

平易な文における日本語母語話者グループ間の 文処理方法の違いについて

澤 崎 宏 一

日本語の文処理についての研究は、袋小路文と非袋小路文、かき混ぜ文と通常語順文、同格連続文と同格非連続文といったような対立ペアを比較させることにより、文処理の方略や制約に関して論じられることが多い(cf., 坂本, 2001 a; 2001 b; 2001 c; 2001 d; 2001 e)。これらの実験では、日本語母語話者の被験者を同質の集団として扱うのが通常である。しかしながら、被験者の年齢やワーキングメモリ、実験文に使われる語彙の親密度などの違いによって、文処理データにも違いが現れることは報告されており (Daneman and Carpenter, 1980; Just and Carpenter, 1992; Kondo and Mazuka, 2006; 苧阪, 2002)、同じ言語の母語話者であるというだけで言語データをいつも等しく取り扱うことには問題がある。

本稿では、先行実験とその追実験を通して、ふたつの日本語母語話者グループの読み時間を比較する。その結果、たとえ平易な日本語文を読む場合であっても、被験者集団によって文処理に違いの現れる事象とそうでない事象のあることを示す。さらに、同一被験者が文を読む場合、必ずしも一貫性のある読み方をいつも行っているわけではないことも併せて論ずる。

先行実験

ここでは、Sawasaki (2004, 2007) で行われた、移動窓読み (Moving Window) の実験内容と結果の一部を報告する。¹ 実験の目的は、ごく平易な単文を読む際に、(i)目的語などの項と副詞などの付加句の処理速度に差があるのか、(ii)文全体を通してどのような読み時間の配分が見られるかの2点を調べることであった。

実験方法

被験者は、神戸・大阪地区にある5つの大学・大学院生から成る、日本語母語話者

1 後述するように、本稿で提示する実験結果は新しい方法で分析されているので、Sawasaki (2004, 2007) と全く同一ではない (脚注4、脚注5も参照)。

53人であった。² 課題文の文型タイプは(1)に示す5つである。

- (1) a. Int 1: 田中さんがびっくりした (主語+自動詞)
- b. Int 2: 田中さんが昨日びっくりした (主語+副詞1+自動詞)
- c. Int 3: 田中さんが昨日とてもびっくりした (主語+副詞1+副詞2+自動詞)
- d. Tran 2: 田中さんがお客を手伝った (主語+目的語+他動詞)
- e. Tran 3: 田中さんが昨日お客を手伝った (主語+副詞1+目的語+他動詞)

Int 1からInt 3の3タイプは主語と自動詞だけ共通しており、Int 2タイプには副詞が1つ、Int 3タイプには副詞が2つ主語のあとに現れている。副詞1は「昨日」「先週」のような時間を示す副詞句で、副詞2は「とても」「ちょっと」のような程度や様式を示す副詞句であった。次にTran 2とTran 3タイプは、主語と他動詞が共通しており、Tran 2タイプには主語と目的語の間に時間を示す副詞1が置かれた。

課題文の読み時間が適正に計られるよう、(1)の語句が文頭や文末に表れないように配慮された結果、(1)は場所を表す副詞句と後続助動詞（または動詞）で挟まれ、(2)のような文が作られた。

- (2) 図書館で / 田中さんが昨日とてもびっくりした / と言いました。(Int 3)

課題文は全部で20セットあり、さらにフィラーとしての40文が加えられた。最終的に、同じ被験者が同一語句を繰り返し読むことのないように、これらの刺激文を5リストに分けて実験材料が用意された。

実験の手続きとして、被験者ペースによる移動窓の読み方法が採られた。各文が文節毎にコンピュータスクリーン上に現れ、被験者が所定のキーを押すことで次の文節へと順次読み進めていくというものである。キーが押されるたびに、文節毎の読み時間がコンピュータ内に自動的に記録された。また、刺激文の半数については、文終了後に内容確認のための簡単な質問を伴った。

先に述べたように、この実験の目的の1つは、目的語と副詞の読み時間に違いがあるかを調べることであった。課題文タイプ(1b)から(1e)の内、第2領域（文節）を見ると、Tran 2だけが目的語「お客を」で、あとの3つ（Int 2, Int 3, Tran 3）には同じ副詞1「昨日」が現れている。また、Int 3とTran 3タイプの第3領域でも副詞2「とても」と目的語「お客を」の対立がある。Abney(1986)の解析モデルやFrazier

2 Sawasaki (2004, 2007) では、被験者は神戸地区の4つの大学・大学院とあるが、実際には神戸・大阪地区の5つの大学・大学院の間違い。尚、この実験は本来日本語学習者を対象に行われたもので、日本語母語話者は統制群として参加した。また、日本語学習者の漢字レベルに合わせて、刺激文は必ずしも全て漢字で表示されているわけではなかった。

and Clifton(1996)のConstrual理論では、項の処理は付加句より優先されるとしており、英語母語話者のデータを見ても、項の処理時間が付加句よりも短く表れているものが多い (Ahrens, 2003; Boland and Blodgett, 2006; Schutze and Gibson, 1999; Speer and Clifton, 1998)。³ 先行研究に見られるこのような項の優位性が、日本語にも同様にあてはまるならば、(1)のような刺激文タイプにおいても、副詞に較べて目的語の読み時間が速くなることが予想される。

結果

図1は、被験者53人の内、内容確認の質問への正答率が90%を上回った計50人の結果をグラフに表したものである。縦軸にモーラ毎の平均読み時間（ミリ秒単位）をとり、横軸には領域を示した。便宜上、以下では文型タイプ(1 a)を除く(1 b)から(1 e)までの4つのタイプについてのみ結果を報告する。⁴ また、後で紹介される実験の被験者グループと区別するため、この被験者グループをグループAと呼ぶことにする。

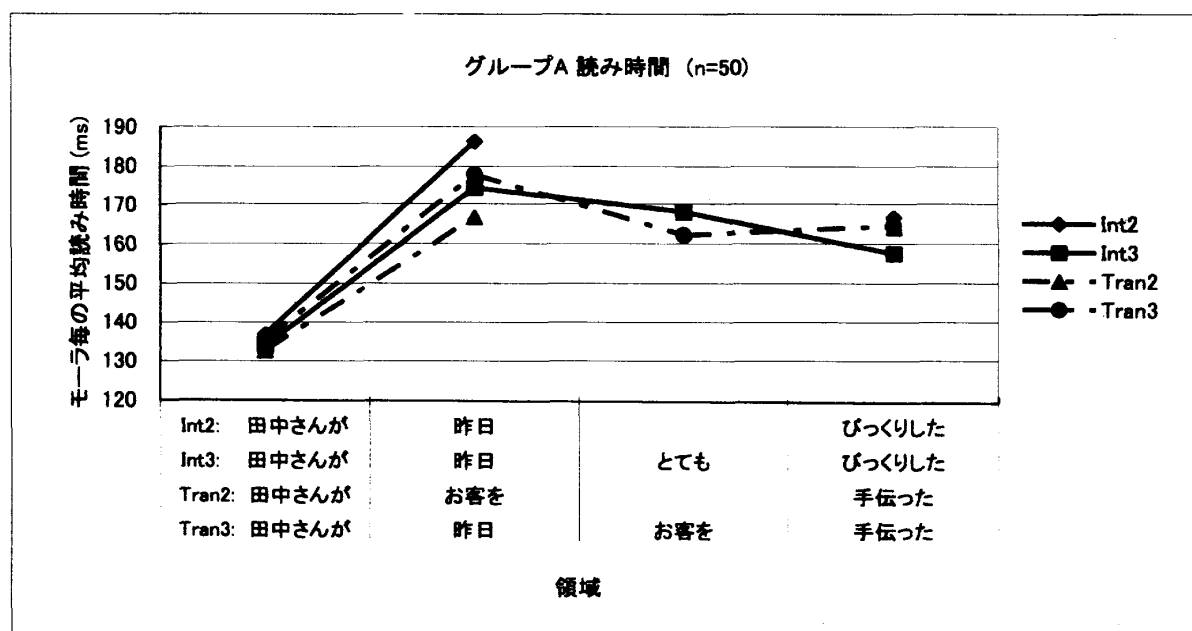


図1. グループA 読み時間

図1からは、目的語と副詞の読み時間は、第2領域と第3領域の両方でわずかな違いが認められる。この違いについて、2つの領域でそれぞれ文型タイプを主要因とす

³ Schutze and Gibson (1999)やSpeer and Clifton (1998)では、処理時間の差は該当語句の後続領域で確認されていた。

⁴ Sawasaki (2004, 2007)では、結果はモーラ毎の残差読み時間(Residual Reading Times)で計算されている。本稿では、残差読み時間ではなく、モーラ毎の単純平均読み時間で計算し直した上で分析結果を提示する。その他、統計方法も若干変更した。2つのデータ提示方法による統計分析結果を比較したところ、平均読み時間による分析にやや強い効果が表れる箇所があったものの、全体的には両者に殆ど差は見られなかった。

る分散分析を行った。その結果、第2領域では「お客を」の読み時間が「昨日」の読み時間に比べて有意に速かった($F(1,147)=2.769, p<.05$; $F(2,57)=2.236, p=.094$)。⁵

しかし、Bonferroniによる多重比較を行ったところ、読み時間の違いはInt 2とTran 2の間でのみ有意差が確認され($F(1, p<.05$; $F(2, p=.155)$)、他のペアに違いはなかった。次に第3領域における分析では、目的語と副詞の読み時間の違いに有意差は認められなかった($F(1,149)=1.537, p=.221$; $F(2,19)=.215, p=.648$)。尚、その他の領域については、読み時間に有意差はなかった。

先行実験の考察と問題点

第2領域において目的語と副詞句に読み時間の違いが認められたことから、Sawasaki (2007) では、項と付加句の処理方法の違いが主張された。しかし、多重比較の結果からわかるとおり、目的語と副詞句の読み時間の違いはTran 2とInt 2の間でしか起こっておらず、他のペア比較では有意差が現れていない。また、第3領域における比較では目的語と副詞句の差が見られなかったことから、このままでは説得力のある結論とは言い難い。⁶

十分な有意差が得られなかった理由として、Sawasaki(2007)は課題文があまりに平易すぎたため(処理負担が軽く)、被験者の個人差が表れやすくなったのでであろうと推測している。しかしながら、個人差が理由であるとすれば、問題文の平易さ以外にも、被験者グループが均質ではなかったこともまた個人差表出の理由として考えられるだろう。⁷ 先行研究でも、ワーキングメモリの個人差が実験結果を左右することが報告されている(Daneman and Carpenter, 1980; Just and Carpenter, 1992)。例えば、ワーキングメモリと大学入試センター試験などに見られるような国語長文読解問題の成績との相関関係が指摘されており(苧阪, 2002)、その意味で被験者となる学生の構成要素を均質に保つことは重要である。

前述したとおり、この実験の被験者は、神戸・大阪地区にある複数の大学のいずれかに通う学生であった。表1はその内訳である。

5 本稿では被験者分析($F(1)$)と項目分析($F(2)$)の両方の分析結果を提示するが、被験者分析による結果を主な判断材料とする(cf., Hino and Lupker, 1996; 宮本, 2002)。

6 このことから、Sawasaki(2007)では50人の被験者グループを、データが均質と思われる30人に絞り込んで再分析している。その結果、第2領域における結果の改善は見られなかったが、第3領域においては項の読み時間が副詞句に比べて有意に長かったことが指摘された。しかし、これは実験の主眼の1つであった項の優位性には反する結果である。このため、文処理方略として「項の優位性」が常に母語話者に作用するという見方は退けられ、「目的語」、「時間の副詞」、「程度・様式の副詞」の持つそれぞれの語の性質がどの程度後続語句を予測し易くさせているかが、読み時間の差につながったと考察している。本稿では、あくまでも50人の被験者について考察することにする。

7 先行実験の問題点は他にもあるが、ここでは本稿の趣旨に沿った点についてのみ述べる。

平易な文における日本語母語話者グループ間の文処理方法の違いについて

学 校	区 分	専 攻 分 野	人数(男,女)	平均年令(SD)
神戸松蔭女子学院大学	学部生	文学・言語学	20 (0,20)	21 (0.76)
流通科学大学	学部生	商学・経済学	10 (9,1)	20 (1.65)
神戸市外国語大学	学部生	外国語・国際関係学	10 (0,10)	19 (0.79)
神戸大学	大学院生	看護学	7 (0,7)	30 (5.28)
大阪学院大学	学部生	経済学	1	19 (5.28)
不明	不明	不明	2 (1,1)	28 (12)
合 計			50 (11,39)	22 (4.55)

表 1. Sawasaki(2004,2007)における被験者内訳

表 1 からわかるとおり、50人の被験者は所属大学も専攻分野も散らばりがあり均一ではない。年令分布も全ての大学で共通しているとは言えない。表には示されていないが、出身地を見てみると、大半が京都・大阪・兵庫の 3 県であるが(30名)、わずかながら広島(3)、鳥取(1)、香川(1)、愛媛(1)や石川(1)などにも分布が見られた(不明13)。以上のように、被験者の50人は、関西地区の大学(院)生であることの 1 点を除き、特定の素性で括ることが非常に困難なグループであったことがわかる。

表 1 にある各大学の被験者同士を比較してみることは有用だが、各大学の被験者数が一定でない上に数も少ないので現実的ではない。そこで次節以降では、新たに設定したより均質な被験者グループに対して追実験を行い、その結果をSawasaki(2004, 2007)の結果と比べてみたい。

新しい被験者（グループ B）による追実験

本節では、前節で示されたSawasaki(2004, 2007)の追実験の結果を示す。前節の被験者（グループ A）に対し、新しい被験者をグループ B とする。

被験者

グループ B の被験者は静岡県立大学に通う大学生で日本語母語話者40人（内男性 3 名）であった。専攻は 1 人（国際関係学）を除き言語学・文学・文化などの人文科学分野である。平均年令は19才(SD: 0.79)で、グループ A と比べて若干値が小さい。出身地は半数以上が静岡県内(24名)であるが、栃木(1)、茨城(1)、神奈川(1)、愛知(1)、岐阜(1)、三重(2)、福井(2)、石川(1)、不明(6)といった散らばりも見られた。出身地の上では若干のばらつきがあるが、所属大学、専攻、年令などの点では均質なグループと言える。

尚、実験の材料と手続きは前節で示した先行実験と同一とした。

結果

図 2 は、被験者40人の内、内容確認の質問への正答率が90%を上回った計38人の結

果をグラフに表したものである。

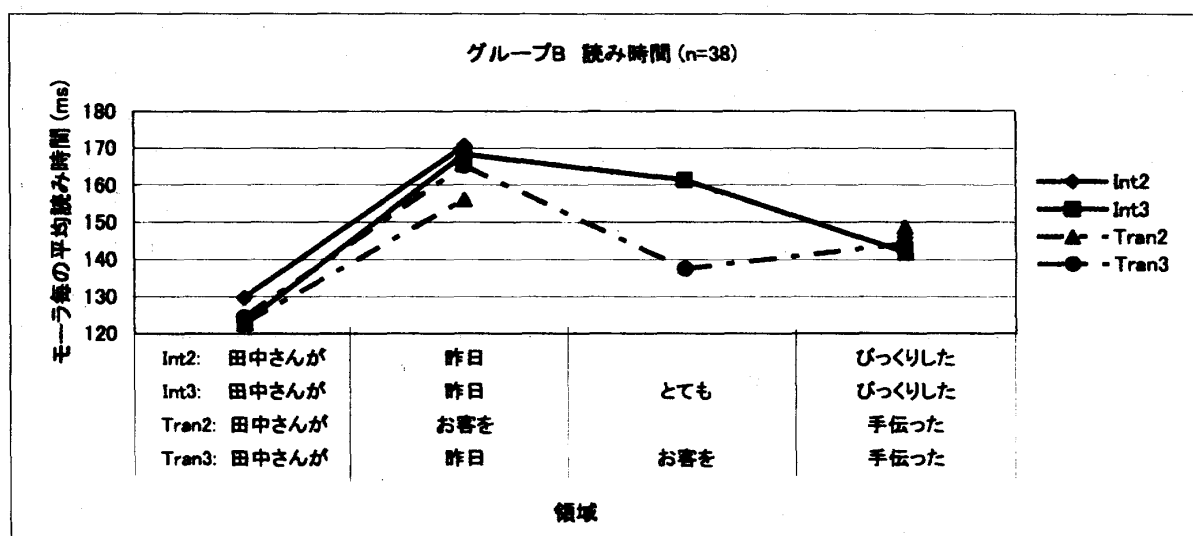


図2. グループB 読み時間

図2から、目的語と副詞の読み時間は、第2領域でわずかな違いが、第3領域では明確な違いが認められる。この違いについて、各領域で文型タイプを主要因とする分散分析を行った結果、第2領域では有意差が認められなかったが($F_{1(3,111)}=1.001$, $p=.395$; $F_{2(3,57)}=.704$, $p=.554$)、第3領域では目的語が有意に速く読まれたことがわかった($F_{1(1,37)}=8.119$, $p<.01$; $F_{2(1,19)}=3.860$, $p=.064$)。その他の領域では読み時間に違いは見られなかった。

グループAとの比較で考えると、第2領域では、グループAは若干の違いを示したが、グループBでは違いが認められなかった。また第3領域では、グループAは違いがなかったのに対してグループBは有意差があった。どちらの被験者グループも、目的語が速く処理されたという点では同じであったが、その違いが表れる領域が被験者グループによって異なっていた。

次節以下では、両グループを更に詳しく比較して考察を加える。その際、(i)各領域での読み時間、(ii)各文型タイプの読み方のパターン、(iii)被験者内の一貫性についての3つの観点から比較を試みる。

各領域の読み時間のグループ間比較

グループAとグループBの読み時間を、領域毎に比較するため、被験者グループと文型タイプを要因とする2 x 4 (但し第3領域は2 x 2) の分散分析を各領域で行った。表2に両グループの読み時間の平均値と標準偏差を示す。

平易な文における日本語母語話者グループ間の文処理方法の違いについて

タイプ	グループ	田中さんが		昨日/お客を		とても/お客を		驚いた/手伝った	
		平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
Int 2	A	136.94	64.44	186.49	90.41			167.24	75.94
	B	129.84	73.95	170.65	90.66			145.97	65.71
Int 3	A	133.90	69.39	174.55	80.98	168.53	78.68	158.14	78.36
	B	122.94	62.84	168.36	82.27	161.42	84.15	142.00	69.18
Tran 2	A	132.74	64.28	167.10	81.26			164.92	78.75
	B	122.83	66.20	156.21	80.50			148.90	68.81
Tran 3	A	135.64	69.31	177.98	82.69	162.69	74.48	165.15	77.70
	B	124.51	57.11	165.35	84.57	137.52	58.80	144.34	65.46

表 2. 両グループの読み時間平均値と標準偏差

まず第一領域（「田中さんが」）では、被験者グループ($F(1,86)=1.258$, $p=.265$; $F(2,138)=2.388$, $p=.131$)、文型タイプ($F(1,3,258)=1.118$, $p=.340$; $F(2,3,114)=.612$, $p=.608$)共に違いが認められず、交互作用もなかった($F(1,3,258)=.069$, $p=.969$; $F(2,3,114)=.108$, $p=.955$)。

次に第2領域（「お客を」/「昨日」）では、被験者グループの主効果はなかったが($F(1,86)=.962$, $p=.329$; $F(2,138)=1.380$, $p=.247$)、文型タイプ的主効果が見られ($F(1,3,258)=3.176$, $p<.05$; $F(2,3,114)=2.230$, $p=.089$)、交互作用はなかった($F(1,3,258)=.304$, $p=.823$; $F(2,3,114)=.152$, $p=.928$)。Bonferroniによる多重比較の結果、Tran 2タイプ（「お客を」）がInt 2（「昨日」）に比べ有意に速く処理されたことが確認された($F(1, p<.01$; $F(2, p=.156)$)。被験者グループ間の差がないにも拘わらず、全体として目的語(Tran 2)と副詞句(Int 2)の違いがあったことから、グループBもグループA同様、項の優位性を潜在的に有している可能性があると言える。

続く第3領域（「お客を」/「とても」）では、被験者グループによる主効果はなく($F(1,86)=2.640$, $p=.108$; $F(2,138)=2.818$, $p=.101$)、文型タイプによる主効果が見られた($F(1,86)=19.106$, $p<.001$; $F(2,138)=3.085$, $p=.087$)。また交互作用も確認された($F(1,86)=7.976$, $p<.01$; $F(2,138)=1.268$, $p=.267$)。交互作用をさらに分析したところ、この結果はグループBが目的語をより速く処理しているためであることがわかった。このことから、第3領域における項の優位性は、グループBに強く表れていると言える。

最後に第4領域（「びっくりした」/「手伝った」）では、被験者グループによる主効果が見られ、グループBの動詞の読み時間が有意に速かったことがわかった($F(1,86)=4.374$, $p<.05$; $F(2,138)=2.381$, $p=.131$)。文型タイプでは主効果が現れず($F(1,3,258)=1.020$, $p=.384$; $F(2,3,114)=.680$, $p=.566$)、交互作用も確認できなかった($F(1,3,258)=.255$, $p=.858$; $F(2,3,114)=.133$, $p=.940$)。つまり、自他動詞や文型タイプの

別に拘わらず、動詞を処理する速度はグループBが一貫して速かったと言える。

考察

上の比較結果をまとめると、項の優位性は、グループBでより強く現れたことになる(Tran 3「お客を」/ Int 3「とても」)。また動詞においてもグループBが速い処理時間を示したことなどから、同じ母語話者であっても、ふたつの被験者グループは文処理の方法に違いがあると言える。グループBは、動詞や格情報を含む領域が速く読まれたので、基本的な文構造に関する情報の処理が、より自動的に行われたと言えるかも知れない。

グループAに比べ、グループBは所属大学、専門分野、年令の点でより均質な母語話者グループであった。2つのグループを分かつ原因が具体的に何であるかは、今回の実験だけでは特定できないが、被験者を均質にしたことが何らかの影響を与えている可能性は高い。

同じ母語話者間でも文処理方法が異なる原因のひとつとして、ワーキングメモリの個人差がある。例えば、Just and Carpenter (1992)は、(3)のような袋小路文を用いて、英語母語話者の読み時間を測定する実験を行った(眼球運動の測定)。その際、被験者をリーディングスパンテスト(RST)の高得点者と低得点者のグループに分けたところ、高得点グループでは、下線部の"by the lawyer"において、(3a)の読み時間が(3b)に比べて有意に長かったことがわかった。⁸ しかし低得点グループではそのような違いは認められなかった。

- (3) a. The defendant / examined / by the lawyer / turned out / to be unreliable.
b. The evidence / examined / by the lawyer / turned out / to be unreliable.

ふたつの文は文構造は同じだが、主語が有生名詞か無生名詞かにおいて異なる。主語が有生名詞("the defendant")の場合は、後続動詞が"examine"の過去形としても過去分詞(関係節の一部)としても解釈が可能となり、"by the lawyer"を読むまでは文構造が曖昧である。しかし主語が無生名詞("the evidence")の場合は、曖昧性は消滅し、"examined"が過去分詞だとする解釈だけが残る。実験の結果、Just and Carpenterは、RSTの高得点者のみがこういった名詞の有生性の情報を有効に使って文処理をしたとし、同じ母語話者でも文処理方法に違いがあることを示した。

ワーキングメモリの個人差は、英語だけでなく、日本語の文処理においても関連性が指摘されている。Nakano, et al.(2000)は、(4)のようなかき混ぜ文を被験者に聞かせ、下線部分とその500 ms前の位置でそれぞれ「レモン」の語を提示して語彙判断

8 リーディングスパンテストとは、ワーキングメモリの個人差を測る目的で開発されたテストである。(3)の実験文は、元々Ferreira and Clifton (1986)で用いられたのと殆ど同じもので、ワーキングメモリと文処理の相関関係を検証するため、Just and Carpenter (1992)で追実験用に使われた。

テストを行った。

- (4) レモン_iを[二人目の人が司会者に[その子供が女の人に___iねだっている]と答えた。

その結果、RSTの高得点グループは、下線部で「レモン」が提示されたときの方が、500 ms前に提示されたときよりも有意に速く語彙判断を行っていた。それに対し、低得点グループは、どちらのタイミングで「レモン」を提示されても判断にかかる時間は同じであった。このことからNakano, et al.は、高得点者グループの方が、文頭で聞いた「レモン」が、下線の位置でより活性化され易かったと結論づけている。

本稿で見た2つの母語話者グループの違いが、ワーキングメモリの個人差に由来しているかどうかは、さらなる研究を待たなければわからない。しかし、曖昧性を持つ袋小路文や複雑なかき混ぜ文といった文に限らず、本実験で扱ったようなごく平易な単文を読む場合であっても、同じ母語話者同士で違いが出るということは、大変興味深い発見であったと言える。

各文型タイプの読み方のグループ間比較

前節では各領域における読み時間のグループ比較を行ったが、ここでは、読み時間が課題文の最初から最後までどのように推移していったかに着目して、各文型の読み時間のパターンをグループ間で比較してみたい。文型毎に、被験者グループと領域を要因とする2 x 4（但しInt 2とTran 2タイプは2 x 3）の分散分析を行った。領域の主効果は全ての文型タイプで認められたが、以下では、結果としてより重要となる、被験者グループによる主効果と、被験者グループと領域の相互作用についてのみ分析結果を記すことにする。比較を容易にするために、下の図3から図6に、各文型タイプの読み時間の推移をグラフで示す。

まずInt 2タイプ（図3）では、被験者グループによる主効果は見られなかった（ $F(1,86)=2.260$, $p=.136$; $F(2,138)=2.443$, $p=.126$ ）。被験者グループと領域の交互作用も現れなかった（ $F(1,2,172)=.883$, $p=.404$; $F(2,2,76)=.583$, $p=.516$ ）。

次にInt 3タイプ（図4）でも、被験者グループによる主効果はなく（ $F(1,86)=1.013$, $p=.317$; $F(2,138)=1.584$, $p=.216$ ）、相互作用も確認できなかった（ $F(1,3,258)=.478$, $p=.671$; $F(2,3,114)=.174$, $p=.870$ ）。

Tran 2タイプ（図5）でも、被験者グループによる主効果はなかった（ $F(1,86)=1.542$, $p=.218$; $F(2,138)=2.077$, $p=.158$ ）。また被験者グループと領域の相互作用も確認できなかった（ $F(1,2,172)=.297$, $p=.729$; $F(2,2,76)=.096$, $p=.908$ ）。

最後にTran 3タイプ（図6）では、被験者グループ間での主効果に有意傾向が見られた（ $F(1,86)=3.401$, $p=.069$; $F(2,138)=4.380$, $p<.05$ ）。しかし交互作用は現れなかつ

た($F(1,258)=1.528$, $p=.211$; $F(2,114)=.585$, $p=.595$)。図6にあるとおり、被験者グループAに比べて、グループBの方が一貫して処理速度が速い。各領域において、 t 検定を用いて2つのグループの平均値を比較したところ、グループBの読み時間が有意に速かったのは第3領域の目的語と($t(186)=1.048$, $p<.01$; $t(2(38)=2.522$, $p<.01$)と第4領域の動詞においてだった($t(186)=1.946$, $p<.05$; $t(2(38)=1.426$, $p=.081$)。この結果は、前節で示された、領域別の比較結果と対応していると言える。

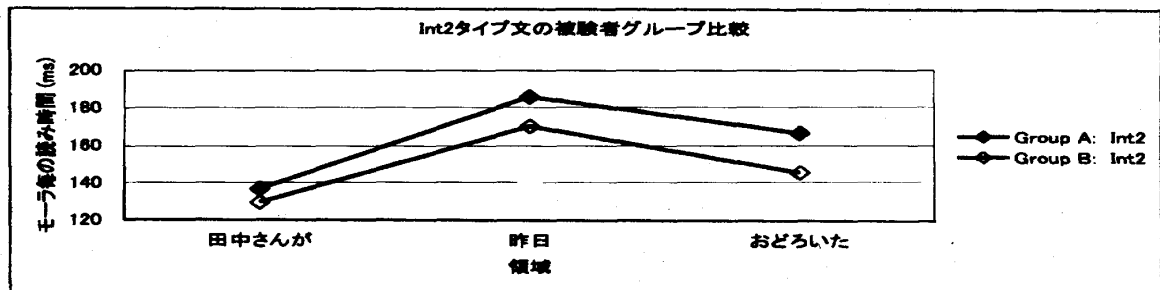


図3. Int 2タイプのグループ間比較

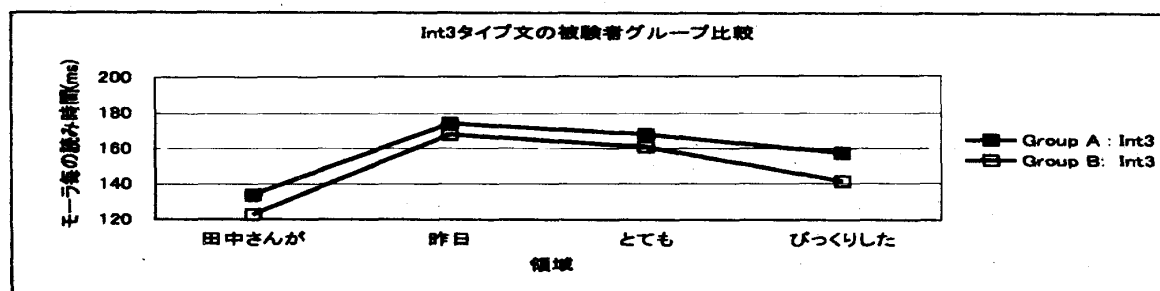


図4. Int 3タイプのグループ間比較

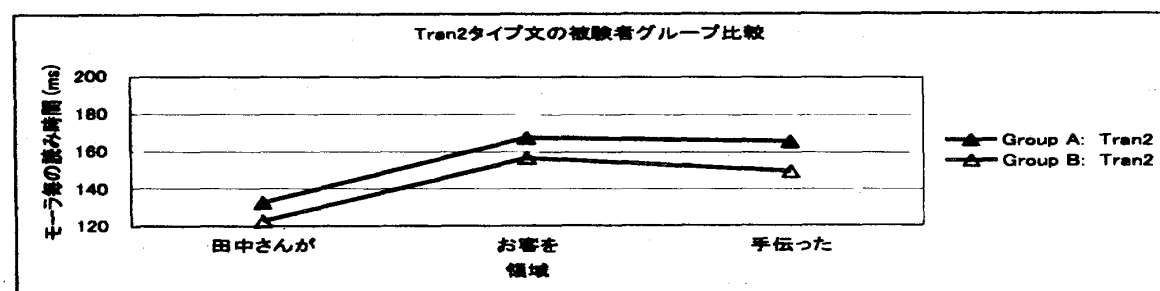


図5. Tran 2タイプのグループ間比較

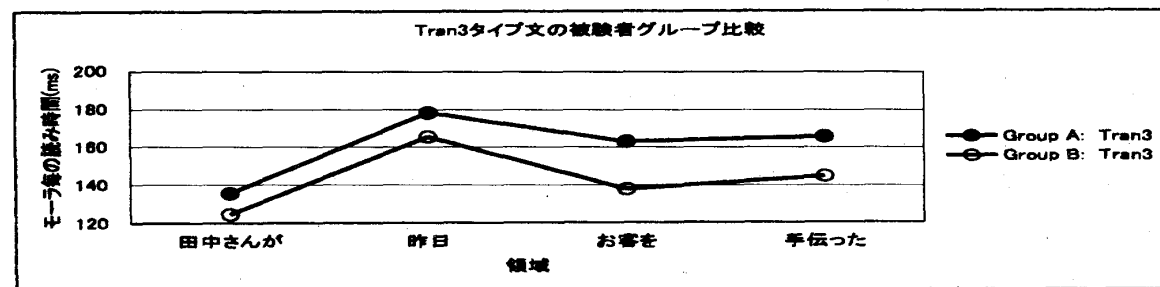


図6. Tran 3タイプのグループ間比較

平均的読み方（プロトタイプ）についての比較

上の図3～図6に示したパターンは、それぞれの領域での読み時間の平均値を線でつなげたものに過ぎない。いわばプロトタイプの読み方であり、各被験者の実際の読み方と必ずしも同一であるとは限らない。例えば、意味のプロトタイプ論でよく引き合いに出される、鳥のプロトタイプのイメージが実際の鳥の姿と常に合致しているわけではなく、鳥の種類によってプロトタイプから離れたりより近くなったりするという考えがあるが(cf., Gleason and Ratner, 1993; Harley, 2001)、図3～図6の読み方もこれに当てはめることができるかも知れない。では、各被験者グループのプロトタイプの読み方は、個々の被験者の読み方とどの程度の重なり（または隔たり）をみせているのであろうか。また、被験者グループAとグループBでは、プロトタイプの読み方との重なり具合に違いがあるのだろうか。このことを検証するため以下では、プロトタイプの読み方を、被験者ひとりひとりのデータと比較することによって、さらに詳しい分析を行うことにする。

グループ間で共通する読み方のパターンとして、第2領域で読み時間が一番増大し、その後は読み時間が減少していく傾向が見て取れるが、図3～図6を見ると、文型タイプにより若干のずれがあることがわかる。そのため、プロトタイプの読み方を文型タイプ毎に定義していくことにした。その結果が表3で、それを単純にイメージ化したものが図7である。

タイプ	プロトタイプの定義（グループA・グループB共通）
Int 2	読み時間が、第2領域（副詞1）で増大し、その後動詞に向かって減少する。
Int 3	同上
Tran 2	同上
Tran 3	読み時間が、第2領域（副詞1）で増大し、その後第3領域（目的語）にかけて減少し、最後は動詞に向かって増加する。

表3. プロトタイプの読み方の定義

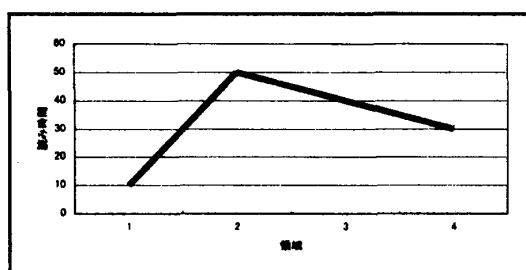


図7 a. Int 2, Int 3, Tran 2のプロトタイプ

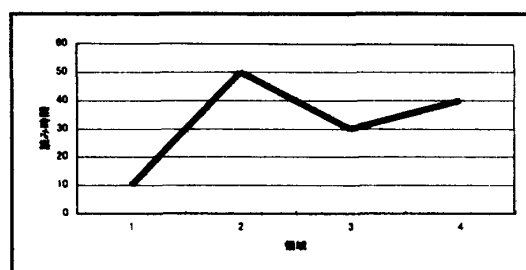


図7 b. Tran 3のプロトタイプ

文型タイプ		完全に一致	部分的に一致	完全に不一致	計	適合度 検定 p 値
グループ A	Int 2	31 (62%)	14 (28%)	5 (10%)	50 (100%)	**
	Int 3	18 (36%)	29 (58%)	3 (6%)	50 (100%)	**
	Tran 2	23 (46%)	20 (40%)	7 (14%)	50 (100%)	**
	Tran 3	18 (36%)	31 (62%)	1 (2%)	50 (100%)	**
小 計		90 (45%)	94 (47%)	16 (8%)	200 (100%)	
グループ B	Int 2	24 (63%)	12 (32%)	2 (5%)	38 (100%)	**
	Int 3	12 (32%)	24 (63%)	2 (5%)	38 (100%)	**
	Tran 2	21 (55%)	13 (34%)	4 (11%)	38 (100%)	**
	Tran 3	22 (58%)	14 (37%)	2 (5%)	38 (100%)	**
小 計		79 (52%)	63 (41%)	10 (7%)	152 (100%)	
合 計		156	97	106	352	

(** p<.001)

表 4. プロトタイプの読み方との一致度 (文型タイプ別)

次に、表 3 と図 7 で示したプロトタイプに沿って、文型タイプ毎に、被験者ひとりひとりの平均読み時間を比較した。その際、「プロトタイプのパターンに完全に一致する被験者」、「プロトタイプのパターンに部分的に一致する被験者」、「プロトタイプのパターンに全く一致しない被験者」の 3 つのカテゴリーを設け、そのいずれかに被験者を振り分けた。⁹ その振り分けの結果をまとめたものが表 4 である。振り分けを行う際、ある領域から次の領域に至るまでに被験者の読み時間が増加しているか減少しているかだけを判断材料とし、便宜上その度合いに関しては考慮から外している。

表 4 が示すように、プロトタイプの読み方と完全に一致している割合は、グループ A で 45% (36%-62%)、グループ B で 52% (32%-63%) と、文型によっては過半数を下回るものもある。データの半数前後しかプロトタイプの読み方をしていないというこの結果は、偶然の一致の範囲を超えるものだろうか。ちなみに、でたらめに読んだときに「完全に一致」となるチャンスレベルの割合は、Int 2 と Tran 2 では 25% で、Int 3 と Tran 3 では 12.5% である。¹⁰ 逆に、でたらめに読んだときに「完全に不一致」となるチャンスレベルの割合は、どのタイプも 25% である。 χ^2 検定により、表 4 の実測数とチャンスレベルの割合との適合度を検証したところ、どのグループのどの文型タイプにおいても、両者は 1% の水準で有意な違いを見せた。さらに残差分析の結果、全ての文型タイプで、「完全に一致」の実測数が有意に多く、また「完全に不一致」の

9 ひとりの被験者が実験で読んだ課題文の数は、文型タイプにつき 4 文ずつだった。この 4 文を平均したものが各被験者の平均読み時間として算出され、グループのプロトタイプと比較された。各被験者の平均ではなく、一文一文の読み時間をプロトタイプと比較することも必要であるが、それは今後の課題としたい。

実測数が有意に少ないことも確認された。つまり、両グループとも、偶然よりも高い割合で、表4におけるプロトタイプの読み方が現れていると言える。

次に、表4で示された一致度の偏りが、被験者グループ間で違いがあるかどうかを検証した。グループと一致度の関連について2 x 3の χ^2 検定を行ったところ、件数の偏りは両者間で違いがなかった($\chi^2(2)=1.708$, $p=.426$)。文型タイプ別に検定した場合も、Tran 3で有意傾向が確認された以外は($\chi^2(2)=5.624$, $p=.060$)、他のタイプではやはり違いが認められなかった。残差分析の結果、Tran 3における有意傾向は、「完全に一致」する件数がグループAでは少なく、グループBでは多いことと、「部分的に一致」する件数がグループAでは多くグループBでは少ないことに起因していることがわかった。つまり、Tran 3に限って見れば、グループBの方がグループAに比べてよりばらつきが少ないということが言える。

考察

以上、各文型タイプの読み方を両グループで比較した上、さらに各文型のプロトタイプの読み方を、被験者ひとりひとりの読み時間と比較して論じた。要約すると、各文型の読み進め方の違いは、グループAとグループBではあまり大きな違いがないということがわかった。これは、前節でグループ間で読み時間の差が現れた結果とは対照的である。

ただしTran 3タイプに関しては、グループBの方がより速く文を読み進んでいることが（特に文の後半にかけて）指摘された。このことは、グループBがTran 3タイプで強い項の優位性を呈したことや、速い読み時間を示したことなど、本稿でこれまで見てきた結果と矛盾なく対応するものである。つまり、節の最初から最後までをどのような時間配分で読むかという点において、動詞の処理や項の優位性といった点を除けば、同じ母語話者間では似たような文の読み進め方であったと言える。

また、プロトタイプと各被験者の隔たり具合は、両グループで大きな違いがなく、偶然以上の確率でプロトタイプとの一致度を示していることも確認された。しかしその一方で、プロトタイプ通りの読み方をしているのはどの文型でも約半数前後であり、この一致率が高いか低いかの判断は難しい。いずれにせよ、グラフで表された平均読み時間のイメージは、あくまでも指標として捉えられるべきであって、実測値を直接反映しているのでは必ずしもないということを含みおく必要があるだろう。

10 チャンスレベルの計算は、可能な読み方のパターンの総数を求めることにより算出した。読み方のパターンは、ある領域から次の領域に至るまでに増加しているか減少しているかの2つに1つであると想定した場合、3つの領域からなるInt 2やTran 2の文型タイプでは2 x 2の4通りが総パターン数となる。プロトタイプの読み方は、その中の1通りなので、「完全に一致」する場合は25%と計算できる。他方、Int 3やTran 3のような、4つの領域からなる文型タイプでは、可能な読み方のパターンは2 x 2 x 2の8通りが総数である。「完全に一致」する場合はその中の1通り、つまり12.5%と計算できる。しかしながら、実際には、読み時間の増減はこれほど単純ではない。2つの領域間で殆ど違いが見られない場合から、かなり大きく変化している場合まで、千差万別であるが、本稿では便宜上このように単純化して考えた。

プロトタイプの読み方の被験者内の一貫性について

前節では文型タイプ別にプロトタイプの読み方を検証した。しかし表4で見た一致度は、例えばInt 2で「完全に一致」した同じ被験者が、他の文型タイプでも一致しているかどうかまでは保証していない。そこで、同一被験者が読んだ4つの文型タイプが、プロトタイプとどの程度一致しているかについてもさらに検証することにした。表5がその結果である。

	4タイプ 全て合致	3タイプ 合 致	2タイプ 合 致	1タイプ 合 致	合 致 な し	計	適 合 度 検定 p 値
グループ A	4 (8%)	10 (20%)	15 (30%)	14 (28%)	7 (14%)	50 (100%)	**
グループ B	3 (8%)	11 (29%)	13 (34%)	8 (21%)	3 (8%)	38 (100%)	**

(** p<.001)

表5. 同一被験者におけるプロトタイプの読み方との一致度 (人数)

表5からまずわかるのは、グループA・グループB共に、4つの文型タイプ全てでプロトタイプとの一致を見た被験者の割合が約8%しかないということである。3つの文型タイプが一致する被験者数と合わせても、どちらのグループも40%に満たない一致度である。つまり、半数以上の被験者は、4つの文型タイプのうち、多くて2つまでしかプロトタイプと一致する読み方をしていないということだ。グループと一致度の関連について2x5の χ^2 検定を行ったところ、件数の偏りは両者間で有意差がなかった($\chi^2(4)=1.970$, $p=.741$)。このことから、平均の読み時間を線でつなげたプロトタイプの読み方のパターンは、両グループ共に、必ずしも同一被験者が高い一貫性をもって読み進めた結果ではないことがわかる。

同一被験者が高い一貫性を示してはいないながらも、表5に示した一致度は、チャンスレベル以上の確率で起こっていることも確かである。表4で計算したチャンスレベルの測定法に従って算出すると(脚注10参照)、4つの文型を読む場合の読み方パターンは、全部で1024通りあることになる。¹¹ この内4つの文型タイプ全てがプロトタイプの読み方と一致するようなパターンは1通りだけであるので、チャンスレベルは1/1024となる。同様に3文型タイプでプロトタイプの読み方と一致するようなパターンは20通りあり、チャンスレベルは20/1024 (5/256)である。¹² これらのチャンスレベルを表5の実測値と比較して χ^2 検定により適合度を検証したところ、どちらのグループも1%の水準で有意であることがわかった。つまり、表5に見られる一致度は、偶然よりも高い割合で導かれたものであり、少なくともでたらめに読まれた結果

11 脚注10で示したように、Int 2の総パターン数が4通り、Int 3が8通り、Tran 2が4通り、Tran 3が8通りある。よって、4つの文型を読むときのパターンの総数(順列)は、 $4 \times 8 \times 4 \times 8$ で1024通りである。

12 以下、2文型タイプの一致が142/1024 (71/512)、1文型タイプの一致が420/1024 (105/256)、全く一致しない場合が441/1024となる。

ではないということが言える。

考察

プロトタイプの読み方と、被験者内の一貫性という観点では、両グループとも、偶然以上の確率で一致を示していることがわかった。しかしながら、被験者の過半数以上がプロトタイプ通りの読み方をしていないことも同時に示された。前節の結果と合わせて考えると、全体として平均化すればある読み方のパターンが浮かび上がるが、ひとりひとりで見れば、必ずしも毎回同じ読み方をしているとは限らないということになる。これは、人によって読み方が違うことの他に、多くの被験者が“ぶれ”のある読み方をしていたことを示唆している。

何故このようなぶれが生じるのかについては、今後の研究をまたなければならないが、可能性として次の2つが挙げられるかも知れない。ひとつは、ある程度の曖昧さを許す読み方をするのも文処理方略のひとつであるという考えである。通常文処理の規則や方略といえば、ある一定の方向性をもって行う文理解の方法を指し、Minimal AttachmentやLate Closureなどは有名である(Frazier, 1987; Kimball, 1973)。しかしその一方で、文理解はいつも正しく行われるわけではなく、解釈に矛盾や誤りを含む場合のあることも報告されている(Christianson, Hollingworth, and Ferreira, 2001; Ferreira, 2003; Ferreira, Bailey, and Ferraro, 2002)。例えば、Christianson, et al. (2001)やFerreira (2003)では、以下の文を誤って解釈した被験者の例が紹介されている。

(5) While Anna dressed the baby played in the crib. (Christianson, et al., 2001)

(6) The dog was bitten by the man. (Ferreira, 2003)

(5)は、“dressed”の後にカンマを挟まない形式で提示した袋小路文である。被験者に文の意味について質問したところ、「赤ちゃんはベッドで遊んだ」のであり、かつ「アンナは赤ちゃんに服を着せた」と矛盾のある解釈をした場合が少なからずあった(Christianson, et al., 2001)。また(6)は単純な受身文であるが、被験者の25%が、この文を「犬が人を噛んだ」のごとくに誤って解釈したことが報告されている(Ferreira, 2003)。

Ferreira, et al. (2002)はこういった現象について、たとえ文法情報が文解析を可能にしている、人間はいろいろな制約や状況のもとでも文処理に対応できるように、解釈の曖昧さを許した処理を行っているとし、日常生活で起こる話の聞き違いや思い違いといったこととも関連づけて論じている。Ferreira, et al.のこのコメントは、特に(5)の袋小路文や、(6)のような一般的常識とは逆行するような文意の場合について述べられたものであるが、文処理は必ずしも全く正しく行われるとは限らないという指摘は大変示唆的である。本稿で扱ったごく平易な文などは、誤った解釈をされる

ことはないであろうが、時間配分など文の読み進め方に関しては、プロトタイプの読み方がある一方で、ある程度逸脱した読み方もまた許しているのかも知れない。

被験者内の読み方にばらつきが見られた理由として考えられるもうひとつの可能性は、今回の実験での課題文が平易すぎたということである。平易な文の場合、それだけ処理負荷が軽いので、ワーキングメモリを能率良く使って読みこなすために一定の読み方に頼る必要性が低くなる。その結果、意味理解に影響がでない範囲で多様な読み方を許し、被験者内の読み方にばらつきが生まれたとも考えられるだろう。しかしこのことを実証するためには、より複雑な文による読み時間との比較を通して検証することが今後必要であろう。

まとめ

本稿では、ふたつの日本語母語話者グループの文処理方法について、読み時間の比較を試みた。比較は、各領域の読み時間、各文型タイプの読み方のパターン、被験者内の一貫性についての3つの観点から行われた。その結果、平易な文を読む場合であっても、被験者集団によって違いの現れる事象とそうでない事象のあることがわかった。違いが現れるのは、目的語と副詞句の読み分けや、動詞の処理速度であった。一方、被験者グループが違っても読み方が変わらないのは、文頭から文末にいたる読み時間配分（読み方のパターン）である。さらに、同一被験者による読み方パターンを調べたところ、必ずしもいつも同じ読み方を重ねているわけではないことがわかった。

被験者グループにより違いが起こる理由としてはワーキングメモリの個人差が、そして同一被験者における一貫性の不在については、文処理方略が曖昧性を許すこと等が示唆された。しかしこれらはまだ推測の域を超えるものではない。本稿の結果をもとに、多様な実験文を用いての読み時間の比較や、リーディングスパンテストとの相関関係を探るなどの、さらなる今後の分析に期待したい。

参考文献

- Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research* 18, 129-144.
- Ahrens, K. (2003). Verbal integration: The interaction of participant roles and sentential argument structures. *Journal of Psycholinguistic Research*, 32(5), 497-516.
- Boland, J. E., & Blodgett, A. (2006). Argument status and PP-attachment. *Journal of Psycholinguistic Research*, 35, 385-403.
- Christianson, K., Hollingworth, A., Halliwell, J.F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive*

- Psychology, 42, 368-407.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164-203.
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11-15.
- Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.
- Frazier, L. (1987). Sentence processing: Review. In M. Coltheart (Ed.), *Attention and performance: Vol. XII. The psychology of reading* (pp. 559-586). Hillsdale, NJ: Erlbaum.
- Frazier, L., & Clifton, C., Jr. (1996). *Construal*. MA: MIT Press.
- Gleason, J. B., & Ratner, N. B. (Eds.). (1993). *Psycholinguistics*. Fort Worth, TX: Harcourt Brace College Publishers.
- Harley, T. A. (2001). *The psychology of language*. Hove, England: Psychology Press.
- Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1331-1356.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- 茅本百合子. (2002). 語彙判断課題と命題課題における中国語母語話者の日本語漢字アクセス. *教育心理学研究*, 50, 436-445.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2, 15-47.
- Kondo, T., & Mazuka, R. (2006). Effects of word properties on Japanese sentence processing. In M. Nakayama, R. Mazuka, & Y. Shirai (Eds.), *The handbook of East Asian psycholinguistics*. Cambridge, UK: Cambridge University Press.
- Nakano, Y., Felser, C., & Clahsen, H. (2000). Antecedent priming at trace positions in Japanese long-distance scrambling. *Essex Research Reports in*

Linguistics, 31, 45-76.

芋阪満里子. (2002). 脳のメモ帳ワーキングメモリ. 新曜社.

坂本勉. (2001 a). 文の理解 1 : 空主語文. 言語, 30(9), 92-97.

坂本勉. (2001 b). 文の理解 2 : かき混ぜ文. 言語, 30(10), 106-111.

坂本勉. (2001 c). 文の理解 3 : 同格連続文. 言語, 30(11), 100-105

坂本勉. (2001 d). 文の理解 4 : 中央埋め込み文. 言語, 30(12), 98-103.

坂本勉. (2001 e). 文の理解 5 : 袋小路文から再分析へ. 言語, 30(13), 110-117.

Sawasaki, K. (2004). L2 sentence processing by English-speaking learners of Japanese. *Ars Linguistica*, 11, 136-155.

Sawasaki, K. (2007). L2 reading by learners of Japanese: A comparison of different L1s. Unpublished doctoral dissertation. The Ohio State University, Columbus, OH, USA.

Schutze, C, T., & Gibson, E. (1999). Argumenthood and English propositional phrase attachment. *Journal of Memory and Language*, 40, 409-431.

Speer, S. R., & Clifton, C., Jr. (1998). Plausibility and argument structure in sentence comprehension. *Memory and Cognition*, 26, 965-978.