

The Effect of Plausibility on Eye Movements in Reading

Keith Rayner (University of Massachusetts, Amherst)

Tessa Warren (University of Pittsburgh)

Barbara J. Juhasz (University of Massachusetts, Amherst)

Simon P. Liversedge (University of Durham)

Journal of Experimental Psychology: Learning, Memory, and Cognition, 2004, Vol. 30, No.6, 1290-1301

- ある単語を見る時間は、その単語の処理しやすさに影響される (Liversedge & Findlay, 2000; Rayner, 1998)
 - 低頻出語 > 高頻出語 (Inhoff & Rayner, 1986; Rayner & Duffy, 1986)
 - 文脈による制限なし > 制限あり (Ehrlich & Rayner, 1981; Rayner & Well, 1996)
- 眼球運動のモデルでは、単語の頻出度と予測可能性を入力すると、固視時間やスキップ率の予測を出力することができる
 - E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003) と SWIFT model (Engbert, Longtin, & Kliegl, 2002)
- 本研究では、妥当性 (plausibility) が眼球運動にどんな影響を与えるのかを検証する
 - 妥当性が重要である2つの根拠
 1. 上記のモデルの精度を上げるには、様々な要因の検討が必要
 2. 妥当性がどの段階から影響してくるのかが不明
- 2.については多くの先行研究があるが、結論は一致していない
 - 傍中心窩で (中心窩に入る前に) すでに影響がある (*parafoveal-on-foveal effect*) (Murray, 1998; Murray & Rowan, 1998)
 - 固視したその瞬間から影響がある (Cook & Myers, 2004)
 - 固視してからしばらく経たないと影響はない (Boland & Blodgett, 2001; Garrod & Terras, 2000; Traxler, Pickering, & Clifton, 1998)
- Braze et al. (2002)

2a. <i>The wall will surely /crack after/ a few years in this harsh climate.</i> (Control)
2b. <i>The wall will surely /bite after/ a few years in this harsh climate.</i> (Pragmatic Anomalous)
2c. <i>The wall will surely /cracking after/ a few years in this harsh climate.</i> (Syntactic Anomalous)

 - ターゲット領域 (スラッシュ内) の読み時間において、両 Anomalous > Control
 - 逆行確率において、Syntactic Anomalous > 他2条件

- この先行研究における問題点
 - ターゲット語 (*crack / bite / cracking*) が各条件で異なり, 長さも異なる
 - ターゲット領域が狭いときには, 文字数補正はあまり効果がない
 - ターゲット語の次の単語 (*after*) の長さがテーマにより異なる
 - 最短 3 文字, 最長 7 文字
 - Kennedy et al. (2004); Murray (1998); Murray & Rowan (1998)
 - 動作主と動詞の関係, 動詞とテーマ (目的語) の関係の妥当性をそれぞれ操作
- 3a. *The savages smacked the child.* (Plausible–Plausible, PP: Control)
 - 3b. *The savages smacked the money.* (Plausible–Implausible, PI)
 - 3c. *The uranium smacked the child.* (Implausible–Plausible, IP)
 - 3d. *The uranium smacked the money.* (Implausible–Implausible, II)
- parafoveal-on-foveal effect が観察された
 - すべての単語を固視する以前に, 単語の語彙同定・単語の意味的处理・文章の (浅い) 意味的評価を行っている
- この先行研究における問題点
 - 刺激が適当ではない
 - 追試によると, 全 24 テーマ中 6 テーマは IP と PP の好ましき評定が同程度
 - また, 別の 5 テーマは 4 条件間の評定に差なし
 - parafoveal-on-foveal effect を再現できない
- これらを踏まえ, より統制された刺激を用いて妥当性を検証すべき
- 4a. *John used a knife to chop the large carrots for dinner.* (Plausible: Control)
 - 4b. *John used an axe to chop the large carrots for dinner.* (Implausible)
 - 4c. *John used a pump to inflate the large carrots for dinner.* (Anomalous)
- 構成
 - 主語の名詞句 (*John*)
 - 他動詞 (*used*)
 - 道具の名詞句 (*a knife / an axe / a pump*) ※条件により変化
 - 不定詞 (*to chop / to inflate*) ※条件により変化
 - 形容詞つきの名詞句 (*the large carrots*)
 - 道具の名詞句の, 動詞に対する主題役割 (*thematic role*)¹ の妥当性で条件分け
 - 非妥当な場合 (*inflate-carrots*) Anomalous
 - 妥当な場合 (*chop-carrots*) は, 文全体としての妥当性でさらに条件分け

¹ 述語と項の意味関係を分類したもの. たとえば「太郎が次郎をなぐった」なら, 主題 (*theme*) は「次郎を」, 動作主 (*agent*) は「太郎が」にあたる.

- ◇ 妥当な場合 (knife-carrots) Plausible
- ◇ 非妥当な場合 (axe-carrots) Implausible
- 道具の名詞句と不定詞の関係 (knife-chop / axe-chop / pump-inflate) は常に妥当
 - よって、条件間で差が出るだろうターゲット語 (領域) は“(the large) carrots (for dinner)”
- Plausible 道具 (knife) は不定詞 (chop) と主題 (carrots) に対して妥当
 - 主題役割の割り当てとスキーマにおいて、解釈が一致する
- Implausible 道具 (axe) は不定詞と主題に対して明示的には妥当、意味的には非妥当
 - 主題役割の割り当てと意味的統合を行った後でしか、非妥当性が検出されない
- Anomalous 道具 (pump) は不定詞 (inflate) に対して妥当、主題に対して非妥当
 - 主題役割の割り当てとスキーマにおいて、解釈が一致しない
- 今回の実験では、ターゲット領域における妥当性の効果について、眼球運動の指標を用いて検証する
 - Anomalous は妥当性の崩れ始めるのが早いだろう
 - Anomalous はすぐに遅れが観察されるだろう
 - Implausible は Anomalous ほど早くないにしても、いつかは妥当性が崩れるだろう
 - Implausible は時間経過とともに、遅れが観察始めるだろう

Method

Participants

- マサチューセッツ大学の学部生 36 名

Apparatus

- 右目に、Fourward Technologies Dual-Purkinje Image Generation VI を装着
 - 空間分解能は 10 分 (どこを見ているかを知るのに十分な性能)
- 目から 61 cm のところに刺激を提示するモニターを設置
- 前頭部を固定し、bite bar を使用

Materials

- 上記のような文を 30 テーマ用意 (Appendix 参照)
 - 各テーマに Control / Implausible / Anomalous の 3 条件あるので、全 90 種類
 - ターゲット語を固視する可能性を上げるため、常に 5 文字以上

- Control / Implausible / Anomalous の条件分けが妥当かを予備実験で確かめた
 - 本実験に参加していない 12 名が文を 5 段階で評定
 - 5 (normal) ~ 3 (unlikely but possible) ~ 1 (highly unlikely)
 - “anomalous”や“implausible”という語は使わないようにした
- その結果, Anomalous は 1.4, Implausible は 2.9, Control は 4.8 と評定された
 - 全テーマにおいて, Anomalous < Implausible < Control となった
- 上記の本刺激とは構文構造が異なるフィラー刺激 90 文を用意
 - 本刺激は構文構造がすべて同じであるため, 予測されるのを防ぐ役割
- 本刺激 10 とフィラー刺激 30 を組み合わせたものを 1 セットとして, 参加者に提示
 - 本刺激提示後には, 真偽で答える理解度テストあり (操作チェック)

Procedure

- 実験時間は 30 ~ 45 分
- 教示
 - 提示される文章はそれぞれ長さ・難易度が異なる
 - 中には少し奇妙なものも含まれている
 - 提示される文を普通に理解しながら読む (終わったらキーを押す?)
 - 全体の 25% の文には理解度テストがあり, 真偽をキーで答える
- 参加者に器具を取り付け, キャリブレーション
- 1 文読むごとにキャリブレーションチェック
 - 点がずれた場合は再キャリブレーション

Results and Discussion

- 除外データ (全体の 10%)
 - 理解度テストの成績が 80% 未満の参加者
 - ターゲット語に固視しなかった参加者
 - 装置の不具合により採れなかったデータ
 - 80 ms 以下かつ他の固視から 3 文字以上離れた固視
 - 平均から 2 *SD* 以上の固視
- 80 ms 以下の固視は隣接する文字の固視と統合
- ターゲット領域 (the large carrots for dinner) を 3 つに分割する
 - プレターゲット領域 限定詞と形容詞 (the large)
 - ターゲット語 名詞 (carrots)

- ポストターゲット領域 副詞など (for dinner)
- 各領域で指標ごとに ANOVA を行う
 - 1 テーマについて、参加者の異なる 3 条件間を参加者間で (F_1)
 - 1 参加者について、テーマの異なる 3 条件間を参加者内で (F_2)
- 眼球運動測定の指標 (下に行くほど広範な定義・深い処理)
 - First-fixation duration ある単語について最初の固視時間
 - Single-fixation duration ある単語の固視回数が 1 回のときの固視時間
 - 固視回数が 2 回以上なら、分析対象外
 - Gaze duration ある単語から離れるまでの合計固視時間
 - その単語内なら逆行も含む
 - Go-past time ある単語とそれ以前の部分との合計固視時間
 - 以前の部分への逆行を含む
 - 語彙処理過程と統合過程を反映している
- 予測
 - Anomalous における妥当性の崩れの影響は、すぐに眼球運動に現れるだろう
 - Implausible における妥当性の崩れの影響は、徐々に眼球運動に現れるだろう

Pretarget Region (the large)

- 結果は Table 1 の上半分
- 各指標において ANOVA を行った
 - First-fixation duration ・ Single-fixation duration ・ Gaze duration において有意差なし
- プレターゲット領域について詳しく調べるため、指標 Last 3 を追加 (Murray, 1998)
 - Last 3 ある単語の最後 3 文字における合計固視時間
- Anomalous > Implausible \asymp Control ($t_1(32) = 2.32, p < .05; t_2(24) = 2.55, p < .05$)
- Last 3 ではまだ妥当性は崩れないはずだが、条件間で差が見られた
 - 「先にある単語の意味が、現在の固視に影響を与える」という parafoveal-on-foveal effect が観察された？
- それもありうるが、単語 $n - 1$ に固視しながら単語 n を並行処理していたと思われる
 - この現象についての 3 つの説明
 1. 固視の位置と注意の位置が異なる (Rayner, 1998)
 2. 読者は n を見ていたにもかかわらず、装置が $n - 1$ を示していた
 3. 読者は n に跳躍したかったが、距離が足りずに $n - 1$ に落ちた (undershoots)

Table 1
Fixation Times on the Pretarget Region (in ms)

Condition	FFD		Single		Gaze		Last 3		Go-past		Total	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
All items												
Anomalous	255	34	275	43	348	65	261	68	373	75	453	150
Implausible	243	32	270	45	334	65	231	49	352	67	369	66
Control	248	32	270	40	331	66	241	35	345	68	359	74
Supplemental analysis (18 items)												
Anomalous	254	39	269	44	356	73	254	59	384	83	457	157
Implausible	241	36	267	43	333	73	227	63	360	78	371	72
Control	249	36	264	46	338	81	228	51	349	84	370	86

Note. FFD = first-fixation duration.

(McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1998)

- プレターゲット領域の全固視のうち 22% は Last 3 領域に含まれる
 - undershoots が起こり、すでに *n* (ターゲット語) を処理し始めていた
- parafoveal-on-foveal effect ではなく、すでに *n* を処理していただい
- Go-past time において有意差あり ($F_1(2, 70) = 7.10, p < .01$; $F_2(2, 58) = 3.31, p < .05$)
 - 両側 *t* 検定より, Anomalous > Implausible \approx Control
($t_1(35) = 3.43, p < .01$; $t_2(29) = 2.00, p < .055$)
- Go-past time には、プレターゲット領域の以前の部分への逆行も含まれている
 - 深い処理が行われたことにより逆行が起き、条件間で差が見られた²
- この結論についての反論
 - Anomalous で固視時間が長くなった原因は、名詞ではなく形容詞なのでは？
 - chop-large より、inflate-large の好ましさのほうが低いことが影響？
- この可能性を排除するため、10 名に文の好ましさを 5 段階で評定させた
 - 1 (not fit at all) ~ 5 (fit very well)
- その結果、条件間で有意差あり ($F(2, 58) = 10.19, p < .01$)
 - Anomalous (4.1) \approx Implausible (3.9) < Control (4.4)
- そこで、条件間で好ましさに差がなかった 18 テーマ ($p > .28$) のみで追加分析
 - 結果は Table 1 の下半分
 - 全 30 テーマが対象の上半分と同じ傾向で、有意差も確認された
 - Anomalous の効果は、形容詞ではなく名詞により引き起こされた

² ただし、逆行確率に有意差なし (Anomalous 6.3%, implausible 4.4, control 3.9%; $ps < .35$).

Target Word (carrtos)

- 結果は Table 2 (下半分は追加分析)

Table 2
Fixation Times on the Target Word (in ms)

Condition	FFD		Single		Gaze		Go-past		Total	
	M	SE	M	SE	M	SE	M	SE	M	SE
All items										
Anomalous	268	40	277	46	306	53	332	73	390	121
Implausible	261	35	270	39	289	39	310	37	312	48
Control	262	38	265	46	286	49	297	57	297	52
Supplemental analysis (18 items)										
Anomalous	266	47	280	54	308	65	339	85	383	112
Implausible	262	36	268	38	286	37	307	41	312	44
Control	261	41	263	48	282	46	295	60	291	48

Note. FFD = first-fixation duration.

- First-fixation duration • Single-fixation duration において有意差なし (all $ps > .15$)
- Gaze duration において有意差あり ($F_1(2, 70) = 3.44, p < .05$; $F_2(2, 58) = 4.50, p < .05$)
 - 両側 t 検定より, Anomalous > Implausible \approx Control
($t_1(35) = 2.47, p < .05$; $t_2(29) = 2.24, p < .05$) ($t_1(35) = 2.03, p < .05$; $t_2(29) = 2.57, p < .05$)
- Go-past time において有意差あり ($F_1(2, 70) = 6.45, p < .01$; $F_2(2, 58) = 7.82, p < .01$)
 - 両側 t 検定より, Anomalous > Implausible > Control
($t_1(35) = 3.08, p < .01$; $t_2(29) = 3.63, p < .01$) ($t_1(35) = 2.14, p < .05$; $t_2(29) = 2.32, p < .05$) ($t_1(35) = 1.75, p = .089$; $t_2(29) = 1.71, p = .098$)

Posttarget Region (for dinner)

- 結果は Table 3 (下半分は追加分析)
 - ANOVA
 - First-fixation duration • Single-fixation duration • Gaze duration • Go-past time において, 有意差あり
 - 両側 t 検定
 - First-fixation duration • Single-fixation duration • Gaze duration において, Anomalous > Implausible \approx Control
 - Go-past time において, Anomalous > Implausible > Control
($t_1(35) = 6.59, p < .01$; $t_2(29) = 6.37, p < .01$) ($t_1(35) = 4.77, p < .01$; $t_2(29) = 3.24, p < .01$) ($t_1(35) = 2.76, p < .01$; $t_2(29) = 2.36, p < .05$)

Table 3
Fixation Times on the Posttarget Region (in ms)

Condition	FFD		Single		Gaze		Go-past		Total	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
All items										
Anomalous	288	58	299	75	351	84	493	226	416	112
Implausible	261	39	261	45	325	75	362	88	354	85
Control	267	48	268	54	312	64	325	76	325	65
Supplementary analysis (18 items)										
Anomalous	293	75	291	79	344	94	434	183	398	121
Implausible	255	46	259	51	303	88	341	107	344	105
Control	258	52	266	61	311	84	325	98	320	79

Note. FFD = first-fixation duration.

- ポストターゲット領域からの逆行確率において、有意差あり ($F_1(2, 70) = 19.71, p < .01$; $F_2(2, 58) = 16.62, p < .01$)
 - Anomalous > Implausible \doteq Control

Total Times

- Total reading-time は各 Table の右端に書かれている
 - プレターゲット領域では、Anomalous > Implausible \doteq Control
 - ターゲット語・ポストターゲット領域では、Anomalous > Implausible > Control

General Discussion

- 本実験の目的は、妥当性がどの段階から影響してくるのかを検証することであった
 - Anomalous の効果は「読者の頭を叩くような」もの
 - 妥当性が崩れるとすぐに眼球運動に現れる (Gaze duration)
 - 別の単語へ移行する前にもう一度固視することが主な要因
 - Implausible の効果は、Anomalous に比べると弱く、長続きしない
 - 深い意味的処理をして初めて妥当性が崩れ始める (Go-past time)
- Anomalous の効果はプレターゲット領域の時点でもうすでに観察された
 - parafoveal-on-foveal effect の可能性もあるが、少なくとも単語 $n - 1$ に固視しながら

- n の処理を行っていたのは間違いない
- n の処理の結果、プレターゲット領域でも逆行が増え、**Go-past time** が長くなった
 - **Anomalous** が検出されると、以前の箇所をもう一度調べようとするため、ポストターゲット領域で逆行が増える
 - 奇妙な文であり、解釈の一貫性を保とうと試みる必要があったため
 - 妥当性の崩れをいつ検出できるかは、崩れ具合による
 - 低次な処理では、大きな崩れしか検出できない
 - 高次な処理では、小さな崩れも検出できる
 - **Anomalous** と **Implausible** の違いは崩れ具合
 - θ 役割 (**theta role**)³ を割り当てる動詞と割り当てられた項の関係で生じる崩れ
 - **Anomalous** は θ を割り当て、その語彙情報だけ見れば崩れが検出できる
 - **Implausible** は θ を割り当て、意味的表象の統合まで行えば崩れが検出できる
 - 文章理解時には、質的に異なるタイプの処理が異なるステージで行われている
 - 最初にも言ったとおり、眼球運動のモデルを洗練させることも今回の目的であった
 - しかしながら、今回の状況はあまりに特殊なので、妥当性を変数としてモデルに組み込む必要はないだろう
 - 特に妥当性の崩れが大きい場合のみ考慮すればいいだろう

³ θ 役割とは、動詞が主語や目的語に与える役割のこと。

Appendix

Sentences Used in the Experiment and Their Scores in Two Norming Studies

Sentences	Overall plausibility rating	Adjective fit
John used a pump to inflate the large carrots for dinner last night.	1.0	4.4*
John used an axe to chop the large carrots for dinner last night.	3.0	4.7
John used a knife to chop the large carrots for dinner last night.	5.0	4.8
The man used a grill to cook the steaming asphalt on the road.	1.5	3.8*
The man used a knife to spread the steaming asphalt on the road.	2.0	3.3
The man used a shovel to spread the steaming asphalt on the road.	4.8	3.3
The man used a feather to tickle the thin spaghetti yesterday evening.	1.0	3.6*
The man used a net to drain the thin spaghetti yesterday evening.	3.5	3.3
The man used a strainer to drain the thin spaghetti yesterday evening.	5.0	3.3
The woman used a book to teach the tough bread very carefully.	1.3	4.2*
The woman used a saw to cut the tough bread very carefully.	3.0	4.4
The woman used a knife to cut the tough bread very carefully.	4.5	4.6
The woman used the map to direct the large present yesterday.	1.3	4.0
The woman used the rags to wrap the large present yesterday.	3.5	3.4
The woman used the paper to wrap the large present yesterday.	5.0	4.5
Bill used the calculator to compute the hard cheese from Italy.	1.0	4.8*
Bill used the scissors to cut the hard cheese from Italy.	2.8	4.0
Bill used the knife to cut the hard cheese from Italy.	4.3	4.2
The man used an iron to flatten the stiff patient after surgery.	1.0	3.7
The man used a bribe to relax the stiff patient after surgery.	2.0	2.9
The man used a tranquilizer to relax the stiff patient after surgery.	4.0	4.3
The woman used a dagger to stab the dirty miniatures sitting on the shelf.	1.5	3.6
The woman used a vacuum to clean the dirty miniatures sitting on the shelf.	3.0	4.8
The woman used a duster to clean the dirty miniatures sitting on the shelf.	4.8	4.8
The man used a tank to ambush the front porch for the party.	1.8	3.8
The man used a toothbrush to clean the front porch for the party.	3.3	3.6
The man used a mop to clean the front porch for the party.	5.0	4.7
The hostess used the music to calm the hot beans for dinner.	1.3	3.4
The hostess used the toothpick to serve the hot beans for dinner.	3.3	3.5
The hostess used the dish to serve the hot beans for dinner.	5.0	4.6
The man used the phone to call the old frame together.	1.0	4.5
The man used the chopsticks to hold the old frame together.	3.3	3.2
The man used the clamp to hold the old frame together.	4.8	4.1
Jenny used a hose to water the small butterfly flying by.	2.0	4.6*
Jenny used a mousetrap to catch the small butterfly flying by.	3.0	4.9
Jenny used a net to catch the small butterfly flying by.	5.0	4.7
Patricia used a fork to eat the fresh water extremely carefully.	1.5	5.0
Patricia used a purse to carry the fresh water extremely carefully.	2.5	3.6
Patricia used a bucket to carry the fresh water extremely carefully.	4.3	4.7
Matthew used a broom to sweep the bright comet as it passed by.	1.0	3.4
Matthew used a microscope to watch the bright comet as it passed by.	3.5	3.1
Matthew used a telescope to watch the bright comet as it passed by.	5.0	4.7
George used a harness to restrain the many flowers in his garden.	2.5	4.1*
George used a sword to protect the many flowers in his garden.	2.8	4.5
George used a fence to protect the many flowers in his garden.	4.5	4.1
Frank used a hammer to nail the heavy groceries from the store.	1.8	4.5*
Frank used a helicopter to transport the heavy groceries from the store.	3.0	4.3
Frank used a cart to transport the heavy groceries from the store.	4.5	4.7

Appendix

Continued

Sentences	Overall plausibility rating	Adjective fit
Julie used a needle to sew the various children after recess.	1.3	4.4*
Julie used a flare to summon the various children after recess.	3.5	4.0
Julie used a bell to summon the various children after recess.	5.0	4.4
Melinda used a spatula to flip the little house at night.	1.0	4.3*
Melinda used an anchor to secure the little house at night.	1.8	3.8
Melinda used a lock to secure the little house at night.	4.8	4.5
Donald used a jeep to climb the highest spire of the cathedral.	1.5	4.1*
Donald used a stool to reach the highest spire of the cathedral.	2.5	4.5
Donald used a ladder to reach the highest spire of the cathedral.	4.0	5.0
The man used a rag to polish the precious liquid for the sauce.	1.8	4.2*
The man used a basket to hold the precious liquid for the sauce.	3.0	4.0
The man used a bowl to hold the precious liquid for the sauce.	4.5	3.9
The woman used a band-aid to cover the irritated child in the waiting room.	1.5	4.2
The woman used a parody to entertain the irritated child in the waiting room.	4.0	3.2
The woman used a doll to entertain the irritated child in the waiting room.	5.0	4.7
Beatrice used a towel to dry the important program on the computer.	1.5	3.6
Beatrice used a key to open the important program on the computer.	2.0	4.8
Beatrice used a password to open the important program on the computer.	5.0	4.3
Stuart used a blender to mix the various dimensions of the poster.	1.0	4.9*
Stuart used a stopwatch to measure the various dimensions of the poster.	1.5	4.3
Stuart used a ruler to measure the various dimensions of the poster.	4.8	4.6
Nancy used a sponge to wash the long cigarette that a friend gave her.	1.3	4.1*
Nancy used a torch to light the long cigarette that a friend gave her.	3.8	4.2
Nancy used a match to light the long cigarette that a friend gave her.	5.0	4.4
The man used a formula to explain the beautiful yacht after the outing.	2.3	3.4*
The man used a ribbon to secure the beautiful yacht after the outing.	2.8	3.6
The man used a rope to secure the beautiful yacht after the outing.	4.8	3.7
Robert used a folder to file the large pheasant that weighed ten pounds.	1.0	4.4*
Robert used a hook to catch the large pheasant that weighed ten pounds.	3.3	4.9
Robert used a trap to catch the large pheasant that weighed ten pounds.	5.0	4.7
The woman used a pitchfork to carry the thick mascara in the morning.	1.0	3.1
The woman used a rag to apply the thick mascara in the morning.	2.5	3.7
The woman used a brush to apply the thick mascara in the morning.	5.0	4.6
Justin used a patch to mend the spotted greyhound that he walked.	1.0	3.8*
Justin used a joystick to control the spotted greyhound that he walked.	2.3	3.3
Justin used a muzzle to control the spotted greyhound that he walked.	4.8	3.7
Alberto used a chemical to treat the brand-new information about inventory.	1.3	3.9*
Alberto used a tunnel to access the brand-new information about inventory.	1.8	3.8
Alberto used a website to access the brand-new information about inventory.	5.0	4.2
Gloria used a mask to scare the annoying traffic on Main Street.	2.5	4.8
Gloria used a loophole to avoid the annoying traffic on Main Street.	2.8	3.5
Gloria used a shortcut to avoid the annoying traffic on Main Street.	5.0	4.7

Note. For each set of three sentences, the first sentence is the anomalous condition, the second is the implausible condition, and the third is the plausible (control) condition. Adjective fit refers to how well the adjective (e.g., *large* in the first set) fit with the words that preceded it.

* Sets of sentences used in the supplementary analysis.