



The International Journal of Language Studies

ISSN: 3078-2244 (Online)

2026, Vol. 3, No. 2, pp. 148–160

DOI: <https://doi.org/10.60087/ijls.vol3.n2.009>

Journal homepage: <https://ijlangstudies.org>



Is a Distant Language Harder to Interpret? Evidence from a Comparative Empirical Study

Yuxuan Zheng^{1*}; Boning Zhang¹; Qianyu Zhao²

¹ Beijing International Studies University, Beijing, China

² Beijing Foreign Studies University, Beijing, China

*Corresponding author: Yuxuan Zheng — e-mail: zyx@bisu.edu.cn

Abstract

Whether language distance is a built-in influencing factor for interpreting performance has rarely been empirically investigated. The current research aimed to examine the language distance effect in consecutive interpreting with two experiments. Experiment 1 involved 45 Chinese-English bilinguals and Experiment 2 recruited 46 German-English bilinguals. Both experiments required participants to interpret 10 groups of English sentences into their first languages respectively. Demographic and linguistic background data such as participants' age and age of second language acquisition were collected, and were analyzed together with their interpreting performance. Statistical analyses revealed that Chinese-English bilinguals spent significantly longer time learning English, but fared significantly worse than German-English bilinguals in both English proficiency test and consecutive interpreting task, that language distance is a significant predictor for L2 proficiency and interpreting performance. However, multiple linear regression model showed that when L2 proficiency is taken into account, the language distance factor becomes insignificant. We concluded that language distance mainly exerts influence on L2 proficiency, and it is the latter that plays a pivotal role in influencing interpreting performance. Future studies are advised to expand the current paradigm to more groups of participants and language pairs.

Keywords: language distance effect; interpreting; empirical research; L2 proficiency

Article information: Received: 09/03/2026;

Accepted: 26/04/2026;

Published: 25/06/2026

Copyright © 2026 The Author(s). Published by the International Journal of Language Studies. This is an open-access article distributed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

1. Introduction

Interpreting performance is subject to the influence of various factors, among which an important but significantly less probed one is language (or linguistic) distance, or language pair specificity/language similarity as some other researchers may prefer to use (Gile, 1992), i.e., the degree of difference between languages, mainly reflected in vocabulary, grammar, pronunciation and spelling rules (Renfrew et al., 2000). Expectedly, the characteristics of one's first language (L1) necessarily impact his/her learning of other languages in one way or another. For example, Cui et al. (2018) investigated the effect of L1 difference on Chinese learning, with a large sample of 507 foreign students in China speaking 45 languages as their L1. Multiple linear regression model showed that language distance was negatively correlated with their Chinese language test scores, and such correlation was significant. It follows therefore, that the longer the distance between L1 and the target language, the more difficult to learn that language well.

Yet, does such effect also hold when it comes to language processing activities such as interpreting? Early attempts of the interpreting research community partly answered this question in the 1970s, when Seleskovitch (1976) first distinguished transcoding (i.e., word-for-word direct transposition) from interpreting proper (i.e., translating the meaning of the source text regardless of the linguistic form). Later discussions on this topic soon evolved into the debate over form-based translation versus meaning-based translation. Dam (1998) conducted the first quantitative study to probe the distribution of form-based and meaning-based outputs in a Spanish-Danish consecutive interpreting corpus. The rationale of her research was, if there were more lexical similarities (between the source and target texts) than dissimilarities, then the interpreting was more form-based than meaning-based. An analysis of 482 interpreted segments by professional interpreters showed that similarity, rather than dissimilarity, dominated the interpreting outputs, leading the author to conclude that consecutive interpreting is not necessarily based on a non-verbal representation of the source text. This pioneering research, albeit offering valuable insights, was far from scientific in terms of study design, and provided very little information to disentangle the relation between language distance and interpreting.

A more relevant investigation came from Wang and Gu (2016), in which they examined the impact of language structural asymmetry on English-Chinese simultaneous interpreting performance. To that end, they analyzed the performance of three professional interpreters when they were interpreting a live speech on TV, with a focus on such indexes as pause, information loss, misinterpreting and coping strategies. Results showed that the adoption of waiting as a coping strategy occasioned much more problems than that of segmentation (31 versus 4), manifested as unnatural long pauses and information loss. Considering the fact that English is dominated by right-branching structures whereas Chinese features left-branching structures, the authors concluded that it was the language-pair specificity that forced interpreters to adopt waiting as a coping tactic. This research, despite its small sample and observational paradigm, eloquently proposed language-pair as an independent influencing factor for English-Chinese simultaneous interpreting, apart from the traditionally recognized factors such as interpreter's competence and cognitive load. Shortly after, a more dedicated corpus-driven study was implemented, exclusively focusing on non-fluencies in Chinese-English consecutive interpreting and potential language-pair related causes (Zou & Wang, 2023). A quantitative analysis was conducted on a 97,239-word English corpus produced by 7 professional Chinese interpreters, with pauses and self-repairs being categorized as the major indexes for analysis. Results revealed that syntactic differences between English and Chinese were the main cause for pauses and self-repairs in the interpreting corpus, which again confirmed language-pair specificity as an important independent influencing factor in interpreting. Expanding the validity of the language-pair specificity theory was another corpus-based investigation of English-Arabic simultaneous interpreting by Al Zahran (2021), which presumed that, when doing simultaneous interpreting, interpreters would be hugely influenced by the subject-verb-object (SVO) structure of English as they produced target speech in Arabic, which is verb-subject-object (VSO) structured. Zahran (ibid) analyzed three English speeches and their professional Arabic interpretations in real-life settings, finding that in all three interpreted speeches SVO-structured expressions took up at least 73% of total renderings, which was in

great violation of the Arabic grammar. Conclusion of the study was that language-pair specificity compelled interpreters to follow structures in source language, and that such phenomenon was an indication for form-based interpreting. Summarizing the three aforementioned corpus-driven research, we might reach the conclusion that language-pair specificity does pose certain challenges for interpreters, but the details of such influence need further explorations, especially on-site empirical investigations.

So far the only on-site experiment on language-pair specificity was done by Korpál (2012). With the aim to examine the effect that morphosyntactic features of the source language might have on the sight interpreting process, Korpál (ibid) recruited two groups of Polish trainee interpreters (15 in total) and implemented a comparative experiment, i.e., asking 7 of them to interpret an English speech into Polish (E-P group) and the other 8 to interpret the German version of the same speech into Polish (G-P group). The researcher hypothesized that, due to complex morphosyntactic structures of the German language, the G-P group should have more and longer fixations, and a lower overall score compared with the E-P group. The fixation results however, were quite surprising in that only 1 out of 6 *t*-tests reached the level of significance, meaning that participants' processing of German verb-final structures was not always significantly harder than that of English SVO structures. In terms of the general scores of sight interpreting, E-P group fared significantly better than G-P group. Korpál (ibid) concluded that, considering the non-significant eye-tracking results, the group-wise discrepancy of interpreting performance should be attributed to the differentiated interpreting skills of the participants, rather than language-pair specificity or other objective language-related characteristics.

Though focusing on a limited sample, the research of Korpál (2012) was in a sense groundbreaking because it for the first time implemented a contrastive experimenting paradigm to explore language distance factor in interpreting, transcending the corpus-driven tradition in this research avenue. Moreover, it expanded the subject matter from simultaneous and consecutive interpreting to sight interpreting. As to the results however, it would be more plausible to rationalize the discrepancy from the following two aspects: the sample size and sample characteristics. For one thing, the small sample size may have biased the statistical results, for another, the trainee interpreters might be a confounding group to explain the results compared with experienced professional interpreters. In other words Korpál (2012) thought the differences of interpreting performance were associated with participants' interpreting skills rather than language-pair characteristics, but that may not be true. Why did the language-pair effect that was so manifest in previous corpus-driven studies fade out among novice interpreters?

The current research aims to answer such question. To begin with, we prefer to justify the result of Korpál (2012) from the perspective of language distance, for if interpreting skills take precedence over the language distance factor, then non-interpreter bilinguals should not be influenced by characteristics of the language per se and therefore should give similar interpreting performance when interpreting the same materials. To examine this, we implemented two experiments involving two groups of non-interpreter bilinguals. Our hypothesis is that, language distance is an objective and robust built-in influencing factor in interpreting, and that the closer the language distance, the easier to interpret that language into

L1, which should hold in bilinguals without any interpreting skills, just as it does in professional interpreters. But before the current study is done a piece of fact needs to be clarified, that is, there is no consensus until now on a quantified measurement of language distance. Criteria for gauging language distance vary from phonetic convergence (Kim et al., 2011) to orthographic transparency (Dong et al., 2021). Especially in the past decade, researchers from the field of natural language processing (NLP) also offered their insights on defining language distance from the angle of computer science (e.g., Gamallo et al., 2017; 2020). After a thorough examination of the existing literature from both fields, we would like to offer an operational definition of language distance tailored to the current research, i.e., the total amount of time spent in learning a language to the level of proficiency, which is similar to the measurement of Chiswick and Miller (2005). According to such definition, German-English and Chinese-English were selected as the close and distant language pairs respectively.

The current research consists of two experiments, i.e., an English-Chinese one-way consecutive interpreting experiment on Chinese-English bilinguals, and an English-German one-way consecutive interpreting experiment on German-English bilinguals. Both groups were asked to interpret exactly the same materials into their L1 such that no other confounding variable was included. Additionally, we analyzed the correlation between demographic and linguistic background of the participants and their interpreting performance, which is a progress compared to existing research. Note that our focus of analysis here is the macro-level comparison between language pairs, therefore we will not delve into the specific features such as lexical or syntactic differences either of the source and target languages, or of the interpreting produced by participants, which might be to the interest of a corpus-driven study.

2. Methods

2.1. Experiment 1

2.1.1. Participants

Participants were recruited using flyers posted around the University of Auckland city campus. To be eligible for experiment 1, any potential participant had to meet the following requirements: 1) born in a Chinese-speaking country or region and speaking Chinese since birth, thus native Chinese speakers; 2) have some English literacy; 3) aged between 18 and 40 years; 4) did not receive any form of interpreting training before; and 5) have normal hearing and speaking capabilities.

Altogether there were 46 Chinese-English bilingual participants eligible for and then involved in experiment 1. One participant did not finish the experiment and was thus excluded. The remaining 45 participants were used for the current data analysis. Participants were Chinese (L1) – English (L2) bilinguals with varying degrees of L2 proficiency. All participants gave written, informed consent and each received a 20-dollar grocery voucher as compensation for their time. The demographic details of the 45 participants are shown in table 1.

Table 1. *Demographic and language background for the initial 45 Chinese-English bilingual participants involved in Experiment 1*

Gender	N	Age (years) Mean (SD)	Age of L2 acquisition (years) Mean (SD)	Years learning L2 Mean (SD)	Stay in English- speaking countries (years) Mean (SD)
Male	23	27.00 (5.16)	8.88 (3.19)	17.92 (4.93)	2.67 (2.73)
Female	22	25.23 (3.37)	7.36 (3.11)	17.36 (4.04)	4.11 (4.59)
Total	45	26.15 (4.44)	8.15 (3.21)	17.65 (4.48)	3.36 (3.77)

2.1.2. Materials and Task Procedures

A written version of the Oxford Placement Test (<https://www.oxfordenglishtesting.com/>) was used for the testing of participants' English proficiency before the experiment. Then the experiment consisted of a custom-made English audio clip made up of ten groups of English sentences with varying length (11 to 33 words, mean=20.5, $SD=8.18$), played at a mean speed of 153 words per minute. All sentences (see Appendix) were in the simple declarative form so that syntactic complexity as a confound variable can be ruled out.

Before the experiment session, participants were given time to read through the Participant's Information Sheet, and then signed Consent Form on a voluntary basis. Then they were asked to complete the Oxford Placement Test (within 30 minutes) and a demographic questionnaire that mainly covered background information about their language learning (shown in Table 1). During the experiment participants were seated 70 cm before a computer screen, and were asked to listen to a short English sentence, after which the sounding of a beep would occur to remind them to interpret the sentence into Chinese, and then a second beep would ring to remind them to stop. The time duration for participants to interpret was 1.5 times the length of the source speech. Such procedure was repeated ten times for ten consecutive but different sentence groups. Participants were not allowed to take notes or use other tools to aid their memory, and there were no visual cues anywhere in the lab. The experiment was conducted using E-Prime 2.0 (Psychology Tools, Inc., Pittsburgh, PA, United States) such that the timing of each sentence was controlled at the level of milliseconds. Participants' verbal outputs were recorded with a voice recorder. To familiarize participants with these tasks a practice session was administered before the real experiment. The real experiment began after the participant was clear about the procedures and fully prepared. The whole experiment lasted for around 20 minutes. The experiment was implemented in accordance with the Declaration of Helsinki, and was approved by The University of Auckland Human Participants Ethics Committee (Reference number 022991).

2.1.3. Data analysis

Data were analyzed using IBM SPSS version 26 (SPSS Inc, Chicago, IL, USA), and all statistical analyses were applied with an alpha level for significance at $p < .05$. Participants' L2 proficiency was calculated as per the sample answer of the test, and their oral interpreting output transcribed and marked as the main behavioral index. The accuracy of sentence interpreting was obtained through dividing correctly interpreted sense groups by the total number of sense

groups in that sentence. This was repeated for each of the ten sentences and finally a mean score was obtained by averaging the accuracy scores across ten sentence groups for that particular participant. Marking was completed by a native Chinese speaker with a high level of English proficiency (IELTS overall score > 8). The average score of the 45 participants in the Oxford Quick English Placement test was 44.42 (full marks = 60), $SD = 7.17$. The mean accuracy of their interpreting performance was 54.76%, $SD = .15$. Pearson correlations were then performed on the interpreting accuracy and participants' language background, including their L2 proficiency, time spent living in English-speaking countries, age, age of L2 acquisition, and years of L2 learning.

2.1.4. Results

Pearson correlation results show that, there was a significant correlation between participants' L2 proficiency and their interpreting accuracy, $r(43) = .67, p < .01$. There was also a significant correlation between the time length that participants spent living in English-speaking countries and their interpreting performance, $r(43) = .47, p < .01$. There were no significant correlations between interpreting performance and age, age of L2 acquisition and years of L2 learning, all $ps > .05$. Partial correlation between L2 proficiency and interpreting accuracy controlling for time spent in English-speaking countries showed that participants' L2 proficiency was still significantly correlated with their interpreting accuracy, $r = .59, p < .001$. R square changed from 44.89% to 34.57% after controlling for the factor of time spent living in English-speaking countries.

2.2. Experiment 2

2.2.1. Participants

The participant screening and recruiting criteria for experiment 2 was exactly the same to those for experiment 1, except that the participants' L1 has to be German. As a result, 46 participants were recruited for the second experiment. All participants have successfully finished the experiment and produced usable data. Participants were German (L1) – English (L2) bilinguals with varied degrees of L2 proficiency. Participants gave written, informed consent and each received a 20-dollar grocery voucher as compensation for their time. The demographic details of the 46 participants are shown in table 2.

Table 2. Demographic and language background for the initial 46 German-English bilingual participants involved in Experiment 2

Gender	N	Age (years) Mean (SD)	Age of L2 acquisition (years) Mean (SD)	Years learning L2 Mean (SD)	Stay in English- speaking countries (years) Mean (SD)
Male	14	30.29 (8.21)	8.43 (4.38)	17.18 (8.94)	5.39 (4.71)
Female	32	27.53 (5.40)	8.06 (3.50)	14.16 (6.14)	4.95 (5.82)
Total	46	28.37 (6.42)	8.17 (3.74)	15.08 (7.15)	5.09 (5.46)

Materials and Task Procedures

Both materials and task procedures used in experiment 2 were identical to those in experiment 1, except that participants were asked to interpret the English sentences into German.

2.2.3. Data analysis

Data analysis was identical to that in experiment 1, except that participants' interpreting output was marked by a native German speaker who is highly proficient in English (IELTS score > 8).

2.2.4. Results

The average score of the 46 German-English bilingual participants in the Oxford Placement Test was 52.43 (full marks = 60), $SD = 6.52$. The mean accuracy of their interpreting performance was 64.29%, $SD = .09$. Pearson correlation results showed that, there was a significant correlation between participants' scores in the Oxford Placement Test and their accuracy of interpreting, $r(44) = .31, p < .05$. There was also a significant correlation between participants' time spent living in English-speaking countries and their interpreting accuracy, $r(44) = -.31, p < .05$. Partial correlation between L2 proficiency and interpreting accuracy controlling for time spent in English-speaking countries showed that participants' L2 proficiency was even more significantly correlated with their interpreting accuracy, $r = .54, p < .001$. R square changed from 9.80% to 28.83% after controlling for the factor of time spent living in English-speaking countries.

2.3 Comparative analysis and results

Apart from the two sessions of statistical analyses implemented within each group, we also compared the results of two experiments in a group-wise fashion. Specifically, participants' demographic indexes were compared using separate independent samples *t*-tests. This included participants' age when participating the research (Age), the age of L2 acquisition (AoA), the time spent living in English-speaking countries or regions (Exposure), and the time spent learning English (Years of learning). The results are shown in Figure 1. More importantly, scores in the Oxford Placement Test (Proficiency) and average accuracy in the interpreting task (Accuracy) were put into independent samples *t*-tests respectively. The comparison results between Chinese-English and German-English bilinguals are presented in Figure 2.

Figure 1 shows the comparison results of demographic data between Chinese-English and German-English bilinguals. On average, Chinese-English bilinguals spent significantly longer time learning English ($M = 17.65, SD = 4.48$) than German-English bilinguals ($M = 15.08, SD = 7.15$), $t(75.68) = 2.07, p < .05$. Other than this, the two groups did not significantly differ from each other in age, age of L2 acquisition and time spent living in English-speaking countries or regions, all $ps > .05$. Note that however, German participants were older than Chinese participants on average, and also lived longer in English-speaking countries than Chinese participants, although such differences were not significant.

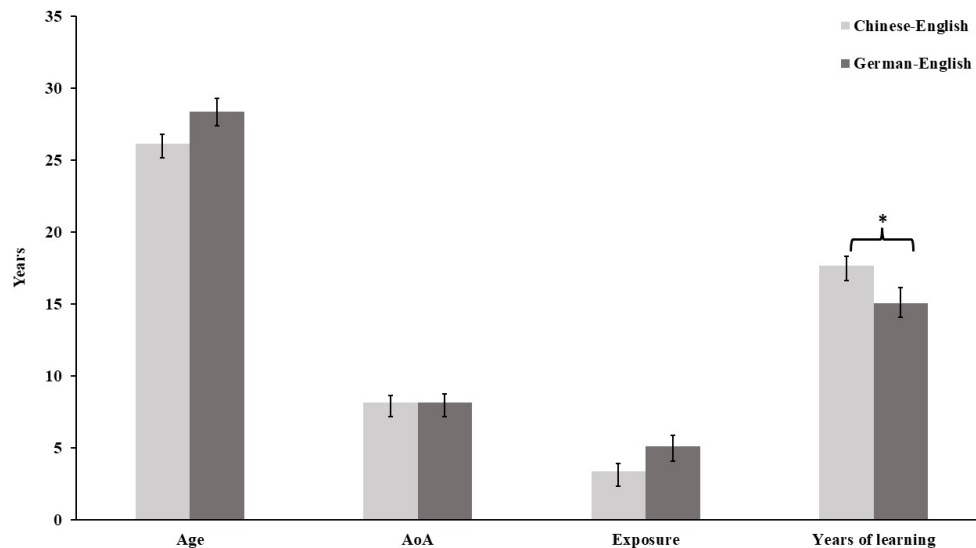


Fig. 1. Four independent samples *t*-test results of the demographic data between Chinese-English and German-English bilinguals. On the X-axis: age is the participants' age when participating; AoA is participants' age of L2 acquisition; Exposure is the time participants have spent living in English-speaking countries; Years of learning is how long participants have been learning English. Error bars represent standard errors and significant results are marked with asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

As shown in Figure 2, on average, German-English participants achieved significantly higher scores in the Oxford English Placement Test ($M = 87.39\%$, $SD = .11$) than Chinese-English participants ($M = 73.73\%$, $SD = .12$), $t(87) = 5.49$, $p < .001$. Moreover, the mean accuracy in the interpreting task was significantly lower for Chinese-English bilinguals ($M = 54.51\%$, $SD = .15$) than for German-English bilinguals ($M = 64.29\%$, $SD = .09$), $t(65.32) = -3.61$, $p < .001$.

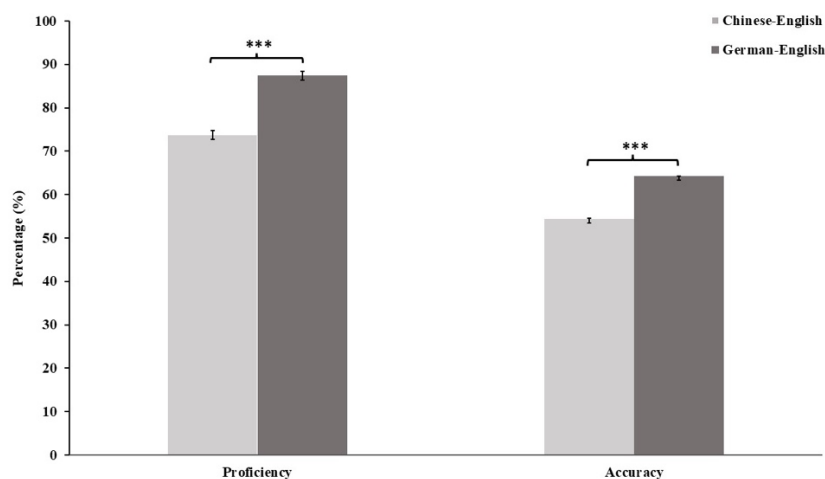


Fig. 2. Two independent samples *t*-test results of the behavioral data between Chinese-English and German-English bilinguals. On the X-axis: Proficiency is participants' score in Oxford Placement Test and Accuracy is their performance in the interpreting task. Error bars represent

standard errors and significant results are marked with asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

3. General discussions

To examine whether language distance is a robust influencing factor for interpreting performance, we recruited two groups of non-interpreter bilinguals to form distant-vs-close language pairs and implemented on them a consecutive interpreting test without note-taking. Both demographic information like age and language learning background information such as L2 proficiency were measured and integrated into the data analysis. To our knowledge this is the first empirical research that has compared the effects of close and distant language pairs on consecutive interpreting performance with a fairly large sample.

Our hypothesis was that interpreting performance should be more prone to the influence of language distance than to that of interpreting skills. The results confirmed our hypothesis. With exactly the same materials, German-English bilinguals fared significantly better than Chinese-English bilinguals in the task of interpreting English sentences into their first languages. It is well grounded to attribute such result to the language distance effect, as both groups of participants in our study were randomly sampled, devoid of any interpreting training, and did not significantly differ from each other in terms of age, age of English acquisition, and time of exposure to the English-speaking environment. In other words, if we randomly invite an adult German-English bilingual (<40 years old) to interpret English speeches into German, we would very likely obtain better results than that of inviting a Chinese-English counterpart to interpret English speeches into Chinese, because German is inherently closer to English than Chinese is. Adding to the evidence of language distance effect is the result showing that German-English bilinguals spent significantly less time learning English, but scored significantly higher than Chinese-English bilinguals in the English proficiency test. Such result is in agreement with the language assessment data. For example an English proficiency global ranking released by EF EPI shows that German is rated as high proficiency country with an English index of 598, whereas China is at low proficiency level with an index of 455 (<https://www.ef.com/assetscdn/WIBlwq6RdJvcD9bc8RMd/cefc.com-epi-site/reports/2024/ef-epi-2024-english.pdf>). Thus, the proficiency results strongly support the theory that other factors being the same, a distant language (e.g., English to Chinese) is harder to learn and master than a close language (e.g., English to German), which echoes the proposal of Ringbom and Jarvis (2009).

Apart from the aforementioned direct evidence for the language distance effect, a detailed analysis of the correlation results is even more revealing. In both groups participants' English proficiency was significantly correlated with their interpreting performance, i.e., the higher the participants' English proficiency, the higher their English-L1 consecutive interpreting accuracy. Such result is in agreement with our general impression and corroborates previous findings (e.g., Tzou et al., 2012). However, when it comes to the effect of time spent living in English-speaking countries on interpreting performance, the two groups showed opposite trends. In the Chinese-English bilingual group, the significant positive correlation indicates that the longer the participants live in English-speaking countries, the better their English-Chinese interpreting performance. But for German-English bilinguals, living in an English-speaking environment longer actually reduces their ability to interpret English speech into German. This

is counterintuitive and not in line with previous findings (e.g., Walczyński, 2017). The discrepancy might be accounted for in two ways. First, language exposure in previous research counted both the time spent on formal language learning domestically and the time spent living in a native speaking environment of the target language, but in our research we distinguished these two concepts and specified language exposure exclusively as the time spent living abroad in an English-speaking country. Such division brings about another broad and complex topic of the definition and measurement of, language exposure. While it is not the aim of the current study to have an in-depth discussion on language exposure, one thing that must be clarified is that exposure is not equal to input, and the same goes to the distinction between living in the native-speaking environment and receiving formal education or training (cf. Carroll, 2017). In fact, stay-abroad language exposure and formal institutionalized language training, though seemingly belonging to a common umbrella concept, have different impact on simultaneous interpreting (Kapranov, 2008) and necessitate a differentiation. The other reason is that prior studies reported the results of L1-L2 interpreting performance, which is opposite to our current study. This pertains to another issue of directionality of interpreting which is not the focus of the current analysis (see Lu et al., 2023 for a comprehensive review). One may naturally expect that living in the L2 environment for a long time helps the bilingual's expression of L2, thus L1-L2 interpreting. In the reverse way however, exposure to L2 may do harm to L1 expression when the two languages are very similar or close to each other (cf. Pavlenko, 2000). In our instance, it could be well justified that German speakers, when asked to express their ideas in German, are more prone to the influence of English, as a long time of living in an English-speaking country made them accustomed to the English way of expression, because the two languages are quite close and share many homographs, cognates or even grammar. In contrast, Chinese speakers are much less interfered by English even if they live in a native English speaking environment for long, because the two languages are polar opposite in terms of writing systems, pronunciation, grammar etc. Such differential cross-linguistic interference of English on German and Chinese respectively is another implication for the language distance effect in interpreting.

Lastly, the language distance effect might offer a novel perspective to reconcile the form-versus-meaning based interpreting debate. Instead of the form-or-meaning dichotomy, a possible scenario in L2-L1 interpreting could be like this: when the source and target languages are close and similar to each other and it is relatively easy for the interpreters to reach proficiency in L2, they would either intentionally or unconsciously adopt a form-based approach in interpreting as it is cognitively easy and alluring to borrow the words or even syntax from a similar language; when the two languages are distant or dissimilar however, the interpreters have to devote a large share of their cognitive resources into comprehending the speech, and are forced to produce meaning-based interpreting as they can not find homographs or any other shared linguistic units between L1 and L2.

4. Conclusions

Our empirical research compared the performance of 91 Chinese-English and German-English bilinguals in a task requiring them to consecutively interpret English sentences into their first language. Results revealed that German-English bilinguals performed significantly better than Chinese-English bilinguals. Considering the fact that our participants have not received any

training on interpreting, we conclude that the different interpreting performance was caused by the language distance effect. In other words, German is close to English, therefore interpreting English into German is relatively easy. In contrast, Chinese is distant from English, which adds to the difficulty of interpreting English into Chinese. The fact that German participants spent less time learning English but achieved much higher scores in the English test also demonstrates the language distance effect. Finally, the language distance effect might provide a common ground for the long debate over form-versus-meaning based interpreting.

Data availability statement:

Data are available upon request.

Funding

The current research was supported by Beijing Social Science Fund (23YYC016).

References

- Al Zahran, A. (2021). Structural challenges in English> Arabic simultaneous interpreting. *Translation & Interpreting: The International Journal of Translation and Interpreting Research*, 13(1), 51-70.
- Carroll, S. E. (2017). Exposure and input in bilingual development. *Bilingualism: Language and Cognition*, 20(1), 3-16.
- Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of multilingual and multicultural development*, 26(1), 1-11.
- Cui, M., Zhang, W., & Sun, T. (2018). 语言距离、母语差异与汉语习得：基于语言经济学的实证研究[Linguistic distance, language learners' mother tongue, and Chinese acquisition: An empirical study from the perspective of language economics]. *Chinese Teaching in the World*, 32(2), 280-288.
- Dam, H. V. (1998). Lexical similarity vs lexical dissimilarity in consecutive interpreting: A product-oriented study of form-based vs meaning-based interpreting. *The Translator*, 4(1), 49-68.
- Dong, J., Li, A., Chen, C., Qu, J., Jiang, N., Sun, Y., ... & Mei, L. (2021). Language distance in orthographic transparency affects cross-language pattern similarity between native and non-native languages. *Human Brain Mapping*, 42(4), 893-907.
- Gamallo, P., Pichel, J. R., & Alegria, I. (2017). From language identification to language distance. *Physica A: Statistical Mechanics and its Applications*, 484, 152-162.
- Gamallo, P., Pichel, J. R., & Alegria, I. (2020). Measuring language distance of isolated European Languages. *Information*, 11(4), 181.
- Gile, D. (1992). Predictable sentence endings in Japanese and conference interpretation. *The Interpreters' Newsletter*(Special Issue No. 1).

- Kapranov, O. (2008). *The impact of language exposure on fluency in simultaneous interpreting* (Doctoral dissertation, University of Western Australia).
- Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology 2* (1), 125-156.
- Korpál, P. (2012). On language-pair specificity in sight translation: An eye-tracking study. *Übersetzen in die Zukunft: Tagungsband der 2. Internationalen Fachkonferenz des Bundesverbandes der Dolmetscher und Übersetzer eV (BDÜ)*, 522-530.
- Lu, R., Abdullah, M. A. R., & Ang, L. H. (2023). Into-A or into-B, that is a question: A systematic literature review of directionality and performance in consecutive interpreting. *Sage Open*, 13(4), 21582440231202294.
- Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in applied linguistics*, 11(2).
- Renfrew, C., McMahon, A., & Trask, L. (2000). *Time depth in historical linguistics*. Cambridge: McDonald Institute for Archaeological Research.
- Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign language learning. *The handbook of language teaching*, 106-118.
- Seleskovitch, D. (1976). Traduire: de l'expérience aux concepts. *Études de linguistique appliquée*, 24, 64.
- Tzou, Y. Z., Eslami, Z. R., Chen, H. C., & Vaid, J. (2012). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers. *International Journal of Bilingualism*, 16(2), 213-227.
- Walczyński, M. (2017). The length of language exposure and its influence on individual factors experienced by interpreting trainees during an in-class consecutive interpreting test. In Walczyński Marcin (Ed.), *Selected Aspects of Interpreting in the 21st Century* (pp. 91-105). Wydawnictwo Uniwersytetu Wrocławskiego
- Wang, B., & Gu, Y. (2016). An evidence-based exploration into the effect of language-pair specificity in English-Chinese simultaneous interpreting. *Asia Pacific Translation and Intercultural Studies*, 3(2), 146-160.
- Zou, B., & Wang, B. (2023). Non-fluency and language-pair specificity in Chinese-English consecutive interpreting: A corpus-driven study. *Research in Corpus Linguistics*, 11(2), 30-49.

Appendix

Sentence stimuli used in this study:

S1: The United Kingdom research councils are establishing their first overseas office in Beijing.

S2: Over the next several months, let us do what Americans have always done, and build a better world for our children and our grandchildren.

S3: America's economy is the fastest growing of any major industrialized nation.

S4: To make our economy stronger and more dynamic, we must prepare a rising generation to fill the jobs of the twenty first century.

S5: And I do not mistrust the future, I do not fear what is ahead, for our problems are large, but our heart is larger, our challenges are great, but our will is greater.

S6: To make our economy stronger and more productive, we must make healthcare more affordable.

S7: Baseball is the most popular sport in the United States. It is played throughout the spring and summer, and professional baseball teams play well into the fall.

S8: The year was 1982. The IBM personal computer had only been born the year before.

S9: In 1942, only one out of every 25 high schools in America offers science courses.

S10: With a healthy growing economy, with more Americans going back to work, with our nation and active force for good in the world, the state of our union is confident and strong.