Appendix A: Descriptions of the predictors considered in this study

Japanese-English-bilingual-specific lexical predictors. Phonological similarity (*PhonologicalSimilarityJPN)* is a rated cross-language phonological similarity measure obtained by 10 Japanese speakers. Participants completed, in a spread sheet, phonological similarity between English words (1st column) and Japanese words (2nd column), using a 7-point scale. In order to safeguard *PhonologicalSimilarityJPN* from potential confound arising from other lexical knowledge, we also considered objective Levenshtein distance coding phonological similarity (*PhonologicalDistance*). *PhonologicalDistance* gauges the number of operations required to transform Japanese words into the corresponding English translation equivalents in their phonologically transcribed form (Levenshtein, 1966; Yarkoni, Balota, & Yap, 2008; Dijkstra et al., 2010; Gooskens & Heeringa, 2004; Schepens, Dijkstra, & Grootjen, 2011; Schepens, Dijkstra, Grootjen, & van Heuven, 2013) based on the `sdists` function available in the `R` package `cba` (Buchta & Hahsler, 2009). In order to compare words of different lengths, we normalized the phonological Levenshtein distance based on the length of target English words (M = 4.3, SD = 1.4).

SemanticSimilarity was based on 10 Japanese-English bilingual readers' ratings on cross-language conceptual similarity. English and Japanese words were presented to the raters side-by-side in two columns in a spreadsheet (English words in the 1st column and Japanese words in the 2nd column). Using a seven-point scale (1 = very different, 7 = identical), the raters assessed the extent to which *katakana* loanwords in Japanese were similar in meaning to the corresponding English target words and whether any Japanese katakana words were completely unfamiliar to them.

FreqJPN reflected how many times Japanese *katakana* words appeared in a Japanese newspaper corpus containing over three million words and covering the 14-year period from 1985 to 1998 (Amano & Kondo, 2003). *FreqJPN* was log-transformed, as its distribution had a long right tail. Note that it is often also possible to translate English words to logographic *kanji* or moraic *kana* words, as well as *katakana* loanwords. However, because the log-transformed frequency of *kanji* or *kana* translations obtained from the same corpus was not a significant predictor, we do not further discuss it.

Although *FreqJPN* comprises two distributions due to zero frequency of occurrence for some words, the corresponding log-transformed Google document frequency measure (*GoogleFreqJPN*) does not indicate such qualitative difference among the set of *katakana* words (see Appendix B). This suggests that the zero frequencies in *FreqJPN* are not due to qualitative differences with respect to words' lexical status, such as transliterations and translations, but due to the fact that the written word corpus is conservative for the purpose of the present study (i.e., it provides frequency counts for *katakana* words only up to the year 1998 and only in the context of newspaper texts).

Finally, a factor *Cognate* (levels: *Cognate* and *NotCognate*) was considered in addition to the above mentioned numerical predictors. Because considerable semantic and phonological overlap and sufficient exposure to words are expected for such special cognate representations to emerge in the first place, words with larger-than-the-average values in all *FreqJPN*, *PhonologicalSimilarityJPN*, and the conceptual similarity were categorized into *Cognate* ($N = 58$).

Lexical distributional predictors of the target English words. As orthographic predictors, we considered word length (*Length*) and orthographic Levenshtein distance

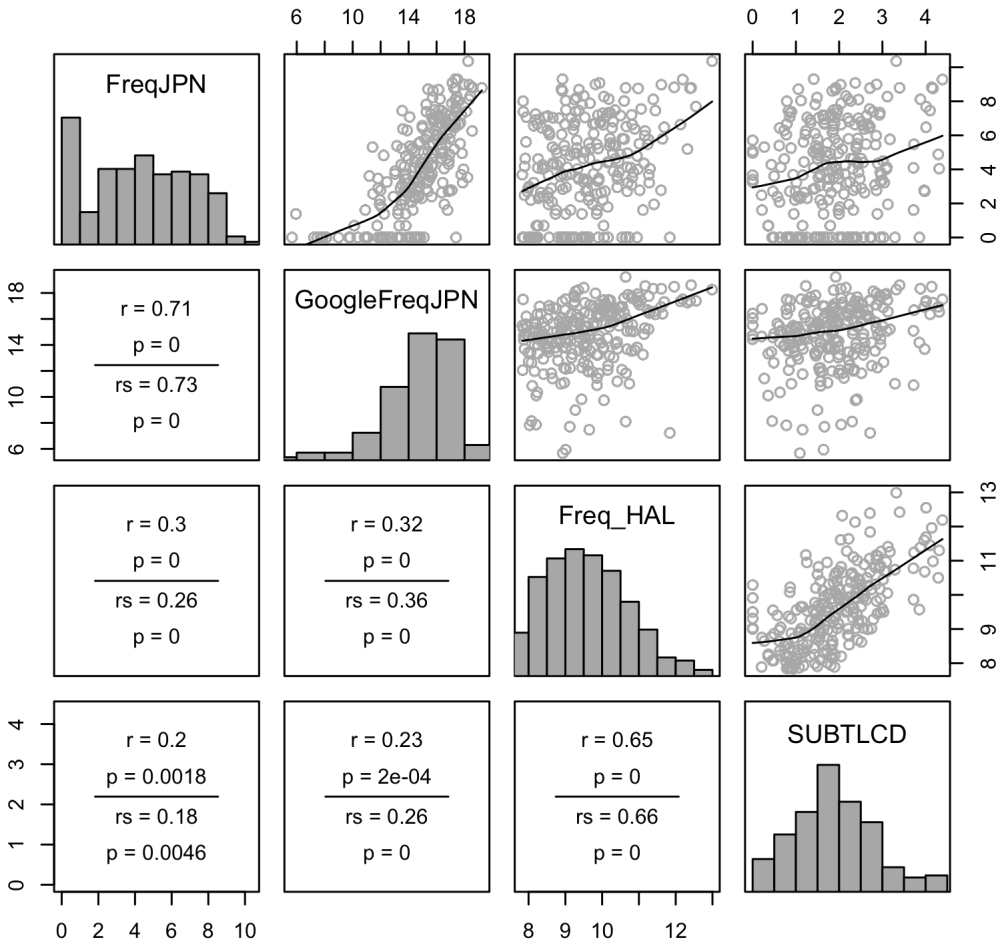
(*OLD20*, Yarkoni et al., 2008). A low *OLD20* score indicates that a given word is located within a dense orthographic space. To measure English word frequency, we used log-transformed *FreqHAL* (HAL: Hyperspace Analogue to Language, Lund & Burgess, 1996; Burgess, & Livesay, 1998, as available in Balota et al., 2007). *SUBTLCD* is a log-transformed context diversity measure based on a number of films in which a given word had been used (Adelman, Brown, & Quesada, 2006; Brysbaert & New, 2009a, 2009b). We also considered ratings of word *Imageability*. Because *SemanticSimilarity* is expected to vary with imageability, with a larger cross-linguistic variance for abstract concepts relative to concrete concepts, *Imageability* safeguards our rated *SemanticSimilarity* measure. We obtained *Imageability* scores rated by 10 native English readers, using a seven-point scale (1 = not imageable, 7 = very imageable).

Task-related predictors and individual differences. In the response time analyses, we considered the following variables: *PreviousRT*, inversely transformed RT in the previous trial; *Trial*, the number of preceding trials; and *PreviousResponseCorrect*, whether the responses in the preceding two trials were correct (see Baayen & Millin, 2010 for autocorrelation in the time-series of response times). In the eye-movement analyses, we also considered *PreviousFixationDuration* for second fixation duration analyses to account for potential spillover effects from the previous fixation. *PreviousRT Trial,* and *PreviousResponseCorrect* were also considered in the eye-movement analyses.

Consideration of readers' L2 proficiency is also important because such individual differences potentially lead to distinct processing mechanisms (Kroll & Stewart, 1994; Potter, So, von Eckhardt, & Feldman, 1984). In our sample, we considered log-transformed participants' months of stay away from Japan for each participant

(*LengthOfStayCanada*) in our regression analyses as a measure of L2 English proficiency. *LengthOfStayCanada* naturally brings in several other components of language proficiency; it highly correlated with age ($r = 0.68$, $p = 0.001$), with their vocabulary size in English measured by X_Lex The Swansea Levels Test (Meara, 2005, $r = 0.48$, $p = 0.03$). We leave the specific advantages and disadvantages of other related measures to future research.

Appendix B: A matrix of scatterplots for Japanese and English frequencies.



Appendix C: A residualization procedure

In order to study independent contributions of lexical distributional properties, we opted for a residulization procedure to resolve multicollineariies. For example, *OLD20* highly correlates with *Length*, *SUBTLCD*, and the number of meanings in WordNet (Miller, 1990). *OLD20* was therefore regressed on these three variables, and the residuals were used as a new predictor *OLD20_resid*. The new predictor correlated significantly and strongly with the original predictor ($r = 0.67$, $p < 0.01$ with *OLD20*). To safeguard our measures, the same residualization procedure was applied to the following variables with highly significant inter-correlations: *SUBTLCD* (regressed on *FreqHAL* and a number of meanings); *Imageability* (regressed on *FreqHAL*, *Length*, number of meanings); *FreqJPN* (regressed on *FreqHAL*, *SUBTLCD*, and number of meanings); *SemanticSimilarity* (regressed on *FreqJPN* and *GoogleFreqJPN*). After the residualization procedure, all the new predictors correlated significantly with the respective original predictors: $r = 0.67$ for *OLD20_resid* and *OLD20,* $r = 0.75$ for *SUBTLCD_resid* and *SUBTLCD*, and $r = 0.95$ for *FreqJPN_resid* and *FreqJPN*. *GoogleFreqJPN* and *PhonologicalDistance* were not included in statistical models together with *FreqJPN* and *PhonologicalSimilarityJPN* but considered separately to assess whether the pattern of results remains unchanged when one predictor is replaced with another. Task-related variables *PreviousRT* and *FirstSubgazeDuration* were similarly regressed on correlated predictors (*Trial* for the former and *Trial*, *PreviousRT*, *FreqHAL*, *Length*, and *SUBTLCD* for the latter), resulting in *PreviousRT_resid* and *FirstSubgazeDuration_resid*.

Appendix D: Materials used in this study

English words used in the present study and their Japanese *katakana* translation and phonology. The flap /ɽ/ was used to encode English approximants /r/ and /l/. /ɸ/ represents a voiceless bilabial fricative. Vowels and consonants were repeated in the Japanese phonological transcriptions to encode the Japanese-specific moraic long vowels, moraic nasals, and moraic obstruents. Words marked with * were excluded from the analyses.

English word	Japanese katakana	Japanese phonology	English word	Japanese katakana	Japanese phonology
accent	アクセント	akusennto	lesson	レッスン	ɽessunn
account	アカウント	akaunnto	letter	レター	ɽetaa
advance	アドバンス	adobannsu	library	ライブラリー	ɽaibuɽaɽii
advantage	アドバンテージ	adobannteedʒi	license	ライセンス	ɽaisennsu
advice	アドバイス	adobaisu	magazine	マガジン	magadʒinn
agenda	アジェンダ	adʒennda	manifest	マニフェスト	maniɸesuto
amateur*	アマチュア	amatʃua	manner	マナー	manaa
anchor	アンカー	annkaa	marble	マーブル	maabuɽu
answer	アンサー	annsaa	margin	マージン	maadʒinn
appeal	アピール	apiiɽu	massage	マッサージ	massaadʒi
arcade*	アーケード	aakeedo	matrix	マトリックス	matoɽikkusu
architect	アーキテクト	aakitekuto	measure	メジャー	medʒaa
aspect	アスペクト	asupekuto	medicine	メディスン	medisunn
attempt	アテンプト	atemmputo	merchant	マーチャント	maatʃannto
auction	オークション	ookuʃonn	message	メッセージ	messeedʒi
autumn	オータム	ootamu	method	メソッド	mesoddo
avenue	アベニュー	abenjuu	minister	ミニスター	minisutaa
balloon	バルーン	baɽuunn	miracle	ミラクル	miɽakuɽu
ballot*	バロット	baɽotto	mirror	ミラー	miɽaa
basket	バスケット	basuketto	mission	ミッション	miʃʃonn
blanket	ブランケット	buɽannketto	moment	モーメント	moomennto
bottom	ボトム	botomu	monster	モンスター	monnsutaa
bracket	ブラケット	buɽaketto	morning	モーニング	mooniŋŋgu
breast	ブレスト	buɽesuto	motion	モーション	mooʃonn
breath	ブレス	buɽesu	mountain	マウンテン	maunntenn
bronze	ブロンズ	buɽonnzu	muscle	マッスル	massuɽu
buffalo	バッファロー	baɸɸaɽoo	museum	ミュージアム	mjuudʒiamu
buffer	バッファー	baɸɸaa	nature	ネーチャー	neetʃaa

bullet	ブレット	buɾetto	needle	ニードル	niidoɾu
bulletin	ブリテン	buɾitenn	notice	ノーティス	nootisu
bundle	バンドル	banndoɾu	notion	ノーション	nooʃonn
burden	バードン	baadonn	number	ナンバー	nammbaa
business	ビジネス	bidʒinesu	occasion	オケーション	okeeʃonn
butter	バター	bataa	office	オフィス	oɸisu
cabinet	キャビネット	kjabinetto	opinion	オピニオン	opinionn
camera	カメラ	kameɾa	opponent	オポーネント	opoonennto
candle	キャンドル	kjanndoɾu	option	オプション	opuʃonn
cannon	キャノン	kjanonn	palace	パレス	paɾesu
career	キャリア	kjaɾia	parade	パレード	paɾeedo
cartoon	カートゥーン	kaatuunn	paradise	パラダイス	paɾadaisu
castle	キャッスル	kjassuɾu	paradox	パラドックス	paɾadokkusu
catalog	カタログ	kataɾogu	pencil	ペンシル	pennʃiɾu
cathedral*	キャシードラル	kjaʃiidoɾaɾu	peninsula*	ペニンスラ	peninnsuɾa
cattle	キャトル	kjatoɾu	period	ピリオド	piɾiodo
ceiling*	シーリング	ʃiiɾiŋŋgu	personnel	パーソネル	paasoneɾu
century	センチュリー	senntʃuɾii	phantom	ファントム	ɸanntomu
challenge	チャレンジ	tʃaɾenndʒi	planet	プラネット	puɾanetto
champion	チャンピオン	tʃammpionn	plastic	プラスチック	puɾasutʃikku
chance	チャンス	tʃannsu	pocket	ポケット	poketto
channel	チャンネル	tʃannneɾu	poison	ポイズン	poizunn
chapter	チャプター	tʃaputaa	police	ポリス	poɾisu
character	キャラクター	kjaɾakutaa	politics	ポリティクス	poɾitikkusu
charter	チャーター	tʃaataa	poverty	パーバティー	paabatii
cherry	チェリー	tʃeɾii	priest	プリースト	puɾiisuto
chocolate	チョコレート	tʃokoɾeeto	prince	プリンス	puɾinnsu
church	チャーチ	tʃaatʃi	principle	プリンシプル	puɾinnʃipuɾu
circuit*	サーキット	saakitto	prison	プリズン	puɾizunn
circus	サーカス	saakasu	privilege	ブリビレッジ	puɾibiɾeddʒi
cluster	クラスター	kuɾasutaa	profile	プロフィール	puɾoɸiiɾu
college	カレッジ	kaɾeddʒi	program	プログラム	puɾoguɾamu
comment	コメント	komennto	promise	プロミス	puɾomisu
complaint	コンプレイント	kommpuɾeinnto	protest	プロテスト	puɾotesuto
component	コンポーネント	kommpoonennto	rabbit	ラビット	ɾabitto
condition	コンディション	konndiʃonn	receipt	レシート	ɾeʃiito
conflict	コンフリクト	konnɸuɾikuto	recipe	レシピ	ɾeʃipi
content	コンテント	konntennto	rescue	レスキュー	ɾesukjuu
corner	コーナー	koonaa	result	リザルト	ɾizaɾuto
couple	カップル	kappuɾu	rocket	ロケット	ɾoketto
course	コース	koosu	salary	サラリー	saɾaɾii
courtesy*	カーテシー	kaateʃii	sample	サンプル	sammpuɾu
credit	クレジット	kuɾedʒitto	satellite*	サテライト	sateɾaito

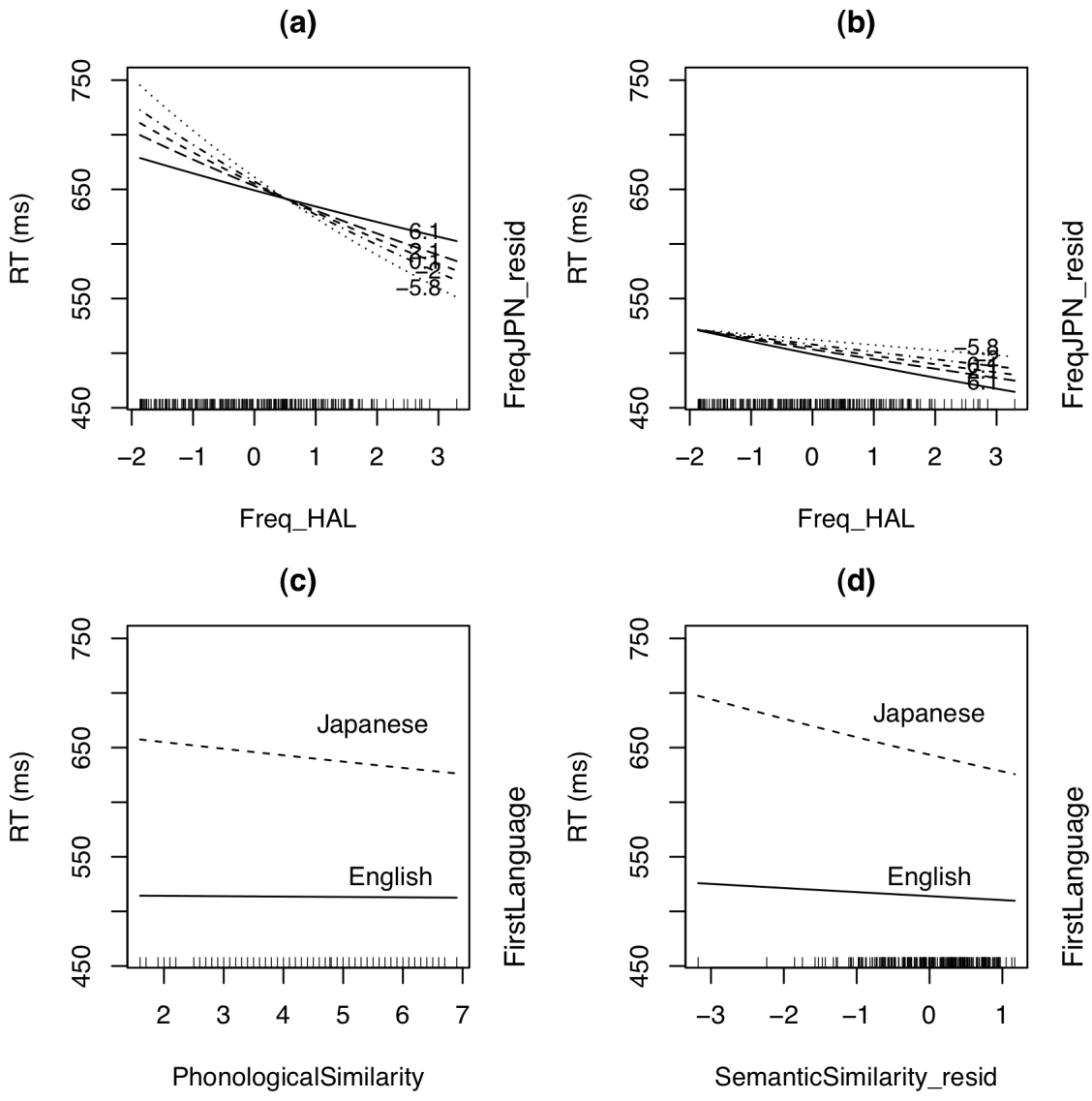
crystal	クリスタル	kuɾisutaɾu	scheme*	スキーム	sukiimu
culture	カルチャー	kaɾutʃaa	school	スクール	sukuuɾu
damage	ダメージ	dameedʒi	search	サーチ	saatʃi
danger	デンジャー	denndʒaa	secretary	セクレタリー	sekuɾetaɾii
debate	ディベート	dibeeto	sentence	センテンス	senntennsu
defense	ディフェンス	diɸennsu	session	セッション	seʃʃonn
degree	ディグリー	diguɾii	shadow	シャドー	ʃadoo
design	デザイン	dezainn	shield*	シールド	ʃiiɾudo
diagram	ダイアグラム	daiaguɾamu	soccer	サッカー	sakkaa
diamond	ダイアモンド	daiamonndo	socket	ソケット	soketto
dilemma*	ジレンマ	dʒiɾennma	soldier	ソルジャー	soɾudʒaa
disaster	ディザスター	dizasutaa	source	ソース	soosu
disease	ディジーズ	didʒiizu	speech	スピーチ	supiitʃi
district	ディストリクト	disutoɾikuto	sponsor	スポンサー	suponnsaa
doctrine*	ドクトリン	dokutoɾinn	square	スクエア	sukuea
domain	ドメイン	domeinn	stance	スタンス	sutannsu
donkey	ドンキー	doŋŋkii	statue*	スタチュー	sutatʃuu
dragon	ドラゴン	doɾagonn	status	ステータス	suteetasu
dungeon*	ダンジョン	danndʒonn	street	ストリート	sutoɾiito
effort	エフォート	eɸooto	strength	ストレングス	sutoɾenngusu
elephant	エレファント	eɾeɸannto	string	ストリング	sutoɾiŋŋgu
embassy	エンバシー	emmbaʃii	studio	スタジオ	sutadʒio
emergency	イマージェンシー	imadʒennʃii	summer	サマー	samaa
emperor*	エンペラー	emmpeɾaa	surface	サーフェス	saaɸesu
episode	エピソード	episoodo	syndrome	シンドローム	ʃinndoɾoomu
example	エグザンプル	eguzammpuɾu	system	システム	ʃisutemu
expert	エキスパート	ekisupaato	talent	タレント	taɾennto
fashion	ファッション	ɸaʃʃonn	target	ターゲット	taagetto
fatigue*	ファティーグ	ɸatiigu	technique	テクニック	tekunikku
fellow	フェロー	ɸeɾoo	template	テンプレート	temmpuɾeeto
finance	ファイナンス	ɸainansu	temple	テンプル	temmpuɾu
flavor	フレーバー	ɸuɾeebaa	territory	テリトリー	teɾitoɾii
flight	フライト	ɸuɾaito	texture	テクスチャ	tekusutʃa
friend	フレンド	ɸuɾenndo	theatre	シアター	ʃiataa
garbage	ガービッジ	gaabiddʒi	thread	スレッド	suɾeddo
garlic*	ガーリック	gaaɾikku	threshold*	スレッシュホールド	suɾeʃʃuhooɾudo
gender	ジェンダー	dʒenndaa	toilet	トイレット	toiɾetto
grease*	グリース	guɾiisu	traffic	トラフィック	toɾaɸikku
guitar	ギター	gitaa	tragedy	トラジェディー	toɾadʒedii
hazard	ハザード	hazaado	treaty	トリーティー	toɾiitii
helmet	ヘルメット	heɾumetto	tunnel	トンネル	tonnneɾu
heroin*	ヘロイン	heɾoinn	twilight	トワイライト	towaiɾaito
horizon	ホライズン	hoɾaizunn	vanilla	バニラ	baniɾa

husband	ハズバンド	hazubanndo	vehicle	ビークル	biikuɽu
impact	インパクト	immpakuto	venture	ベンチャー	benntʃaa
incentive	インセンティブ	innsenntibu	version	バージョン	baadʒonn
industry	インダストリー	inndasutoɽii	veteran*	ベテラン	beteɽann
insect	インセクト	innsekuto	village	ビレッジ	biɽeddʒi
instinct	インスティンクト	innsutinnkuto	violin	バイオリン	baioɽinn
interest	インタレスト	inntaɽesuto	vitamin	ビタミン	bitaminn
interval	インターバル	inntaabaɽu	volume	ボリューム	boɽjuumu
interview	インタビュー	inntabjuu	weather	ウェザー	wezaa
jacket	ジャケット	dʒaketto	whistle	ホイッスル	hoissuɽu
leisure	レジャー	ɽedʒaa	witness	ウィットネス	wittonesu

Appendix E: A direct comparison between monolinguals and bilinguals

Estimate, standard error, t-value, p-value, and effect size (ms) of the fixed effects in the model of Japanese-English bilinguals and English monolingual readers' lexical decision response times analyzed together for a direct comparison of the two groups.

Response time	Type	Estimate	Std.Error	t-value	p-value	Effect size
(Intercept)		-1.910	0.044	-43.72	< 0.0001	
PreviousResponseCorrect (Error)	Task	0.065	0.009	7.23	< 0.0001	17
Trial	Task	-0.054	0.006	-8.79	< 0.0001	-63
PreviousRT	Task	0.158	0.014	11.67	< 0.0001	115
FirstLanguage (Japanese)	Individual	0.456	0.054	8.39	< 0.0001	127
Length	Engl	0.030	0.008	3.90	0.0001	24
FreqHAL	Engl	-0.031	0.007	-4.45	< 0.0001	-41
SUBTLCD_resid	Engl	-0.059	0.008	-6.99	< 0.0001	-66
Imageability_resid	Engl	-0.009	0.004	-1.94	0.0520	-14
FreqJPN_resid	Jpn-Engl	-0.004	0.003	-1.77	0.0763	-13
PhonologicalSimilarity	Jpn-Engl	-0.001	0.006	-0.22	0.8282	-2
SemanticSimilarity_resid	Jpn-Engl	-0.014	0.011	-1.31	0.1901	-16
Length	Engl					66 (Jpn)
* FirstLanguage (Japanese)	* Individual	0.019	0.007	2.69	0.0071	24 (Engl)
Imageability_resid	Engl					15 (Jpn)
* FirstLanguage (Japanese)	* Individual	0.013	0.004	3.49	0.0005	-14 (Engl)
Freq_HAL	Engl					-134 (Jpn)
* FirstLanguage (Japanese)	* Individual	-0.036	0.007	-5.01	< 0.0001	-41 (Engl)
FreqJPN_resid	Jpn-Engl					-17 (Jpn)
* FirstLanguage (Japanese)	* Individual	0.001	0.002	0.64	0.5223	-14 (Engl)
FreqHAL	Engl					Appendix C
* FreqJPN_resid	* Jpn-Engl	-0.002	0.002	-0.70	0.4828	(a, b)
FreqHAL	Engl					
* FreqJPN_resid	* Jpn-Engl					Appendix C
* FirstLanguage (Japanese)	* Individual	0.007	0.002	3.38	0.0007	(a, b)
PhonologicalSimilarity	Jpn-Engl					Appendix C
* FirstLanguage (Japanese)	* Individual	-0.013	0.005	-2.56	0.0105	(c)
SemanticSimilarity_resid	Jpn-Engl					Appendix C
* FirstLanguage (Japanese)	* Individual	-0.024	0.009	-2.70	0.0070	(d)


(a)
RT (ms)
750
650
550
450
−2
−1
0
1
2
3
Freq_HAL
6.1
0.4
−5.8
FreqJPN_resid
(b)
RT (ms)
750
650
550
450
−2
−1
0
1
2
3
Freq_HAL
−5.8
FreqJPN_resid
(c)
RT (ms)
750
650
550
450
Japanese
English
2
3
4
5
6
7
PhonologicalSimilarity
FirstLanguage
(d)
RT (ms)
750
650
550
450
Japanese
English
−3
−2
−1
0
1
SemanticSimilarity_resid
FirstLanguage

Appendix F: Re-consideration of rated measures

As in many past studies, the present study relied on human rated measures. However, there are many different ways to obtain rated measures, and there is no consensus as to what best simulates on-line processing. The figure below summarizes the cross-language similarity measures in phonology and semantics (indicated by solid and dotted lines respectively).

In this study, the objective edit distance measure (*PhonologicalDistance*) replicated the late effect, but not early effects, of the rated *PhonologicalSimilarityJPN* successfully. This may be due to Japanese-English late bilinguals' incomplete mental phonological representations of English words, which was not assumed for *PhonologicalDistance*. When a rated phonological similarity based on an assessment by 10 native English speakers was considered (*PhonologicalSimilarityENGL, M* = 4.3, *SD* = 1.1), this variable behaved much like the objective *PhonologicalDistance* albeit the fact that *PhonologicalSimilarityENGL* correlated with *PhonologicalSimilarityJPN* ($r = 0.63$, $p < 0.01$) more strongly than *PhonologicalDistance* ($r = 0.4$, $p < 0.01$). In the *PhonologicalSimilarityENGL* rating task, the English speakers saw English words on a computer display, while they heard corresponding Japanese words recorded by a native female Japanese reader.

Similarly, there are different ways to rate cross-language *SemanticSimilarity*. A reviewer pointed out that the result might be different if semantic similarity is rated in the reversed direction (i.e., Japanese words in the first column and English words in the second column). We therefore collected ratings from 9 Japanese-English bilinguals (*SemanticSimilarityReversed*). This new measure correlates with *SemanticSimilarity*

largely ($r = 0.73$, $p < 0.01$) and replicated the semantic similarity effects (with an interaction with *PreviousRT*) in response time and last fixation duration analyses.