

# 英語学習者向け代名詞処理を伴うリーディングスパン テストの材料文開発

## Developing the Materials of a Reading Span Test for English Learners Involving Pronoun Resolution

浅見 紫織 ASAMI, Shiori

- 国際基督教大学大学院アーツ・サイエンス研究科  
Graduate School of Arts and Sciences, International Christian University

ローランド, ダグラス ROLAND, Douglas

- 早稲田大学  
Waseda University

森島 泰則 MORISHIMA, Yasunori

- 国際基督教大学  
International Christian University



**Keywords** 作動記憶容量, 第二言語読解, リーディングスパンテスト, 代名詞処理  
working memory capacity, comprehension in a second language, reading span test,  
pronoun resolution

### ABSTRACT

本研究では、英語学習者向けの代名詞処理を伴うリーディングスパンテスト（Anaphora Reading Span Test）の材料文の開発過程を報告した。読解能力をより精緻に検証する作動記憶容量の指標として、代名詞処理を組み込んだ Carriedo, Elosúa, & García-Madruga（2011）のリーディングスパンテストの問題点を指摘し、テストの材料文の特徴と構造を検討した。そして、材料文は代名詞とその指示対象候補2個を含み、文の意味から正しい指示対象が特定される構造で80文作成された。また、代名詞の指示対象がどの程度明確かについて、7件法の質問紙調査を行った。調査を経て、指示対象が意味的に明確なテスト材料候補文は、指示対象が統語的に明確な文と指示対象の明確さが同程度となるよう統制された。調

査結果を下にテストで使用される70文を選定した。

The current study reported the development of the materials for a new reading span test for English learners, named the Anaphora Reading Span Test, involving pronoun resolution. We first discuss the problems of the original reading span test (Daneman & Carpenter, 1980) and the reading span test that is involved in pronoun resolution (Carriedo, Elosúa, & García-Madruga, 2011). Eighty sentences for the new reading span test were developed. The sentences had two referent-candidates and a pronoun. The proper referent of the pronoun could be contextually identified. Then the comprehensibility of the pronouns in each sentence was investigated. The participants were asked to decide which candidate was the most likely the referent of the pronoun by circling a point on a scale. The results indicated that the referents were semantically as clear as those which were syntactically unambiguous. Finally, 70 sentences were selected for the new reading span test according to the comprehensibility of the pronouns.

## 1. 作動記憶とリーディングスパンテスト

大学ではグローバル化の推進を背景に第二言語、特に英語での授業が増加している。また、そのような授業ではテキストや論文の読解を媒介にした学習が少なくない。しかし、テキストに含まれた情報や知識を獲得するために読解を行う場合、第二言語では記憶できる情報量の減少 (van den Noort, Bosch, & Hugdahl, 2006) や、高度な認知処理、例えば思考や推論、問題解決の抑制 (Takano & Noda, 1993; Harrington & Sawyer, 1992; Morishima, 2013) が懸念される。このような問題が起こる要因として、情報の処理と保持を司る作動記憶の認知資源の有限性が考えられる。読解中の作動記憶では、現在行っている言語情報の処理、処理された情報の保持、さらに記憶された先行情報や既有知識との統合が並列に行われる。作動記憶はこれらの処理に有限の認知資源を配分しているが、第二言語で読解を行う場合、言語情報の処理は第一言語よりも負荷が高くなると考えられるため、比較的より多くの認知資源を必要とすることが推測できる。第二言語での読解中に作動記憶はどの程度効率的に処理と保持に認知資源を配分しているのだろうか。

読解中の作動記憶の効率性を評価する指標の1つに、リーディングスパンテストに基づく作動記憶容量がある。Daneman & Carpenter (1980) の開発したリーディングスパンテストは、2文から5

文セットの文を音読しながら、各文末の単語を記憶し、各セットの最後に記憶した単語を全て再生するテストである。再生された単語が多いほど、作動記憶容量が大きいと評価される。Daneman & Carpenter (1980) は、作動記憶容量と読解テストの成績に強い正の相関が見られることを報告した。しかし、作動記憶容量と読解能力の相関関係は安定していない (Caplan & Waters, 1999)。例えば、Daneman & Carpenter (1980) が示したほど強い相関が得られない研究 (中程度の正の相関: Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Caplan & Waters, 1999, 弱い正の相関: 柴崎・時本・小野・井上, 2015), あるいは相関が見られないとする研究 (Carriedo, Elosúa, & García-Madruga, 2011) も存在する。この理由として、リーディングスパンテストでの言語処理課題である文の音読が必ずしも意味の理解を必要としていない点が指摘された (Carriedo et al., 2011)。つまり、従来のリーディングスパンテストは言語処理の負荷を反映しているとは言えず、実際の読解過程における言語処理と保持の効率性を評価するには不十分であると考えられる。

そこで、本研究では言語処理として代名詞処理という読解を伴うリーディングスパンテスト (Anaphora Reading Span Test) の開発を目的とした。代名詞処理とは、代名詞 (例えばsheやhe) が何を指示しているかを特定する過程である。代名詞処理は照応解決とも呼ばれる推論の一種で、

読解には必須の処理である (Graesser, Singer, & Trabasso, 1994)。したがって、第二言語学習者でも頻繁に行う処理であり、文の意味理解を必要とするため、リーディングスパンテストの言語処理課題に適していると言える。本稿で報告する Anaphora Reading Span Test は、García-Madruga, Gutiérrez, Carriedo, Luzón, & Vila (2005, 2007) と Carriedo et al. (2011) の一連の研究によって発展した Semantic Anaphora Span Test を基盤に作成された。Carriedo et al. (2011) は、Semantic Anaphora Span Test が読解テストの得点だけでなく推論 (reasoning) 問題の正答率とも正の相関を持ち、それらの得点や正答率を有意に説明する指標となることを明らかにした。さらにこのとき、従来のリーディングスパンテストは読解テストの得点や推論問題の正答率を有意に説明する指標とならないことが示されており、Semantic Anaphora Span Test が従来のリーディングスパンテストよりも推論能力を反映した言語処理と情報保持の指標として有効であることが示唆された。しかし、García-Madruga et al. (2005, 2007) と Carriedo et al. (2011) が作成したリーディングスパンテストには、2つの問題点が挙げられる。(a) スペイン語母語話者向けに作成されているため英語学習者には使用できない点、(b) 一般的な代名詞処理の過程とは異なる過程をテストで再現している点である。本稿では、指摘した問題点を改善し、Anaphora Reading Span Test の材料文作成の過程を報告した。

## 2. 代名詞処理を伴うリーディングスパンテスト

本節では、前述の Semantic Anaphora Span Test の2つの問題点について検討した。具体的に、(a) テストの対象者がスペイン語話者から英語学習者になる場合、テストの材料文にはどのような統制が必要であるか。また、(b) Anaphora Reading Span Test の材料文を実際の読解過程と適合させるには、どのような文の構造とすべきかを検討した。

まず、García-Madruga et al. (2005, 2007) と Carriedo et al. (2011) の Semantic Anaphora Span Test はスペイン語母語話者向けであるため、1文の単語数や記憶対象語の文字数、文作成で留意した点をそのまま英語学習者向けの材料に適用することはできない。そこで、英語学習者向けに開発された Osaka & Osaka (1992) と柴崎・時本・小野・井上 (2015) のリーディングスパンテストの特徴を参考にした (表1 参照)。

Anaphora Reading Span Test は大学生を対象とするため、単語数や記憶対象語の文字数は Osaka & Osaka (1992) の材料を参考に統制した。しかし、Anaphora Reading Span Test は後述する特殊な文構造を持つため、中学や高校の教科書から文を採用することは難しい。そこで、柴崎他 (2015) に倣い『大学英語教育学会基本語リスト新 JACET 8000』(大学英語教育学会基本語改訂特別委員会, 2016) を用いて、文中の単語が大学生にとって基本的な単語となるよう留意した。

次に、García-Madruga et al. (2005, 2007) と Carriedo

表1 英語学習者向けリーディングスパンテストの特徴

	Osaka & Osaka (1992)	柴崎他 (2015)
対象者	大学生	高校生
各文の単語数	9-16 ( $M = 10.49, SD = 1.20$ )	5-10 ( $M = 7.20, SD = 1.10$ )
記憶対象語の文字数	2-11 ( $M = 5.64, SD = 1.93$ )	3-7 ( $M = 5.06, SD = 1.24$ )
備考	高校の教科書から文を採用	中学の教科書の文を参照 JACET8000の単語を使用

(注) JACET8000は、『大学英語教育学会基本語リスト』に基づく JACET8000 英単語』(相澤他, 2005) の略称である。

et al. (2011) の開発した材料文の構造を再考した。例1に示す通り、Carriedo et al. (2011) は代名詞を含む文を提示した後に、その代名詞が指す単語の候補2つを文とは独立に提示した。

例1 Carriedo et al. (2011) が使用した英語訳版の材料文

Eladio encouraged **her** a lot to interpret such a demanding role.

Career / **Actress**

(Carriedo et al., 2011, p. 40)

太字のherが指示する単語は、太字のActressである。

文から独立した指示対象を選択する構造は現実の読解において異質であると言えるが、最も重要な問題は例示された代名詞が前方照応であり、一般的に頻繁に用いられる代名詞の後方照応とは異なる点である。つまり、前方照応では代名詞の後に指示対象が導入されるため、読み手に主に求められるのは指示対象を後続の情報から特定することである。しかし、後方照応の場合、先行情報を保持しながら後続の文を処理し、保持された先行情報から適当な指示対象を特定するという過程が想定できる。この2つの照応関係を比較したとき、先行情報の文脈に加えて指示対象候補を保持しなければいけない後方照応で、記憶の負荷がより高いと考えられる。本研究では実際の読解場面における作動記憶の処理と保持の効率性を見ることを重視しているため、後方照応が前方照応と記憶負荷の点で異なることを考慮しなければならないと考えた。そこで、より現実の読解場面に適合する後方照応の代名詞を用いた材料文の構造を組み立てた。Anaphora Reading Span Testの材料文を例2に示した。

例2 Anaphora Reading Span Testで使用する材料文

A **bike** is allowed to be taken on any **train** if it is properly bagged.

指示対象候補の2つは下線を引いて明示され

る。太字のitが指示するのは、太字のbikeである。

各文では下線を引かれた指示対象候補の単語2つが先行して導入され、後続する代名詞 (she, he, it, they) がどちらか一方を指示する。指示対象候補2つは、両方が統語的に可能であり、意味的にどちらか一方にだけ当てはまるよう作成された。つまり、指示対象候補の性別や、有生性、単数・複数から代名詞を特定するのは困難であり、正しい指示対象の特定のために文の理解を必要とする仕組みとなっている。このように、リーディングスパンテストで言語処理に負荷をかけることによって、単純な記憶課題になることを防止した。例2では、bikeとtrainが指示対象候補として挙げられ、後続のitは統語的にbikeとtrainのどちらも取り得るが、「正しく袋に入れられていたら、電車に乗せることができる」という内容からbikeが指示対象であると特定できる。

以上の構造を持つ材料文候補として80文を作成した。材料は次のような特徴を持っている。まず、she, he, it, theyの4種類の各代名詞に20文ずつ用意された。各文に含まれる単語数は、最大19語、最小11語、平均15.39語 (標準偏差1.78) であった。指示対象の文字数は、最大11字、最小3字、平均6.56字 (標準偏差2.01) であった。また、指示対象となる単語が1番目に導入される場合と2番目に導入される場合が40文ずつ用意された。指示対象語と代名詞の間には、最大11語、最小0語、平均5.14語 (標準偏差3.06) の単語が介在した。指示対象候補の単語は、『大学英語教育学会基本語リスト新JACET8000』(大学英語教育学会基本語改訂特別委員会, 2016) から選択された。

このように作成された候補80文に関して、各文の代名詞の指示対象特定の明確さが調査および検討される必要がある。もし代名詞の指示対象が十分に明確でない場合、言語能力に関わらず記憶すべき単語を特定できないという問題が起こる。同様に、作動記憶の効率性とは関係なく記憶課題の成績が低くなる可能性もある。これらの問題を防ぐため、各材料文の代名詞の指示対象が、どの

程度明確に特定（選択）されるか調査を行った。調査は全3回行われ、それらの結果から明確さの統制を行い、材料70文が選定された。以下に3回の調査結果を報告する。

### 3. 代名詞の指示対象の明確さに関する一連の調査

#### 3.1 調査(1)

第1回目の調査の目的は、前述の通り、作成された材料文に含まれる代名詞がどの程度明確に特定されるかである。本調査では7件法の尺度を用いて明確さを評定させたので、代名詞の指示対象が明確な場合、各文の評定平均値は7に近い値となる。また、英語学習者だけでなく、英語母語話者も対象に行った。英語学習者と英語母語話者では基本的な言語理解や推論能力に差があり、内容の解釈に差が生じると考えられる。つまり、英語母語話者にとって明確な指示対象が学習者にとっては曖昧になる可能性があり、その逆も起こり得る。2者の結果を比較することで、指示対象の明確さが読み手の言語能力に大きく左右される文が含まれていないかを検討した。

##### 3.1.1 方法

都内の大学に通う英語母語話者13名と英語学習者65名が調査に参加した。

Anaphora Reading Span Testの候補80文とフィラー40文を使用した。フィラー文のうち20文は、指示対象候補の性別や、単数・複数、有生性から指示対象が統語的に明確となる文で構成されている(例3参照)。残りの20文は、指示対象が統語的にも意味的にも曖昧になるよう作成された(例4参照)。フィラー文を取り入れることで、指示対象が統語的に完全に明確な文と完全に曖昧な文の中での、Anaphora Reading Span Testの材料文の明確さの位置づけを明らかにする。

#### 例3 指示対象が統語的に明確な文

The boyfriend proposed to his **girlfriend**, but **she** thought it was a joke.

太字のsheが指示するのは、太字のgirlfriendである。代名詞の性別から指示対象を特定できる。

#### 例4 指示対象が曖昧な文

When the photographer took a picture of the soldier, **he** got injured.

太字のheはphotographerとsoldierの両方を指し得る。

まず、英語母語話者向けに、Anaphora Reading Span Test材料候補80文とフィラー40文が無作為な順番で1つの冊子にまとめられた。英語学習者向けには、Anaphora Reading Span Test材料候補80文がAとBの2つの冊子に40文ずつ分けられた。各冊子には、実験材料候補40文とフィラー40文が無作為な順番でまとめられた。

「指示対象候補のうち、どちらがより指示対象として適切か」に関して、両端に各指示対象候補、真ん中に両方可能(英語母語話者版はBoth Possible)を配置した7件法で評定させた(図1参照)。

A bike is allowed to be taken on any train if it is properly bagged.



図1 評定項目の提示方法

代名詞の指示対象が曖昧であるほど両方可能に寄り、指示対象が明確であるほど両端に寄って丸を付けなければならない。指示対象が統語的にあるいは意味的に明確な文は、正しい指示対象を7、間違った指示対象を1として評定平均値を算出した。したがって、指示対象を正しく特定できているほど7に近い値となる。指示対象が曖昧な文は、1番目に導入された指示対象候補を1、2番目に導入された指示対象候補を7として評定平均値を算出した。したがって、曖昧であれば評定値は4、1番目の指示対象候補に偏るほど1に近い値、2番目の指示対象候補に偏るほど7に近い値となる。

参加者は配布された冊子の説明を読んで、自分

のペースで回答を行うように教示された。

### 3.1.2 結果と考察

まず、評定結果全体の分析を行った。これによって、材料候補文の代名詞が一般的に明確に解決されていたかを明らかにする。次に、材料候補文ごとの評定平均に着目した分析を行った。この分析から、評定平均の高い文だけを実験材料文として選定する。

前処理として、統語的に明確な文の平均評定値が6.3以下だった英語学習者2名を分析から除外した。英語母語話者と英語学習者の評定平均値と標準偏差を表2に示した。

指示対象が統語的に明確な文の平均評定値について、英語母語話者と学習者はともに7に近い値と言える。指示対象が統語的に明確な文では、指示対象を代名詞の特徴から機械的に特定することが出来るため、明確さが高くなったと考えられる。次に、指示対象が意味的に明確な文について、統語的に明確な文と比較したとき、英語母語話者と学習者ともに平均評定値が0.2から0.4程度低下した。しかし、評定値7を明確さ100%としたとき、6.3以上は明確さ90%以上に当たることを考慮すると、指示対象が意味的に明確な文は英語母語話者と学習者どちらにとっても、指示対象は十分に明確だったと考えられる。最後に、指示対象が曖昧な文について、英語母語話者と学習者はどちらも平均評定値が3台であった。したがって、1番目に導入された指示対象候補を選択する傾向が多少あるが、指示対象は基本的に曖昧であったと考えられる。

これまで、記述統計から材料文の種類ごとに平

均評定値の解釈を行った。次に、材料の種類ごとに評定平均値に差があると言えるか、その差は英語母語話者と学習者で異なるかを検討した。分析には、Rのlme4パッケージ（バージョン1.1.12；Bates, Marchler, Bolker, & Walker, 2015）で線形混合効果モデルを利用した。固定効果について、carパッケージ（バージョン2.1.4；Fox & Weisberg, 2011）のAnova機能を用いてSatterthwaiteの自由度推定を行い、その有意性が決定された。単純主効果はphiaパッケージのtestInteractions機能を用いて算出し、多重比較の結果はlmerTestパッケージのdiffsmeans機能を用いた。

まず、英語学習者において、冊子の種類（A, B）と材料の種類（指示対象が統語的に明確、意味的に明確、曖昧）を固定効果、参加者と材料を変量効果としたモデルを立て、材料の各種類の平均評価得点が冊子AとBで異なることを確かめた（冊子の主効果： $\chi^2(1) = 0.07, p = .79$ ；材料の種類と冊子の交互作用： $\chi^2(2) = 2.46, p = .29$ ）。次に、材料の種類と言語（英語母語話者、英語学習者）を固定効果、参加者と材料を変量効果とする線型混合効果モデルを立てた。分析の結果、材料文の種類の主効果（ $\chi^2(2) = 528.1, p < .001$ ）、言語の主効果（ $\chi^2(1) = 10.4, p = .001$ ）、そして材料文の種類と言語の交互作用（ $\chi^2(2) = 35.6, p < .001$ ）は全て有意であった。各言語の単純主効果の検定を行ったところ、英語母語話者と学習者どちらにおいても、材料の種類の主効果は有意であった（英語母語話者： $\chi^2(2) = 521, p < .001$ 、英語学習者： $\chi^2(2) = 488, p < .001$ ）。英語母語話者における多重比較の結果、指示対象が統語的に明確な文と曖昧な文の間（ $t(189.5) = -18.58, p < .001$ ）および明確

表2 [第1回目] 英語母語話者と英語学習者による代名詞の指示対象の評定結果

		英語母語話者 <i>M (SD)</i>	英語学習者 <i>M (SD)</i>
意味的に明確 (80文)	Total	6.71 (0.97)	6.46 (1.42)
	A	-	6.39 (1.52)
	B	-	6.53 (1.30)
統語的に明確 (20文)		6.95 (0.46)	6.86 (0.76)
曖昧 (20文)		3.16 (1.64)	3.48 (1.94)

な文と曖昧な文の間 ( $t(189.5) = -22.03, p < .001$ ) で評定平均値の有意差が認められた。しかし、指示対象が統語的に明確な文と意味的に明確な文の間には有意差が見られなかった ( $t(189.5) = 1.47, p = .143$ )。同様に、英語学習者における多重比較の結果、全てのペアで有意差が認められた(統語的に明確vs曖昧： $t(114.3) = -18.81, p < .001$ 、意味的に明確vs曖昧： $t(117.8) = -20.84, p < .001$ 、統語的に明確vs意味的に明確： $t(117.8) = 2.77, p = .006$ )。また、材料の各種類の単純主効果の結果、指示対象が意味的に明確な文と曖昧な文でのみ、英語母語話者と学習者に有意差が見られた(指示対象が統語的に明確： $\chi^2(1) = 1.86, p = .173$ 、意味的に明確： $\chi^2(1) = 28.18, p < .001$ 、曖昧： $\chi^2(1) = 10.81, p = .002$ )。以上の結果で特に重要な点は、英語学習者は指示対象の明確さが統語的か意味的かで、その明確さが有意に異なることである。

英語学習者でのみ指示対象が意味的に明確な文が、統語的に明確な文より比較的 low に評定されたことについて検討する。指示対象が意味的に明確な材料文で、文ごとに平均値を算出したところ、英語母語話者と英語学習者の両方、あるいはいずれか一方で評定平均値が6に満たない文が80文中13文あった。これらの文は内容を再検討する必要がある。また、文の長さも平均評定値が低くなった原因として考えられる。Osaka & Osaka (1992) や柴崎他 (2015) の英語学習者向けリーディングスパンテストの材料文では、1文に使用された平均単語数が10個前後であったことから、今回作成された平均単語数15個の文は学習者には長かったのではないかと考えられる。第2回目の調査では、以上の問題点を改善した上で、改訂版の材料文の指示対象の明確さを検討した。

### 3.2 調査 (2)

第1回目の調査で、英語学習者は指示対象を意味的に特定しなければいけない場合、統語的に特定する場合よりも明確さが低下することが分かった。意味的に指示対象がより明確になるように、1文中の単語数を減らすことで比較的単純な文に修正した。さらに、平均評定値が6に満たなかつ

た13文は、英語母語話者と協議し、内容に変更を加えた。本節では、改訂版の材料文を用いて代名詞の指示対象の明確さを調査した結果を報告する。

#### 3.2.1 方法

第1回目の調査に参加していない、都内の大学に通う日本人英語学習者28名が調査に参加した。Anaphora Reading Span Testの材料候補80文とフィルター30文(指示対象が統語的に明確15文・曖昧15文)を使用した。第1回目の調査と異なり、1人の英語学習者に材料候補80文全てを評定させるため、フィルターを10文減らすことで負担を軽減した。

第1回目の調査で使用した材料80文は以下の点が改善された。まず、文中から余分な情報を削ることで1文中の単語数を減らした。そして、平均評定値が6を下回った13文の内容を変更した。その結果、各文に含まれる単語数は、最大15語、最小9語、平均11.93語(標準偏差1.17)となった。指示対象の文字数は、最大11字、最小2字、平均6.13字(標準偏差2.00)であった。

参加者は、Creative Surveyで作成されたウェブページにアクセスし、インターネット上で回答を行った。材料は1画面に1文ずつ、図2のように提示された。材料は無作為に決められた順番で提示された。

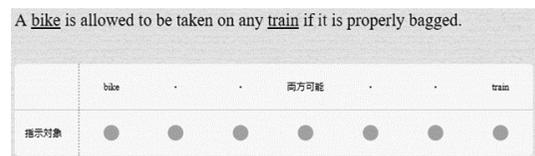


図2 指示対象評価画面

「下線が引かれた指示対象候補2個のうち、どちらがより赤字の代名詞の指示対象として適切か」を7件法で評定するよう教示された。各文の回答制限時間は90秒に設定され、制限時間を超えた回答は無効とした。

### 3.2.2 結果と考察

第1回目の調査と同じように、全体的な評定値の分析と考察を行った後に、各文の考察を行った。制限時間を越えた回答はなかった。指示対象が統語的に明確な文の評定平均値が6.3以下の参加者4名は分析から除外した。評定平均値と標準偏差を表3に示した。

表3 [第2回目] 代名詞の指示対象の評定結果

	<i>M (SD)</i>
意味的に明確 (80文)	6.59 (1.23)
統語的に明確 (15文)	6.85 (0.80)
曖昧 (15文)	3.68 (1.91)

参加者が異なるため直接的な比較は難しいが、指示対象が意味的に明確な文の平均評定値は第1回目の調査よりも多少高まったように見える。また、第1回目の調査と同様に、指示対象が統語的に明確な文では7に近い値、曖昧な文では4に近い値であった。

次に、材料文の種類（意味的に明確・統語的に明確・曖昧）を固定効果、参加者と材料を変量効果とする線形混合効果モデルを立てた。分析の結果、材料文の種類が有意であった ( $\chi^2(2) = 354, p < .001$ )。多重比較を行ったところ、指示対象が統語的に明確な文と曖昧な文の間 ( $t(107) = -15.32, p < .001$ ) と意味的に明確な文と曖昧な文の間 ( $t(107) = -18.25, p < .001$ ) に有意差が見られた。第1回目の調査と異なり、指示対象が統語的に明確な文と意味的に明確な文の間には有意差は見られなかった ( $t(107) = 1.64, p = .100$ )。この結果から、Anaphora Reading Span Testの材料候補である指示対象が意味的に明確な文は、統語的に明確な文と同じ程度に指示対象が明確であったと言える。

指示対象が意味的に明確な文について、文ごとに平均評定値を算出したところ、平均評定値が6に満たなかった文は6個であった。この6文と、残りの74文中で評定平均値の低いものから4文を除いた70文をAnaphora Reading Span Testの材料として暫定的に定めた。この材料70文の特徴

を調べたところ、heの代名詞文が14文と少なく、代名詞の指示対象の導入位置は1番目が37文と偏っていた。そこで、heを使用した文と指示対象の導入位置が2番目の文を増やすために追加調査を行った。

### 3.3 調査 (3)

第2回目の調査で、指示対象が統語的に明確な文と意味的に明確な文では、同程度に指示対象が明確であったことが分かった。この結果をもとに、Anaphora Reading Span Testの材料70文が暫定的に選定された。しかし、選定された70文は代名詞heを用いた文と指示対象の導入位置が2番目の文が少ないことが明らかになった。そこで、代名詞heを用いた文と指示対象の導入位置が2番目の文を追加で作成し、作成された文の指示対象の明確さを調査した。

#### 3.3.1 方法

第1回目と第2回目の調査に参加していない、英語学習者11名が参加した。

Anaphora Reading Span Testの材料候補6文（指示対象が意味的に明確な文）と、指示対象が統語的に明確な6文、指示対象が曖昧な6文を使用した。指示対象が意味的に明確な文は、1文の平均単語数12.00語、標準偏差2.00、最大15語、最小9語であった。指示対象語の平均文字数は6.33字、標準偏差1.86、最大9字、最小4字であった。6文中、代名詞heの文が5個、代名詞sheの文が1個であった。指示対象が2番目に導入される文は5個、1番目に導入される文は1個であった。

第1回目の調査と同様に冊子で調査を行った。材料18文は無作為な順番で1つの冊子にまとめられた。「指示対象候補のうち、どちらがより指示対象として適切か」に関して、両端に各指示対象候補、真ん中に両方可能を配置した7件法で評定させた。

#### 3.3.2 結果と考察

まず、全体的な評定値の分析と考察を行った。指示対象が統語的に明確な文で平均評定値6.3以

下の参加はいなかったため、全ての参加者のデータを分析に用いた。評定平均値と標準偏差を表4に示した。

表4 [第3回目] 代名詞の指示対象の評定結果

	M (SD)
意味的に明確 (6文)	6.67 (0.81)
統語的に明確 (6文)	6.97 (0.17)
曖昧 (6文)	3.64 (2.05)

評定平均値はこれまでの調査と同じパターンの結果を得られた。指示対象が統語的あるいは意味的に明確な文では7に近い値となり、曖昧な文では4に近い値となった。

材料の種類を固定効果、参加者と材料を変量効果とする線型混合効果モデルの分析を行った結果、材料文の種類の効果は有意であった ( $\chi^2(2) = 193, p < .001$ )。多重比較を行ったところ、第2回目の調査結果と同様に、指示対象が統語的に明確な文と曖昧な文の間 ( $t(15) = -12.57, p < .001$ ) および意味的に明確な文と曖昧な文の間 ( $t(15) = -11.43, p < .001$ ) に有意差が見られた。そして、指示対象が統語的に明確な文と意味的に明確な文の間には有意差がなかった ( $t(15) = 1.14, p = .3$ )。

また、指示対象が意味的に明確な文について、文ごとに平均評定値を算出したところ、6文全ての平均評定値が6以上であった。暫定的な材料70文で、代名詞 they または it の文から評定平均値が比較的低く、指示対象の導入が1番目の文を、今回作成された6文と入れ替えた。最終的に完成した Anaphora Reading Span Test の材料70文の特徴は以下の通りであった (表5参照)。

#### 4. 結論

本稿では、代名詞処理という読解作業を伴う新

しいリーディングスパンテスト (Anaphora Reading Span Test) の材料文開発の過程を報告した。まず、García-Madruga et al. (2005, 2007) と Carriedo et al. (2011) による代名詞処理を含む Semantic Anaphora Span Test の問題点を指摘し、それらを改善した材料80文が作成された。3回の調査を経て、各材料文の代名詞の指示対象は、英語学習者にとって明確に特定できると言えることが分かった。最終的に、代名詞の指示対象の明確さが統制されたテストの材料70文が選定された。

しかし、今回作成された材料文が、Semantic Anaphora Span Test (García-Madruga et al., 2005, 2007; Carriedo et al., 2011) のように、従来のリーディングスパンテスト (Daneman & Carpenter, 1980) よりも読解能力を精緻に予測および説明するのに有用なテストであるかを明らかにするためには更なる検討が必要である。今後は、Anaphora Reading Span Test と読解テストとの相関研究によるリーディングスパンテストの信頼性および妥当性の調査を行う。さらに、測定された作動記憶の効率性が、第二言語読解にどのような影響を与えるかを実験的に検討していく。

#### 引用文献

- 相澤一美・石川慎一郎・村田年・磯達夫・上村俊彦・小川貴宏…望月正道 (2005). 「大学英語教育学会基本語リスト」に基づく JACET8000 英単語 桐原書店
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1-48. doi: 10.18637/jss.v067.i01
- Baddeley, A. D., Logie, R., Nimmo-Smith, L., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119-131. doi: 10.1016/0749-596X(85)90019-1
- Caplan, D. & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and brain sciences*, 22, 77-126. doi: 10.1017/

表5 Anaphora Reading Span Test の特徴

代名詞				1文の単語数				指示対象の文字数				代名詞の導入位置	
he	she	it	they	M	SD	Max	Min	M	SD	Max	Min	1番目	2番目
17	18	17	18	11.99	1.25	15	9	6.17	1.97	11	3	35	35

S0140525X99001788

- Carriedo, N., Elosúa, M. R., & García-Madruga, J. A. (2011). Working memory, text comprehension, and propositional reasoning: A new semantic anaphora WM test. *The Spanish Journal of Psychology, 14*, 37-49. doi: 10.5209/rev\_SJOP.2011.v14.n1.3
- 大学英語教育学会基本語改訂特別委員会（編著）（2016）. 大学英語教育学会基本語リスト新JACET 8000 桐原書店
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466. doi: 10.1016/S0022-5371(80)90312-6
- Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (2nd ed.). Thousand Oaks, California: Sage.
- García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Luzón, J. M., & Vila, J. O. (2005). Working memory and propositional reasoning: Searching for new working memory tests. In V. Girotto and P. N. Johnson-Laird (Eds.), *The shape of reason. Essays in honour of Paolo Legrenzi* (pp. 69-89). London: Psychology Press.
- García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Luzón, J. M., & Vila, J. O. (2007). Checking the role of central executive in propositional reasoning. *Thinking and Reasoning, 13*, 370-393.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395. doi: 10.1037/0033-295X.101.3.371
- Harrington, M. W., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skills. *Studies in Second Language Acquisition, 14*, 25-38. doi: 10.1017/S0272263100010457
- Kintsch, W. (1986). Learning from text. *Cognition and Instruction, 3*, 87-108. doi:10.1207/s1532690xci0302\_1
- Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language. *Discourse Processes, 50*, 577-597. doi: 10.1080/0163853X.2013.846964
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society, 30*, 287-289. doi: 710.3758/BF03330466
- 柴崎秀子・時本真吾・小野雄一・井上次夫（2015）. 高校生用集団式日英語リーディングスパンテストの開発および英語における習熟度と作動記憶の関係の検討 認知心理学研究, 12, 101-120. doi: 10.5265/jcogpsy. 12. 101
- Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology, 24*, 445-462. doi: 10.1177/0022022193244005
- van den Noort, M., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*, 289-296. doi: 10.1027/1016-9040.11.4.289