

TR-H-107

筋電位に基づく調音運動開始時点の測定
— 発話潜時との比較 —

西本 卓也 (早稲田大学)
正木 信夫
本多 清志

1994. 11. 17

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

筋電位に基づく調音運動開始時点の測定

— 発話潜時との比較 —

西本 卓也 (早稲田大学)

正木 信夫, 本多 清志

1 はじめに

手の運動などにおいて、複雑な運動ほど動作開始の合図から運動開始までの時間が増加することが知られている [1]。これは、運動が複雑になるにしたがい、中枢における運動のプログラミング要する時間が増加するためだと考えられる。

発話過程においても、類似の現象が観測されている。Sternberg らは数字や曜日などの単語を用いて、発話する単語数の反応時間への影響を調べた [2]。その結果、単語数の増加が発話潜時を線形的に増加させた。そこで彼らは発話企画の処理単位はストレス・グループではないかと考えた。

この実験は、発話の反応時間に影響を与える要因を探ることにより、発話のプログラミングに関するユニットを探ることができる可能性を示唆したものとして興味深い。日本語の場合、英語のストレス・グループに相当するものとして、音節や拍が考えられる。これらの単位はストレス・グループと同様、発話の中でほぼ等時的に現れるからである。

正木らは、音節数の増加により発話潜時が増加する傾向にあることを示した [3]。ただし、子音の種類やその並び方によって、音節数の影響は異なっていた。特に音節数の影響が強いのは、同じ子音が連続する場合であった [4]。

ところで、これらの先行研究では、発話の準備過程のための処理にかかる時間の「指標」として、発話潜時（発話開始合図の提示から語頭音開始時点までの時間）が用いられてきた。しかし、「語頭音開始時点」にはすでに調音器官は語頭音生成のための運動を開始している。さらに、調音器官の運動開始は筋の収縮により実現されるが、その収縮を引き起こす電氣的現象（筋電位）は調音器官の運動に先行して発生する。実際に、発話潜時には、(1) 発話開始合図が出た判断して発話準備を開始するまでの時間、(2) 発話準備の処理、すなわち、調音運動のプログラミングを行い、調音器官を動かす筋電位が発生されるまでの時間、さらに、(3) 筋電位発生時点から語頭音開始時点までの時間、が加算されている。発話の準備過程を対象にする場合、上記の(1)と(3)の時間は取り除きたい部分である。しかしながら、現在の実験手法 ([3],[4]) を用いる限り、(1)を取り除くことは不可能である。しかし、調音器官の筋電位は筋電信号 (EMG) により測定することができるため、(3)の時間を取り除くことは可能である。そこで、本実験では、「発話開始合図の提示から EMG 発火開始時点までの時間」（以下、「EMG 発火潜時」という）を発話の準備過程のための処理にかかる時間の「指標」として用いることができるかどうかを検討することとした。そのために、発話潜時測定時に、語頭音の生成に関与する筋の EMG を同時観測し、EMG 発火潜時の抽出アルゴリズムを検討した。さらに、発話潜時と EMG 発火潜時の相互の時間関係を調べた。

2 実験

2.1 発話試料

表 1 に発話試料を示す。発話試料としては口唇を調音位置とする子音 [p] と、発話の準備のために口唇を開く必要のある子音 [t] を語頭に持つようにした。実際には、「パ」および「タ」からなる 1～4 音節の無意味単語 14 語とした。「パ」で始まるものと「タ」で始まるもの、これらの音節が交互に現れるものと同じ音節が連続するものの組合せによって、4 系列に分類した。

表 1: 発話試料

	1 音節	2 音節	3 音節	4 音節
パタ系列	パ	パタ	パタパ	パタパタ
タバ系列	タ	タバ	タバタ	タバタバ
パバ系列	(パ)	パバ	パババ	パバババ
タタ系列	(タ)	タタ	タタタ	タタタタ

2.2 実験装置

図 1 に装置の構成を示す。文字単語などの刺激は AV タキストスコープのディスプレイより被験者に提示する。タキストスコープはパーソナルコンピュータによりコントロールされる。被験者の音声はマイクロフォンにより採取され、EMG 信号は表面電極から採取される。EMG 信号には 100Hz ～ 3kHz の帯域フィルタと、商用電源に起因する雑音を除去するための 60Hz のノッチ・フィルタをかけた。EMG 信号と音声信号は標本化周波数 12kHz で A/D 変換し、発話開始合図の提示時点から 2 秒ずつ、直接ワークステーションのハードディスクに記録した。

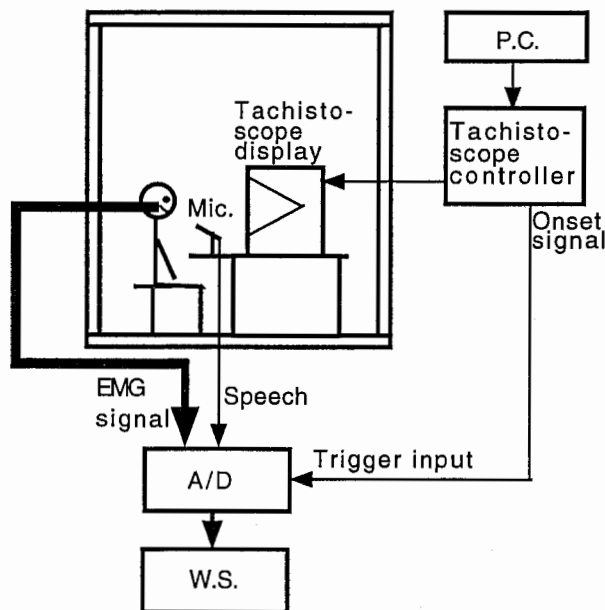


図 1: 実験装置

表面電極の取り付け位置を図2に示す。また、これらの電極により採取されるEMG信号は表2に示す筋に対応する。

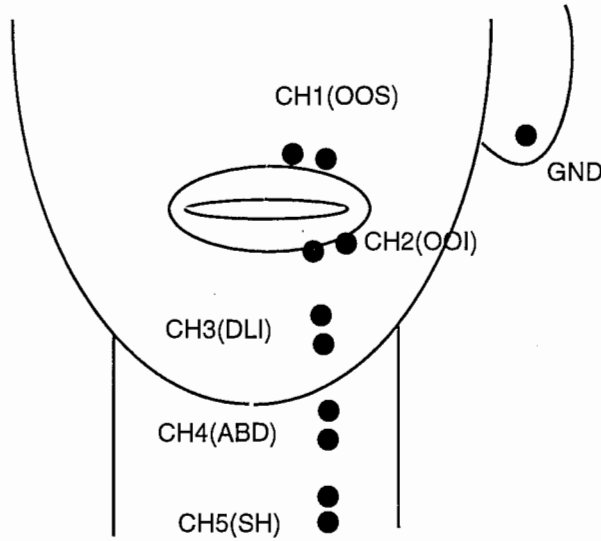


図 2: 表面電極設置部位

表 2: 各チャンネルに対応する筋の名称

電極名	筋の名称	略称	該当筋の機能
CH1	上口輪筋 (Orbicularis oris, superior)	OOS	口唇閉鎖、円唇化、突き出し
CH2	下口輪筋 (Orbicularis oris, inferior)	OOI	口唇閉鎖、円唇化、突き出し
CH3	下唇下制筋 (Depressor labii inferioris)	DLI	下口唇の下降
CH4	顎二腹筋前腹 (Anterior belli of the digastric)	ABD	下顎の開大
CH5	胸骨舌骨筋 (Sternohyoid)	SH	喉頭の下降

2.3 実験手順

被験者は39歳の男性で、東京方言の話者である。発話潜時およびEMG発火潜時の測定は以下の手順で行なった(図3)。

タキストスコープのディスプレイに発話すべきひとつの単語を2.5秒間カタカナで表示する。次に、正方形(□)を凝視位置として表示し、これが十字(+)になったとき被験者に単語を発話させた。被験者が発話開始合図の表示時点を予測することを避けるため、正方形が表示されてから十字に変化するまでの時間を1.4, 1.6, 1.8, および2.0秒のいずれかとし、ランダムに出現させた。また、同じ目的で、全試行のうち20%は十字を表示せず、発話を要しないものとした。

この試行を各単語あたり5回出現するようランダムに配列して提示した。70試行を1セッションとし、全部で7セッションを行った。このうち最初のセッションを練習とし、2~6セッションのデータを分析に用いた。最後のセッションは、実験セッション中の不適切なデータを補うために用いた。ここで不適切なデータとは、発話しなかったり、発話が異常に遅れた場合のことを指す。このような手順で各単語につき20個のデータを採取した。

十字が表示されてから発話のためにEMGが発火を開始するまでの時間をEMG発火潜時とした。EMG発火開始時点は採取した波形から後述するように定義し、半自動的に検出した。また、波形とスペクトログラムの視察・音声聴取によって、十字が表示されてから語頭子音の生成開始までの時間を測定し、これを音声の発話潜時(図3の"RT"の区間)とした。

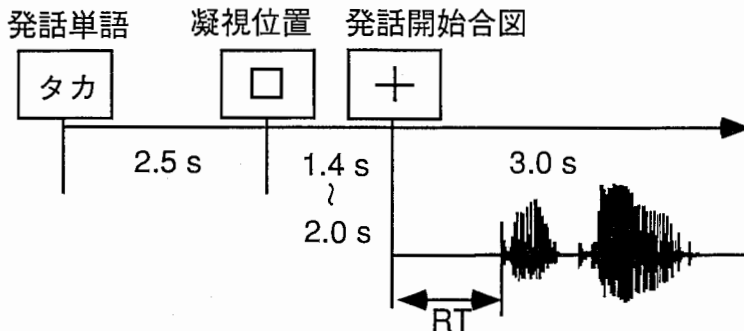


図3: 本実験の試行スケジュールの模式図。発話開始合図(+マーク)提示後、被験者は出来るだけ早く発話する。発話潜時の測定は発話開始合図提示時点から、語頭子音発話開始時点まで、(図の"RT"の区間)とした。なお、各試行毎に十字表示時点から音声と同様にEMG信号も採取した。

3 結果

3.1 EMG データの分析

図 4 に採取した音声波形と EMG 波形の例を示す。

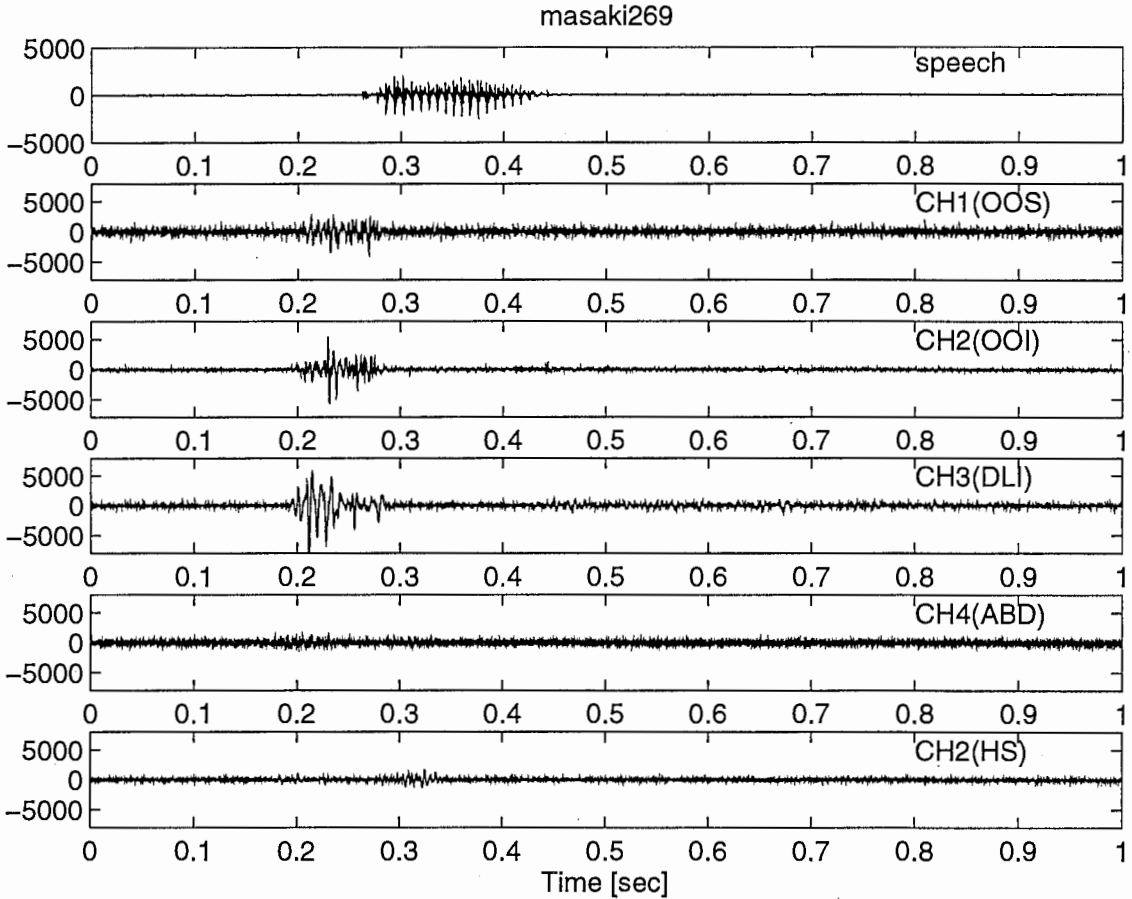


図 4: 採取した音声波形と EMG 波形の例。上から順に音声波形、表面電極 CH1 ~ CH5 (図 2, 表 2 参照) から採取された EMG 波形。発話試料は「パタ」である。

ここでは、語頭の [p] の調音、および [k] の調音のための準備のために、それに先行して動作すると考えられるの口唇と下顎の開大に関わる筋について、EMG 発火開始時点を測定することとした。今回の実験で EMG 信号を採取した筋の中で口唇の開大に関わるのは下口唇下制筋 (DLI) であり、下顎の開大に関わるのは、顎二腹筋前腹 (ABD) である。そこで、この 2 種類の筋からの EMG 信号について分析した (以下では筋の名称は略称を用いる)。DLI と ABD は、今回採取した電極のうち、それぞれ CH3 と CH4 にあたる (図 2, 表 2 参照)。

EMG 発火開始時点の自動検出は次のように行った。まず、採取したデータを 3kHz の低域フィルタに通して、12kHz から 6kHz にダウンサンプリングする。そして、測定値の絶対値を取ることにより振幅包絡を求め、10ms のハミング窓を通して移動平均を行った。さらに、各点を中心として 10ms の窓幅で線形近似を行い、その傾きを各点の傾きと定義した。そして、この傾きの最大値の 15% を越えた時点を EMG 発火開始時点とした。

EMG 発火開始時点の自動検出例を図 5 に示す。以下で「EMG 発火潜時」というのは、発話開始合図の提示時点（図 5 の時刻 0 の時点）から、EMG 発火開始時点までの時間である。

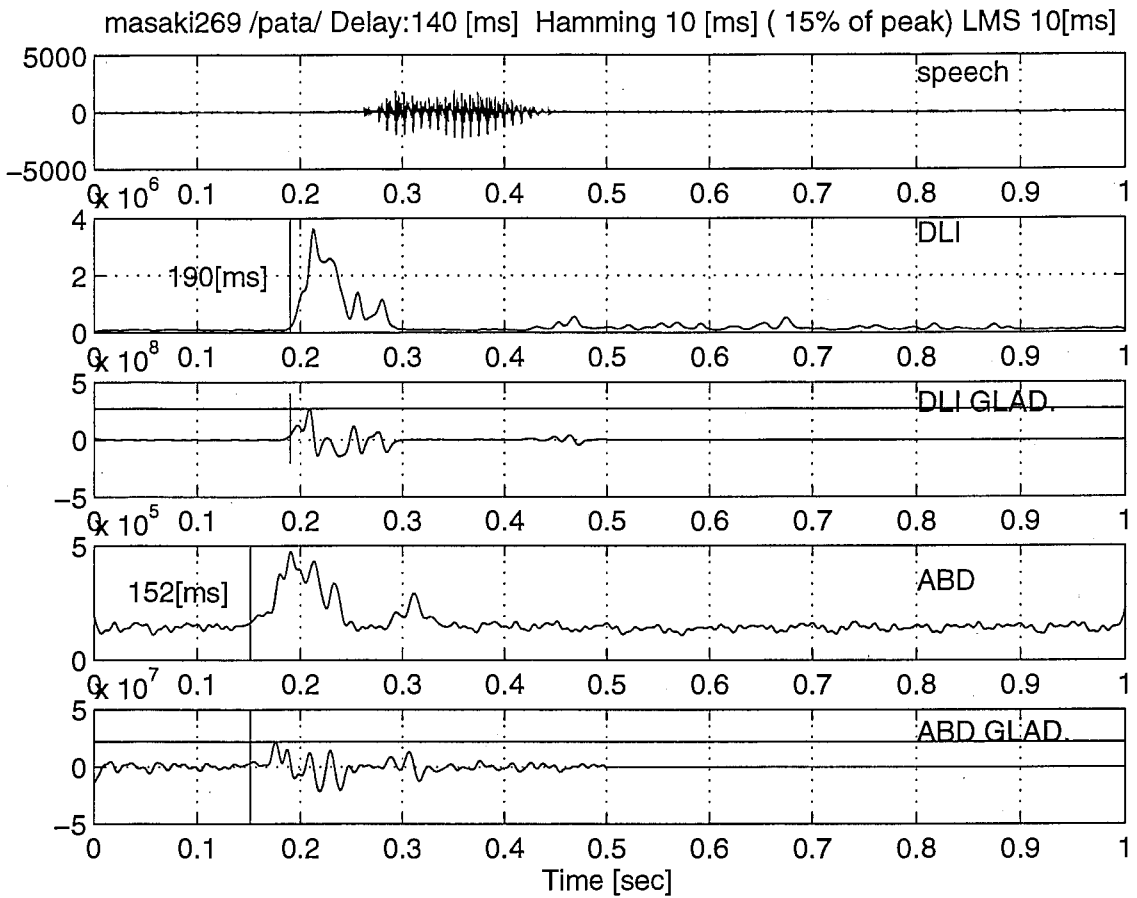


図 5: EMG 発火開始時点の自動検出の例（発話は図 4 と同じ「パタ」）。それぞれの波形は上から順に、音声波形 (SPEECH)、DLI における EMG の移動平均 (DLI)、DLI における EMG の移動平均の傾き (DLI GRAD.)、ABD における EMG の移動平均 (ABD)、ABD における EMG の移動平均の傾き (ABD GRAD.) を示す。傾きの図 (DLI GRAD. および ABD GRAD.) の水平の線は、発話中の最大値を示す。EMG 波形 (DLI, DLI GRAD., ABD, および ABD GRAD.) に示した縦線は、検出された EMG 発火時点である。

3.2 各単語系列と音節数の EMG 発火開始時間への影響

DLI の EMG 発火潜時を図 6 に、また ABD のそれを図 7 に示す。左から、「パタ系列」、「タパ系列」、「パパ系列」、「タタ系列」の順で、それぞれにおいて左から、1～4 音節の順に並んでいる。EMG 発火潜時の分布は以下の要領で表現されている。中央の箱の上下がそれぞれ 25 パーセントイルと 75 パーセントイルの値を示しその中間の横線が中央値を示す。その上下の縦線の上限と下限がそれぞれ 10 パーセントイルと 90 パーセントイルの値を示している。さらに 10 パーセントイルと 90 パーセントイルを越えるデータが○で示される。

音節数の EMG 発火潜時への影響を調べるために、各系列毎に 5% 水準の分散分析を行った。その結果、図の網掛け表示した系列が有意であった。さらに、これらの系列について Student-Newman-Keuls 検定による多重比較を行ったところ、□ または □ のマークで示した音節数の間の差が有意であった。

この結果、DLI の EMG 発火潜時については、「パパ系列」のみ有意であった。多重比較では、1 音節と 2 音節の間にも有意差が見られた。なお、分散分析では有意ではなかったものの、その他の 3 つの系列において、少なくとも、中央値で見ると、1 音節の単語において他の単語の場合よりも EMG 発火潜時が短い傾向が見られた。

また、ABD の EMG 発火潜時については、「パパ系列」および「タタ系列」において音節の影響が有意であった。多重比較では、いずれの系列でも、1 音節と他のすべての音節数との間で有意差が見られた。分散分析では有意ではなかったものの、「パタ系列」において、少なくとも、中央値で見ると、1 音節の単語において他の単語の場合よりも EMG 発火潜時が短い傾向が見られた。

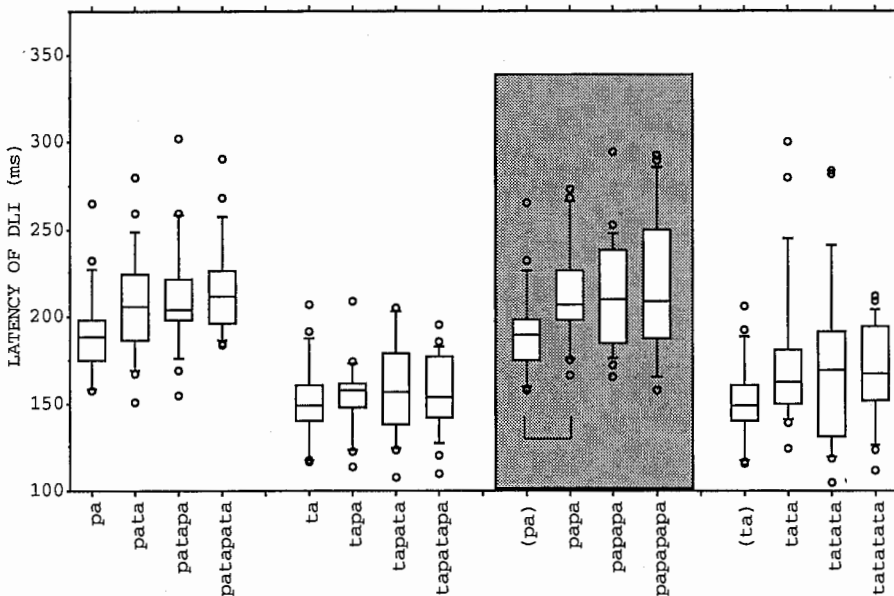


図 6: 各単語系列と音節数における DLI 発火潜時の分布

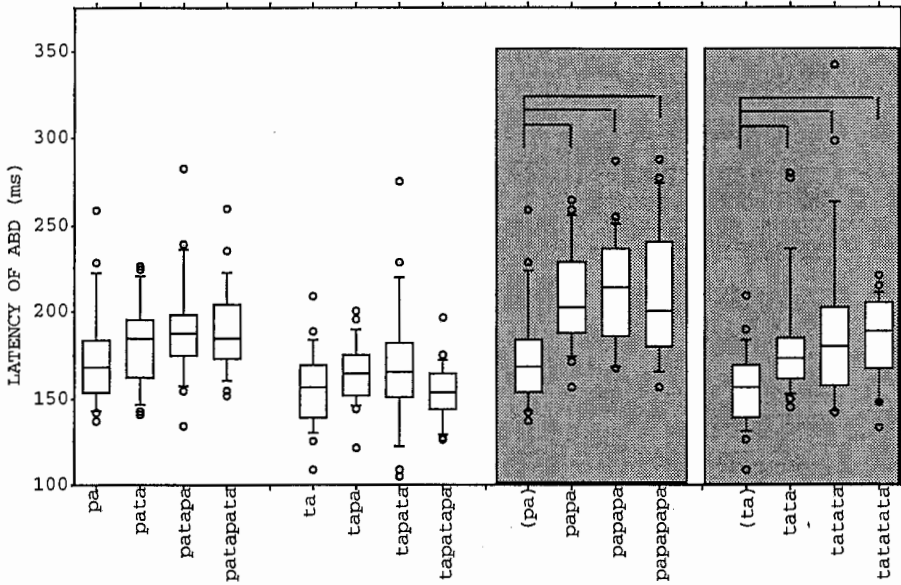


図 7: 各単語系列と音節数における ABD 発火潜時の分布

3.3 EMG 発火潜時と発話潜時の関係

音声における EMG 発火潜時と発話潜時の関係を調べた。ここでは予備的検討として、第3・第4セッションの「パパ系列」と「タタ系列」を分析の対象とした。データの数は、各系列内では1～4音節の各単語8データ、合計32データである。図8, 9それぞれに「パパ系列」と「タタ系列」における DLI, ABD の EMG 発火潜時と発話潜時の分布を示す。ここでは、音節数の区別をせず、32データをプールした分布を示している。潜時の分布を示すボックスグラフの表示形式は、図6および図7と同じである。

図中 25～75 パーセンタイルを示す箱と中央値を示す横線で見ると、「パパ系列」では、ABD の EMG 発火が DLI よりもやや早い傾向にあり、「タタ系列」では、その逆の傾向にある。

ただし、いずれの「系列」においても、EMG の発火潜時のばらつきと発話潜時のそれはほぼ同じ程度である。

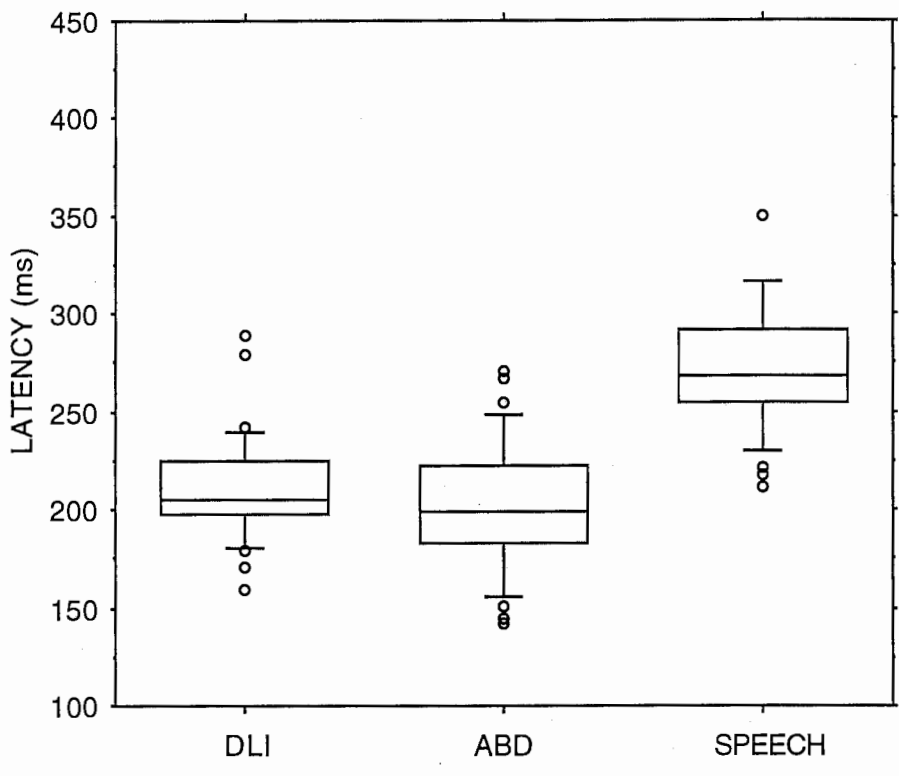


図 8: 「パパ系列」における DLI, ABD の EMG 発火潜時と発話潜時の分布

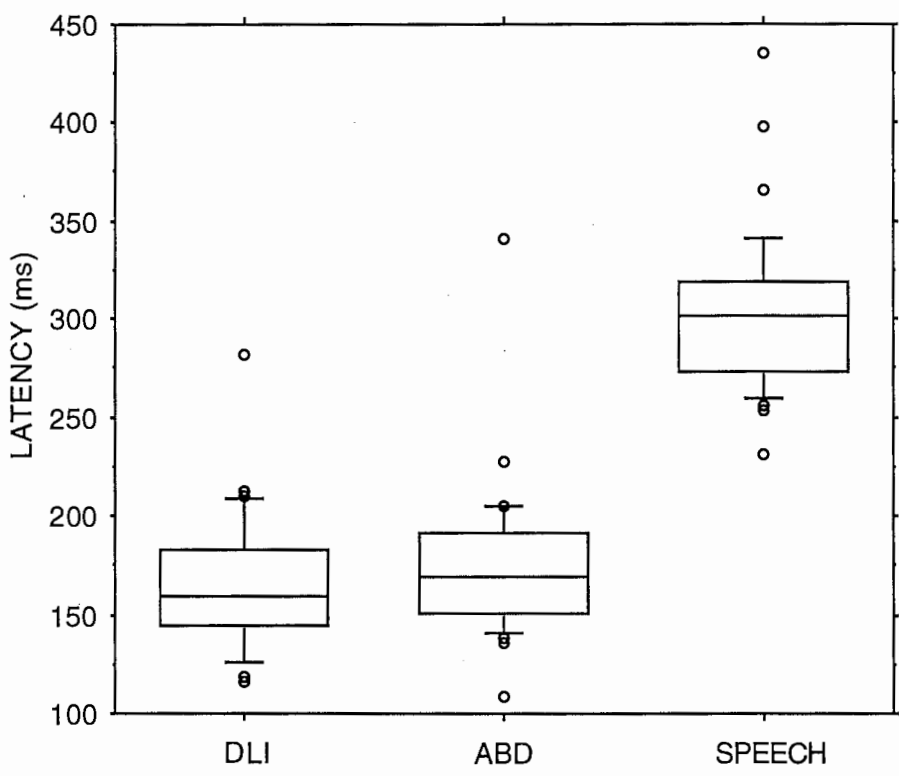


図 9: 「タタ系列」における DLI, ABD の EMG 発火潜時と発話潜時の分布

では、DLI と ABD の EMG 発火潜時と発話潜時の相対的な関係は、どのくらい保たれているのだろうか。これを知るために、2つの EMG 発火潜時と発話潜時の相関関係を上記のデータ（各系列1～4音節の単語をプールした32データ）について調べた。まず、「パパ系列」について、図10に DLI の EMG 発火潜時と発話潜時の相関を、図11に ABD の EMG 発火潜時と発話潜時の相関を示す。横軸が EMG 発火潜時、縦軸が発話潜時である。45度の対角線は両潜時が等しい場合を示す。また、全データについての回帰直線を実線で引いた。いずれの EMG 発火潜時とも発話潜時と 0.8 以上の高い相関がある。回帰直線の傾きは 1 より小さく、発話潜時が遅いときは EMG 発火潜時との差がわずかに、小さくなる傾向がみられた。

次に、「タタ系列」について、図12に DLI の EMG 発火潜時と発話潜時の相関を、図13に ABD の EMG 発火潜時と発話潜時の相関を示す。いずれの EMG 発火潜時とも発話潜時と 0.8 以上の高い相関がある。回帰直線の傾きはについては、DLI においては 1 よりわずかに大きく、ABD においては 1 よりわずかに小さかった。すなわち、DLI では発話潜時が遅いときは EMG 発火潜時との差がわずかに大きくなり、ABD ではわずかに小さくなる傾向がある。しかし、データの1つに、他のものにくらべて大きく離れたものがあり、これが大きく影響している可能性がある。

最後に「パパ系列」と「タタ系列」それぞれについての DLI と ABD の EMG 発火潜時の関係を図14, 15に示す。DLI と ABD の EMG 発火潜時の相互の相関も高く、「パパ系列」では 0.945、「タタ系列」でも 0.895 であった。DLI と ABD の時間差は、「パパ系列」では DLI が ABD よりもやや遅れる傾向にあり、「タタ系列」では逆に ABD が DLI よりもやや遅れる傾向にあった。回帰直線によれば傾きは「パパ系列」では 0.795、「タタ系列」では 0.776 であった。いずれも 1 より小さかったが、その結果、ABD の EMG 発火潜時が遅いほど、「パパ系列」では DLI の EMG 発火潜時が遅れる程度は小さくなり、「タタ系列」では DLI の EMG 発火潜時が先行する程度が大きくなった。

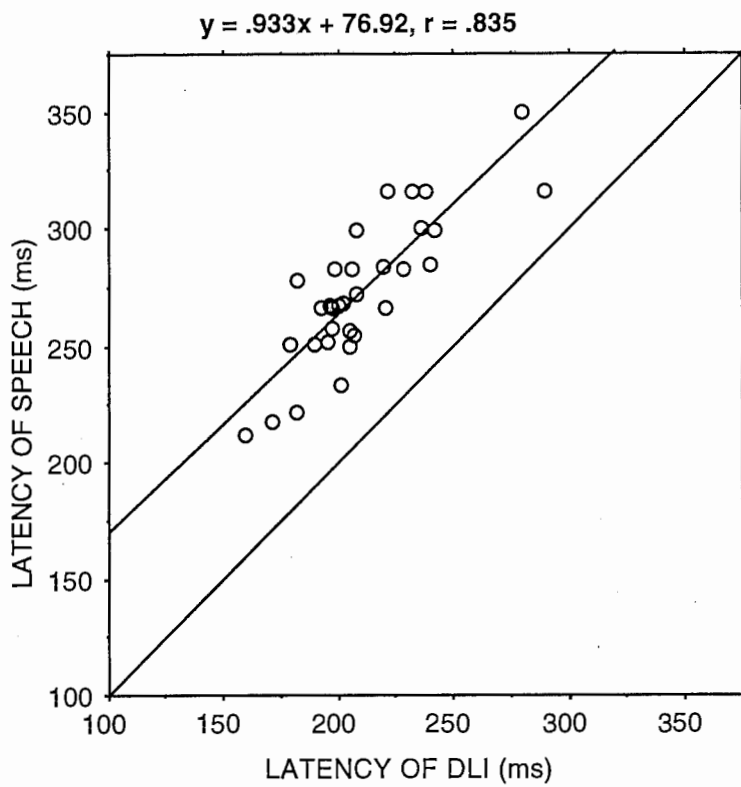


図 10: DLI と音声の相関 (パパ系列)

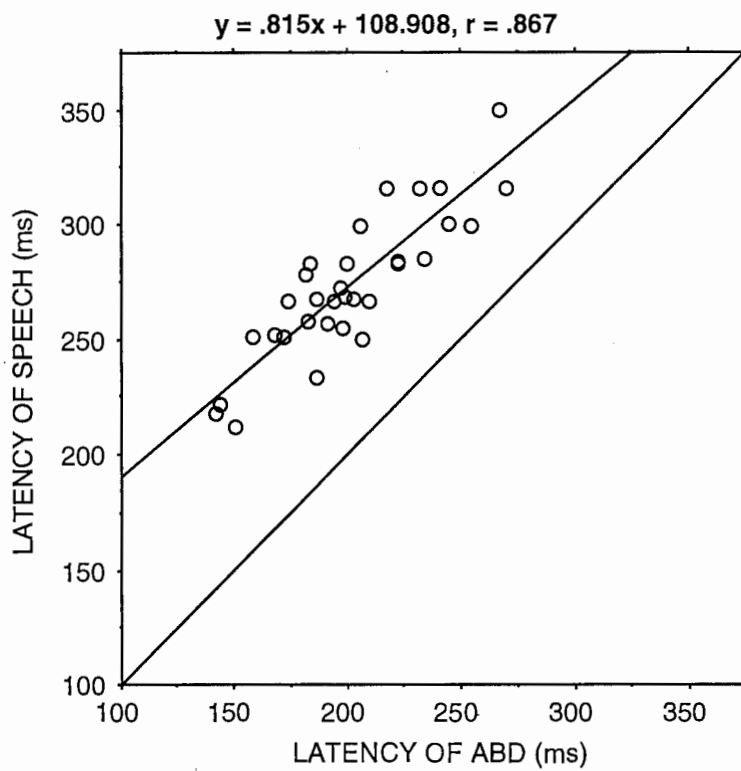


図 11: ABD と音声の相関 (パパ系列)

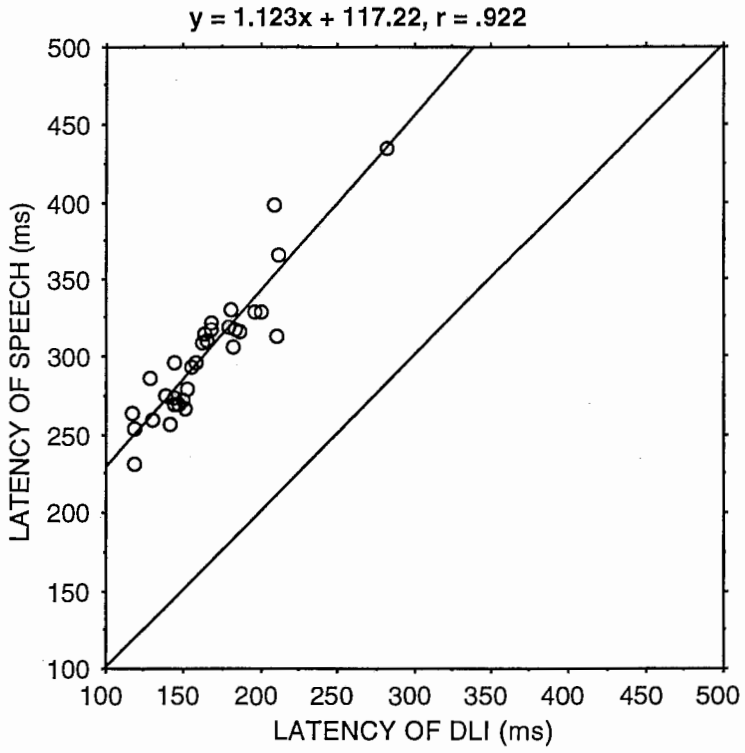


図 12: DLI と音声の相関 (タタ系列)

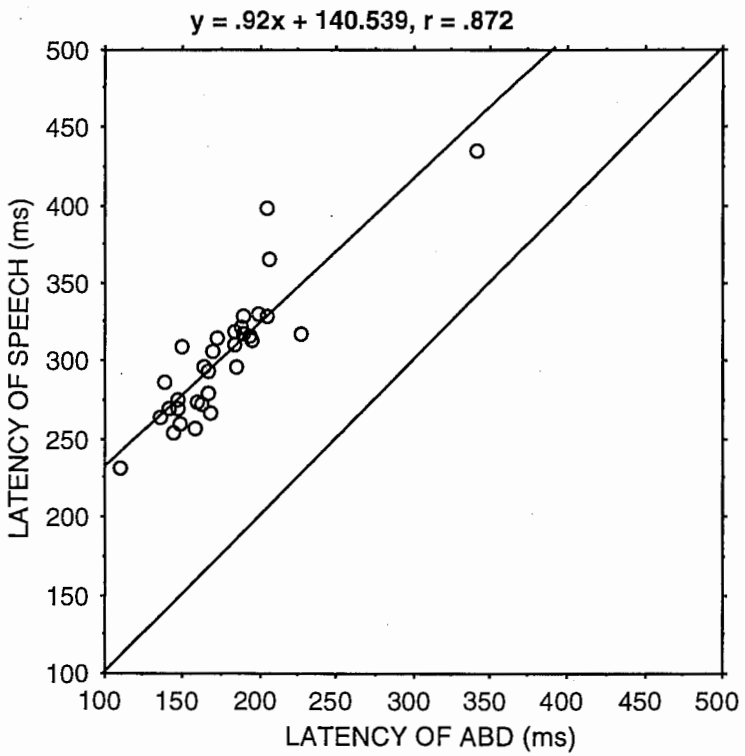


図 13: ABD と音声の相関 (タタ系列)

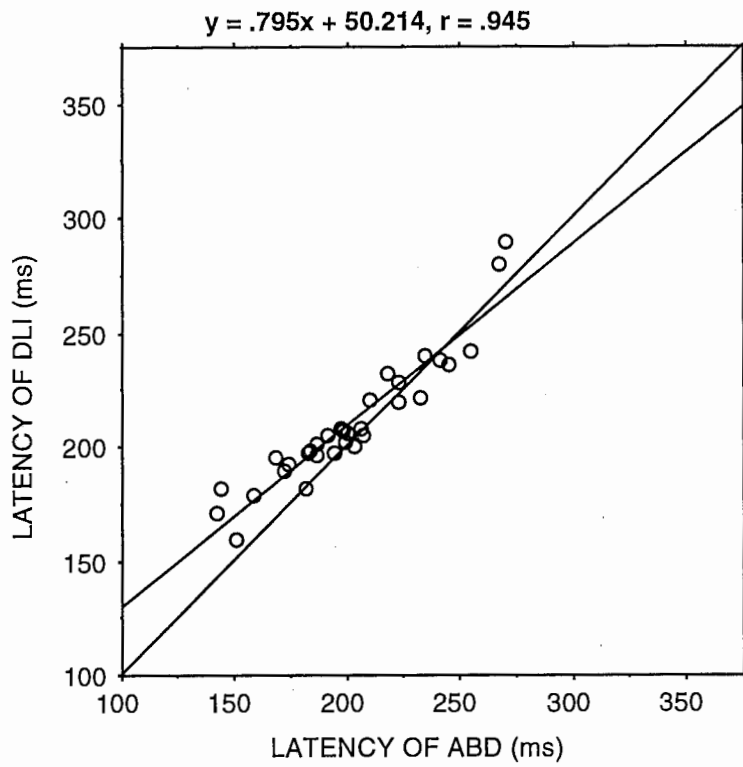


図 14: ABD と DLI の相関 (パパ系列)

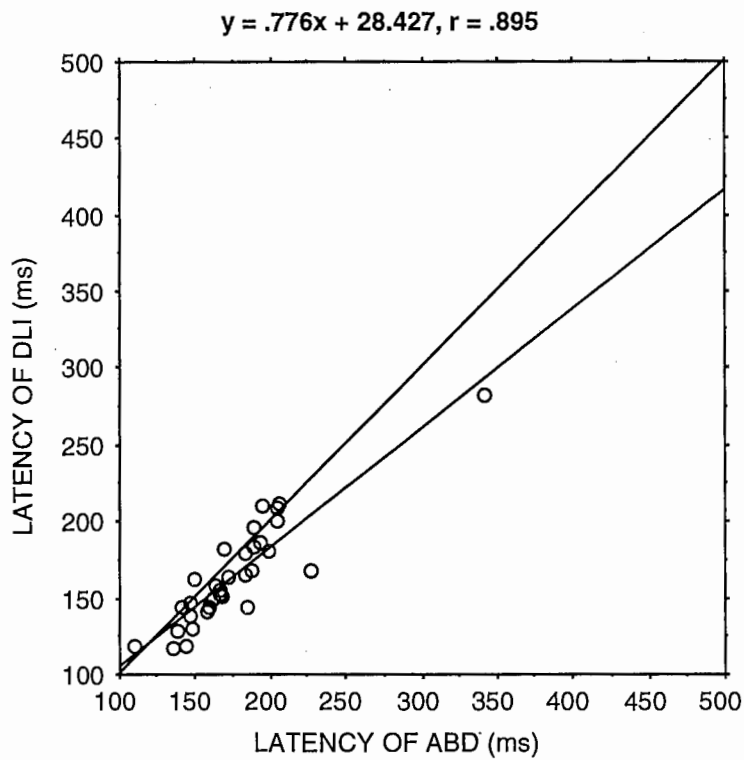


図 15: ABD と DLI の相関 (タタ系列)

4 考察

図6, 7から、「パパ系列」「タタ系列」のように同じ調音位置を持つ音節が連続する発話の場合に、音節数の増加がEMG発火潜時を増加させる傾向にあることがわかる。これは[3], [4]の実験において音声で測定される発話潜時にみられた傾向と同じである。なお、今回の実験では音節数の影響があった系列内でも平均値の差が有意であったのは、1音節の単語と2音節以上の単語の間のみであった。今回と同じ被験者([3],[4]の被験者1)による実験では、音声による発話潜時の測定の結果、それ以外の音節数の間にも平均値に有意な差がみられた。今回の実験データを用いた場合でも、音声による発話潜時の測定を行うと他の音節数の間にも有意な差が出るのか、あるいは、EMG発火潜時と同様に1音節と2音節の間にも有意な差が出るのか、という点については、今後検討する予定である。

ただし、一部(第3, 第4セッション)のデータについては発話潜時の測定を行ない、EMG発火潜時との関係を調べた。その結果、EMG発火潜時と音声における発話潜時のばらつきには大差がなく、EMG発火潜時の測定結果は、そのまま発話潜時の結果に直接反映される可能性が大きいことが示された。つまり、「1音節単語でのみ潜時が短い」というEMG発火潜時で得られた結果は、発話潜時の測定によっても同様に得られる可能性が大きい。同じ被験者でありながら、「音節数の潜時への影響が消失する傾向」があるとするれば、それはおそらく被験者の「慣れ」によるものであろう。

この「慣れ」による解釈はまだ、想像ではあるが、基本的な考えを述べておく。まず、調音の単位あたりにかかる時間が一定であると仮定する。はじめは比較的簡単な動きを含む単位が数多く組合わさって発話されているとする。この場合、発話の複雑さの影響が発話潜時に直接反映される。[3],[4]の実験を行った時点では、これが結果として現れたものと考えられる。それが「慣れ」により、簡単な動きを含む単位が結合されて、一つの複雑な動きを含む単位を構成するようになった。これにより、例えば、2音節から4音節までの単語の発話においては、単位の数が増え、潜時に変化が見られなくなったと考える。

ところで、発話すべき「系列」により発火する筋の順は異なっていた。すなわち、語頭の子音が[p]の場合はABDの発火潜時がDLIの発火潜時に先行するが、語頭の子音が[t]の場合は逆にDLIの発火潜時がABDの発火潜時に先行していた。これは口唇を開くことが主たる調音運動である[p]では、その前に、口腔内の狭めを取り除いておく必要がある。そのため、下顎を下げ、舌を下げるという動作が唇の開大に先行して行われなくてはならない。これが、DLIに先行してABDが発火するという現象として捉えられたのである。同様に、舌尖を歯茎に接触させて閉鎖を形成し、口腔内圧を上昇させ、それを開放することによって生成される[t]の場合、これらの動作の前に、口唇は開いていなくてはならない。舌は下顎に乗っているため、下顎の下降運動の開始時点は、口唇の開大や舌の閉鎖完了時点よりも後になる。このことが、ABDに先行してDLIが発火するという現象として捉えられたのである。

なお、このような発話すべき音素によって拘束されている、各筋の発火時点の時間関係は、EMG発火潜時の増加により変化する傾向にあった。しかしこれが、音節数の増加によって変化するのかどうかについては、分析に用いたデータ数が少ない(第3, 第4セッションについてのみ分析)ため、明確に述べるには至らなかった。

今後、EMGと同時に採取した音声データについて発話潜時を測定し、今回明確にできなかった、音節数の発話潜時への影響の学習効果の有無、音節数のEMG発火潜時と発話潜時の相対的な関係への影響の有無、について検討する予定である。また、さらに被験者の数

を増やして検討することも計画している。

ところで、「はじめに」で述べたように、音声で測定できる発話潜時には、(1) 発話開始合図が出たと判断して発話準備を開始するまでの時間、(2) 発話の準備の処理にかかる時間、および、(3) 筋電位発生時点から語頭音開始時点までの時間、が含まれている。今回の実験で、(1)+(2)+(3)の時間によるばらつきは、(1)+(2)の時間のばらつきはほぼ等しかった。すなわち、(3)自体のばらつきは、(1)+(2)の時間のばらつきよりも、小さいことが示唆された。もし(1)の時間をなんらかの方法で検出できるとすれば、我々が本来測定したい、(2)の時間を正確に測定できる可能性がある。実際、発話潜時のばらつきの最大の原因は、発話開始合図の見落としであることが多いようである（これを確認するための実験が必要であろう）。そこで、たとえば脳磁図(MEG)などを用いることにより、(1)の時間を測定することができないか、検討する予定である。

謝辞

今回の実験を進めるにあたって、協力、助言をいただいた、ATR 人間情報通信研究所、第4研究室の草川直樹氏、大西通博氏、平井啓之氏、党建武氏、第1研究室の河原英紀氏、山田玲子氏に感謝いたします。また、本報告書の第一著者、西本卓也に本実験のための実習の機会を与えてくださった早稲田大学の白井克彦先生、小林哲則先生、ATR 人間情報通信研究所の東倉洋一社長に感謝いたします。

参考文献

- [1] Henry, F.M. & Rogers, D.E.: Increased Response Latency for Complicated Movements and a "Memory Drum" theory of neuromotor reaction, *Research Quarterly*, 31:448-458 (1960)
- [2] Sternberg, S., Knoll, R.L., Monsell, S. and Wright, C.E.: Motor Programs and Hierarchical Organization in the Control of Rapid Speech, *Phonetica*, 45:175-197 (1988)
- [3] 正木信夫、本多清志: 無意味単語の発話企画時間への語頭音とモーラ数の影響, *日本音響学会講演論文集* 1-8-2 (1994.3)
- [4] Masaki, S. & Honda, K.: Estimation of Temporal Processing Unit of Speech Motor Programming for Japanese Words Based on the Measurement of Reaction Time, *Proceedings of ICSLP94* (1994.9)

