Dynamics of L3 lexical representations of Dutch-English-Mandarin trilinguals

Xiaowen Ji and Niels Olaf Schiller
University of Shanghai for Science and Technology | City University of Hong Kong

Features and structures of mental lexicon representation are fundamental in psycholinguistics. However, previous investigations into third language acquisition have varied widely in their results and have seldom involved Mandarin as the L3 of participants. In addition, explorations into how mental lexicon representation develops are far from sufficient. In light of the above, this study investigates 72 Dutch-English-Mandarin trilinguals with beginner and advanced Mandarin levels and combines questionnaires, interviews, and priming experiments to explore the dynamic and static features of L3 lexical representations. Results suggest that lexical representations of their three languages are separate to some extent. Links among the three lexical stores seem to exist, with differing strengths. The link between L2 (the instructional language) and L3 may possibly be slightly stronger than that between L1 and L3 for participants of both Mandarin levels. As trilinguals improve their L3 level, the L1-L3 link strengthens, and the asymmetry may gradually diminish.

Keywords: Dutch-English-Mandarin trilinguals, L3, lexical representation, proficiency, instructional language

1. Introduction

The organizational pattern of information pertaining to words within the mental lexicon is usually called structural representation of the mental lexicon, which generally contains conceptual and lexical representations, i.e., two separate but closely interconnected parts (Schiller 2021).

Weinreich (1979[1953]) hypothesized three possibilities of bilingual mental lexicon, i.e., coordinate, compound, and subordinate. These patterns outline that lexical items of each language are stored separately, yet with or without links between lexical stores. Weinreich's (1979[1953]) subordinate and compound patterns are echoed by word association and concept mediation models (Potter et al.

https://doi.org/10.1075/lali.00258.ji | Published online: 6 February 2026
Language and Linguistics 27:2 (2026), pp. 302–323. ISSN 1606-822X | E-ISSN 2309-5067

1984), respectively. The mixed (memory) model (de Groot 1993) holds that lexical items of two languages are stored separately with ever-existing links between the two stores. Later, the concepts of link strength and asymmetrical relations were introduced, and according to the revised hierarchical model, "[…] lexical associations from L2 to L1 are assumed to be stronger than those from L1 to L2 […]" (Kroll & Stewart 1994: 158).

However, some later studies indicate no asymmetry on the lexical level and attribute the asymmetry found to the strengths of different word-concept links. For instance, La Heij et al. (1996) observed that both forward (L1 to L2) and backward (L2 to L1) translations are influenced by semantic variables. Besides, de Groot & Poot (1997) reported faster reaction times in forward than in backward translation for less proficient relative to more proficient Dutch-English bilinguals. Van Hell & de Groot (1998) detected a slightly larger word-type effect in forward translation. Jiang (1999) inferred that these results support that the concept-L1 link is stronger than the concept-L2 link.

Regarding trilinguals, Herwig (2001) proposed that in the initial stage of L3 acquisition, L3 is strongly associated with either L1 or L2 and that with L3 proficiency advancement, the L3 network gradually gains independence. Singleton (2012) concluded that initially the L3 is subordinate to the language that is typologically closer to it and hypothesized that in the second stage of L3 learning, the above-mentioned two languages are related mainly via the shared concept only. Moreover, in the most advanced stage, all three languages may have separate conceptual and lexical representations and a high level of connectivity. According to the Parasitic Model (Ecke 2015), new lexical representations are integrated with previous lexical knowledge at points of similarity or overlap. Furthermore, Gabryś-Barker (2005) hypothesized that all lexical links exist in trilinguals' mental lexicon representations. Cui & Zhang (2009) mapped out a mental lexicon representational structure for Tibetan-Mandarin-English trilinguals, which assumes that only an L2-L3 lexical link exists.

Recent relevant investigations include such topics as language switching costs (Chen et al. 2022), lexical cross-linguistic influence (Efeoğlu et al. 2020), and trilinguals' cognitive advantages (Guðmundsdóttir & Lesk 2019). Although relevant studies are increasingly being conducted, they still have several limitations. First of all, consensus has not been reached on fundamental issues even in bilinguals' mental lexicon representational structure, let alone more complicated trilingualism with far fewer studies; therefore, more research needs to be implemented. Second, most studies do not include developmental trajectories of lexical representations as their subject of investigation. However, only with such knowledge can one grasp learning processes and the nature of learning a foreign language. Third, relatively few psycholinguistic studies of trilingualism involve

Mandarin in the language repertoires of participants; moreover, even if Mandarin is included in their participants' language repertoires, they have acquired Mandarin either as their mother tongue or from an early age (Tytus & Rundblad 2016; Wang 2016). Nowadays, nevertheless, it is an increasingly common phenomenon that Mandarin is learnt as an L3/Ln. This gap should be bridged.

Given the above review, this study employs Dutch-English-Mandarin trilinguals as participants. It aims to answer the following research questions: Are L3 lexical representations stored together with or separate from L1 and L2 lexical items? Are the link strengths of different language pairs symmetrical? Is there any difference between trilinguals with beginner and advanced Mandarin levels concerning link strength? In this study, a long-term cross-language repetition priming paradigm was adopted, and the word naming task was used. Word naming may be accomplished with little conceptual access (Kroll & Stewart 1994), which well suits the goal of this experiment. If we obtain a long-term cross-language repetition priming effect, and that effect is insignificantly different from the long-term within-language repetition priming case, it may be assumed that trilinguals store lexical items of their three languages together. Besides, for symmetry or asymmetry in lexical representation, priming effects under cross-language conditions are interpreted as signals of lexical link strengths. A large priming effect corresponds to a strong lexical link between the two lexical stores involved.

2. Methods

2.1 Participants

All seventy-two participants were Chinese majors at Leiden University in the Netherlands. Based on the interviews with five teachers from different grades in Chinese education programs at the university, participants with low Mandarin proficiency were recruited from sophomores for the beginner group, whereas participants with advanced Mandarin proficiency were from postgraduate students for the advanced group.

Criteria of participant selection were determined by the purposes of this study, information gained from interviews and the limitations of previous studies. (1) The participants needed to be proficient in Dutch, English, and Mandarin, with Dutch being their mother tongue and Mandarin being the least proficient among the three. (2) They needed to be capable of word naming in all three languages. (3) To minimize the potential cross-linguistic influence, the fewer languages they knew beyond the three concerned, the better, and the less proficient they were in other languages, the better. (4) The more homogenous the partici-

pants in one group were, the better. Homogeneity can be demonstrated by language proficiencies, contexts, and the age of acquisition and use of the three languages, and the duration of stays in countries or regions where the languages in question are spoken as the main communication tools. (5) The distinction between Mandarin levels of the two groups needed to be as prominent as possible so that Mandarin proficiency was indeed observed as a reliable influencing factor in L3 lexical representations. Among all the criteria, the first two had to be met, while efforts were made to satisfy the latter three to the greatest possible extent.

Language-background questionnaires were distributed among the then 109 sophomores and 52 postgraduates at the university, and the 36 candidates who best met the above criteria in each group participated. There were 14 male and 22 female participants in the beginner group, with a mean age of 20.19 years and a standard deviation (SD) of 1.33. They began to learn English at a mean age of 8.72 years with an SD of 2.55. There were 13 male and 23 female participants in the advanced group, with a mean age of 22.94 years and an SD of 2.10. They began to learn English at a mean age of 8.28 years with an SD of 3.20. None of the beginner group participants ever went to a Chinese-speaking country or region, whereas all the advanced group participants had been in a Chinese context for at least a year. None of them had learned Mandarin before starting their undergraduate studies. The questionnaires also involve self-ratings of overall proficiencies and specific ones (pronunciation and reading) in the three languages on 7-point Likert scales. SPSS 26 was used to run cross-group one-way ANOVA tests, and results show that the two groups did not differ from each other significantly regarding the self-ratings of the overall Dutch ($p = 0.310$) and English ($p = 0.505$) levels, while their self-ratings of the overall Mandarin level differed significantly ($p < 0.01$). All participants had normal or corrected-to-normal vision, and they voluntarily took part and were paid for their participation. The study was carried out in accordance with the ethical approval of the local ethics committee.

Table 1. Mean self-ratings for proficiencies in the three languages

		Reading	Pronunciation	Overall
Beginner group	Dutch	6.95	6.89	6.92
	English	6.57	5.89	6.53
	Mandarin	3.36	2.92	3.03
Advanced group	Dutch	6.97	6.92	6.97
	English	6.62	6.06	6.61
	Mandarin	4.93	4.45	4.89

2.2 Materials

To select appropriate triads of Dutch, English, and Mandarin words as experimental materials, the following demands were to be met. (1) Mandarin words needed to be familiar to all participants. (2) Traditional characters should be adopted, as students had a longer history of learning traditional characters than simplified ones. This was especially crucial for the beginner group, whose studies of simplified forms had just begun. (3) Efforts should be made to select lexical items that are concrete, monosemous and semantically unrelated to each other, so that conceptual-level activation could be minimized. (4) To avoid huge differences among different language pairs, as Mandarin has cognates with neither English nor Dutch, English and Dutch words chosen also needed to be non-cognates to minimize a possible influence from cognate status in Dutch and English word pairs. However, this demand means that the experiment only explores lexical representation of non-cognate words.

To select materials that conform to the above constraints, the procedure below was followed. Ninety-five concrete Mandarin words were chosen from their textbook *Integrated Chinese* (Traditional Characters, Level 1, Part 1) (Liu et al. 2009), which is used in the very first semester of undergraduate studies at the institution. Only words requiring mandatory mastery could be chosen. Proper names like *Beijing* were removed. English equivalents were taken directly from their textbook and three Dutch/English translation majors at the university translated the English words into Dutch. Finally, a Mandarin teacher of Dutch nationality at the university read all seventy-four triads of words in three languages and confirmed that these materials are translation equivalents and that Mandarin words should be familiar to all participants. Furthermore, the Dutch word *morgen* for 'tomorrow' can also mean 'morning' and was thus removed together with the English and Mandarin equivalents, with seventy-three triads of concrete and monosemous words left. Moreover, efforts were devoted to select non-cognate Dutch and English words. Word pairs that have the same or similar word forms for one meaning (e.g., *papier* and *paper*) were excluded, along with their Mandarin counterparts, leaving forty-three triads.

To further diagnose the cognate status of the remaining Dutch-English word pairs, a guess-the-meaning-questionnaire was designed in which respondents were asked to guess the meanings of those forty-three Dutch words without reference to any source. Twenty college students and staff members from the School of English of the University of Leeds in the U.K. answered the questionnaire. They voluntarily took part and were paid for their participation. These respondents were native speakers of English who spoke neither Dutch nor German. If most respondents can work out an English equivalent of a Dutch word just

based on its orthographic form, it may be the case that the Dutch word triggers the English equivalent, e.g., because the Dutch and English words are cognates or behave cognate-like. All questionnaires were collected, and the threshold was set that if 40% of all respondents (40%*20 respondents = 8 respondents) gave an English word with similar meanings, the respective Dutch and English words were considered cognates or behaved cognate-like. For example, the Dutch word *geld* means 'money', while the word *gold*, which may share some key functions with 'money', was considered an English word with a similar meaning. Three word pairs did not meet the threshold, namely *geld: money, broertje*: *younger brother*, and *mobieltje*: *cell phone*. These words and the Mandarin equivalents were removed. In the end, forty triads of Dutch-English-Mandarin words entered the test materials actually used.

Practically, it was very difficult to find forty mutually unrelated Mandarin words that also met other criteria. Currently, the forty words in each language belong to a wide range of semantic categories (for the materials, see the Appendix), and in the test, efforts were made so that any two words that are semantically related (e.g., *daughter* and *son*) would be arranged to appear with an interval of at least two intervening stimuli. These forty triads of words satisfy all above-mentioned requirements for materials. An example of one triad is *kleren* (Dutch), *clothes* (English), and 衣服 *yīfú* (Mandarin).

2.3 Procedure

The 36 participants per Mandarin level were randomly distributed across three subgroups with 12 participants each. One-way ANOVA tests displayed that the three subgroups for each Mandarin level showed no significant difference in their self-ratings of overall proficiency and reading as well as pronunciation proficiency in the three languages. Thus, the three subgroups for each Mandarin level were considered homogeneous.

Table 2. One-way ANOVA test results for both groups

		Reading			Pronunciation			Overall		
		D[a]	E	M	D	E	M	D	E	M
p-values	1[b]	0.61	0.48	0.18	0.15	0.57	0.18	1.00	0.49	0.20
	2	0.38	0.17	0.96	1.00	0.08	0.96	0.38	0.16	0.30

a. D = Dutch; E = English; M = Mandarin.
b. 1 = beginner group; 2 = advanced group.

Half of the material, i.e. 20 triads of words, was randomly assigned to the priming block. For each language condition, the priming block contained 20 trials (the studied words); the test block contained these same 20 trials (either in the same or in a different language) plus the other 20 trials (the non-studied words), i.e. 40 trials altogether. The studied words were regarded as homogeneous with respect to the non-studied words, as they came from the same set of materials and went through the same screening procedure.

In the priming block, words in the three languages were assigned to the three subgroups of participants per Mandarin level, with each subgroup naming the 20 words in one language. In the test block, all 40 words in three languages were named by all participants. In this way, the only difference among the three subgroups was in the language of words in the priming blocks.

All participants took part in the test individually in a quiet lab with a soundproof testing room. When a participant arrived, instructions were given that they needed to name four blocks of words altogether. Among the four blocks, the first block was a priming block, and the remaining ones were test blocks. However, differentiation of priming and test blocks was concealed from the participants so as not to divert their attention. They were also made aware of the naming task and which buttons to press to pause and resume the test, and they were told that the voice key-trigger was adopted and that they needed to accurately read aloud the words as quickly as possible. They were required not to make any other sound, which might interfere with the mechanism. At the same time, they were advised to say "Nee", "No" or "不 (*bu*)" if they happened not to be able to name a word, which was to minimize the undesirable influence of participants' possible use of strategies. After completing the priming block, the experimenter corrected the participants if they made mistakes or did not provide a response within 3,000 ms.

Participants sat in the testing room at a distance of approximately 60 cm from a 22-inch wide LCD monitor (Philips 220BLP) with a 1,680*1,050 pixel resolution. E-Prime was used to present all stimuli. There were nine training trials with three in each language. During the test, languages of words that appeared in all blocks were made known to participants through the relevant screen pages that offered instructions, to reduce unwanted errors caused by unexpected language switches. For each subgroup, the 20 words in priming block were presented in the planned language, while the order of the three languages of the 40 triads in the test blocks was counterbalanced across the twelve participants. In each of the four blocks, the trials were arranged to appear in a pseudo-random order, with the experimenter making necessary changes to the trial list randomly generated by E-Prime, so that any two words that are semantically related appeared at least two words away. Moreover, three filler words in the same language were added to

the beginning of each trial list. Additionally, there was a mandatory rest between two blocks.

The training procedure and the priming and test blocks followed the same procedure for each trial. A fixation cross was first presented for 1,000 ms to draw participants' attention, followed by a blank screen for 20 ms. After that, the target word appeared and remained in view until it was named, yet it would disappear automatically if no response was given within 3,000 ms. Then another blank screen (500 ms) followed, marking the end of the trial. E-Prime recorded RT data, while the experimenter took notes of errors and invalid cases, caused by uncalled-for noises or other sounds accidentally made. Recordings were made to assist with later analyses.

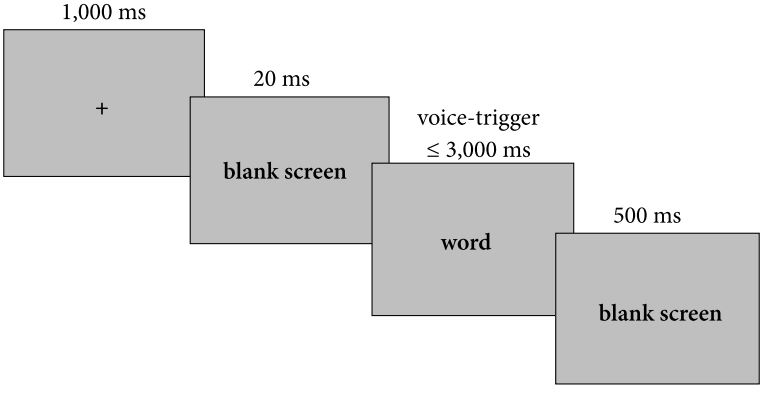

Figure 1. Trial sequence in the experiment

3. Results

Data collected fall into two categories, namely those for baseline testing and those for main experiments. Baseline testing concerns two questions. (1) Were the three subgroups in each group homogeneous in terms of language proficiency? (2) Were the randomly assigned two sets of 20 triads of words homogeneous according to baseline error rate (hereinafter "ER") and reaction time (hereinafter "RT") data?

Data were preprocessed first. The authors marked invalid cases of participant performance with the help of recordings, such as situations when participants were only able to trigger the corresponding words to disappear the second time they read them aloud. Also, the authors spotted outliers three SD units above and below the mean RTs of all participants. The invalid cases and outliers were all

removed, and they made up 2.3% of the total data of the beginner group, and the number was 2.2% for the advanced group.

3.1 Results for baseline testing

3.1.1 ER data and analyses

For the first baseline question, ER data for naming non-studied words (*viz.* the 20 triads of words only appearing in the test block) in three languages were compared across the three subgroups within each group. Dutch was first considered. No error was spotted with any of the participants. The same was true for English. Mean ER data for Mandarin were calculated, and one-way ANOVA tests with "group" as the independent variable and "ER" as the dependent variable were run. The results show that the three subgroups in each group were not significantly different from each other, concerning the ER levels in naming words in all three languages.

For the second baseline question, ER data for naming studied and non-studied words in three languages were compared within each group. Again, no error was spotted in naming Dutch or English words for any participant. Mean ER data for Mandarin words were calculated, and paired-samples tests were run. The results show that the two sets of words in all three languages did not show any significant difference in baseline ER in either group.

Table 3. Mean ER (%) data for naming non-studied Mandarin words and one-way ANOVA test results

	Mean ER (%) data			*p*-values in ANOVA tests
	Subgroup 1	Subgroup 2	Subgroup 3	
Beginner group	9.58	10.42	10.42	0.89
Advanced group	4.58	3.33	4.58	0.67

Table 4. Mean ER (%) data for naming studied and non-studied Mandarin words and paired-samples test results

	Mean ER (%) data		*p*-values in paired-samples tests
	Non-studied	Studied	
Beginner group	10.42	11.67	0.57
Advanced group	4.58	3.33	0.52

3.1.2 RT data and analyses

For the first baseline question, relevant mean RT data were calculated, and one-way ANOVA tests with "group" as the independent variable and "RT" as the dependent variable were run. RT data reveal that the three subgroups from each group displayed no significant difference in RT levels in naming words in all three languages.

For the second baseline question, RT data for studied and non-studied words were calculated, and paired-samples tests were run. The results indicate that words for study and words not for study in all three languages did not significantly differ in baseline RT, i.e., when they were named for the first time.

Table 5. Mean RT (ms) data for naming non-studied words in the three languages across subgroups and one-way ANOVA test results

		Mean RT (ms) data			*p*-values of ANOVA tests
		Subgroup 1	Subgroup 2	Subgroup 3	
1[a]	Dutch	509	516	500	0.89
	English	551	541	541	0.88
	Mandarin	1078	1077	1070	0.98
2	Dutch	504	501	507	0.95
	English	542	540	538	0.95
	Mandarin	940	939	941	1.00

a. 1 = beginner group; 2 = advanced group.

Table 6. Mean RT (ms) data for naming non-studied and studied words and paired-samples test results

		Mean RT (ms) data		*p*-values in paired-samples tests
		Non-studied	Studied	
1[a]	Dutch	509	512	0.89
	English	541	543	0.88
	Mandarin	1070	1083	0.78
2	Dutch	504	517	0.49
	English	540	539	0.94
	Mandarin	941	990	0.12

a. 1 = beginner group; 2 = advanced group.

3.1.3 Summary

ER and RT data support the validity of participant and material selection and division. For each group, the proficiencies of the three subgroups in naming words in the three languages did not differ significantly from each other, and homogeneity among the three subgroups in this aspect validates cross-group and thus cross-language comparisons. Furthermore, the two sets of words, studied *vs.* non-studied, in all three languages did not differ significantly from each other for the two groups. This suggests that the ER and RT data of the two sets of words can be compared in the following.

3.2 Results for main experiments

3.2.1 ER data and analyses

In the beginner group, language conditions with Dutch target words were analyzed. In the priming block, one subgroup named Dutch words (DD), another subgroup English words (ED), and the third subgroup Mandarin words (MD). The first capitalized letter indicates the language of studied words in the priming block, and the second capitalized letter indicates the language of words tested in the test block. This labeling manner applies hereafter. ER data for DD and ED, and for DD and MD were compared, to examine whether the lexical representation of Dutch vocabulary is shared with that of English or Mandarin vocabulary. No error was spotted. The same was true for language conditions with English target words, i.e., DE, EE, and ME.

Language conditions with Mandarin target words (DM, EM, and MM) were analyzed. ER data for MM and DM, and for MM and EM were compared. Repeated measures ANOVA test results show that both the main effect of word types (studied and non-studied words) ($p = 0.015$) and interaction effect between word types and language conditions (DM, EM, and MM) ($p = 0.028$) were significant. Results also show that the main effect of language conditions ($p = 0.039$) was significant. Therefore, ERs for non-studied words were significantly larger than those for studied words, and overall, there was significant long-term repetition priming. However, according to paired-samples tests results, *p*-values were non-significant for DM and EM (i.e., 0.830 and 0.701, respectively), indicating insignificant differences between ER data for studied and non-studied words under the language conditions and hence no significant long-term cross-language repetition priming. Furthermore, results show that the *p*-value was 0.001 for MM, indicating a significant difference between ERs for studied and non-studied words and hence significant long-term within-language repetition priming.

Table 7. Mean ER (%) data for Mandarin target words

		Non-studied	Studied
Beginner group	DM	9.58	9.17
	EM	10.42	9.58
	MM	10.42	3.33
Advanced group	DM	4.58	4.58
	EM	3.33	2.92
	MM	4.58	2.50

Similar analyses were carried out in the advanced group. Repeated measures ANOVA test results show that neither the main effect of word types ($p = 0.316$) nor the interaction effect between word types and language conditions ($p = 0.552$) was significant. The main effect of language conditions ($p = 0.326$) was also insignificant. These results indicate that although ER data varied per language condition with Mandarin target words, differences within all the comparable pairs were insignificant, showing no significant long-term repetition priming in either cross-language or within-language cases. Therefore, there was no need to carry out further data analyses.

3.2.2 RT data and analyses

In each group, data were divided into three parts for separate analyses, i.e., language conditions with Dutch target words, those with English target words, and those with Mandarin target words. Mean RT data for naming non-studied and studied words in test blocks are given in Table 8.

Table 8. Mean RT (ms) data (M±SD) in both groups

	Language conditions	Non-studied	Studied
1[a]	DD	509 ± 56	502 ± 38
	ED	516 ± 73	518 ± 59
	MD	500 ± 101	506 ± 59
	DE	551 ± 57	549 ± 58
	EE	541 ± 63	535 ± 66
	ME	541 ± 47	542 ± 47
	DM	1078 ± 67	1064 ± 152
	EM	1077 ± 135	1051 ± 198
	MM	1070 ± 134	807 ± 32

Table 8. *(continued)*

	Language conditions	Non-studied	Studied
	DD	504 ± 34	498 ± 60
	ED	501 ± 27	501 ± 39
	MD	507 ± 69	511 ± 59
	DE	542 ± 35	533 ± 51
2	EE	540 ± 36	528 ± 52
	ME	538 ± 20	529 ± 17
	DM	940 ± 196	922 ± 179
	EM	939 ± 80	917 ± 87
	MM	941 ± 158	645 ± 17

a. 1 = beginner group; 2 = advanced group.

Repeated measures ANOVA tests were conducted to delve into the roles of word types and language conditions. Note that "word types" is a within-subjects factor and that "language conditions" is a between-subjects factor.

For language conditions with Dutch target words, neither the main effect of word types [$p = 0.985$ (beginner); $p = 0.899$ (advanced)] nor the interaction effect between word types and language conditions [$p = 0.792$ (beginner); $p = 0.784$ (advanced)] was significant. The main effect of language conditions [$p = 0.849$ (beginner); $p = 0.890$ (advanced)] was not significant either. RT data across all three language conditions in each group were almost identical, with p-values all being 1. These results show that neither within-language nor cross-language long-term repetition priming was significant. The same was true for language conditions with English target words.

For language conditions with Mandarin target words, both the main effect of word types ($p < 0.001$ for both groups) and the interaction effect between word types and language conditions ($p < 0.001$ for both groups) were significant. The main effect of language conditions [($p = 0.014$ (beginner); $p = 0.015$ (advanced)] was also significant. RT data under the language conditions DM and MM were significantly different [$p = 0.027$ (beginner); $p = 0.032$ (advanced)], and those under EM and MM also differed significantly [$p = 0.038$ (beginner); $p = 0.037$ (advanced)]. Thus, RT data under MM were significantly shorter than those under DM or EM.

Pairwise comparisons by word types show that p-value was smaller than 0.001 for each group, which indicates that RTs for studied words were significantly shorter than those for non-studied words. Paired-samples tests were performed under the three language conditions in each group, and results show that p-values were 0.692, 0.394, and < 0.001 for DM, EM, and MM, respectively, in the beginner

group, and that *p*-values were 0.274, 0.290, and < 0.001 for DM, EM, and MM, respectively, in the advanced group.

All results put together present the following points for language conditions with Mandarin target words. (1) There was a main effect of word types, indicating the existence of overall long-term repetition priming and the usefulness of the current data. (2) There was a main effect of language conditions, which shows a difference between within-language and cross-language conditions. (3) There was an interaction effect between word types and language conditions, and further analyses show that only under MM, RTs for studied words were significantly shorter than those for non-studied words. In other words, significant long-term repetition priming existed only in the within-language condition but never in the cross-language condition.

In addition, as to whether lexical links are symmetrical, despite insignificant cross-language priming, relevant priming effects were calculated for potential reference. The priming effect (PE) is obtained from the priming amount (the RT for naming studied words subtracted from those for non-studied words) divided by the RT for naming non-studied words, as illustrated below.

$$\text{PE} = \frac{\text{RT(non-studied words)} - \text{RT(studied words)}}{\text{RT(non-studied words)}}$$

PE data under language conditions DM and EM were studied and are presented in Table 9, which displays that PE data under DM are smaller than those under EM. Though the actual difference is very small, possibly the result still suggests a slightly stronger English-Mandarin link than the Dutch-Mandarin link.

Table 9. PE (%) data under language conditions DM and EM

Language conditions		**DM**	**EM**
PE	beginner-level	1.37	2.39
	advanced-level	1.90	2.34

Furthermore, we observed that PE data under EM have remained almost the same with Mandarin proficiency improvement, whereas those under DM have been greatly boosted by around 39%.

4. General discussion

The current experiment yielded the following results. ER and RT data from both the beginner and advanced groups failed to show any significant long-term repetition priming under the language conditions with Dutch and English target words. The main reason may be that Dutch and English equivalents to those Mandarin words are so basic that the priming processes did not make a difference. However, ER and RT data per language condition with Mandarin target words from the beginner group displayed significant within-language long-term repetition priming but insignificant cross-language long-term repetition priming. This result demonstrates that whether the language of the words in priming blocks is the same as that of the words in test blocks significantly influences priming. This suggests that Mandarin lexical items are stored separately from Dutch and English lexical items for participants with lower-level Mandarin. In the advanced group, although no conclusion could be drawn from the ER data, the RT data also pointed towards a separation of the Mandarin lexical store from each of the other two.

Although separation of non-cognate Dutch and English lexical representations can neither be validated nor falsified in this experiment, some inferences can be drawn based on existing models, many of which were proposed based on studies of Dutch and English (Kroll & Stewart 1994; van Hell & de Groot 1998). One of the common results from these studies is the separation of Dutch and English lexical representations, which is considered to apply to participants in the current experiment. Therefore, we suggest that Dutch-English-Mandarin trilinguals seem to have separate stores for non-cognate lexical entries in their three languages.

Furthermore, the current experiment suggested an asymmetry in lexical representation, with the Dutch-Mandarin link deemed slightly weaker than the English-Mandarin link for both the beginner and advanced groups. Besides, with Mandarin proficiency improvement, the Dutch-Mandarin link seemed to experience a much greater change regarding link strength than the English-Mandarin link did, resulting in a more obvious asymmetry for the beginner group.

These findings highlight the critical roles of language proficiency and instructional language. The rapid enhancement of the L1-L3 link strength with Mandarin level advancement can be found, and this observation seems to indicate that the more proficient the L3 is, the stronger the L1-L3 link is. It is consistent with previous studies such as Kroll & Stewart (1994). Meantime, instructional language also matters. The L2-L3 connection is interpreted to be slightly stronger than the L1-L3 link for both the beginner and advanced groups. Besides, the L2-L3 link in the beginner group is already as strong as in the advanced group.

This belief echoes the conclusions of previous studies such as Cui & Zhang (2009).

On one hand, results of this study are in agreement with certain aspects of the mixed (memory) model (de Groot 1993) and the revised hierarchical model (Kroll & Stewart 1994) in that all lexical links are hypothesized to exist in participants' lexical representations. The shift between the word association model and the concept mediation model (Potter et al. 1984) resembles the developmental trajectory of link strengths in the lexical representations of trilinguals shown in this study. The two agree in that the link between the least (less) proficient language and another (the other) language, which is the instructional language from the language repertoire is already very strong initially and that any possible link connecting the least (less) proficient language and the concept or still another language is later strengthened or established. Herwig (2001) supports this point by holding that L3 is strongly related to either L1 or L2 initially and gradually gains its independence as the L3 level increases. Singleton's model (2012) also strives to provide a more detailed description of the dynamic transformation of trilinguals' mental lexicon representation. The current study suggests that language proficiency and instructional language, among many others contribute to trilinguals' lexical representations, thus supporting Gabryś-Barker's (2005) conclusion. Moreover, this paper has similar implications as Cui & Zhang's model (2009) by acknowledging the role of instructional language in L3 lexical representations.

On the other hand, this study's results show some distinctly deviating features from all previous relevant models. Unlike the concept mediation and word association models (Potter et al. 1984), which suggest that bilinguals may not have all links established in their mental lexicon representations, this study hypothesizes that all possible lexical links exist. The mixed (memory) model (de Groot 1993) contains all possible lexical links, but it does not incorporate developmental factors, which are part of the current study. The revised hierarchical model (Kroll & Stewart 1994) is developmental in nature. However, it distinguishes forward and backward lexical links, whereas the current study does not. Furthermore, the current study also builds on Gabryś-Barker's (2005) work, as the latter incorporates only some influencing factors, with no detailed study of representational structure. Herwig's model (2001) mentions only some factors that may influence the mental lexicon structure of trilinguals. The Parasitic Model (Ecke 2015) makes no relevant detailed comments. Singleton's model (2012) describes the pattern of L3 conceptual access as a word association model in the initial stage of L3 learning, as a concept mediation model in the second stage and as a separate storage model in the last stage, whereas the current paper features the existence of all lexical links for L3 conceptual access throughout all stages of L3 learning.

Another model previously proposed is Cui & Zhang's (2009) model. Unlike the current study, it makes no comments on developmental factors or link strengths, and it considers L2 and L3 to be linked directly, with no L1-L3 lexical link. It is possible that the L1-L3 link, in Cui & Zhang's (2009) case, actually exists, which Cui & Zhang only failed to discover owing to its weak link strength. Moreover, Cui & Zhang's (2009) participants were trilinguals with low L3 levels, and the model she proposed can thus explain only novice trilinguals' lexical representations.

The methodology in this study has the following limitations. (1) Although English was used mostly as the instructional language of Mandarin for the participants, every week there was still one lesson hour of Mandarin taught by Dutch teachers with Dutch as the instructional language, in each semester of the first two years of undergraduate studies. Therefore, this is not an extreme case of the contrasting roles of the mother tongue and the instructional language. (2) Materials used were limited, mainly by the Mandarin level of the beginner group. The current Mandarin materials were selected to ensure the participants' ability to name them and were deemed to correspond to the most frequently used vocabulary. In terms of word length, mainly two-character Mandarin words were used. However, word lengths of Dutch and English equivalents were left unattended. Besides, behavioral methods adopted in the current study are to be complemented by a brain imaging approach. Additionally, the mirror-image design was not used because it was extremely difficult to find English-Dutch-Mandarin trilinguals. A reasonable alternative (Cabrelli Amaro & Rothman 2010) is to compare acquisition of Mandarin as L2 and L3 when L1 is held constant. However, it was still hard to find Dutch-Mandarin bilinguals with lower English levels than those of their Mandarin.

The current study may also shed light on third language acquisition and language transfer studies. To start with, many previous studies proposing and testing the Typological Primacy Model (Rothman et al. 2019), the Abbreviated Grappling Period Model (Sprouse & Schwartz 2023), and the Linguistic Proximity Model (Westergaard et al. 2017), among other models of third language transfer, generally witness a very limited number of language repertoires that involve the L3 as genetically, typologically, or orthographically related to neither L1 nor L2. Moreover, the more advanced stages of L3 acquisition have not received sufficient attention. This study contributes to the understanding of both the beginner and advanced stages of L3 acquisition, with the L3 having completely distinct writing systems from either of the previously acquired languages, thus offering consideration for relevant research.

Further studies may delve deeper into the characteristics of the various aspects of the Mandarin interlanguage, and what roles certain abstract properties of Dutch and English may play in the processes of L3 acquisition.

5. Conclusion

Dutch-English-Mandarin trilinguals seem to have separate stores for non-cognate lexical items in their three languages. Links among the three lexical stores all seem to exist. For trilingual participants with both beginner and advanced Mandarin levels, the link between L2 English, the instructional language, and L3 Mandarin is deemed slightly stronger than the L1-L3 link.

From a dynamic perspective, as trilinguals improve L3 levels, the lexical link between L1 and L3 is greatly consolidated. On the other hand, the lexical link between L2 English, the instructional language, and L3 Mandarin does not experience a great change in terms of the link strength. Language proficiency and instructional language are deemed to exert profound influence over link strength and thus symmetry patterns in the lexical representation of trilinguals.

Acknowledgements

We thank all those who have helped us improve the paper, including Professor Deming Mei, the reviewers and editors. This work is supported by the 2026 Project of University of Shanghai for Science and Technology for Cultivating Young Scholars in Humanities and Social Sciences.

Appendix. Experimental materials

No.	Dutch	English	Mandarin
1	Dhr.	Mr.	先生
2	kleren	clothes	衣服
3	gisteren	yesterday	昨天
4	ochtend	morning	上午
5	gerecht	dish	菜
6	juffrouw	Miss	小姐
7	maand	month	月
8	dochter	daughter	女兒

Appendix. *(continued)*

No.	Dutch	English	Mandarin
9	klasgenoot	classmate	同學
10	donderdag	Thursday	星期四
11	beker	cup	杯
12	vijf uur	five o'clock	五點
13	zoon	son	兒子
14	baan	job	工作
15	nu	now	現在
16	verjaardag	birthday	生日
17	vliegveld	airport	機場
18	vandaag	today	今天
19	ontbijt	breakfast	早飯
20	maaltijd	meal	餐
21	buitenland	foreign country	外國
22	docent	teacher	老師
23	vliegtuig	airplane	飛機
24	winkel	store	商店
25	kind	child	孩子
26	achternaam	surname	姓
27	vraag	question	問題
28	tijd	time	時間
29	vrouwelijk	female	女
30	thuis	home	家
31	arts	doctor	醫生
32	zaak	thing	事
33	mensen	people	人
34	maat	size	大小
35	klaslokaal	classroom	教室
36	jaar	year	年
37	zusje	younger sister	妹妹
38	avond	evening	晚上
39	zwart	black	黑
40	opstaan	get up	起床

References

Cabrelli Amaro, Jennifer & Rothman, Jason. 2010. On L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems. *International Review of Applied Linguistics in Language Teaching* 48(2-3). 275–296.

Chen, Jianlin & Wang, Junzhao & Jiang, Ci & Gong, Yifei & Tsering Tashi. 2022. The proactive language control mechanism of language switching in Tibetan-Chinese-English trilinguals. *Waiyu Jiaoxue yu Yanjiu* 2022(3). 433–443.

Cui, Zhanling & Zhang, Jijia. 2009. A research on lexical and conceptual representations in Tibetan-Mandarin-English trilinguals. *Xinli Kexue* 2009(3). 559–562.

de Groot, Annette M. B. 1993. Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In Schreuder, Robert & Weltens, Bert (eds.), *The bilingual lexicon*, 27–52. Amsterdam: John Benjamins.

de Groot, Annette M. B. & Poot, Rik. 1997. Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning* 47(2). 215–264.

Ecke, Peter. 2015. Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language and Cognition* 18(2). 145–162.

Efeoğlu, Gülümser & Yüksel, Hatice Gülru & Baran, Suat. 2020. Lexical cross-linguistic influence: A study of three multilingual learners of L3 English. *International Journal of Multilingualism* 17(4). 535–551.

Gabryś-Barker, Danuta. 2005. *Aspects of multilingual storage, processing and retrieval.* Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Guðmundsdóttir, Margrét D. & Lesk, Valerie E. 2019. Does the bilingual advantage extend to trilingualism?. *International Journal of Multilingualism* 16(4). 549–562.

Herwig, Anna. 2001. Plurilingual lexical organisation: Evidence from lexical processing in L1-L2-L3-L4 translation. In Cenoz, Jasone & Hufeisen, Britta & Jessner, Ulrike (eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*, 115–137. Clevedon: Multilingual Matters.

Jiang, Nan. 1999. Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition* 2(1). 59–75.

Kroll, Judith F. & Stewart, Erika. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33(2). 149–174.

La Heij, Wido & Hooglander, Andre & Kerling, Robert & van der Velden, Esther. 1996. Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language* 35(5). 648–665.

Liu, Yuehua & Yao, Tao-chung & Bi, Nyan-Ping & Ge, Liangyan & Shi, Yaohua. 2009. *Integrated Chinese, level 1, part 1*, 3rd edn. Boston: Cheng & Tsui Company, Inc.

Potter, Mary C. & So, Kwok-Fai & Von Eckardt, Barbara & Feldman, Laurie B. 1984. Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior* 23(1). 23–38.

Rothman, Jason & González Alonso, Jorge & Puig-Mayenco, Eloi. 2019. *Third language acquisition and linguistic transfer*. Cambridge: Cambridge University Press.

Schiller, Niels Olaf. 2021. The mental lexicon. In Aronoff, Mark (ed.), *Oxford bibliographies in linguistics*, 1–32. Oxford: Oxford University Press. (https://www.academia.edu/60326727/The_Mental_Lexicon) (Accessed 2022-09-22.)

Singleton, David. 2012. Multilingual lexical operations: Keeping it all together…and apart. In Cabrelli, Jennifer & Flynn, Suzanne & Rothman, Jason (eds.), *Third language acquisition in adulthood*, 95–114. Amsterdam: John Benjamins.

Sprouse, Rex A. & Schwartz, Bonnie D. 2023. L3 cognitive states and the Abbreviated Grappling Period Model. In Sheu, Vanessa & Weirick, Joshua D. & Zhou, Alexis (eds.), *Variation in linguistics: Second language acquisition, discourse studies, sociolinguistics, syntax*, 2–25. Newcastle: Cambridge Scholars Publishing.

Tytus, Agnieszka Ewa & Rundblad, Gabriella. 2016. Cross-language priming as a means of investigating bilingual conceptual representations: A comparison of visual and auditory modality. *Linguistic Approaches to Bilingualism* 6(4). 440–466.

van Hell, Janet G. & de Groot, Annette M. B. 1998. Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology* 51(1). 41–63.

Wang, Zhen. 2016. Investigating the long-term cross-language repetition priming effects in Chinese-English-Japanese trilinguals. *Xiandai Waiyu* 2016(1). 64–73.

Weinreich, Uriel. 1979[1953]. *Languages in contact: Findings and problems*. The Hague: Mouton. (Originally published in 1953 by the Linguistic Circle of New York.)

Westergaard, Marit & Mitrofanova, Natalia & Mykhaylyk, Roksolana & Rodina, Yulia. 2017. Crosslinguistic influence in the acquisition of a third language: The linguistic proximity model. *International Journal of Bilingualism* 21(6). 666–682.

Address for correspondence

Niels Olaf Schiller
Department of Linguistics and Translation
City University of Hong Kong
83 Tat Chee Avenue, Kowloon Tong
Hong Kong
China

niels.schiller@cityu.edu.hk
https://orcid.org/0000-0002-0392-7608

Co-author information

Xiaowen Ji
College of Foreign Languages
University of Shanghai for Science and Technology
jixiaowen@usst.edu.cn
https://orcid.org/0000-0002-0244-2529

Publication history

Date received: 2 January 2024
Date accepted: 24 April 2025
Published online: 6 February 2026