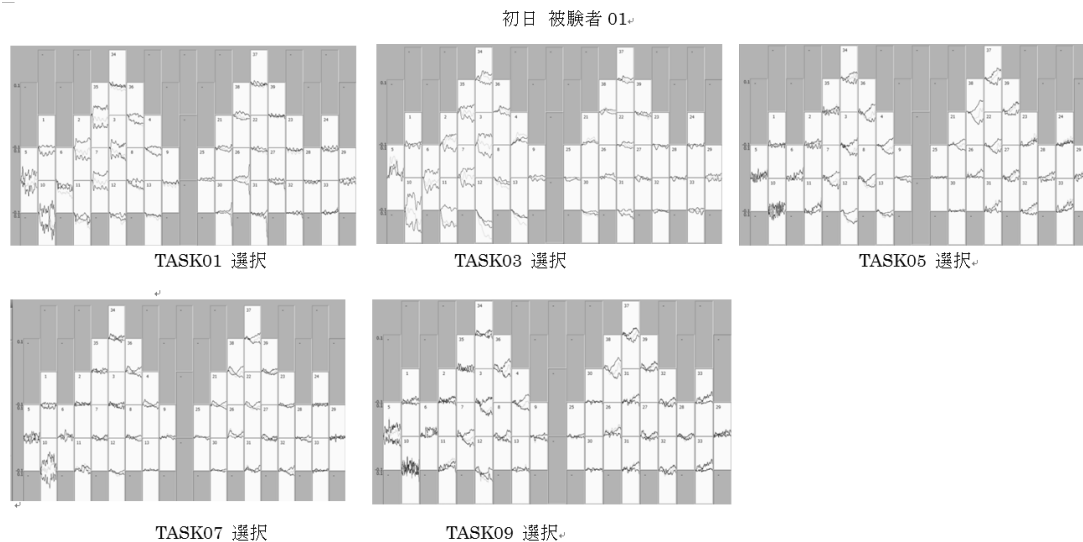


図6 トレンドグラフ画面一覧

統計分析は、まず FOIRE-3000 から分析対象のデータをテキスト出力し、エクセルのデータに読み替えた後、必要な処理を施してから行う。統計ソフトは IBM の SPSS 22 を使用し、タスクタイプ を要因とした一元配置、6水準(選択/選択パズル/訳/会話式/日本語会話/母語会話)の分散分析(ANOVA)を行う。同一個人内での検証なので、対応ありのデータである。対象とするチャンネルは、これまでに分かっている代表的な言語野ということで、今回の分析ではブローカー野およびウィルニッケ野に一番近いと考えられる2つのチャンネルに限定する。ブローカー野とウィルニッケ野を比較することが目的ではないので、それぞれ別に分析にかけることとする。

福長ほか(2011)によると、NIRSの解析方法には未だスタンダードはなく、fMRIで使用されるBOLD (blood oxygenation level dependent)信号との対応付けとしてどの指標を用

いるかは研究者の考え方によって異なっている。計測手法の妥当性という点では、これはfNIRSだけの課題ではない。fMRIもfNIRSも脳の神経活動から発せられる信号を直接測定するものではなく、神経活動の結果生じた血流量の変化を測定し、それによって脳神経活動を間接的に推定するものであるので、測定値によって脳活動の何が分かるのかについては研究と議論が継続している。（例えば、Huettel *et al.* (2009)は、fMRIの測定において抑制作用の伝達信号と刺激作用の伝達信号という相反するインプットによる神経活動とBOLD信号の関係の曖昧性を指摘している。）

筆者の前年度の研究では、oxy-Hbの使用が妥当だとする先行研究に倣い(田村 2002; Shimoyama *et al.* 2006; Toronov *et al.* 2007; 梁ほか2008; Malonek *et al.* 1997; Strangman *et al.* 2002; 福田 2009など)、oxy-Hbを分析対象とした。今回の分析でもoxy-Hbを対象とするが、トレンドグラフでの傾向把握ではdeoxy-Hbが大きく反応しているものもあった。具体的な脳の神経活動とoxy-Hb/deoxy-Hbの関係については神経脳科学分野の研究を広くあたり、oxy-Hbとdeoxy-Hbを合わせて分析対象にする妥当性の検討など今後の課題としたい。

下の表3にデータを整理した形のイメージを示す。ただし、あくまでイメージであるので計測値は実測値とは全く関係がない。

表3 統計処理対象データイメージ

選択			選択パズル			訳			会話式			日本語会話			母語会話														
計測時間	a	b	c	計測時間	d	e	f	計測時間	g	h	i	計測時間	j	k	l	計測時間	m	n	o	計測時間	p	q	r						
0秒～最大55秒	0	0	0	0秒～最大52秒	0	0	0	0秒～最大58秒	0	0	0	0秒～全て60秒	0	0	0	0秒～全て60秒	0	0	0	0秒～全て60秒	0	0	0						
	1.1	0.9	1.1		3	1.5	2.5		1.1	0.9	1.1		1.1	0.9	1.1		1.1	0.9	1.1		1.1	0.9	1.1	1.1	1.1	0.9	1.1		
	1.5	1.2	2.8		2.1	1.8	3.2		1.5	1.2	2.8		1.5	1.2	2.8		1.5	1.2	2.8		1.5	1.2	2.8	1.5	1.2	2.8	1.5	1.2	2.8
	中略				中略				中略				中略				中略				中略			中略					
	0.8	2.5	2		2	2.7	2		1.5	2.8	2.7		1.5	2.8	2.7		1.5	2.8	2.7		1.5	2.8	2.7	1.5	2.8	2.7	1.5	2.8	2.7
	1.2	3.2	1.2		0.8		2		0.8	2.5	2		0.8	2.5	2		0.8	2.5	2		0.8	2.5	2	0.8	2.5	2	0.8	2.5	2
	2		2						1.2	3.2			1.2	3.2			0.8	2.5	2		0.8	2.5	2	0.8	2.5	2	0.8	2.5	2
	2.2													1.5			1.2	3.2	1.2		1.2	3.2	1.2	1.2	3.2	1.2	1.2	3.2	1.2
														1.8			2	2.7	2		2	2.7	2	2	2.7	2	2	2.7	2
																	0.8	2.5	2		0.8	2.5	2	0.8	2.5	2	0.8	2.5	2

表の縦方向は計測時間が入り、今回の設定では0.1秒毎に計測値が入る。それぞれのタスクタイプで5つから6つのタスクがあり、1分よりも短い時間でタスクが終了しているものもある。難易度が低いタスクほど短時間で終わるが、本研究ではそれぞれのタスク種別の特徴が脳活動という形でできるだけ顕著に出たものを比較したいので、時間が長くかかったものから3つ選んでその平均値を使った。表中の横方向、aからrが個別のタスクにあたり、a-cは初日の奇数番号タスクから3つ選んだもの、d-fは初日の偶数番号タスクから3つ選んだもの、以下、j-lまで同様である。日本語会話タスクと母語会話タスクは初日と2日目それぞれ3つずつ、合計6つずつのタスクが



あり、全て 60 秒間計測しているが、トレンドグラフを見て初日と 2 日目でブローカー野およびウィルニック野の反応が大きいと考えられる日の 3 つを対象とした。

それぞれのタスクの種類で 3 つを選んだことになるが、個別のタスクで実際に計測した時間の長さが違う、すなわち測定したデータ数が違うことになる。測定数が違う部分に欠損値を入れるのも 1 つの対応方法であるが、今回の実験では 1 秒間に 10 回の測定で欠損値が多くなってしまいうので、分析対象の 18 個のタスクで一番短い時間のものにデータの個数を合わせることにした。上の表に示した例でいうと、選択パズルの e が一番測定時間が短く、データ数が少ないので、全部のデータ数をこれに合わせるということになる。下の表 4 に被験者毎の分析対象タスクと対象チャンネルを示す(1 番目がブローカー野、2 番目がウィルニック野)。タスクについては、それぞれの種別内の 3 つの平均値を使う。被験者 01 の方でブローカー野対応チャンネルが 21 と 25 というように違うものがあるのは、プローブホルダーの装着状態が初日と 2 日目で違っていたためである。

表 4 分析対象のタスクとチャンネル

	選択	選択パズル	訳	会話式	日本語会話	母語会話
被験者01	TASK05, 07, 09 CH21, 27	TASK06, 08, 10 CH21, 27	TASK05, 07, 09 CH25, 27	TASK02, 08, 10 CH25, 27	初日TASK11, 12, 13 CH21, 27	初日TASK14, 15, 16 CH21, 27
被験者02	TASK03, 07, 09 CH25, 27	TASK04, 06, 08 CH25, 27	TASK03, 05, 09 CH25, 27	TASK02, 04, 08 CH25, 27	2日目TASK11, 12, 13 CH25, 27	2日目TASK14, 15, 16 CH25, 27

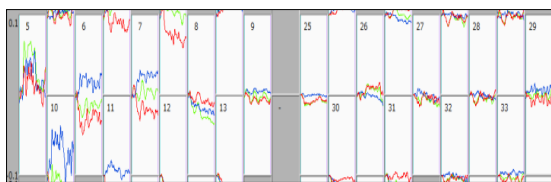
今回の実験データの分析方法を決める要素としては、対象のチャンネル選別方法、特定のチャンネルか複数チャンネルの平均か、対象タスクの選別方法、同種類の複数タスクの平均か合算か、計測時間（データ数）の取り扱い、などが考えられるが、それぞれの妥当性の比較検証も今後の課題としたい。

## 4. 実験結果と考察

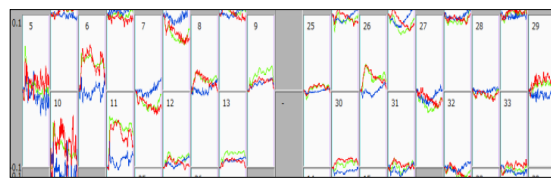
### 4.1 トrendグラフによる分析

トレンドグラフ画面の一覧は全部で 12 枚(被験者 2×3[筆記奇数/偶数/会話]×2 日)となり、紙幅の関係で全ては掲載できないので、例としてタスクを選び画面の一部(上下 3 列の中央のライン)を抽出したものを示す。視覚的に把握した傾向を数字として検証する方法は、全チャンネルの積算値比較、特定のチャンネルの積算値比較、特定のチャンネルの分散分析などが考えられるが、今回は全ての分析をするに至らなかった。今後の課題とする。(ただし、今回の分析でもブローカー野とウィルニック野に対応すると考えられる 2 つのチャンネルについては、ANOVA の際に平均値などが分かる。)

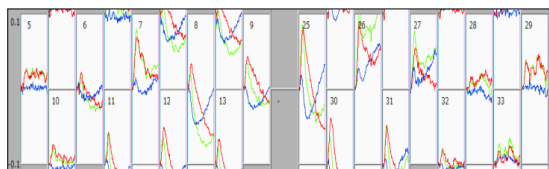
被験者毎に 6 つのタスクのトレンドグラフ(抽出版)を次ページの図 7 に示す。以下、トレンドグラフから分かることを箇条書きでまとめる。



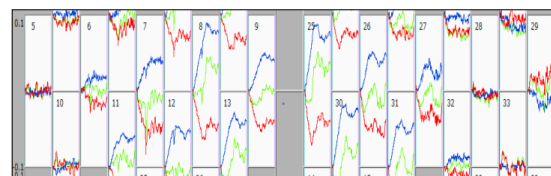
[1] 被験者 01 選択 (初日 TASK03)



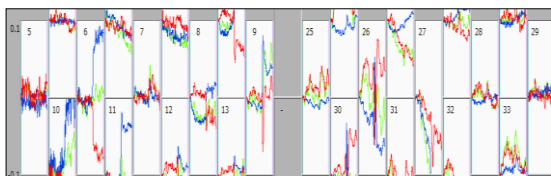
[2] 被験者 01 選択パズル(初日 TASK08)



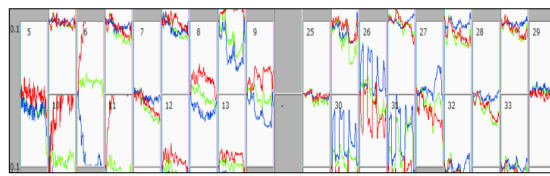
[3] 被験者 01 訳 (2 日目 TASK07)



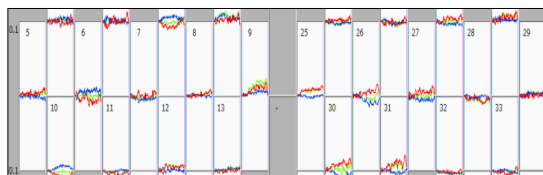
[4] 被験者 01 会話式 (2 日目 TASK08)



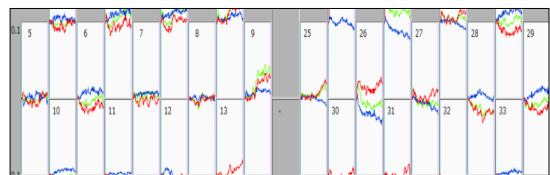
[5] 被験者 01 日本語会話 (初日 TASK11)



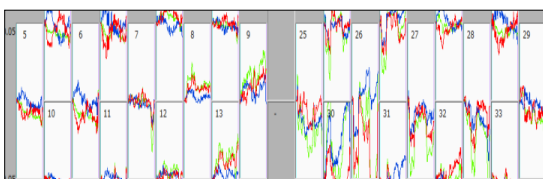
[6] 被験者 01 母語会話 (初日 TASK14)



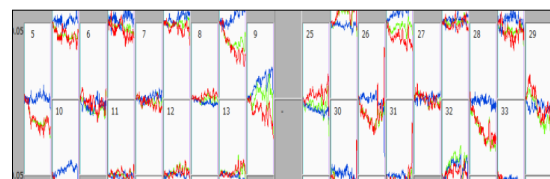
[7] 被験者 02 選択 (初日 TASK03)



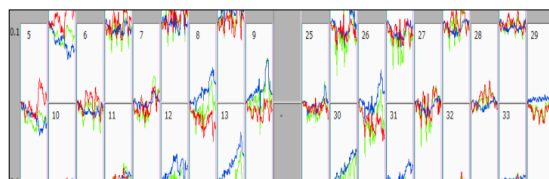
[8] 被験者 02 選択パズル (初日 TASK08)



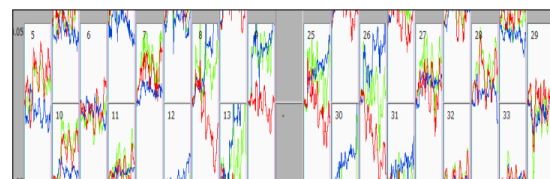
[9] 被験者 02 訳 (2 日目 TASK03)



[10] 被験者 02 会話式 (2 日目 TASK08)



[11] 被験者 02 日本語会話(初日 TASK12)



[12] 被験者 02 母語会話(初日 TASK16)

図7 トレンドチャート一覧サンプル

1. 2名の被験者ともに筆記タスクよりも会話タスクの方が脳の活性化の度合いが大きいことが多い(被験者 01: 次ページトレンドグラフ[1]-[4]筆記タスク vs. [5][6]会話タスク; 被験者 02: [7]-[10]筆記タスク vs. [11][12]会話タスク)。筆記タスクの中で、[3][4][9]は会話タスクと同じぐらい活性化しているように見えるが、同種類タスクの中で最も活性化したものを例示しており、同種類内の他のタスク全て活性化の度合いが大きいとは言えない。
2. 筆記タスクの中でも、通常の言語使用行動と明らかに違うタイプの「選択」「選択パズル」と、第二言語での通常使用により近い「訳」「会話式」を比べると、2名の被験者ともに後者の方が脳の賦活度が高い([1][2]vs.[3][4]; [7][8]vs.[9][10])。
3. 被験者 01 の筆記タスクでは、左脳側よりも右脳側(紙面左側)の賦活度が高い([1]-[4]参照)。つまり、言語野であるブローカー野とウィルニッケ野がない側の賦活度が高いと考えられる。それに対し、会話タスクでは、日本語の場合も母語の場合も左脳側(紙面右側)の賦活度が高い([5][6]参照)。
4. 被験者 02 は被験者 01 と違い、全てのタスクでブローカー野とウィルニッケ野があると考えられる左脳側(紙面右側)の賦活度が高い([7]-[12]参照)。
5. 脳の活性化は、常に赤で示される oxy-Hb の増加となって表れるのではない可能性がある。被験者 01 の[1]の右脳側(紙面左側)、チャンネル 6 と 7(画面が切れていないコマで左から 2 番目と 3 番目)を見てみると、青色の deoxy-Hb が顕著に増加し、赤色の oxy-Hb が減少していることが分かる。同様のパターンは、[4]の中の多くのコマでも観察される。被験者 02 の場合は、[9]、[10]、[11]、[12]の中の複数のチャンネルで同様のパターンが観察される。これは、前年度の実験でも特に初級の学習者のマップ図に見られた傾向である。
6. 今回の筆記タスクと会話タスクで、ブローカー野とウィルニッケ野に対応すると考えられるチャンネル(被験者 01 の初日のみチャンネル 21/27 で、それ以外は 25/27)が示す活性化パターンは様々である。つまり、oxy-Hb と deoxy-Hb それぞれがプラスに振れたりマイナスに振れたりする場合がある。
7. 言語野に対応する複数のチャンネルをまとめて考える(例えば、平均値を取る)のではなく、分けて考えた方がよいことを示唆するケースがある。例えば、[3]や[4]のチャンネル 25/26/27(左脳側、左端から 3 つのコマ)は、活性化の程度は違うものの、oxy-Hb と deoxy-Hb がプラスに振れるかマイナスに振れるかという点では同傾向を示しているので、まとめて考えてよいと思われる。しかし、[5]のチャンネル 25/26 の oxy-Hb はプラスに振れているのに対し、チャンネル 27 では oxy-Hb は大きくマイナスに振れている。同様に、[10]のチャンネル 25 と 27 はプラス傾向、その間の 26 では明らかにマイナスに振れている。プラスマイナスで相殺される場合は、脳活動を平均値で扱うのは適切でないように思われる。

上記の3と4について補足すると、言語能力のレベルと左脳・右脳の違いについては、バイリンガリズムや第二言語習得の分野で研究されていることが多い。例えば、Illes *et al.* (1999)は、英語とスペイン語の2言語上級話者の場合、2言語間で脳の賦活部位に有意な差はなかったとしている。また、Perani *et al.* (2005)もバイリンガル話者の脳では第一言語と第二言語は違った部位で処理されているという仮説を否定している。外国語学習に関しては、例えば大石(2006)によると、英語上級学習者は右脳より左脳の賦活度が高く、初級学習者は右脳と左脳の賦活度に差がないとし、習熟度が高い学習者の方が左脳優位となる説を支持している。大石・木下(2008)では右脳と左脳の比較ではなく母語と第二言語の比較において、第二言語のリスニングの方がより大きく左脳が賦活するとしている(但し、被験者は2名で、TOEFLの点数は480点と570点である)。

今回の被験者2名に関しては、筆者が実験当日に日本語での会話を行ったところ、被験者01の方が会話力は高かった。トレンドグラフでの視覚的傾向把握という前提条件はつくが、今回の結果はとても興味深い。初級だがある程度の会話力がある被験者01の方は、筆記タスクでは右脳優位(初級学習者)で会話タスクでは左脳優位(会話力あり)という傾向であり、先行研究が示唆する内容に沿って説明することが可能だと考えられる。それに対し、日本語での会話力がそれほど高くない被験者02は全てのタスクで左脳優位の傾向となっている。これまでの研究では、上級・中級・初級などの大きなくくりで学習者の言語能力を分類していたため、脳活動の個人差が見えにくくなっていた面もあるように思われる。本研究だけでなく、今後も被験者個人別の分析を継続的な課題としたい。

トレンドグラフによる分析は仮説形成段階だと言えるが、本研究のテーマである「筆記テスト時と日本語での会話時の脳活動の近似性・相違性」に沿って結果をまとめると、まず、筆記テスト時と実際の会話時では脳の賦活度という点で大きな相違があると考えられる(上記1)。次に、賦活度は実際の会話との近似性を反映し、[選択/選択パズル] < [訳/会話式] < [会話]の順に大きくなっていると考えられる(上記2)。個人差に着目して言語習熟度との関係は更に調べないといけませんが、筆記タスクは右脳優位で処理される可能性が示唆された(上記3、4)。

## 4.2 ANOVAによる検証

被験者2名につきそれぞれブローカー野とウィルニッケ野に関して検証を行うので、SPSSのデータファイルは4つになる。それぞれのファイル内で対応ありの形でタスク種別毎のデータが6列入ることになる。まず、それぞれのデータに関して正規分布を示しているかどうかの検証を行ったが、有意確率 .000 で全て正規分布ではないという結果になった。しかし、データ数については被験者01の場合は全タスク種別が441個、被験者02の場合は581個というように因子間で同数に設定しているので、ANOVA

による検定に頑強性があると判断した。ANOVA で設定した有意水準は .05 で、被験者内因子の多重分析は Bonferroni の方法を用いた。

検定の結果、球面性の仮説が成り立っていなかったため、Greenhouse-Geisser の自由度の修正を行った。下に 4 つの検定結果をまとめたものを示す。

表 5 被験者 01 ブローカー野 分散分析表

ソース	SS	df	MS	F	p	partial $\eta^2$
タスク種別	1.404	1.371	1.024	817.551	.000	.650
経過時間	.114	440	.000			
誤差	.756	603.096	.001			
Total	2.274	1044.467				

表 6 被験者 01 ウィルニック野 分散分析表

ソース	SS	df	MS	F	p	partial $\eta^2$
タスク種別	.269	2.585	.104	1305.553	.000	.748
経過時間	.118	440	.000			
誤差	.091	1137.221	7.97E-5			
Total	0.478	1579.806				

表 7 被験者 02 ブローカー野 分散分析表

ソース	SS	df	MS	F	p	partial $\eta^2$
タスク種別	.389	1.893	.205	2636.610	.000	.820
経過時間	.012	580	2.16E-5			
誤差	.086	1101.431	7.79E-5			
Total	0.487	1683.324				

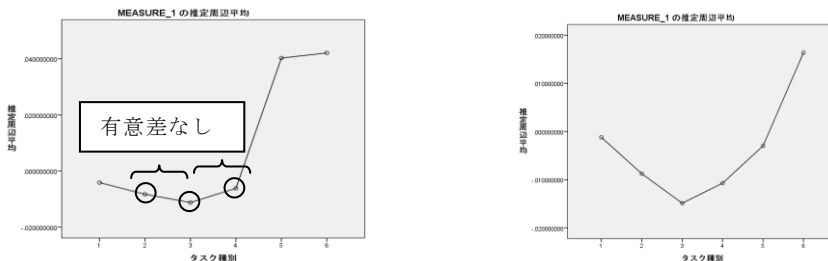
表 8 被験者 02 ウィルニック野 分散分析表

ソース	SS	df	MS	F	p	partial $\eta^2$
タスク種別	.050	3.238	.015	530.637	.000	.478
経過時間	.011	580	1.86E-5			
誤差	.054	1878.184	2.88E-5			
Total	0.115	2461.422				

この表から分かるように、全ての検定で脳賦活度の平均に有意差があることがわかった( $p < .05$ )。(球面性検定のイプシロンの値は 4 つの検定の最少値で .274 だったため、多変量検定の結果はここでは示していない。ちなみに、多変量検定の全ての結果も  $p = .000$  であった。) Bonferroni の方法を用いて多重比較を行った結果、被験者 01 はブローカー野の反応において、「選択パズル vs. 訳」、「訳 vs. 会話式」の 2 つのタスク種別間に有意差がなく、被験者 02 もブローカー野において「選択パズル vs. 日本語会話」の間に有意差がなかった。他の組み合わせは全て 5% の水準で有意差があるとい

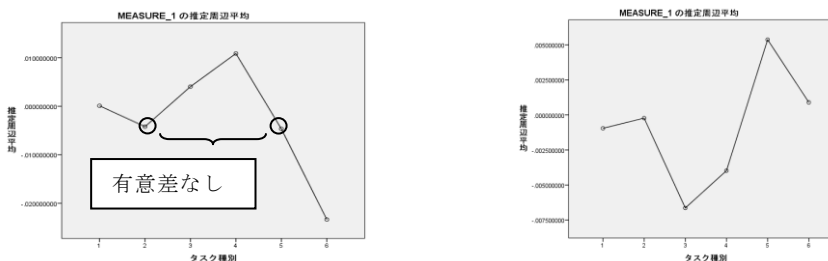
う結果であった。

視覚的に分かりやすいように、SPSSのプロファイルプロット図を下に示す。



タスク種別 1:選択 2:選択パズル 3:訳 4:会話式 5:日本語会話 6:母語会話

図8 被験者01のプロット図 (左：ブローカー野 右：ウィルニック野)



タスク種別 1:選択 2:選択パズル 3:訳 4:会話式 5:日本語会話 6:母語会話

図9 被験者02のプロット図 (左：ブローカー野 右：ウィルニック野)

あくまでグラフによる傾向把握であるが、タスク種別毎のパターンということで見ると、これらの図でブローカー野近辺・ウィルニック野近辺のチャンネルを見る限り、特定のパターンは見られず、個人差も大きいと言える。また、ブローカー野近辺とウィルニック野近辺の違いということで見ると、この2つの領域で平均値がプラスとマイナスに分かれているタスクが多い。具体的には、被験者 01 では「5:日本語会話」がプラスとマイナスに分かれ、被験者 02 では「2:選択パズル」を除く5つのタスクでプラスとマイナスに分かれている(下の図 8 と 9 で同一タスクを左右のグラフで比較)。このような場合、トレンドグラフでの分析の際にも述べたことであるが、分析の際にこの近辺のチャンネル測定値の平均を使うことには慎重になる必要があるのではないだろうか。他に気が付く傾向としては、被験者 01 の場合、ブローカー野の反応は筆記テストと会話(日本語・母語とも)で大きく違う(図 8 左側)。しかし、ウィルニック野の方では(図 8 右側)母語会話時のみ賦活度が大きく、日本語会話時は低いという結果になっている。また、4つの図中3つで「3:訳」の脳賦活度が一番小さいことも分かる(被験者 02 のブローカー野を除く3つ)。

今回のANOVAを使った検証では、上述のようにいくつかのタスク種別間で脳の活性

化の度合いが同じであると考えられる、すなわち近似性があるという結果になった。しかし、この結果はあくまで「分析対象チャンネルを限定し、脳賦活度を示すと考えられるfNIRSのoxy-Hbの平均値において」という但し書きがつくものである。今後、分析対象チャンネルの検討やデータの処理方法、個人差対応なども含め、更に多角的にデータ収集方法および分析方法の検討が必要である。

## 5. 結論

本研究は前年度からの継続的研究であるが、前年度明らかになった課題に対応するため、大きいところでは以下の3点を変更した。まず、[1] 測定部位を変え、ウィルニック野を含めた。次に、[2] 個人内でのタスク種別毎の比較を目的とし、被験者を2名に絞った。最後に、[3] タスク毎の特性を把握しやすいように、同種類のタスクを繰り返す回数を増やし、実験を2日に分けた。分析については、本稿は今回の実験の一次分析として、トレンドグラフによる傾向把握と、対象チャンネルをブローカー野近辺・ウィルニック野近辺の2チャンネルに限定した形で統計分析を行った。本稿でカバーできなかった分析には、チャンネル毎の積算値での検証、前頭前野データの検証、右脳と左脳という切り口での検証などがあり、直近の課題としたい。

トレンドグラフによる分析はあくまで目による傾向把握であり、それをどのように量的に検証するかは継続的な課題であるが、今回の分析では以下のような傾向があることが分かった。まず、筆記テスト時と実際の会話時では脳の賦活度という点で大きな相違があると考えられる。比較上の話しであるが、会話時の方が脳を活発に使うのに対し、筆記テスト時の脳活性化度はそれほど高くないという結果であった。次に、賦活度は実際の会話との近似性を反映し、[選択/選択パズル] < [訳/会話式] < [会話]の順に大きくなっていると考えられる。また、個人差に着目して言語習熟度との関係を更に調べないといけないが、筆記タスクは右脳優位で処理される可能性が示唆された。

次に、ブローカー野近辺・ウィルニック野近辺の2チャンネルのデータの統計分析では、ブローカー野における賦活度で「選択パズル vs. 訳」、「訳 vs. 会話式」、「選択パズル vs. 日本語会話」が近似性を示すという結果になった。但し、前者2つは被験者01の場合、最後の1つは被験者02の場合であり、タスクの属性として一般化できる結果とは言えない。個人内データでの検証を増やす必要がある。なぜ近似性を示すのかを明らかにするためには、複雑な言語機能の構成要素と脳部位の対応を明らかにしていく必要がある、本研究のスコープを越えるものであるが、研究を続けていく上で常に念頭に置いておきたい。

研究方法の検討、分析方法の検討という点では、[1] 分析対象チャンネルの選別方法(測定部位の検討も含む)、[2] 同種内での分析対象タスクの選別方法、[4] データの取り扱い方法(特定のチャンネル単独値か複数チャンネル平均値か、同種の複数タスクの平均か合算か、また、分析対象時間(データ数)をどうするか)などが継続的検討課題として

挙げられる。具体的には、トレンドグラフでの傾向把握でoxy-Hbではなくdeoxy-Hbが大きく反応しているものもあった。また、oxy-Hbの測定値がタスク時にマイナスに振れるという現象も前年度同様に観察された。具体的な脳の神経活動とoxy-Hb/deoxy-Hbの関係については神経脳科学分野の研究を広くあたり、分析においてoxy-Hbとdeoxy-Hbをどのように扱うのか検討する必要がある

前年度同様、今回の分析でも研究テーマの複雑さ、脳活動の複雑さがより具体的に明らかになったと言える。個人差、筆記テストの種類による違い、同一種類内の難易度による違い、測定チャンネルによる違い(隣接するものでも)などが傾向把握に影響するが、これらをうまく統制する研究デザインを模索し、「筆記テスト時と日本語での会話時の脳活動の近似性・相違性」という複雑な研究テーマにアプローチしていきたい。

#### 注

1. 本研究は、科学研究補助金（挑戦的萌芽(H23~H25)、研究代表者: 平田裕、課題番号: 24652108)「形式別、筆記テスト時と会話時の脳活動の近似性の比較検証」の助成を受けて行っているものである。
2. 今回のプローブ配置図は筆者がエクセルで作ったものである。今回のパターンでも FOIRE-3000 の機能できれいな配置図を作れるが、形が不規則で手間としては余計にかかるので、今回はエクセルで代用した。
3. タスク時の脳の反応としては、厳密に言語タスクによる反応だけでなく、「消しゴムで誤答を消す」や「缶をたたく」など、様々な動作を行うことによる反応も測定値に入ることになる。しかし、本研究のテーマは「筆記テスト時と日本語での会話時の脳活動の近似性・相違性」であるので、実際の筆記テストおよび会話時に近いということを優先し、言語タスクのみの反応を抽出する実験デザインにはしていない。缶をたたくことによるデータへの実際の影響についてはトレンドグラフでは確認できなかった。つまり、タスク終了時に不自然に測定値が上がるといったことはなかった。これについては、今回のプローブ配置が理由であることも考えられるが、そこまでの確認は行っていない。また、統計分析を行う際は、データ数を揃えるためにタスク末尾をカットしていることが多いので、その点でも分析に対する影響は少ないと考えられる。

#### 参考文献

- 大石晴美(2001)「インプットからインテイクへの言語情報処理過程—言語の脳科学的視点より英語教育への応用—」『ことばの科学』第 14 号, 321-340.
- 大石晴美(2002)「リスニングとリーディングにおける言語情報処理過程を探る—光トポグラフィにおける脳科学的解明に向けて—」『金城学院大学論集 英米文学編』第 43 号, 25-47. (英語学論説資料第 36 号収録)
- 大石晴美(2006)『脳科学からの第 2 言語習得論』昭和堂
- 大石晴美・木下徹(2008)「第一言語処理と第二言語処理における脳活性状態の違い—日本語と英語のリスニングにおいて—」『ことばの科学』第 21 号, 143-154.



- 大友賢二(監修)・中村洋一(著)(2002)『テストで言語能力は測れるか』桐原書店
- 荻阪直行(2007)「オノマトペの脳科学(特集 オノマトペと日本語教育)」『日本語学』  
26(7), 16-23, 明治書院
- 近藤ブラウン妃美(2012)『日本語教師のための評価入門』くろしお出版
- 酒井邦嘉(2002)『言語の脳科学』中央公論新社
- 坂井建雄・久光正(2011)『ぜんぶわかる脳事典』成美堂出版
- 島津製作所 HP(2014)「LABNIRS (ラボニルス) 原理と仕組み」  
<http://www.an.shimadzu.co.jp/bio/nirs/nirs2.htm> (2014年02月22日最終参照)
- 城生佰太郎(1997)「脳波でとらえた日本語教育(3)脳はアクセントをどのように捉えているのかその1」『月刊日本語』10(9), 58-61, アルク
- 田村守(2002)「光を用いた脳機能イメージング(1)」『臨床脳波』44, 389-397.
- 平田裕(2013a)「形式別、筆記テスト時と会話時の脳活動の検証に向けて:fNIRSによるパイロットスタディ」『言語科学研究』立命館大学言語教育情報研究科,  
43-74.
- 平田裕(2013b)「fNIRSによる筆記テスト時と会話時の脳活動の検証—脳賦活量に着目して—」『日本語教育学会2013年度秋季大会研究発表予稿集』, 145-150.
- 福田正人(2009)『精神疾患とNIRS—光トポグラフィー検査による脳機能イメージング』  
中山書店
- 福長一義・大貫雅也・福井裕輝・舟久保昭夫・福井康裕・中島章夫・嶋津秀昭・石山陽  
事・大瀧純一(2011)「NIRSを用いたニューロフィードバックシステムの開発」  
『杏林医会誌』42巻1号, 2-11.
- 山崎佳子・佐々木薫・高橋美和子・町田恵子・石井怜子(2009)『日本語初級〈2〉  
大地—メインテキスト』スリーエーネットワーク
- 梁志鋭・松野和子・杉浦正利(2008)「コロケーションの処理過程における脳内活性部位  
と心的負荷—NIRSによる脳機能測定法に基づく事例研究—」『平成17年度 -  
平成19年度科学研究費補助金研究成果報告書』197-220.
- Andrews E., Frigau L., Voyvodic-Casabo C., Voyvodic J. and Wright J. (2013) Multilingualism  
and fMRI: Longitudinal Study of Second Language Acquisition. *Brain Sciences*, 3,  
849-876.
- Bachman, L. F. (1990) *Fundamental considerations in language testing*. Oxford: Oxford Uni-  
versity Press.
- Bachman, L. F. and Palmer, A. S. (1996) *Language testing in practice*. Oxford: Oxford Univer-  
sity Press.
- DeWitt, I. and Rauschecker, J.P. (2012) Phoneme and word recognition in the auditory ventral  
stream. *Proceedings of the National Academy of Sciences of the United States of  
America (PNAS)*, 109 (8), E505–E514.

- DeWitt, I. and Rauschecker, J.P. (2013) Wernicke's area revisited: Parallel streams and word processing, *Brain and Language*, 127 (2), 181–191.
- Georgetown University HP (2014) Speech Area in Brain Misidentified for Decades, Researchers Say. <http://www.georgetown.edu/news/brain-speech-area-misidentified.html>  
Last access: 2014.02.22.
- Gui X., Qi D., Zhen J., and Chuansheng C. (2004). Mapping of verbal working memory in nonfluent Chinese–English bilinguals with functional MRI. *NeuroImage*, 22, 1-10.
- Herrmann, M. J., Ehlis, A.-C., and Fallgatter, A. J. (2003) Frontal activation during a verbal-fluency task as measured by near-infrared spectroscopy. *Brain Research Bulletin*, vol. 61, issue 1, 51-56.
- Hatta, T., Kanari, A., Mase, M., Nagano, Y., Shirataki, T. and Hibino, S. (2009). Strategy effects on word searching in Japanese letter fluency tests. evidence from the NIRS findings. *Reading and Writing*, 22, 1041-1052.
- Hughes, A. (2002) *Testing for Language Teachers. Second edition*. Cambridge: Cambridge University Press.
- Huettel, S. A., Song, A. W., McCarthy, G. (2009) *Functional Magnetic Resonance Imaging* (2nd ed.), Massachusetts: Sinauer Associates.
- Illes, J., Francis, W.S., Desmond, J.E., Gabrieli, J.D., Glover, G.H., Poldrack, R., Lee, C.J., and Wagner, A.D. (1999) Convergent cortical representation of semantic processing in bilinguals. *Brain and Language*, 70(3), 347-63.
- Kubota, Y., Toichi, M., Shimizu, M., Mason, R. A., Coconcea, C. M., Findling, R. L., Yamamoto, K., and Calabrese, J. R. (2005) Prefrontal activation during verbal fluency tests in schizophrenia – a near-infrared spectroscopy (NIRS) study. *Schizophrenia Research*, 77, 65-73.
- Malonek, D., Dirnagl, U., Lindauer, U., Yamada, K., Kanno, I., and Grinvald, A., (1997) Vascular imprints of neuronal activity: Relationships between the dynamics of cortical blood flow, oxygenation, and volume changes following sensory stimulation. *Proceedings of Natural Science Academy of the United States of America*, vol. 94, no.26. 14826-14831.
- Obler, L. K. and Gjerlow, K. (1999) *Language and the Brain*. Cambridge: Cambridge University Press.
- Perani, D., and Abutalebi, J. (2005) The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202-6.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of cortical anatomy of single-word processing. *Nature*, 331, 585-589.

- Sakai, K. L., Hashimoto, R., Homae, F. (2001) Sentence processing in the cerebral cortex. *Neuroscience Research*, 39(1), 1-10.
- Scherer, L. C., Giroux, F., Lesage, F., Senhadji, N., Benali, H., and Ansaldo, A. (2006) An optical imaging study of semantic and syntactic processing by bilinguals. *Brain and Language* 99(1-2), 185-186.
- Shimoyama, I., Sato, F., Nakazawa, K., and Ono, K. (2006) Single measurement of oxy- and deoxyhemoglobin for a functional near infrared spectroscopy. *Abstracts of the 29<sup>th</sup> Annual Meeting of the Japan Neuroscience Society (Neuroscience Research Special Issue 2006)*, 232-235.
- Strangman, G., Boas, D. A., and Sutton, J. P. (2002) Non-invasive neuroimaging with near-infrared light. *Biological Psychiatry*, 52, 679-693.
- Taura, H. and Nasu, A. (2012) The effects of onset-age and exposure duration on the L2 as observed in brain activation: an fNIRS study. *Studies in Language Science*, 2, 19-42. Ritsumeikan University.
- Toronov, V. Y., Zhang, X., and Webb, A. G. (2007) A spatial and temporal comparison of hemodynamic signals measured using optical and functional magnetic resonance imaging during activation in the human primary visual cortex. *Neuroimage*, 34, 1136-1148.
- Wagner, A. D. (1999). Working memory contributions to human learning and remembering. *Neuron*, 22, 19-22.