**Appendix A: Stimuli**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table A1**  *Congruent and Incongruent Collocations* | | | | | | | |
| Collocation | L2 Collocation Frequency2 | Condition | Collocation Chinese Translation1 | Word1 Chinese Translation | Chinese Word1 Frequency2 | Word2 Chinese Translation | Chinese Word2 Frequency2 |
| good idea | 1851 | Congruent | 好(good) 想法(idea) | 好 | 334698 | 想法 | 4574 |
| white paper | 1106 | Congruent | 白(white) 纸(paper) | 白 | 2684 | 纸 | 1424 |
| cold water | 586 | Congruent | 冷(cold) 水(water) | 冷 | 2343 | 水 | 7581 |
| fresh air | 488 | Congruent | 新鲜(fresh) 空气(air) | 新鲜 | 1145 | 空气 | 1365 |
| young lady | 410 | Congruent | 年轻(young) 女士(lady) | 年轻 | 3776 | 女士 | 10544 |
| best friend | 402 | Congruent | 最好的(best) 朋友(friend) | 最好 | 905 | 朋友 | 26017 |
| main reason | 379 | Congruent | 主要(main) 原因(reason) | 主要 | 1393 | 原因 | 7829 |
| modern art | 339 | Congruent | 现代(modern) 艺术(art) | 现代 | 510 | 艺术 | 1436 |
| small town | 290 | Congruent | 小(small) 镇(town) | 小 | 34967 | 镇 | 1586 |
| retail price | 198 | Congruent | 零售(retail) 价格(price) | 零售 | 54 | 价格 | 639 |
| left foot | 154 | Congruent | 左(left) 脚(foot) | 左 | 1773 | 脚 | 3836 |
| tall man | 141 | Congruent | 高的(tall) 男人(man) | 高 | 7415 | 男人 | 11608 |
| positive attitude | 118 | Congruent | 积极的(positive) 态度(attitude) | 积极 | 537 | 态度 | 1256 |
| sudden change | 118 | Congruent | 突然(sudden) 改变(change) | 突然 | 4366 | 改变 | 7434 |
| future plans | 117 | Congruent | 未来(future) 计划(plan) | 未来 | 3440 | 计划 | 9169 |
| low risk | 112 | Congruent | 低(low) 风险(risk) | 低 | 1820 | 风险 | 908 |
| extra work | 110 | Congruent | 额外(extra) 工作(work) | 额外 | 496 | 工作 | 30191 |
| final version | 100 | Congruent | 最后(final) 版本(version) | 最后 | 15200 | 版本 | 378 |
| sick children | 98 | Congruent | 生病的(sick) 小孩(child) | 生病 | 928 | 小孩 | 3859 |
| simple answer | 91 | Congruent | 简单的(simple) 答案(answer) | 简单 | 5089 | 答案 | 2801 |
| parallel lines | 57 | Congruent | 平行(parallel) 线(lines) | 平行 | 142 | 线 | 3791 |
| true feelings | 53 | Congruent | 真(true) 感觉(feelings) | 真 | 102445 | 感觉 | 21315 |
| long letter | 51 | Congruent | 长(long) 信(letter) | 长 | 12173 | 信 | 7450 |
| major challenge | 44 | Congruent | 主要的(major) 挑战(challenge) | 主要 | 1393 | 挑战 | 1977 |
| red carpet | 43 | Congruent | 红(red) 地毯(carpet) | 红 | 2932 | 地毯 | 599 |
| quiet corner | 40 | Congruent | 安静的(quiet) 角落(corner) | 安静 | 2951 | 角落 | 641 |
| favourite place | 39 | Congruent | 最爱的(favourite) 地方(place) | 最爱 | NA | 地方 | 23358 |
| special role | 39 | Congruent | 特别(special) 角色(role) | 特别 | 7131 | 角色 | 2403 |
| empty chair | 30 | Congruent | 空(empty) 椅子(chair) | 空 | 2242 | 椅子 | 1184 |
| yellow hair | 30 | Congruent | 黄(yellow) 头发(hair) | 黄 | 979 | 头发 | 3955 |
| clean hands | 18 | Congruent | 干净的(clean) 手(hand) | 干净 | 2189 | 手 | 15010 |
| crucial decision | 18 | Congruent | 关键(crucial) 决定(decision) | 关键 | 2248 | 决定 | 10318 |
| fast growth | 18 | Congruent | 快速(fast) 增长(growth) | 快速 | 53510 | 增长 | 238 |
| important document | 18 | Congruent | 重要(important) 文件(document) | 重要 | 11677 | 文件 | 2698 |
| new article | 18 | Congruent | 新(new) 文章(article) | 新 | 17291 | 文章 | 1110 |
| complex society | 17 | Congruent | 复杂(complex) 社会(society) | 复杂 | 1812 | 社会 | 1805 |
| standard test | 14 | Congruent | 标准(standard) 测试(test) | 标准 | 1099 | 测试 | 2453 |
| local dialect | 18 | Congruent | 当地(local) 方言(dialect) | 当地 | 1030 | 方言 | 55 |
| busy summer | 11 | Congruent | 忙的(busy) 夏天(summer) | 忙 | 8342 | 夏天 | 1065 |
| tired voice | 10 | Congruent | 疲惫的(tired) 声音(voice) | 疲惫 | 182 | 声音 | 5523 |
| wide range | 2740 | Incongruent | 大(big) 范围(range) | 宽 | 396 | 范围 | 1423 |
| working class | 1823 | Incongruent | 工(worker) 薪(salary) 阶层 (class) | 工作 | 30191 | 班级 | 116 |
| primary school | 975 | Incongruent | 小(small) 学(school) | 主要 | 1393 | 学校 | 7502 |
| common sense | 969 | Incongruent | 常(often) 识(knowledge) | 普通 | 1765 | 感觉 | 21315 |
| further education | 916 | Incongruent | 继续(continue) 教育(education) | 进一步 | 610 | 教育 | 1172 |
| high performance | 269 | Incongruent | 好的(good) 表现(performance) | 高 | 7415 | 表演 | 6658 |
| heavy rain | 225 | Incongruent | 大(big) 雨(rain) | 重 | 5079 | 雨 | 1312 |
| very beginning | 206 | Incongruent | 一(one) 开始 (beginning) | 非常 | 26976 | 开始 | 32126 |
| real time | 200 | Incongruent | 实(solid) 时(time) | 真 | 102445 | 时间 | 30703 |
| head teacher | 193 | Incongruent | 班(class) 主任(official) | 头 | 11582 | 老师 | 3879 |
| sharp contrast | 139 | Incongruent | 强烈(strong) 对比(contrast) | 尖 | 436 | 对比 | 226 |
| short cut | 135 | Incongruent | 捷(fast) 径(path) | 短 | 2303 | 割 | 877 |
| elder son | 123 | Incongruent | 大(big) 儿子(son) | 老 | 16451 | 儿子 | 11100 |
| native speakers | 115 | Incongruent | 母(mother) 语(language)者 (person) | 本地 | 740 | 讲话者 | NA |
| dead end | 105 | Incongruent | 绝(absolute) 路(road) | 死 | 44056 | 结束 | 10050 |
| late husband | 92 | Incongruent | 去世的(dead) 丈夫(husband) | 迟到 | 2413 | 丈夫 | 6115 |
| passive smoking | 90 | Incongruent | 二手(second hand) 烟(smoke) | 被动 | 86 | 吸烟 | 395 |
| soy sauce | 84 | Incongruent | 酱(sauce) 油(oil) | 豆 | 409 | 酱 | 558 |
| natural parents | 83 | Incongruent | 生(birth) 父母(parents) | 自然 | 2328 | 父母 | 5081 |
| hard evidence | 82 | Incongruent | 确凿(conclusive) 证据(evidence) | 硬 | 1100 | 证据 | 5118 |
| blind date | 57 | Incongruent | 相(mutual) 亲(close) | 瞎 | 1009 | 日期 | 520 |
| hot springs | 54 | Incongruent | 温(warm) 泉(spring) | 热 | 3025 | 泉 | 146 |
| living rooms | 54 | Incongruent | 客(guest) 厅(room) | 生活 | 15918 | 房间 | 7208 |
| terminal illness | 51 | Incongruent | 绝(absolute) 症(symptom) | 航站楼 | 3 | 病 | 3793 |
| deep sigh | 49 | Incongruent | 长(long) 叹(sigh) | 深 | 2324 | 叹气 | 66 |
| clear days | 45 | Incongruent | 晴(sunny) 天(day) | 清晰 | 337 | 天 | 50660 |
| severe weather | 43 | Incongruent | 糟糕(bad) 天气(weather) | 严重 | 3110 | 天气 | 1260 |
| bad breath | 39 | Incongruent | 口(mouth) 臭(odor) | 不好 | 4482 | 呼吸 | 2435 |
| thick fog | 36 | Incongruent | 浓(concentrated) 雾(fog) | 厚 | 416 | 雾 | 411 |
| double check | 32 | Incongruent | 再次(again) 检查(check) | 双倍 | NA | 检查 | 4967 |
| silent film | 20 | Incongruent | 默(silent) 片(piece) | 沉默 | 714 | 电影 | 7249 |
| strong coffee | 19 | Incongruent | 浓(concentrated) 咖啡 (coffee) | 强 | 4814 | 咖啡 | 4195 |
| sweet home | 18 | Incongruent | 温馨的(warm) 家(home) | 甜 | 4400 | 家 | 27026 |
| fair skin | 17 | Incongruent | 白(white) 皮肤(skin) | 公平 | 2336 | 皮肤 | 1153 |
| black tea | 14 | Incongruent | 红(red) 茶(tea) | 黑 | 4563 | 茶 | 1245 |
| rich chocolate | 14 | Incongruent | 浓郁 (concentrated) 巧克力(chocolate) | 富有 | 698 | 巧克力 | 1693 |
| light meal | 13 | Incongruent | 简 (simple) 餐(meal) | 轻 | 1236 | 餐 | 650 |
| forced entry | 18 | Incongruent | 强行(by force)进入(enter) | 强迫 | 836 | 进去 | 5490 |
| dark grey | 11 | Incongruent | 深(deep) 灰(grey) | 暗 | 881 | 灰 | 660 |
| direct flight | 18 | Incongruent | 直(straight) 飞(fly) 航班(flight) | 直接 | 4383 | 航班 | 881 |
| Note. 1. Approximate English translation equivalents of individual words in parentheses.  2. L2 collocation, Chinese Word1, Chinese Word2 frequencies are raw frequencies. | | | | | | | |

|  |
| --- |
| **Table A2**  *Control Items* |
| quiet article, favourite illness, special contrast, empty evidence, yellow work, clean attitude, crucial chair fast parents, important grey, new fog, complex speakers, standard skin, retail voice, busy film, tired idea, severe performance, bad entry, thick time, double education, silent range, strong rooms, sweet air, fair document, black meal, rich breath, wide water, working cut, primary weather, common lines, further teacher, high sauce, heavy school, very springs, real corner, head sense, sharp town, short society, elder end, native carpet, dead chocolate, light date, forced place, dark class, direct man, good sigh, white role, cold risk, fresh test, young answer, best dialect, main rain, modern foot, small decision, local beginning, left friend, tall challenge, positive tea, sudden lady, future summer, low plans, late feelings, passive coffee, soy paper, natural letter, hard growth, blind version, hot check, living reason, terminal air, deep flight, clear son, extra husband, final children, sick home, simple hands, parallel art, true days, long price, major smoking, red change |

**Appendix B: Manipulation Check**

24 L1 English speakers were recruited for the manipulation check. Table B1 presents the RT and accuracy data of the participants. The first two models compared L1 English speakers’ RT (Table B2) and accuracy (Table B3) for congruent, incongruent, and control items. Each model included condition (congruent, incongruent, and control) as the main effect of interest and item length, MI, L2 Word1 and Word2 frequencies as covariates. Treatment coding was used for condition, with control items as the reference. Data trimming and modelling follow the procedures described in the manuscript. Results showed that L1 English speakers responded faster and more accurately to collocations than to control items, suggesting that collocations were treated as formulaic and control items as non-formulaic.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table B1**  *Descriptive Statistics for L1 English speakers’ RT and Accuracy* | | | | |
| Item type | RT *(SD)* |  | Accuracy |  |
| Congruent | 923.08 (376.91) |  | 92.05 |  |
| Incongruent | 920.88 (394.36) |  | 87.63 |  |
| Control | 1085.28 (430.09) |  | 84.22 |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B2**  *Mixed-effects Model for RT of Collocations vs. Control* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate*  [95% CI] | *SE* | *t* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | 6.93 [6.88, 6.98] | .02 | 273.54 | <.001\*\*\* |  | .01 | .11 |  | .01 | .08 |
| Congruent | -.16 [-.19, -.12] | .02 | -7.83 | <.001\*\*\* |  | - | - |  | - | - |
| Incongruent | -.16 [-.19, -.12] | .02 | -8.18 | <.001\*\*\* |  | - | - |  | - | - |
| Length | .02 [.01, .02] | .003 | 4.53 | <.001\*\*\* |  | - | - |  | - | - |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B3**  *Mixed-effects Model for Accuracy of Collocations vs. Control* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate*  [95% CI] | *SE* | *z* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | 2.28 [1.84, 2.73] | .23 | 10.06 | <.001\*\*\* |  | .46 | .68 |  | 1.87 | 1.37 |
| Incongruent | .91 [.27, 1.56] | .33 | 2.77 | .01\* |  | - | - |  | - | - |
| Congruent | 1.42 [.74, 2.10] | .35 | 4.09 | <.001\*\*\* |  | - | - |  | - | - |
| Length | -.15 [-.27, -.04] | .06 | -2.63 | .01\* |  | - | - |  | - | - |
| L2 Word2 frequency | .92 [.32, 1.52] | .31 | 2.99 | .002\*\* |  | - | - |  | - | - |

In the next models, I examined L1 English speakers’ RT and accuracy to congruent and incongruent items for the congruency effect. Condition was treatment coded with the congruent condition serving as the reference. Main effects of interest were L2 collocation frequency and condition, with item length, L2 Word1 and Word2 frequencies, and MI as covariates. Results revealed no congruency effect in either RT or accuracy.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B4**  *Mixed-effects Model for L1 English speakers’ RT of Congruent vs. Incongruent* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate*  [95% CI] | *SE* | *t* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | 6.77 [6.72, 6.83] | .03 | 239.72 | <.001\*\*\* |  | .01 | .11 |  | .01 | .08 |
| L2 collocation frequency | -.06 [-.11, -.02] | .02 | -3.04 | .004\*\* |  | .002 | .04 |  | - | - |
| Incongruent | .004 [-.04, .05] | .02 | .16 | .87 |  | - | - |  | - | - |
| Length | .01 [.001, .02] | .005 | 2.21 | .03\* |  | - | - |  | - | - |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B5**  *Mixed-effects Model for L1 English speakers’ Accuracy of Congruent vs. Incongruent* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate*  [95% CI] | *SE* | *z* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | 3.35 [2.60, 4.09] | .38 | 8.81 | <.001\*\*\* |  | .14 | .38 |  | 2.91 | 1.70 |
| L2 collocation frequency | .64 [-.37, 1.65] | .52 | 1.24 | .21 |  | 1.29 | 1.14 |  | - | - |
| Incongruent | .48 [-.47, 1.44] | .49 | .99 | .32 |  | - | - |  | - | - |
| L2 Word2 frequency | 1.24 [.07, 2.41] | .60 | 2.08 | .04\* |  |  |  |  |  |  |

The final step was to examine the effect of Chinese lexical frequency. I fitted two mixed models to RT and accuracy respectively of L1 English speakers, with Chinese Word1 and Word2 frequencies, and condition as the main effects of interest, and item length, L2 collocation frequency, Chinese Word1 and Word2 frequencies, and MI as covariates. Results showed that L1 English speakers were not sensitive to Chinese Word1 and Word2 frequencies.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B6**  *Mixed-effects Model for the Effect of Chinese Lexical Frequency on L1 English speakers’ RT* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate*  [95% CI] | *SE* | *t* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | 6.77 [6.72, 6.83] | .03 | 242.18 | <.001\*\*\* |  | .01 | .11 |  | .01 | .08 |
| Incongruent | .005 [-.04, .05] | .02 | .21 | .84 |  | - | - |  | - | - |
| Chinese Word1 | .004 [-.03, .03] | .02 | .26 | .80 |  | - | - |  | - | - |
| Chinese Word2 | -.03 [-.07, .01] | .02 | -1.66 | .10 |  | - | - |  | - | - |
| Length | .01 [.002, .02] | .01 | 2.32 | .02\* |  | - | - |  | - | - |
| L2 collocation frequency | -.06 [-.10, -.02 ] | .02 | -3.07 | .003\*\* |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table B7**  *Mixed-effects Model for the Effect of Chinese Lexical Frequency on L1 English speakers’ Accuracy* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | | |
|  | By participant | |  | By item | | |
| *Estimate*  [95% CI] | *SE* | *z* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* | |
| Intercept | 2.99 [2.29, 3.70] | .36 | 8.30 | <.001\*\*\* |  | .31 | .56 |  | 2.69 | 1.64 | |
| Incongruent | .89 [-.04, 1.82] | .47 | 1.88 | .06 |  | - | - |  | - | - | |
| Chinese Word1 | .32 [-.29, .94] | .31 | 1.03 | .30 |  | - | - |  | - | - | |
| Chinese Word2 | .40 [-.32, 1.12] | .37 | 1.09 | .28 |  | - | - |  | - | - | |
| L2 collocation frequency | 1.18 [.32, 2.03] | .43 | 2.71 | .01\* |  |  |  |  |  |  | |

**Appendix C: Acceptability Judgement Task Instructions (for left-handers)**

English:

Welcome to the acceptability judgement task. In this task, you will first see a fixation cross +, followed by an expression. Decide whether the expression is acceptable in English or not. If you think it is acceptable in English, press "s" for Yes; if not, press "k" for No. Respond as QUICKLY and ACCURATELY as possible. Turn off your caps lock before you start.

Chinese:

你将会首先看到一个 +， 接着看到一个词组。请判断组合在英语中是否构成一个可行的符合英语的词组。如果可行，按 s 键； 不可行，按 k 键。要求以最快速度，并保持准确。开始前请关闭大写。

**Appendix D: Follow-up Analysis on the Effect of Proficiency**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table D1**  *Mixed-effects Model for the Effect of Proficiency on EFL learners’ Accuracy of Control Items* | | | | | | | | | | |
|  | Fixed effects | | | |  | Random effects | | | | |
|  | By participant | |  | By item | |
| *Estimate* [95% CI] | *SE* | *z* | *p* |  | *Variance* | *SD* |  | *Variance* | *SD* |
| Intercept | .80 [.51, 1.08] | .14 | 5.53 | <.001\*\*\* |  | .40 | .63 |  | .35 | .59 |
| Proficiency | .01 [.006, .01] | .002 | 4.60 | <.001\*\*\* |  | - | - |  | - | - |
| *Note.* This table presents the final model. The original model included proficiency and length of instruction as main effects of interest and length, L2 Word1 and Word2 frequencies as covariates. This analysis followed the procedures in variable transformation and modelling described in the manuscript. | | | | | | | | | | |

**Appendix E: Model Comparison of Footnote 1**

|  |  |  |
| --- | --- | --- |
| Model specification | | AIC |
| Dependent variable | Independent variables |  |
| ESL RT | L1 Word1 frequency, L1 Word2 frequency, L2 collocation frequency, condition (congruent vs. incongruent), age of arrival, length of residence, percentage of L2 use, collocation length, MI | 727.44 |
| ESL RT | L2 Word1 frequency, L2 Word2 frequency, L2 collocation frequency, condition (congruent vs. incongruent), age of arrival, length of residence, percentage of L2 use, collocation length, MI | 737.67 |
| ESL RT | (the final model in the manuscript) L2 collocation frequency, condition (congruent vs. incongruent), L1 Word1 frequency, L1 Word2 frequency, age of arrival, length of residence, percentage of L2 use, condition \* length of residence, collocation length | 721.57 |
| ESL RT | L2 collocation frequency, condition (congruent vs. incongruent), L2 Word1 frequency, L2 Word2 frequency, age of arrival, length of residence, percentage of L2 use, condition \* length of residence, collocation length | 731.50 |
| EFL RT | L1 Word1 frequency, L1 Word2 frequency, L2 collocation frequency, condition (congruent vs. incongruent), proficiency, length of instruction, collocation length, MI | 962.03 |
| EFL RT | L2 Word1 frequency, L2 Word2 frequency, L2 collocation frequency, condition (congruent vs. incongruent), proficiency, length of instruction, collocation length, MI | 967.41 |
| EFL RT | (the final model in the manuscript) L2 collocation frequency, condition (congruent vs. incongruent), L1 Word1 frequency, L1 Word2 frequency, proficiency, L2 collocation frequency \* condition, L1 Word1 frequency \* length of instruction, collocation length | 957.49 |
| EFL RT | (the final model in the manuscript) L2 collocation frequency, condition (congruent vs. incongruent), L2 Word1 frequency, L2 Word2 frequency, proficiency, L2 collocation frequency \* condition, L1 Word1 frequency \* length of instruction, collocation length | 966.74 |