Behavior Research Methods (2024) 56:5732–5753
https://doi.org/10.3758/s13428-023-02305-z

ORIGINAL MANUSCRIPT

The Chinese lexicon of deaf readers: A database of character decisions and a comparison between deaf and hearing readers

Philip Thierfelder[1] · Zhenguang G. Cai[1] · Shuting Huang[1] · Hao Lin[2]

Accepted: 22 November 2023 / Published online: 20 December 2023
© The Psychonomic Society, Inc. 2023

Abstract

We present a psycholinguistic study investigating lexical effects on simplified Chinese character recognition by deaf readers. Prior research suggests that deaf readers exhibit efficient orthographic processing and decreased reliance on speech-based phonology in word recognition compared to hearing readers. In this large-scale character decision study (25 participants, each evaluating 2500 real characters and 2500 pseudo-characters), we analyzed various factors influencing character recognition accuracy and speed in deaf readers. Deaf participants demonstrated greater accuracy and faster recognition when characters were more frequent, were acquired earlier, had more strokes, displayed higher orthographic complexity, were more imageable in reference, or were less concrete in reference. Comparison with a previous study of hearing readers revealed that the facilitative effect of frequency on character decision accuracy was stronger for deaf readers than hearing readers. The effect of orthographic-phonological regularity differed significantly for the two groups, indicating that deaf readers rely more on orthographic structure and less on phonological information during character recognition. Notably, increased stroke counts (i.e., higher orthographic complexity) hindered hearing readers but facilitated recognition processes in deaf readers, suggesting that deaf readers excel at recognizing characters based on orthographic structure. The database generated from this large-scale character decision study offers a valuable resource for further research and practical applications in deaf education and literacy.

Keywords Deafness · Lexicon project · Chinese · Character recognition · Lexical variables

Introduction

As users of a visual-manual language who do not have full access to the phonology of spoken language, deaf signers are distinct from the majority hearing population in terms of linguistic experience. Research has shown that this has an impact on how they process written language (see Emmorey & Lee, 2021, for overview). However, most research on written language processing in deaf signers has focused on reading in alphabetic languages, which leaves us with less understanding of how deaf signers process nonalphabetic languages, such as Chinese. Understanding the unique characteristics of how deaf signers recognize Chinese words/characters can help us understand how experience with a visual-manual language may uniquely influence the way deaf individuals perceive and process visual information. In this paper, we report a large-scale investigation of Chinese character recognition by deaf readers. We examined the effects of an array of character-level lexical variables on character recognition processes by deaf readers and compared these patterns to those found in hearing readers (using data from another study, Sze et al., 2014) in an attempt to reveal how lexical variables differentially impact visual word recognition by deaf and hearing readers.

Philip Thierfelder and Zhenguang G. Cai are Co-first authors.

✉ Zhenguang G. Cai
zhenguangcai@cuhk.edu.hk

✉ Hao Lin
linhao@shisu.edu.cn

Philip Thierfelder
dpthierfelder@cuhk.edu.hk

[1] Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Sha Tin, N.T., Hong Kong, SAR

[2] Shanghai International Studies University, 550 Dalian Road(W), Shanghai, People's Republic of China

Chinese characters

Written Chinese is the most widely used nonalphabetic script. In Chinese, characters are the smallest freestanding written forms that carry meaning. Chinese characters can either stand as words on their own (e.g., 清, *qing*$_1$, meaning "clear") or combine with other characters to form multi-character words (e.g., 清风, *qing*$_1$*feng*$_1$, meaning "breeze"). They contain one or more constituent radicals, which are made from strokes (e.g., 氵 is made with the three strokes, 丶丶一), spatially arranged in a particular layout, normally left-to-right (e.g., 清 consisting of 氵 on the left and 青on the right) or top-down (e.g., 箐, with 竹 on top of 青). Sometimes these radicals can signal the pronunciation or meaning of the character. For instance, the character 清 has a semantic radical 氵 (meaning "water") and a phonetic radical (青, which can also function as a character pronounced as *qing*$_1$).

Unlike words in alphabetic scripts, such as written English, Chinese characters cannot be divided into phonemic units. For example, no component in the character 清 (*qing*$_1$, or / tɕʰiŋ/ in IPA, meaning "clear") individually maps to the phonemes /tɕʰ/, /i/, or /ŋ/. Rather, each character represents a full syllable. While some characters do contain phonetic radicals that provide cues about their pronunciation, these cues are generally unreliable. For example, many characters that contain the radical 青 (*qing*$_1$, "green) are pronounced *qing* (e.g., 请, *qing*$_3$, "to invite"; 情, *qing*$_2$, "emotion"), but this is not always the case (e.g., 猜, *cai*$_1$, "guess"). Only around 26% of Chinese characters are said to be "regular" in that their pronunciations are identical (regardless of tone) to their phonetic radical (Gao et al., 1993). Because written Chinese has distinctive properties which make it so different from alphabetic scripts, research on Chinese reading provides an important point of contrast with that conducted on the reading of alphabetic scripts.

Studies of character processing

Psycholinguists have long been interested in the cognitive processes that underlie word recognition during reading. One approach to understanding the intricacies of lexical processing is conducting large-scale studies that collect behavioral response data for comprehensive sets of words or characters. This method has been employed to examine word processing in various languages, including Chinese (Sze et al., 2014; Wang et al., 2020), Dutch (Keuleers et al., 2010), English (Balota et al., 2007), and French (Ferrand et al., 2010), and is advantageous in that it allows researchers to look at interactions between response measures (e.g., response times and accuracy rates) and a wide range of predictors (e.g., age of acquisition/AoA, orthographic complexity, and homophone density) across an extensive, representative sample of the lexicon.

Sze et al. (2014) carried out the first large-scale lexical (character) decision study on Chinese character recognition, gathering behavioral responses for 2500 individual Chinese characters from a sample of 35 native speakers of Mandarin Chinese from mainland China. Using this database, they were able to replicate the results of previous studies (see Chen et al., 2009a; Leong et al., 1987; Peng et al., 2003), finding that (1) Chinese characters with more strokes (i.e., greater orthographic complexity) took readers longer to recognize; (2) characters with earlier AoAs were responded to more quickly and accurately; and (3) response times were faster for low-frequency characters that had more meanings.

Developing the first extensive naming database for traditional Chinese characters, Chang et al. (2016) collected naming reaction times (RTs) from 140 participants for a total of 3314 characters. The regression analyses revealed that characters with higher frequency, consistency, regularity, and familiarity were processed more quickly, as evidenced by reduced RTs. Furthermore, characters with multiple meanings or those used in a greater number of disyllabic compound words also tended to have shorter naming times. Homophone density, however, did not significantly influence naming RTs.

Tse et al. (2017) conducted an extensive lexical (word) decision study, collecting RTs for over 25,000 traditional Chinese two-character compound words from 33 native Cantonese speakers. The research explored the impact of several lexical variables, including character and word frequencies, stroke numbers, and semantic transparency. The results demonstrated that higher-frequency words and those with greater semantic transparency yielded faster RTs. Furthermore, the study revealed that the frequency and semantic transparency of the first character in compounds were slightly stronger predictors than those of the second character. Stroke numbers, often used as an indicator of visual complexity, affected character recognition, with increased strokes leading to slower lexical decisions.

Focusing on written production, Wang et al. (2020) conducted a large-scale study using a spelling-to-dictation task in which participants were directed to write individual Chinese characters as dictated by spoken phrases. They collected data for 1600 Chinese characters from a sample of 204 participants, each of whom produced 200 handwritten characters for the study. They measured accuracy as well as writing latencies and durations. They found that when character frequency and familiarity of the context word (used within the spoken elicitation phrase) were higher, participants were quicker and more accurate in producing the target character. Similarly, early-acquired characters (i.e., those with lower AoAs) were also produced more quickly and accurately. Imageability correlated with writing accuracy but did not affect writing latencies or durations. Greater orthographic complexity, as measured by stroke number, led to participants taking longer to initiate and complete writing and to a decrease in accuracy. Character composition also affected the speed of orthographic access, with

writing latencies and durations being shorter for characters of a left-to-right structure than for characters of other structures. Phonology was also found to play a role in orthographic access, as evidenced by the finding that characters with a higher number of homophones took longer for participants to begin writing. Furthermore, orthographic regularity facilitated orthographic access, such that characters with phonetic radicals that more closely reflected their pronunciation had shorter writing latencies.

The large-scale studies just reviewed demonstrate that lexical access during the recognition and production of Chinese characters is influenced by a range of factors including frequency, AoA, imageability, orthographic complexity, homophone density, and orthographic-phonological regularity. Thus far, no large-scale studies of Chinese character processing have been conducted with deaf individuals who use Chinese as their primary written language. Because deaf users of sign language, as we will see in the following sections, have been shown to be unique in terms of visual cognition and written word processing, we expect that they will diverge from hearing readers in certain respects with regard to how their character recognition processes are influenced by lexical variables.

Deaf signers and mental imagery

Sign languages differ from spoken languages in that sign languages make extensive use of space to express grammatical relations, and many signs exhibit a high degree of iconicity. Because of this, researchers have investigated whether deaf signers differ in their processing of mental visual imagery. In a study by Emmorey et al. (1993), the authors found that both deaf and hearing signers, compared to hearing non-signers, had an enhanced ability to mentally generate and rotate visual imagery. Emmorey and Corina (1993) found that deaf signers show a right hemisphere advantage when processing imageable signs, though only half of these signs were iconic, suggesting that it was imageability, rather than iconicity, that elicited the laterality effect. This right hemisphere advantage was also found for deaf signers in image generation (Emmorey & Kosslyn, 1996), which may indicate that deaf signers' enhanced mental image generation abilities may be linked to a greater allocation of linguistic processing to the right hemisphere. Imageability has been found to facilitate faster responses in lexical decision (Sze et al., 2015) and naming (Liu et al., 2007) and also increase accuracy in character writing (Wang et al., 2020) in non-signing hearing speakers of Chinese. The facilitatory effect of imageability in character recognition may be greater in deaf individuals given the aforementioned advantages they have in processing mental imagery.

Visual word recognition in deaf readers

Deaf individuals appear to exhibit some degree of enhanced peripheral processing (Bosworth & Dobkins, 2002; Codina et al., 2017; Proksch & Bavelier, 2002). Reading researchers have investigated whether deaf readers may have similarly enhanced perception in parafoveal word processing (i.e., the processing of words within the parafoveal range of vision, which generally encompasses one to two words beyond the fixated word). Indeed, deaf readers of Chinese and English have been shown to have a wider perceptual span than hearing readers and are thus able to process more words in a single fixation (Bélanger, Slattery, et al., 2012b; Liu et al., 2021).

Deaf readers' wider perceptual span allows them to process parafoveal orthographic information more efficiently. An eye-tracking study by Bélanger et al. (2013) using a gaze-contingent boundary paradigm found that deaf readers re-fixated on target words less often than hearing readers when those targets had orthographically related preview words, suggesting that the deaf readers were able to process orthographic information in the parafoveal region with greater efficiency than the hearing readers. In a study on Chinese reading that also used a gaze-contingent boundary paradigm, Yan et al. (2015) found that semantic and orthographic previews elicited strong and early-emerging effects, suggesting that deaf readers process orthographic information more efficiently and have stronger ortho-semantic connections relative to reading-level matched hearing readers. Some researchers have proposed that these stronger ortho-semantic connections are a result of deaf readers' reduced reliance on phonology (Gutierrez-Sigut et al., 2019; Morford et al., 2017). There is some evidence that enhanced parafoveal processing may be linked to linguistic experience in sign language, as Thierfelder, Wigglesworth, et al. (2020b) found that deaf readers who acquired sign language earlier in life were also more sensitive to the presence of parafoveal previews.

Speech-based phonological activation in deaf readers

Deaf readers have often been observed to fall behind their hearing peers in literacy skills (Kelly & Barac-Cikoja, 2007; Traxler, 2000), and researchers have debated whether this is due to a lack of phonological coding abilities (Allen et al., 2009; Paul et al., 2009; Perfetti & Sandak, 2000; Wang et al., 2008). A meta-analysis by Mayberry et al. (2011) found that phonological coding abilities had a moderate effect, accounting for about 11% of the variance in deaf individuals' reading ability. However, overall ability in a spoken or signed language was the most significant factor, accounting for 35% of the variance. In a review of the literature, Emmorey

and Lee (2021) argued that skilled deaf readers develop strengthened pathways between orthography and semantics, allowing them to bypass phonological mediation and efficiently access word meanings directly via orthography, and that deaf people exhibit altered neural organization of the brain regions involved in word recognition. Cates et al. (2022) examined reading comprehension among adult deaf and hearing readers of English, focusing on the impact of language variables, cognitive factors, and language experience. Their findings revealed that for monolingual speakers of English and non-native signers, both working memory and phonological awareness contributed significantly to reading comprehension. However, the role of phonological awareness varied across groups: while it served as a significant predictor for the monolingual and non-native signing groups, it did not have the same predictive power for native signers or hearing Chinese-English bilinguals.

Psycholinguistic researchers have looked into the role of phonological code in deaf reading processing using on-line experiments. Results from studies on deaf readers of alphabetic scripts have been mixed. For example, Gutierrez-Sigut et al. (2017) found evidence of speech-based phonological activation, while Bélanger, Baum, et al. (2012a) and Bélanger et al. (2013) found none. Eye-tracking studies focused on deaf readers of Chinese have found some evidence of limited use of speech-based phonological code. Yan et al. (2020) and Yan et al. (2015) found evidence of late phonological activation in deaf readers of Chinese but only among those with higher reading fluency levels. Thierfelder, Wigglesworth, et al. (2020a) found that deaf readers of Chinese from Hong Kong relied primarily on orthography to directly activate word meanings, but they also found evidence that they could activate speech-based phonological representations when contextual predictability was high. It should be noted that, in contrast, phonological effects in hearing readers of Chinese have been found to emerge consistently in later processing measures regardless of factors like reading ability and contextual predictability (Feng et al., 2001; Thierfelder, Durantin, et al., 2020; Thierfelder, Wigglesworth, et al., 2020a; Yan et al., 2015).

In a structural priming study using written Chinese, Cai et al. (2022b) found that orthographic, but not phonological, identities of target words in sentence primes led to a boost in structural priming among deaf participants. In contrast, lexical boost was only enhanced for hearing participants when primes and targets shared a phonological identity (i.e., they had the same pronunciation). The authors concluded that deaf writers are particularly sensitive to orthographic information in text and retrieve lexical items based on their orthographic forms, while hearing readers do so based primarily on their phonological forms. Based on a series of experiments investigating phonological activation in deaf readers of English, Rowley (2018) argued that orthographic-semantic and orthographic-phonological connections function similarly in deaf and hearing readers, but connections between semantics and phonology are weaker in deaf readers.

From the studies on Chinese reading processing, we can see that deaf readers of Chinese rely primarily on orthography for accessing word meanings during silent reading and written production and may have enhanced orthographic processing abilities. As such, we expect that lexical variables related to speech-based phonology (i.e., homophone density and regularity) will have little or no influence on lexical decision accuracy and RTs for deaf participants in the current study.

The present study

Large-scale studies of Chinese character recognition and writing have been conducted to investigate lexical access processes in hearing speakers of Chinese (Chang et al., 2016; Sze et al., 2014, 2015; Tse et al., 2017; Wang et al., 2020). These studies collected behavioral responses for a large number of target characters, allowing researchers to analyze the influence of various lexical variables (e.g., frequency, AoA, etc.) on lexical access processes using comprehensive sets of Chinese characters. However, no large-scale studies of Chinese character recognition have investigated lexical processing in deaf readers, who have been shown to be distinct from hearing readers in how they process orthographic information (and visual information more generally). To gain deeper insight into how Chinese character recognition may differ between deaf and hearing readers, we carried out a large-scale character decision study (using 2500 real characters and 2500 pseudo-characters) with deaf participants from mainland China.

Method

Participants

A total of 26 deaf readers from Shanghai in mainland China took part in the experiment. One of them was excluded for failing to complete the whole experiment. The remaining 25 deaf readers (10 male, 15 female) had a mean age of 30.44 years (ranging from 24 to 37). All had attended deaf schools starting in primary school. Twenty-one had an undergraduate degree or above. Of the remaining four, one had an associate's degree and three had diplomas. Self-reports indicated that all participants became deaf before 3 years of age, were users of Chinese Sign Language (CSL), and were not cochlear implant users. All participants reported being right-handed and having normal or corrected-to-normal vision.

We assessed Chinese language proficiency of the participants through two measures. First, we administered the

LexCHI (Wen et al., 2023), an untimed lexical decision test which was developed to measure Chinese proficiency among both native and non-native speakers. Despite being primarily a vocabulary knowledge test, test performance has been shown to significantly correlate with second-language (L2) performance particularly in translation and cloze tasks. Furthermore, it can serve as a useful screening tool to assess whether an individual possesses a native or near-native level of Chinese proficiency, with a score of 70 or more being indicative of native-level Chinese proficiency. The scores of the deaf participants in our study ranged from 83 to 100, with a mean score of 96.5 ($SD = 4.2$). This suggests that the level of Chinese proficiency, at least in terms of word knowledge, among all deaf participants is comparable to that of typical hearing adults in China.

We used the Chinese Author Recognition Test, identical to the one used by Sze et al. (2014), as a metric for text exposure. The deaf participants in our study had an average score of 13.2 ($SD = 9.9$). This is below the 17.7 average ($SD = 5.6$) achieved by the hearing participants in Sze et al. (2014) but aligns more closely with the 11.4 average score from the original norming participants in Sze et al.'s study. Our results suggest that our deaf participants' exposure to Chinese text is comparable to that of a typical native Chinese speaker from mainland China, albeit lower than the hearing participants in Sze et al.'s study.

Stimuli

The experimental materials consisted of 2500 simplified Chinese characters from Sze et al. (2014) and 2500 pseudo-characters produced using the method described by Sze et al. (2014). All stimuli, together with data and analytical codes, are available at the Open Science Framework (https://osf.io/qthfm/).

Procedure

The experiment was conducted using E-Prime 2.0, and each participant was tested individually in a quiet room. Following Sze et al. (2014), each participant was tested in three blocks (1700, 1700, and 1600 trials, respectively) on separate days. The number of real characters was equal to the number of pseudo-characters in each block. Before the main experiment, an extra 40 trials (half real characters and half pseudo-characters) were provided for practice in each block. The blocks were counterbalanced, and the trials within each block were also randomized.

Each trial began with a fixation cross at the center of the monitor for 500 ms, followed by a blank screen for 120 ms, then a character or pseudo-character in 36-point Song font. Participants were asked to judge whether it was a character or a pseudo-character by pressing [”] or [A]. They were encouraged to respond as quickly as possible, but not at the cost of accuracy. The character or pseudo-character remained onscreen until the participant made a response. If they made an incorrect response, a red cross (×) appeared for 1000 ms. There was no feedback for a correct response. The time interval between each trial was 1000 ms. Participants were asked to rest for 2 minutes after every 100 trials. Our procedure differed from Sze et al. (2014) only in the feedback program. Sze et al. (2014) used a cue sound to indicate incorrect responses, while we used a visually salient red cross.

Lexical variables

Character frequency Frequencies for 2482 characters were obtained from the SUBTLEX-CHR corpus (Cai & Brysbaert, 2010), which has been observed to be the strongest frequency measure against the other six frequency measures in Chinese (Sze et al., 2014). Frequencies for the remaining 18 characters were estimated using the following steps: (1) we obtained frequency information for each of these 18 characters from the *Character Frequency List of Modern Chinese* (Da, 2004); (2) for each of these, 10 characters of a similar frequency were selected from the frequency list; (3) we obtained frequencies for these characters using the SUBTLEX-CHR corpus and calculated their average. These averages were used as the frequency estimations for each of the aforementioned 18 characters.

Age of acquisition (AoA) Cai et al. (2022b) developed two objective AoA norms (the 2001 norm and the 2011 norm) for 3300+ characters in simplified Chinese according to the time in which the character is formally learned in two sets of leading textbooks of Chinese in compulsory education, published respectively on the basis of the 2001 and 2011 national curriculum. They showed that the two objective AoA norms outperformed previous AoA norms when used to predict lexical processing in large-scale databases. For the 2500 characters used in the current study, we found AoAs for 2191 of them in the 2011 norm and for a further 99 characters in the 2001 norm.

In our study, we estimated the AoA for 210 characters based on subjective ratings from 20 participants. Each participant was asked to recall the school grade in which they learned each character, rating them on a scale from 1 to 10 (1 for grade 1, 9 for grade 9, and 10 for any grade after grade 9). To familiarize participants with this task, we first conducted a practice session. During this session, we presented four characters typically learned in each grade (except for grade 6, as no characters were associated with this grade). Participants were then tasked with categorizing each target character into a grade. They were instructed to do this based on whether they believed the character was learned at the

same time as the four characters previously shown from that grade. In the main test, participants were asked to assign a grade level to each character according to when they remembered learning it. We calculated the mean rating for each of 210 characters and transformed these means into the AoA format of the 2011 norm (e.g., a rated mean of 1.4 would be 6.5 in terms of AoA).

Imageability and concreteness Imageability and concreteness measures for 1143 characters were obtained from the Chinese handwriting database (Wang et al., 2020). Then we recruited 51 participants from the same populations as those in Wang et al. (2020) and asked them to rate the imageability and concreteness of the remaining 1353 characters. Each participant was asked to rate the imageability of 400 characters randomly selected from the 1353 characters, and the concreteness of another 400 characters randomly selected from the 1353 characters. The order of these two tasks was counterbalanced. Each trial began with a character at the center of the screen, and the participant was asked to rate imageability/concreteness on a seven-point scale (1 = least imageable/concrete; 7 = most imageable/concrete). To eliminate participant differences we transformed ratings into *z*-scores for each participant.

Orthographic-phonological regularity We obtained regularity information for 1143 of the 2500 characters from Wang et al. (2020). Fifty-one participants from the same populations as those in Wang et al. were recruited, and each of them was asked to rate the orthographic-phonological regularity of 400 characters randomly sampled from the remaining 1353 characters. In each trial, a character was presented at the center of the screen, and the participant was asked to rate the orthographic-phonological regularity of the character according to the degree to which the pronunciation of the character matched its phonetic radical (0 = the character has no sound radical, 7 = pronunciations of the character and its phonetic radical are identical).

Homophone density Following Wang et al. (2020), we calculated the number of characters identical in pronunciation for each pronunciation using the SUBTLEX-CHR corpus (Cai & Brysbaert, 2010). We then determined the pronunciations of all 2500 characters and found the number of homophonous characters for each character according to its pronunciation. For characters with multiple pronunciations, the number of homophonous characters referred to the sum of homophonous characters for all pronunciations. Finally, the log of the number of homophonous characters was referred to as the measure of homophone density.

Number of strokes Every character is constructed with a series of strokes in a specific order, and the number of strokes required to write a character indicates its degree of visual complexity. In the present lexicon project, the stroke number of characters ranged from 2 (e.g., 刁, diao1) to 24 (e.g., 矗, chu4). Descriptive statistics for each of the lexical variables are presented in Table 1.

Table 1 Descriptive statistics for all variables

Variable	Mean	Min	Max	SD
Age of acquisition	9.02	6.50	15.00	2.03
Log frequency	3.13	0.00	6.31	1.02
Number of strokes	9.68	2.00	24.00	3.34
Log homophone density	0.93	0.00	1.86	0.35
Regularity ratings	2.70	1.00	7.00	2.10
Imageability ratings	3.60	1.00	7.00	1.17
Concreteness ratings	3.77	1.00	7.00	1.27

Data analysis

We calculated the accuracy rate and RT for each real character (collapsed over participants), removing all pseudo-characters from further analyses. As the accuracy rate for each character is bound between 0 and 1, and is thus not ideal for parametric tests, following Donnelly and Verkuilen (2017), we transformed each character's accuracy rate into an empirical logit using the following formula, where y is the accuracy rate for a character and n is the number of observations (i.e., 25 here) used to calculate y:

$$elogit\ (y) = \log\left(\frac{y + 1/2n}{1 - y + 1/2n}\right)$$

To calculate the mean RT for a character, we first discarded incorrect responses and then removed RTs faster than 200 ms or slower than 3000 ms as outliers. Following Sze et al. (2014), we then transformed the raw RTs into *z*-scores within each participant. RTs whose *z*-scores were either below −2.5 or above 2.5 were then removed. A total of 12.33% of the RTs were excluded in these trimming steps. We then calculated the mean RTs for each character. Because none of the deaf readers provided a correct response for two real characters, 褫 and 谖, these two characters did not have log RTs and were excluded from the RT analyses (similarly, none of the hearing participants in Sze et al., 2014, provided a correct response for the character 谖). As the mean RTs were right-tailed (as in most RT data), we then took the natural log of the mean RT for each character. Thus, our analyses were performed on the empirical logit of each character in the accuracy analysis and on the log RT of each character in the RT analysis.

While we utilized the RTs and accuracy rates for each target character from the hearing group data provided by Sze et al. (2014), our analysis diverges from their approach in two significant ways. Firstly, we applied transformations to the data, converting the RTs and accuracy rates to their logarithmic and empirical logit forms (as explained above), respectively. Secondly, instead of using the predictors and predictor values employed by Sze et al. (2014), we utilized our own set. These are detailed in the "Lexical variables" section above. Therefore, to ensure consistency and comparability, we reran the original analysis of Sze et al.'s (2014) data using our adapted methodology.

We employed regression analyses using the mgcv package (Wood & Wood, 2015) within the R statistical environment (R Core Team, 2023). Our dependent variables were empirical logit-transformed accuracy rates and log-transformed RTs. Separate models were built for each group and response variable. In the full models, which included the data from both the hearing and deaf groups, we included main effects for group and the lexical variables. Additionally, we modeled interaction effects between group and each lexical variable, which allowed us to investigate how the relationship between the lexical variables and the response variables differed between the hearing and deaf groups. As visual examination of our data indicated potential nonlinear relationships between predictors and response variables, following recommendations from Baayen et al. (2006), we addressed this by incorporating restricted cubic splines into our models. These splines, defined by "knots" or change points, allow for the creation of flexible functions, known as smooth terms, which adapt to fit different patterns across predictor variables.

We initially built generalized additive models using only linear terms (effectively equivalent to generalized linear models/GLMs) and then progressively introduced smooth terms for each of the seven lexical variables under examination, starting with frequency. The decision to include a smooth term in the final model was guided by comparisons using the Akaike information criterion (AIC; Akaike, 1998). If the model's fit improved with the inclusion of a smooth term as indicated by a lower AIC, we retained the smooth term in the final model. Conversely, if the model's fit did not improve with a smooth term, we retained only the linear term for that particular variable. We limited the smooth terms in our models to three knots, as increasing the number of knots beyond three for certain variables introduced collinearity issues.

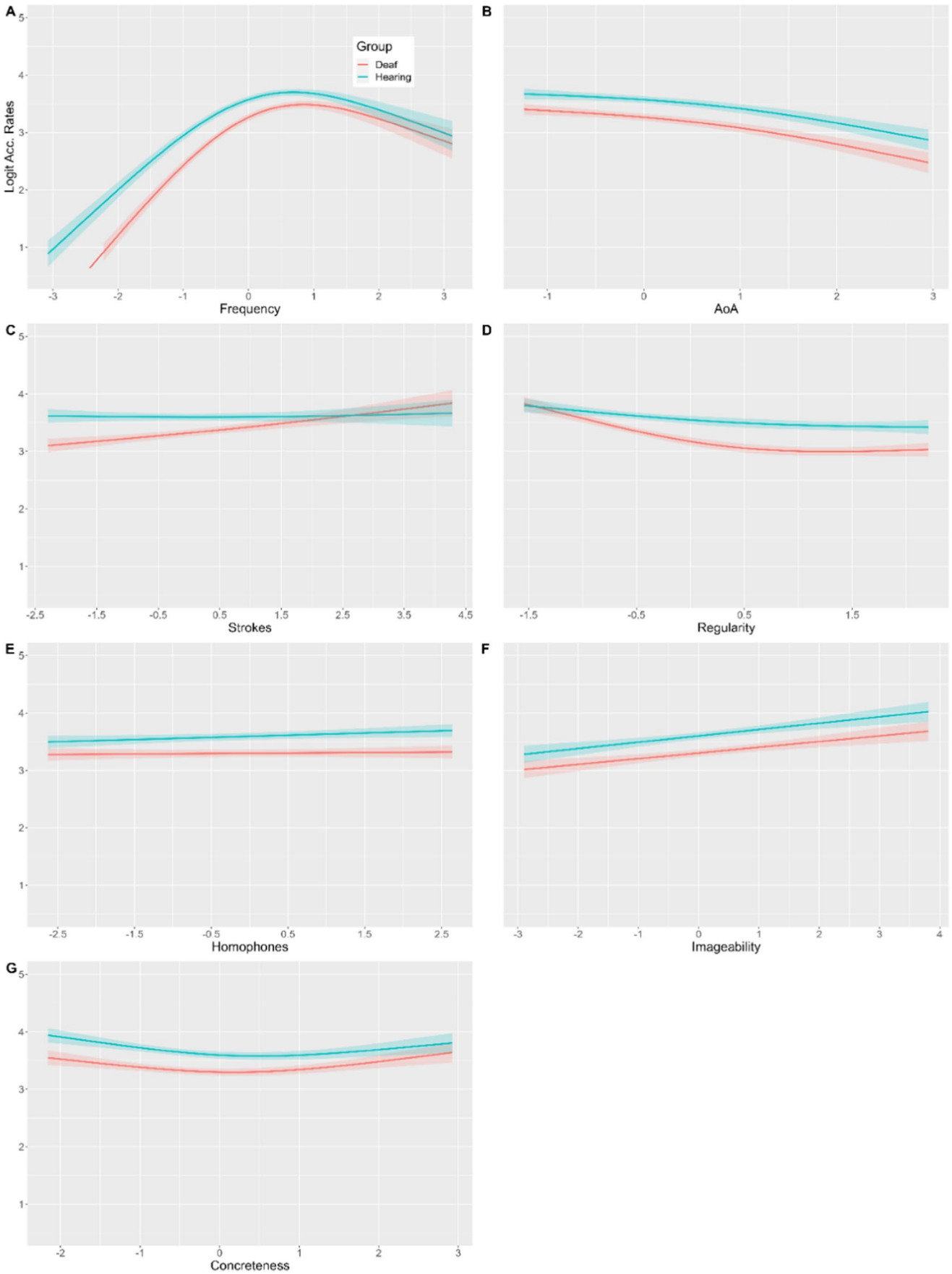
The patterns of statistical significance were consistent between the linear-terms-only GAMs and those with both linear and smooth terms. However, the latter showed a significantly improved model fit. While the inclusion of smooth terms in these models allows us to capture more complex, potentially nonlinear relationships, the interpretation of smooth terms requires visual inspection of spline plots generated from the model. In contrast, linear terms provide coefficients that offer a direct interpretation of the direction and magnitude of the effect of each predictor. In this report, we present detailed statistics from the linear-only models for ease of interpretation (see Tables 3, 4, and 5), while also referencing the spline plots [Figs. 1 (accuracy) and 5 (reaction time)] to give further clarity to the nonlinear relationships revealed by the smooth terms in the models. Additionally, we used difference smooths from these models to identify the specific intervals in which the two groups significantly diverged (see Figs. 2, 3, 4, 6, and 7). Model outputs for the GAMs with spline terms are available for further reference in the appendix (see Tables 6, 7, and 8).

We calculated the relative importance statistic (*lmg*) for each predictor (separately for the deaf and hearing groups) using the relaimpo package (Grömping, 2006) in R. This allowed us to quantify each predictor's contribution to the variance explained by each model. These values can be found in the *lmg* column in Tables 3, 5 and Tables 7 and 8 in the appendix.

Multicollinearity, which is when two or more predictor variables in a regression model are highly correlated, can be a problem in regression analysis because it can inflate the variance of regression coefficients (Hocking, 2013), making them unstable and difficult to interpret. Correlation tests across our predictors suggested that there is a moderate correlation between imageability and concreteness and a strong correlation between frequency and AoA (see Table 2 and scatterplot matrix, Fig. 8). While pairwise correlations of predictors did not exceed .7 (a threshold often used when assessing collinearity, Baayen et al., 2006), we conducted an additional check by calculating their variance inflation factors (VIFs). The VIF values ranged from 1.00 to 4.35 ($M = 1.94$, $SD = 0.90$). Given that all values were below the generally accepted limit of 5 (Sheather, 2009), we can state that multicollinearity is unlikely to have significantly influenced our models.

Results

In the deaf group (see *Deaf accuracy*, Table 3), higher accuracy was notably associated with characters that had a higher frequency, more strokes, and higher levels of imageability. Conversely, lower accuracy was observed with characters that had higher AoAs, higher regularity, and higher concreteness. These results suggest that the frequency, stroke count, and imageability of a character


A
Group
Deaf
Hearing
Logit Acc. Rates
Frequency
B
AoA
C
Strokes
D
Regularity
E
Homophones
F
Imageability
G
Concreteness

◂**Fig. 1** Empirical logits of accuracy rates as a function of the lexical variables, as derived from the full GAM (refer to Table 6) with both smooth and linear terms. The fitted values for each term are depicted, with the corresponding 95% confidence intervals shown as shaded bands. The lexical variables, transformed into *z*-scores, include (**A**) log frequency, (**B**) age of acquisition, (**C**) number of strokes, (**D**) orthographic-phonological regularity, (**E**) log count of homophones, (**F**) imageability, and (**G**) concreteness

positively affect decision accuracy, while the character's AoA, regularity, and concreteness tend to adversely influence accuracy in this deaf cohort. The effect of homophone density was not significant.

In the hearing group (see *Hearing accuracy*, Table 3), a higher level of accuracy was associated with characters that had a higher frequency, higher homophone counts, and higher imageability. Meanwhile, lower accuracy was observed with characters that had higher AoAs, higher regularity, and higher concreteness. The stroke count of a character did not significantly influence decision accuracy within the hearing group. These results suggest that while frequency, homophone density, and imageability positively affect decision accuracy for hearing individuals, the AoA, regularity, and concreteness of a character appear to have a negative impact.

In the full model comparing the hearing and deaf groups (see *Accuracy*, Table 4), the hearing group demonstrated

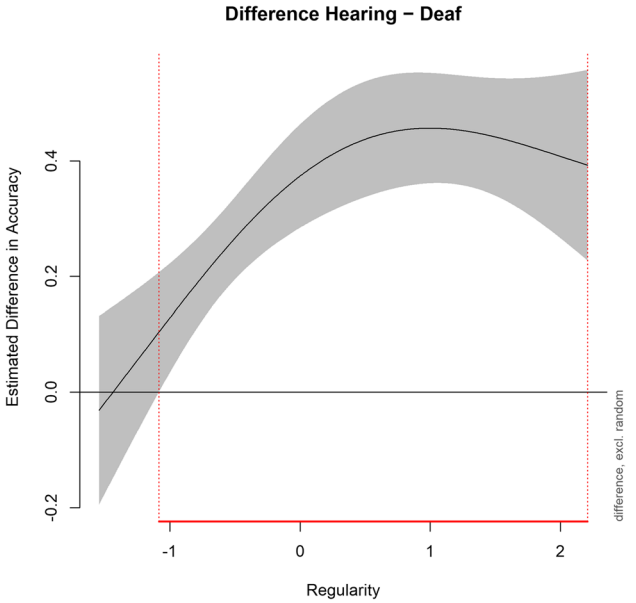


Fig. 4 Differential impact of orthographic-phonological regularity on empirical logit accuracy between deaf and hearing groups. The difference smooth shows estimated differences, highlighting significant differences (where the 95% confidence interval does not include zero) with dotted red lines. The *z*-transformed regularity values span from −1.54 to 2.21, with a significant difference window from −1.08 to 2.21

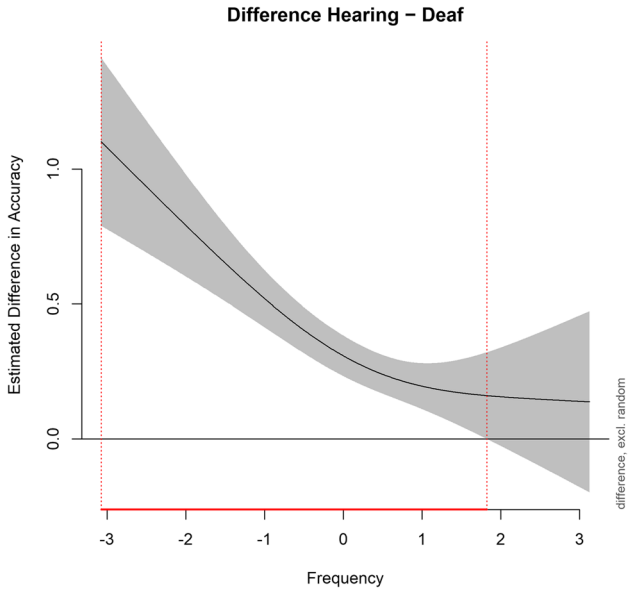


Fig. 2 Differential impact of frequency on empirical logit accuracy between deaf and hearing groups. The difference smooth shows estimated differences, highlighting significant differences (where the 95% confidence interval does not include zero) with dotted red lines. The *z*-transformed frequency values span from −3.08 to 3.13, with significant differences in the range of −3.08 to 1.82

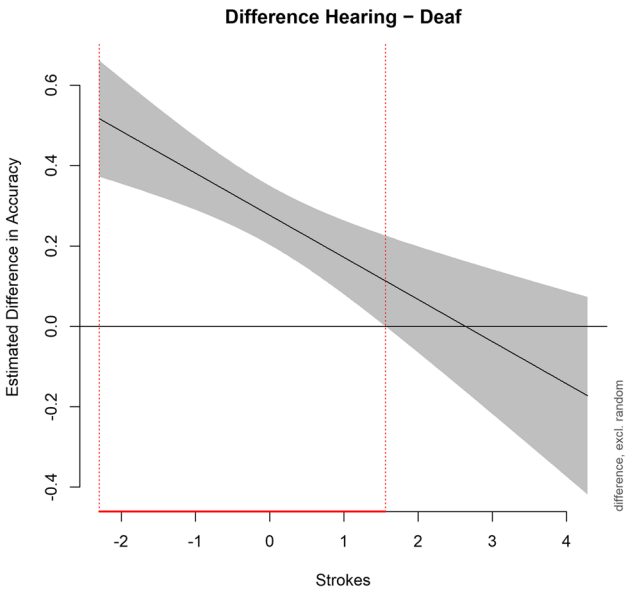
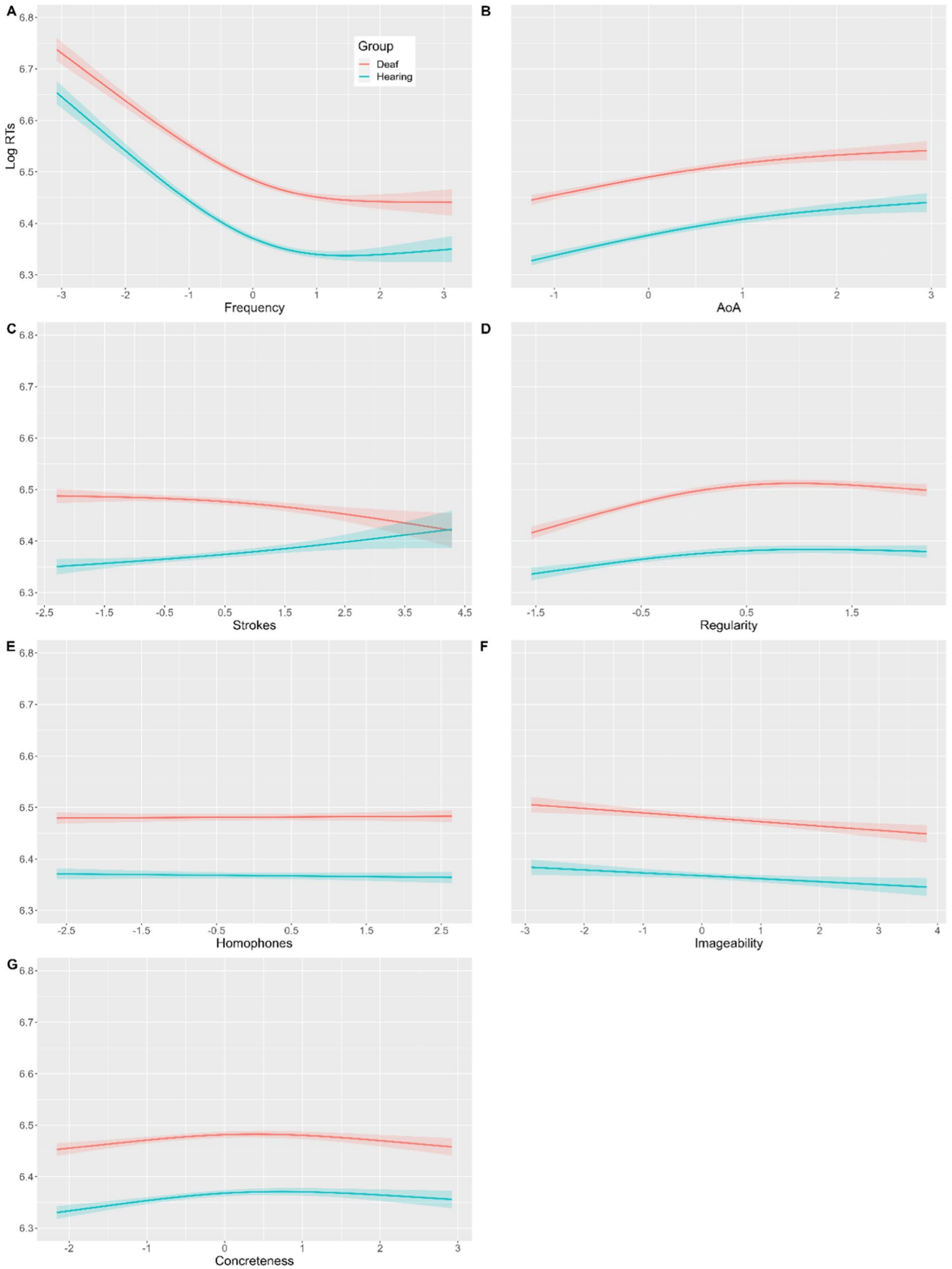


Fig. 3 Differential impact of stroke number on empirical logit accuracy between deaf and hearing groups. The difference smooth shows estimated differences, highlighting significant differences (where the 95% confidence interval does not include zero) with dotted red lines. The *z*-transformed stroke-number values span from −2.30 to 4.29, with a significant difference window from −2.30 to 1.56


A
B
C
D
E
F
G
Log RTs
Group
Deaf
Hearing
Frequency
AoA
Strokes
Regularity
Homophones
Imageability
Concreteness

◂**Fig. 5** Log-transformed reaction times as a function of the lexical variables, as derived from the full GAM (refer to Table 6) with both smooth and linear terms. The fitted values for each term are depicted, with the corresponding 95% confidence intervals shown as shaded bands. The lexical variables, transformed into *z*-scores, include (**A**) log frequency, (**B**) age of acquisition, (**C**) number of strokes, (**D**) orthographic-phonological regularity, (**E**) log count of homophones, (**F**) imageability, and (**G**) concreteness

significantly higher overall accuracy than the deaf group. The interaction effects revealed that the accuracy differential between the hearing and deaf participants was influenced by character frequency, strokes, and regularity. Specifically, the advantage of hearing participants diminished as character frequency increased and as characters contained more strokes. On the contrary, the accuracy advantage of hearing participants over deaf participants was amplified with

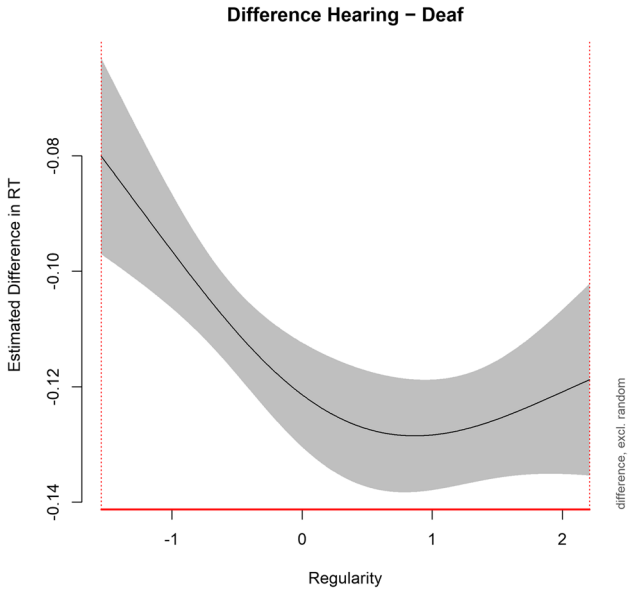


Fig. 6 Differential impact of orthographic-phonological regularity on log RT between deaf and hearing groups. The difference smooth shows estimated differences, highlighting significant differences (where the 95% confidence interval does not include zero) with dotted red lines. The *z*-transformed regularity values range from −1.54 to 2.21, with a significant difference window from −1.54 to 2.21

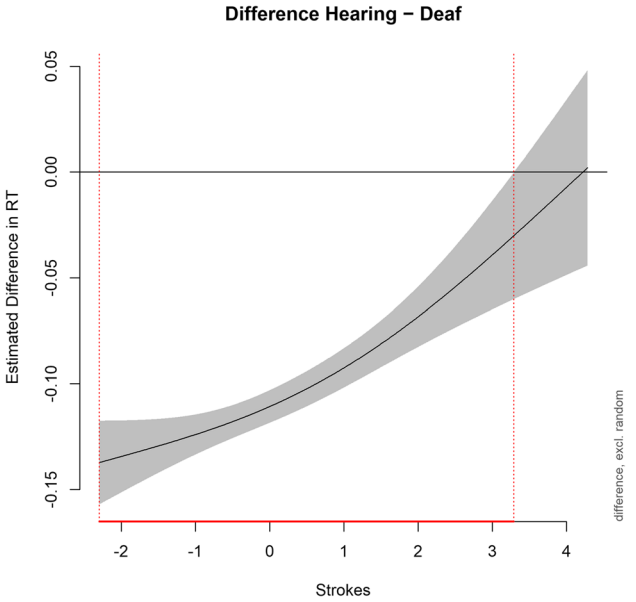


Fig. 7 Differential impact of stroke number on log RT between deaf and hearing groups. The difference smooth shows estimated differences, highlighting significant differences (where the 95% confidence interval does not include zero) with dotted red lines. The *z*-transformed stroke-number values range from −2.30 to 4.29, with a significant difference window from −2.30 to 3.29

characters of higher regularity. Other interaction effects were not found to be statistically significant.

In the analysis of the hearing group (see *Hearing RTs*, Table 5), several factors were found to significantly influence character decision RTs. Specifically, faster character decision RTs in the hearing group were associated with higher character frequency and higher character imageability. Conversely, slower decision RTs were observed when characters had higher AoAs, had a greater number of strokes, were higher in regularity, and had greater concreteness. The effect of homophone density on character decision RTs was not significant.

For the deaf group (see *Deaf RTs*, Table 5), the results revealed similar patterns, with some notable differences. Faster character decision RTs were reported when

Table 2 Correlation matrix for all variables

Predictors	Frequency	AoA	Strokes	Regularity	Homophones	Imageability	Concreteness
Frequency	1.00	−.66	−.33	−.28	−.02	.02	−.22
AoA	−.66	1.00	.29	.17	.06	−.24	.03
Strokes	−.33	.29	1.00	.27	−.02	.09	.05
Regularity	−.28	.17	.27	1.00	.08	.07	.11
Homophones	−.02	.06	−.02	.08	1.00	−.15	−.06
Imageability	.02	−.24	.09	.07	−.15	1.00	.51
Concreteness	−.22	.03	.05	.11	−.06	.51	1.00

Table 3 Results from separate generalized additive models (linear terms only) for accuracy of deaf and hearing participants

	Deaf accuracy						Hearing accuracy					
Linear terms	*β*	*SE*	*t*	*p*	VIF	lmg	*β*	*SE*	*t*	*p*	VIF	lmg
(Intercept)	2.98	0.02	154.17	**< .001**	-	-	3.32	0.02	182.01	**<.001**	-	-
Frequency	0.55	0.03	20.02	**< .001**	2.03	20.6	0.38	0.03	14.7	**<.001**	2.03	14.6
AoA	−0.21	0.03	−7.83	**< .001**	1.98	11.9	−0.18	0.03	−6.96	**<.001**	1.98	9.6
Strokes	0.09	0.02	4.42	**< .001**	1.21	0.8	0.00	0.02	0.09	.926	1.21	0.9
Regularity	−0.18	0.02	−8.93	**< .001**	1.15	3.8	−0.08	0.02	−4.31	**<.001**	1.15	1.6
Homophones	0.03	0.02	1.28	.202	1.03	0.0	0.05	0.02	2.68	**.008**	1.03	0.1
Imageability	0.20	0.02	8.25	**< .001**	1.55	2.4	0.19	0.02	8.50	**<.001**	1.55	2.5
Concreteness	−0.08	0.02	−3.55	**< .001**	1.47	0.8	−0.11	0.02	−5.02	**<.001**	1.47	1.1

Significant *p*-values indicated in bold

Key: *β* = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor, lmg = Lindeman, Merenda, and Gold (relative importance)

characters had higher frequency, more strokes, and higher imageability. Slower character decision RTs were associated with characters of higher AoA, higher regularity, and higher concreteness. As with the hearing group, the effect of homophone density on character decision RTs was also not significant.

When comparing the RTs between the hearing and deaf groups (see *RTs*, Table 4), the hearing group demonstrated significantly shorter RTs overall. Additionally, significant interaction effects were found between group membership (hearing vs. deaf) and certain lexical factors. Specifically, the number of strokes showed a significant interaction effect with group, suggesting that the hearing group had a greater increase in RT for characters with more strokes. Furthermore, there was a significant interaction between group and the orthographic-phonological regularity of the characters. This indicates that the hearing group demonstrated a greater decrease in RT for characters of higher regularity than the deaf group. Other group interactions were not significant.

Table 4 Results from the comprehensive generalized additive model (linear terms only) assessing accuracy and response time (RT) data for both deaf (D) and hearing (H) participants

	Accuracy					*RTs*				
Linear terms	*β*	*SE*	*t*	*p*	VIF	*β*	*SE*	*t*	*p*	VIF
(Intercept)	2.98	0.02	158.75	**<.001**	-	6.49	0.002	3601.52	**<.001**	-
Group H – D	0.33	0.03	12.51	**<.001**	1.00	−0.10	0.003	−39.95	**<.001**	1.00
Frequency	0.55	0.03	20.61	**<.001**	4.06	−0.05	0.003	−21.19	**<.001**	4.05
AoA	−0.21	0.03	−8.06	**<.001**	3.96	0.03	0.003	10.58	**<.001**	3.95
Stroke	0.09	0.02	4.55	**<.001**	2.43	−0.01	0.002	−2.62	**.009**	2.42
Regularity	−0.19	0.02	−9.19	**<.001**	2.29	0.02	0.002	10.38	**<.001**	2.29
Homophones	0.03	0.02	1.32	.189	2.07	0.00	0.002	−0.22	.828	2.07
Imageability	0.20	0.02	8.49	**<.001**	3.10	−0.01	0.002	−6.14	**<.001**	3.10
Concreteness	−0.08	0.02	−3.65	**<.001**	2.94	0.01	0.002	3.08	**.002**	2.94
Group H – D × Frequency	−0.17	0.04	−4.50	**<.001**	-	−0.00	0.004	−0.36	.716	-
Group H – D × AoA	0.04	0.04	0.93	.351	-	0.00	0.004	1.03	.302	-
Group H – D × Strokes	−0.09	0.03	−3.15	**.002**	-	0.02	0.003	5.70	**<.001**	-
Group H – D × Regularity	0.10	0.03	3.55	**<.001**	-	−0.01	0.003	−3.61	**<.001**	-
Group H – D × Homophones	0.02	0.03	0.91	.366	-	−0.00	0.003	−0.70	.482	-
Group H – D × Imageability	−0.01	0.03	−0.17	.862	-	0.00	0.003	0.29	.774	-
Group H – D × Concreteness	−0.03	0.03	−0.86	.392	-	0.01	0.003	1.65	.100	-

Significant *p*-values indicated in bold

Key: *β* = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor

Table 5 Results from separate generalized additive models (linear terms only) for reaction time (RT) of deaf and hearing participants

	Hearing RTs						*Deaf RTs*					
Linear terms	*β*	*SE*	*t*	*p*	VIF	lmg	*β*	*SE*	*t*	*p*	VIF	lmg
(Intercept)	6.392	0.002	3762.48	**<.001**	-	-	6.494	0.002	3415.44	**<.001**	-	-
Frequency	−0.056	0.002	−23.03	**<.001**	2.03	25.9	−0.054	0.003	−20.10	**<.001**	2.02	21.7
AoA	0.031	0.002	12.78	**<.001**	1.98	17.7	0.027	0.003	10.04	**<.001**	1.97	13.9
Strokes	0.011	0.002	5.78	**<.001**	1.21	3.9	−0.005	0.002	−2.49	**.013**	1.21	1.2
Regularity	0.010	0.002	5.59	**<.001**	1.15	2.9	0.020	0.002	9.84	**<.001**	1.15	4.6
Homophones	−0.002	0.002	−1.29	.199	1.03	0.0	0.000	0.002	−0.21	.837	1.03	0.1
Imageability	−0.013	0.002	−6.08	**<.001**	1.55	1.2	−0.014	0.002	−5.82	**.000**	1.55	1.5
Concreteness	0.012	0.002	5.74	**<.001**	1.47	1.7	0.007	0.002	2.92	**.004**	1.47	0.9

Significant *p*-values indicated in bold

Key: *β* = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor, lmg = Lindeman, Merenda, and Gold (relative importance)

Discussion

We conducted a large-scale character decision study to investigate Chinese character processing in deaf readers. We measured accuracy rates and RTs and looked at how these interacted with a variety of lexical variables, including frequency, AoA, number of strokes, regularity, homophone density, imageability, and concreteness. Comparing our deaf participants with hearing participants from another large-scale character decision study (i.e., Sze et al., 2014), we were able to examine differences in lexical processing between the two groups.

Age of acquisition and frequency

Age of acquisition (AoA) and frequency are often correlated but may reflect different aspects of language acquisition. While both capture cumulative exposure to words over a reader's lifetime, AoA effects, if independent from frequency effects, may be attributed to the fact that lexical items acquired at an early age become more entrenched in the lexical system due to the high level of neuroplasticity present in the brain during early life (Brysbaert & Ghyselinck, 2006; Ghyselinck et al., 2004). In our models, frequency followed by AoA accounted for the most variance based on the relative importance metric, suggesting these variables have the greatest impact on lexical decision accuracy and RT in both groups.

In the current study, deaf participants had higher error rates and longer response times for characters acquired at later ages, consistent with previous studies on hearing readers of Chinese (Cai et al., 2022b; Sze et al., 2014, 2015; Wang et al., 2020). The interaction between group and frequency for accuracy rates suggested that the accuracy advantage that hearing readers tended to have over deaf readers decreased as character frequency increased. Difference smooths indicated that the two groups differed significantly in their accuracy for characters that had frequencies from 3.08 standard deviations below the mean to 1.82 standard deviations above the mean. For characters with frequencies greater than this (1.82 to 3.13 standard deviations above the mean), their accuracy was comparable. Evidence has shown that deaf readers can build strong, direct connections between lexical-semantic and orthographic levels and even outperform hearing readers on certain lexical retrieval tasks (Gutierrez-Sigut et al., 2019; Morford et al., 2017). However, the development of these connections relies on repeated exposure to orthographic forms (Coltheart, 2006; Coltheart et al., 2001; Harm & Seidenberg, 2004). The lower text exposure of the deaf readers, as evidenced by their Author Recognition Test scores, suggests that lack of exposure may underlie the pattern observed for frequency and highlights the importance of extensive exposure to Chinese characters for developing character recognition skills.

Imageability and concreteness

Imageability and concreteness are two closely related variables, as imageable words are generally concrete rather than abstract. Kousta et al. (2011) demonstrated that these variables differ in their distribution: Concreteness ratings exhibit a bimodal distribution, with one mode corresponding to abstract words and the other to concrete words. In contrast, the distribution of imageability ratings is unimodal, reflecting the strength of associations between words and the mental images they elicit.

In the present study, concreteness was linked to decreased accuracy rates and longer response times for both deaf and hearing groups. Kousta et al. (2011) observed a processing advantage for abstract words over concrete ones, which they attributed to the stronger affective associations of abstract

words.[1] If their interpretation is correct, it implies that Chinese characters with more robust affective associations are processed more rapidly by both deaf and hearing individuals. However, further analyses incorporating measures that capture emotional valence are necessary to confirm this.

Conversely, imageability positively influenced lexical access, correlating with higher accuracy rates and shorter response latencies for both groups. Contrary to our hypothesis, imageability did not facilitate lexical access more for deaf readers than for hearing readers. Nonetheless, the results suggest that concreteness and imageability exert similar effects on both deaf and hearing readers.

Homophone density and regularity

Previous research has suggested that the speed at which a character is processed may be influenced by the number of homophones that character has. For example, H.-C. Chen et al. (2009b) and Ziegler et al. (2000) found that characters with more homophone mates were named more quickly, and Wang et al. (2020) found that lexical access was slower during writing for characters with more homophones. In our analysis of Sze et al.'s (2014) data, character decision accuracy was higher when characters had more homophones, suggesting a facilitatory effect of homophone density consistent with prior studies. Although the influence of homophone density was not evident among deaf readers—possibly indicating a diminished role of phonology in their character recognition process—the difference observed between the two groups did not reach statistical significance. This precludes a definitive conclusion that homophone density differentially affects deaf and hearing readers.

The impact of orthographic-phonological regularity was somewhat unexpected. Contrary to the anticipated higher accuracy and decreased response times based on prior research (Chang et al., 2016; Hue, 1992; Yang et al., 2009), the current study's deaf group and the hearing comparison group from Sze et al. (2014) exhibited lower accuracy and increased response times.[2] This unexpected pattern may be attributable to the experimental design—specifically, the use of pseudo-characters that all shared a left-to-right composition. This design may have inadvertently biased participants to either reject or hesitate when presented with any character of the same structure. Given that characters with high orthographic-phonological regularity also frequently exhibit a left-to-right composition, this could have unintentionally resulted in slower and less accurate responses for regular characters.

Comparing the deaf readers in our study to the hearing readers from Sze et al. (2014), the deaf group exhibited higher error rates and response latencies for characters with more regular orthographic-phonological correspondence. If we consider the unexpected pattern of orthographic-phonological regularity hindering character decision processes—potentially due to the bias from the similarly structured pseudo-characters—it is plausible that this pattern reflects deaf readers' greater reliance on orthographic structure in character processing. This pattern may also suggest that phonological information in highly regular characters facilitates lexical access, reducing the impact of the pseudo-character bias, for hearing readers but not for deaf readers. Interestingly, examination of the difference smooth for orthographic-phonological regularity, which has a significant difference window from −1.08 to 2.21, suggests that deaf and hearing readers performed similarly in terms of accuracy when regularity was low (between 1.54 and 1.08 standard deviations below the mean).

In conjunction with the lack of effect of homophone density in deaf readers, these findings suggest that speech-based phonological code may not play a major role in lexical access during character recognition for deaf readers. This aligns with the view that deaf readers tend to rely more on direct access to word meanings via orthography and bypass speech-based phonological mediation (Emmorey & Lee, 2021; Sehyr & Emmorey, 2022).

Stroke number

The relationship between stroke count and the accuracy of character decision, as well as the lexical decision times, showed differing patterns between deaf and hearing readers. Within the range of characters that had two to 15 strokes (z-scores from −2.30 to 1.56), deaf readers exhibited increased accuracy as the stroke number rose, while accuracy for hearing readers declined with increasing stroke count. However, this difference in accuracy trends between the two groups was no longer statistically significant when the characters' stroke count exceeded around 15, equating to z-scores greater than 1.56 standard deviations above the mean. In a similar vein, for characters with two to 21 strokes (z-scores from −2.30 to 3.30), character decision times displayed a contrasting pattern. As the number of strokes increased, the RTs of the deaf readers became shorter, while those of hearing readers became longer. The differing RT

[1] Note that while other studies (Bonin et al., 2018; Khanna & Cortese, 2021; Zhang et al., 2006) have similarly found a processing advantage for abstract words, some have found the opposite pattern—i.e., a facilitative concreteness effect (Allen & Hulme, 2006; Bottini et al., 2022; Fliessbach et al., 2006).

[2] Sze et al. (2014, 2015) did not report on regularity effects but rather looked at the effect of a related measure, *consistency*. Sze et al. (2015) found that consistent characters were responded to faster than less consistent characters, with the effect being magnified for high-frequency characters.

trends between the two groups were no longer statistically significant when stroke count surpassed around 21, corresponding to *z*-scores greater than 3.30 standard deviations above the mean. These patterns suggest that higher orthographic complexity has a facilitative effect on character recognition for deaf readers, enabling them to identify characters more quickly and precisely. This finding is particularly striking because it contrasts with the pattern observed in hearing readers in related studies, which have consistently reported a correlation between stroke count and higher error rates, as well as longer response times (Sze et al., 2015; Wang et al., 2020; Yang et al., 2016).

Deaf readers, in comparison with hearing readers, demonstrate greater efficiency in processing orthographic forms (as discussed in Bélanger et al., 2013; Emmorey and Lee, 2021; and Yan et al., 2015), which may be a result of their tendency to bypass speech-based phonological code and greater reliance on direct orthographic-semantic connections. As shown by the current study, beyond simply having more efficient orthographic processing, as reported in previous studies, deaf readers also appear to uniquely benefit from the presence of complex orthographic configurations. This finding implies that intrinsic differences may exist in the mechanisms through which deaf readers process orthographic information, potentially enhancing their ability to efficiently access word meanings. Reinforcing this notion, recent research has started to reveal potential differences in how deaf and hearing individuals process orthographic information. For instance, Emmorey et al. (2021) reported that deaf readers exhibited heightened sensitivity to mismatches in letter casing between prime and target words relative to their hearing counterparts. They hypothesized that the system of abstract letter coding in deaf readers might be inherently different from that used by hearing readers. In a similar vein, Gutierrez-Sigut et al. (2022) conducted an event-related potential (ERP) study that compared visual word recognition abilities in deaf and hearing readers of Spanish. The study found that pseudowords mimicking the outline-shape of their base words elicited less-negative ERP responses in deaf readers. This was interpreted as indicating increased sensitivity among deaf readers to subtle visual similarities, suggesting a high reliance on visual features. The authors concluded that this might be evidence of a fine-grained visual-orthographic processing system in deaf readers, which aids in efficient word recognition. Taken together, these studies suggest that there may be inherent differences in the mechanisms deaf readers employ to process orthographic information, which may underpin the different patterns observed between deaf and hearing readers with regard to orthographic complexity in our current study.

The unique orthographic processing abilities of deaf readers may be linked to enhanced visual attention skills. Certain aspects of visual attention are modified in deaf individuals, such that deaf individuals raised with access to a natural language (e.g., a natural sign language) from birth can develop enhanced spatial attention skills (Dye & Bavelier, 2010). Furthermore, deaf individuals may give greater attention to visual detail, as evidenced by research that suggests that deaf users of sign language are more accurate in detecting subtle facials features (McCullough & Emmorey, 1997) and are able to retain more detailed visual memories than hearing individuals (Craig et al., 2022). As complex Chinese characters tend to occupy more visual space and contain more detailed visual information relative to simpler ones, the ability of deaf readers to allocate their visual attention in space more effectively, coupled with their heightened sensitivity to visual detail, might give them special advantages in recognizing characters with higher stroke numbers.

Another factor to consider is the possible use of specific cognitive strategies for character recognition, which, to our knowledge, have not been systematically investigated. Deaf individuals' increased reliance on visual information likely contributes to the development of unique cognitive strategies for learning[3] (Marschark, 2006). Consequently, deaf readers might employ different cognitive approaches when learning to recognize Chinese characters. For instance, they may depend more on the overall visual shape or structure of characters, which could be more pronounced in complex characters with a higher stroke count. Utilizing such cognitive strategies may provide deaf readers with an advantage in identifying orthographically complex characters, in contrast to hearing readers, who may prefer to use other recognition strategies. These alternative strategies might involve, for example, the use of phonological code. Further research is needed to better understand the underlying factors contributing to this deaf advantage in complex character recognition, as well as to explore the cognitive strategies that may be employed by deaf readers. This could provide valuable insights into reading and character recognition processes in deaf individuals.

Theoretical implications

Our findings hold significant theoretical implications for understanding Chinese character processing in deaf readers. The lexical constituency model (Perfetti et al., 2005; Perfetti & Liu, 2006) posits that a word comprises interconnected orthographic, phonological, and semantic

[3] Learning strategies refer to actions and mental processes that learners employ to modulate how they encode and internalize information (Weinstein & Mayer, 1986).

components, which are co-activated during word retrieval. The lexical quality hypothesis (Perfetti, 2017; Perfetti & Hart, 2002) builds upon this, asserting that reading efficiency relies on the quality of a reader's lexical representations, determined by the connection strength between these components. Through repeated exposure to words, most readers develop high-quality lexical representations, enabling efficient word meaning retrieval. Thierfelder, Wigglesworth, et al. (2020a) suggested that Chinese deaf readers with lower reading fluency may lack these high-quality representations, leading to reduced word retrieval and integration efficiency. In a study investigating the impact of lexical quality variables on reading comprehension in reading-level-matched deaf and hearing adults, Sehyr and Emmorey (2022) argued that deaf readers, having underspecified phonological representations, develop more precise orthographic representations and stronger direct links between orthography and semantics.

Our findings support the lexical quality hypothesis, firstly, in that we observed that both AoA and frequency had significant impacts on the character decision performance of deaf readers, highlighting the importance of early and extensive character exposure for developing high-quality lexical representations. Our results also indicate that deaf readers can develop high-quality orthographic representations, possibly due to a lower reliance on phonology. This is evident in the enhanced orthographic identification abilities observed in deaf readers when presented with complex characters and their reduced sensitivity to orthographic-phonological regularity in character recognition. In line with Sehyr and Emmorey's (2022) findings, it is possible to efficiently retrieve word meanings via orthography even when phonological components are underdeveloped. This can be achieved if strong connections between orthography and semantics have been established through sufficient exposure to written word forms.

Practical implications, limitations, and conclusion

Our findings revealed notable differences between deaf and hearing readers, which may have significant implications for deaf education, particularly in regions that use written Chinese. Firstly, we observed that deaf and hearing readers recognized high-frequency Chinese characters and those learned at an early age with similar speeds and accuracy. However, for characters acquired later in life or with lower frequencies, deaf participants exhibited higher error rates and response latencies than hearing participants. These results potentially reflect the inadequate written language input and age-appropriate instruction for deaf students in some schools, possibly due to low expectations of students (Lytle et al., 2006; Wang & Andrews, 2017).

Secondly, our results suggested that deaf readers tend to place greater reliance on direct orthographic-semantic connections when recognizing characters. In contrast, hearing readers may exhibit a stronger tendency to rely on indirect orthographic-phonological connections to meaning. Furthermore, the orthographic recognition processes of deaf readers may employ distinct mechanisms that facilitate character recognition, particularly in the presence of greater orthographic complexity. These findings highlight the importance of two key factors for literacy development in deaf students: (1) providing robust and age-appropriate exposure to Chinese characters, and (2) tailoring teaching methods to the strengths of deaf students, such as their enhanced orthographic and visual processing abilities. By focusing on the development of strong orthographic-semantic connections, educators can better support the literacy growth of deaf learners.

The database that has resulted from this project offers numerous potential applications, serving as a valuable resource for researchers investigating Chinese reading processing in deaf readers. By facilitating comparisons between deaf readers and other special populations, the database can contribute to a broader understanding of diverse reading experiences. Additionally, the data can be employed to develop computational models of deaf reading, as well as to inform the creation of reading assessments and educational materials specifically tailored to the needs of deaf readers.

Despite offering valuable contributions, our study has certain limitations that must be acknowledged. First, the relatively small sample size, consisting of 25 deaf readers, could limit the generalizability of our findings. While efforts were made to recruit as many participants as possible, attaining adequate sample sizes is a common challenge encountered when recruiting participants from small populations, such as the deaf community. Second, there was an age disparity between our two groups, with a mean age of 30 for the deaf readers and 22 for the hearing readers. This difference in age could potentially influence our results, as age can have an impact on cognitive and linguistic performance (Hartshorne & Germine, 2015). Third, while the deaf readers scored within the range of the general population of native Chinese speakers in terms of vocabulary knowledge and author recognition, they scored slightly lower in the author recognition test than the hearing comparison group from Sze et al. (2014), who screened out participants with below-average scores. Furthermore, the extent to which signing deaf readers, for whom written language is often akin to a second language (Hoffmeister & Caldwell-Harris, 2014), and hearing L1

readers can be directly compared is not clear. Including comparison groups with second-language readers (e.g., as in Cates et al., 2022) may offer new insights. Given the potential comparability issues between the two groups in the current study, in addition to the other limitations mentioned previously, it is important to interpret these findings cautiously and consider these limitations when assessing the implications of this study. Lastly, we should note that the comparability of the two data sets used in this study may be limited by the considerable length of time that passed between the time the hearing data were collected by Sze et al. (2014) and the data collection period for the current study.

In summary, we conducted a large-scale lexical decision study involving 25 deaf readers, collecting responses for 2500 Chinese characters. This study offers valuable insights into the Chinese character processing of deaf readers and presents a foundation for future research and practical applications aimed at enhancing literacy education for this population.

Appendix

Table 6 Results from the comprehensive generalized additive model (with both linear and smooth terms) assessing accuracy and reaction time (RT) data for both deaf and hearing participants

	Accuracy					*RT*				
Linear terms	*β*	*SE*	*t*	*p*	VIF	*b*	*SE*	*t*	*p*	VIF
(Intercept)	2.98	0.02	170.34	**<.001**	-	6.494	0.002	3717.59	**<.001**	-
Group	0.33	0.03	13.43	**<.001**	1.00	−0.102	0.002	−41.25	**<.001**	1.00
Homophones	0.01	0.02	0.44	.657	2.08	0.001	0.002	0.41	.683	2.08
Imageability	0.10	0.02	4.39	**<.001**	3.35	−0.008	0.002	−3.75	**<.001**	3.35
Group H – D × Homophones	0.03	0.03	1.18	.238	-	−0.002	0.003	−0.83	.409	-
Group H – D × Imageability	0.01	0.03	0.35	.729	-	0.003	0.003	0.87	.387	-
Smooth terms	edf	Ref.df	*F*	*p*	VIF	edf	Ref.df	*F*	*p*	VIF
Frequency	2.00	2.00	436.66	**<.001**	2.34	1.99	2.00	272.27	**<.001**	2.29
AoA	1.93	2.00	37.06	**<.001**	4.34	1.95	2.00	63.04	**<.001**	4.35
Strokes	1.38	1.61	15.95	**<.001**	1.68	1.75	1.92	7.70	**<.001**	1.70
Regularity	1.96	2.00	68.06	**<.001**	4.22	1.98	2.00	79.10	**<.001**	4.22
Concreteness	1.96	2.00	13.12	**<.001**	1.48	1.96	2.00	11.96	**<.001**	3.19
Group H – D × Frequency	1.84	1.98	14.94	**<.001**	-	1.80	1.96	2.29	.125	-
Group H – D × AoA	1.00	1.00	0.76	.383	-	1.00	1.00	1.35	.246	-
Group H – D × Strokes	1.00	1.00	14.15	**<.001**	-	1.75	1.92	19.33	**<.001**	-
Group H – D × Regularity	1.87	1.98	11.26	**<.001**	-	1.88	1.98	10.02	**<.001**	-
Group H – D × Concreteness	1.00	1.00	2.07	.150	-	1.00	1.00	1.67	.197	-

Significant *p*-values indicated in bold

Key: *β* = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor

Table 7 Results from separate generalized additive models (with both linear and smooth terms) for accuracy of deaf and hearing participants

	Deaf accuracy							*Hearing accuracy*					
Linear terms	*β*	*SE*	*t*	*p*	VIF	lmg	Linear terms	*β*	*SE*	*t*	*p*	VIF	lmg
(Intercept)	2.98	0.02	167.33	**<.001**	-	-	(Intercept)	3.32	0.02	192.79	**<.001**	-	-
Homophones	0.01	0.02	0.42	.674	1.04	0	AoA	−0.17	0.02	−7.03	**<.001**	1.99	9.6
Imageability	0.1	0.02	4.34	**<.001**	1.71	2.4	Strokes	0	0.02	−0.21	.835	1.21	0.9
Smooth terms	edf	Ref.df	*F*	*p*	VIF	lmg	Regularity	−0.1	0.02	−5.16	**<.001**	1.15	1.6
Frequency	2	2	405.73	**<.001**	1.34	20.6	Homophones	0.04	0.02	2.31	**0.021**	1.03	0.1
AoA	1.89	1.99	32.57	**<.001**	2.24	11.9	Imageability	0.1	0.02	4.65	**<.001**	1.67	2.5
Strokes	1.82	1.97	14.8	**<.001**	1.65	0.8	Smooth terms	edf	Ref.df	*F*	*p*	VIF	lmg
Regularity	1.96	2	61.13	**<.001**	2.23	3.8	Frequency	2	2	252.683	**<.001**	1.2	14.6
Concreteness	1.94	2	7.37	**.001**	2.05	0.9	Concreteness	1.92	1.993	7.663	**.001**	2.1	1.1

Significant *p*-values indicated in bold

Key: *β* = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor, lmg = Lindeman,Merenda, and Gold (relative importance), edf = estimated degrees of freedom, Ref.df = reference degrees of freedom

Table 8 Results from separate generalized additive models (with both linear and smooth terms) for reaction time (RT) of deaf and hearing participants

	Hearing RTs							*Deaf RTs*					
Linear terms	*b*	*SE*	*t*	*p*	VIF	lmg	Linear terms	*b*	*SE*	*t*	*p*	VIF	lmg
(Intercept)	6.392	0.002	3908.912	**<.001**	-	-	(Intercept)	6.494	0.002	3506.74	**<.001**	-	-
Strokes	0.010	0.002	5.583	**<.001**	1.26	3.9	Homophones	0.001	0.002	0.38	.702	1.04	0.1
Homophones	−0.001	0.002	−0.834	.404	1.04	0.0	Imageability	−0.009	0.002	−3.55	**<.001**	1.71	1.5
Imageability	−0.006	0.002	−2.743	**.006**	1.70	1.2	Smooth terms	edf	Ref.df	*F*	*p*	VIF	lmg
Smooth terms	edf	Ref.df	*F*	*p*	VIF	lmg	Frequency	1.99	2	231.804	**<.001**	1.34	21.7
Frequency	1.99	2.00	354.06	**<.001**	1.34	25.9	AoA	1.90	1.99	52.00	**<.001**	2.25	13.9
AoA	1.88	1.99	84.74	**<.001**	2.24	17.7	Strokes	1.78	1.95	7.43	**<.001**	1.66	1.2
Regularity	1.88	1.99	22.32	**<.001**	1.64	2.9	Regularity	1.98	2.00	68.80	**<.001**	2.22	4.6
Concreteness	1.93	2.00	12.31	**<.001**	2.22	1.7	Concreteness	1.90	1.99	5.22	**.008**	1.99	0.9

Significant *p*-values indicated in bold

Key: β = coefficient, *SE* = standard error, *t* = *t*-value, *p* = *p*-value, VIF = variance inflation factor, lmg = Lindeman, Merenda, and Gold (relative importance), edf = estimated degrees of freedom, Ref.df = reference degrees of freedom

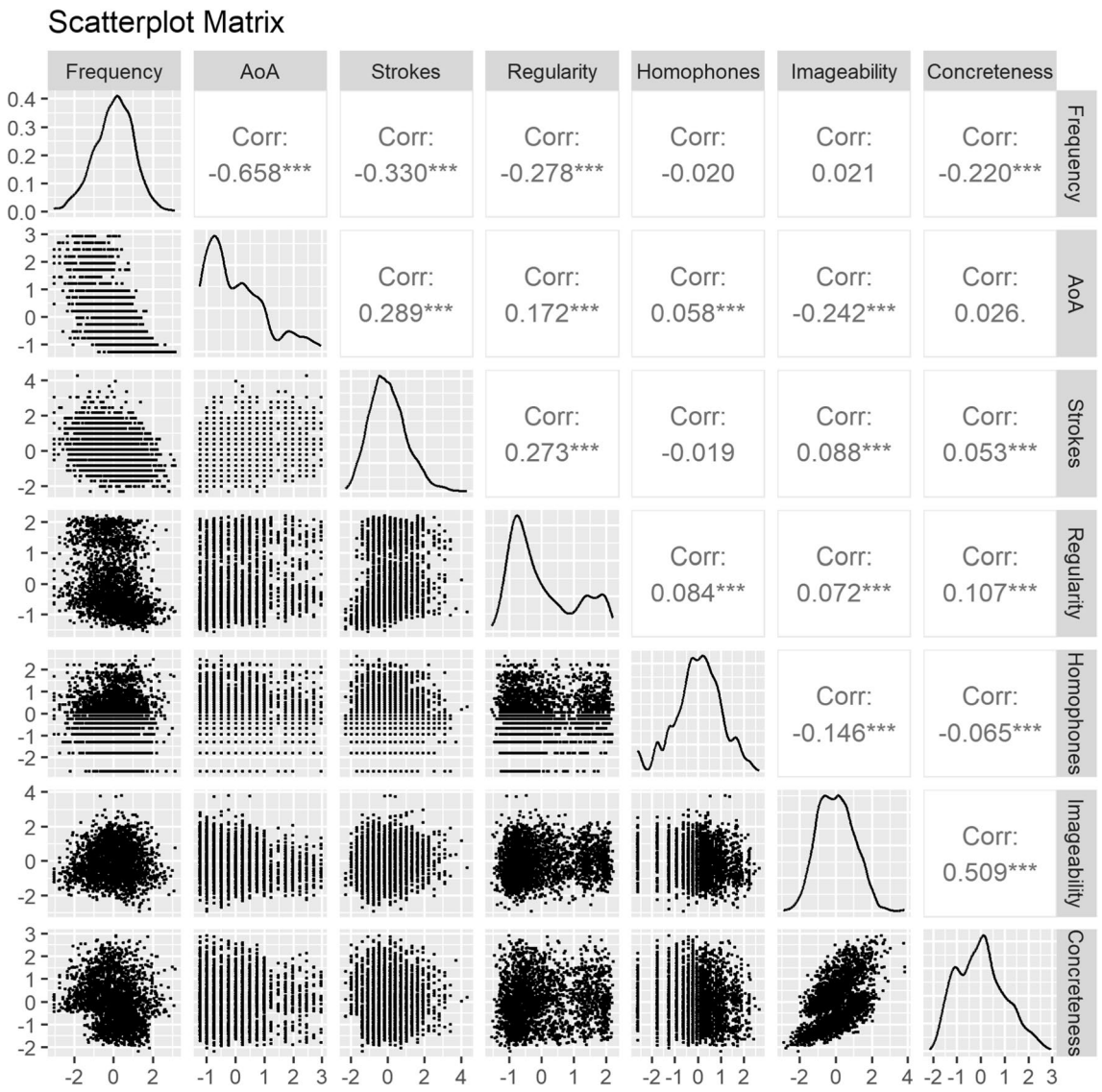


Fig. 8 Scatterplot matrix for all lexical variables

Code availability Codes for statistical analyses are publicly available at https://osf.io/qthfm/.

Author notes All experimental materials, data, and analytical scripts are publicly available at the Open Science Framework (https://osf.io/qthfm/).

This research was supported by a GRF grant from the Research Grants Committee (Project Number: 14600220) and a CUHK Faculty of Arts internal grant, both awarded to Z.G. Cai. We thank Weiyun Cai for help in data collection.

Authors' contributions Conceptualization: ZGC; Methodology: SH, ZGC; Formal analysis: PT, SH, ZGC; Investigation: HL; Writing—Original draft: PT, ZGC, SH; Writing—Review & editing: PT, ZGC, SH, HL. All authors read and approved the final manuscript.

Funding This research was funded by a direct grant from the Faculty of Arts, The Chinese University of Hong Kong, and a grant by the Research Grants Council (14613722), both to ZGC. Additionally, PT is supported by a separate grant from the Research Grants Council (PDFS2122-4H04).

Data availability All data and materials are publicly available at https://osf.io/qthfm/.

Declarations

Conflicts of interest/Competing interests There is no conflict of interest or competing interest to disclose.

Ethics approval The research was ethically approved by an institutional board at The Chinese University of Hong Kong and by an institutional board at the Shanghai International Studies University.

Consent to participate All participants signed a written consent form before taking part in the experiment.

Consent for publication Not applicable.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 199–213). Springer.

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language, 55*(1), 64–88. https://doi.org/10.1016/j.jml.2006.02.002

Allen, T. E., Clark, M. D., del Giudice, A., Koo, D., Lieberman, A., Mayberry, R., & Miller, P. (2009). Phonology and Reading: A Response to Wang, Trezek, Luckner, and Paul. *American Annals of the Deaf, 154*(4), 338–345. https://doi.org/10.1353/aad.0.0109

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55*(2), 290–313. https://doi.org/10.1016/j.jml.2006.03.008

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445–459.

Bélanger, N. N., Baum, S. R., & Mayberry, R. I. (2012a). Reading Difficulties in Adult Deaf Readers of French: Phonological Codes, Not Guilty! *Scientific Studies of Reading, 16*(3), 263–285. https://doi.org/10.1080/10888438.2011.568555

Bélanger, N. N., Mayberry, R. I., & Rayner, K. (2013). Orthographic and phonological preview benefits: Parafoveal processing in skilled and less-skilled deaf readers. *Quarterly Journal of Experimental Psychology, 66*(11), 2237–2252. https://doi.org/10.1080/17470218.2013.780085

Bélanger, N. N., Slattery, T. J., Mayberry, R. I., & Rayner, K. (2012b). Skilled Deaf Readers have an Enhanced Perceptual Span in Reading. *Psychological Science, 23*(7), 816–823. https://doi.org/10.1177/0956797611435130

Bonin, P., Méot, A., & Bugaiska, A. (2018). Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times. *Behavior Research Methods, 50*(6), 2366–2387. https://doi.org/10.3758/s13428-018-1014-y

Bosworth, R. G., & Dobkins, K. R. (2002). The Effects of Spatial Attention on Motion Processing in Deaf Signers, Hearing Signers, and Hearing Nonsigners. *Brain and Cognition, 49*(1), 152–169. https://doi.org/10.1006/brcg.2001.1497

Bottini, R., Morucci, P., D'Urso, A., Collignon, O., & Crepaldi, D. (2022). The concreteness advantage in lexical decision does not depend on perceptual simulations. *Journal of Experimental Psychology: General, 151*(3), 731.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7–8), 992–1011. https://doi.org/10.1080/13506280544000165

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS One, 5*(6), e10729.

Cai, Z. G., Huang, S., Xu, Z., & Zhao, N. (2022a). Objective ages of acquisition for 3300+ simplified Chinese characters. *Behavior Research Methods, 54*(1), 311–323. https://doi.org/10.3758/s13428-021-01626-1

Cai, Z. G., Zhao, N., Lin, H., Xu, Z., & Thierfelder, P. (2022b). Syntactic encoding in written language production by deaf writers: A structural priming study and a comparison with hearing writers. *Journal of Experimental Psychology Learning, Memory, and Cognition*.

Cates, D. M., Traxler, M. J., & Corina, D. P. (2022). Predictors of reading comprehension in deaf and hearing bilinguals. *Applied Psycholinguistics, 43*(1), 81–123.

Chang, Y.-N., Hsu, C.-H., Tsai, J.-L., Chen, C.-L., & Lee, C.-Y. (2016). A psycholinguistic database for traditional Chinese character naming. *Behavior Research Methods, 48*, 112–122.

Chen, B., Dent, K., You, W., & Wu, G. (2009a). Age of acquisition affects early orthographic processing during Chinese character recognition. *Acta Psychologica, 130*(3), 196–203. https://doi.org/10.1016/j.actpsy.2008.12.004

Chen, H.-C., Vaid, J., & Wu, J.-T. (2009b). Homophone density and phonological frequency in Chinese word recognition. *Language and Cognitive Processes, 24*(7–8), 967–982.

Codina, C. J., Pascalis, O., Baseler, H. A., Levine, A. T., & Buckley, D. (2017). Peripheral Visual Reaction Time Is Faster in Deaf Adults and British Sign Language Interpreters than in Hearing Adults. *Frontiers in Psychology, 8*, 50. https://doi.org/10.3389/fpsyg.2017.00050

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education, 4*(1), 5–17. https://doi.org/10.1080/13603110600574322

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256.

Craig, M., Dewar, M., Turner, G., Collier, T., & Kapur, N. (2022). Evidence for superior encoding of detailed visual memories in deaf signers. *Scientific Reports, 12*(1), 9097.

Da, J. (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. *Proceedings of the Fourth International Conference on New Technologies in Teaching and Learning Chinese*, 501–511.

Donnelly, S., & Verkuilen, J. (2017). Empirical logit analysis is not logistic regression. *Journal of Memory and Language, 94*, 28–42.

Dye, M. W., & Bavelier, D. (2010). Attentional enhancements and deficits in deaf populations: An integrative review. *Restorative Neurology and Neuroscience, 28*(2), 181–192.

Emmorey, K., & Corina, D. (1993). Hemispheric specialization for ASL signs and English words: Differences between imageable and abstract forms. *Neuropsychologia, 31*(7), 645–653.

Emmorey, K., Holcomb, P. J., & Midgley, K. J. (2021). Masked ERP repetition priming in deaf and hearing readers. *Brain and Language, 214*, 104903.

Emmorey, K., & Kosslyn, S. M. (1996). Enhanced image generation abilities in deaf signers: A right hemisphere effect. *Brain and Cognition, 32*(1), 28–44.

Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition, 46*(2), 139–181.

Emmorey, K., & Lee, B. (2021). The neurocognitive basis of skilled reading in prelingually and profoundly deaf adults. *Language and Linguistics Compass, 15*(2), e12407. https://doi.org/10.1111/lnc3.12407

Feng, G., Miller, K., Shu, H., & Zhang, H. (2001). Rowed to Recovery: The Use of Phonological and Orthographic Information in Reading Chinese and English. *Journal of Experimental Psychology. Learning. Memory & Cognition, 27*(4), 1079.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42*, 488–496.

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32*(3), 1413–1421. https://doi.org/10.1016/j.neuroimage.2006.06.007

Gao, J., Fan, K., & Fei, J. (1993). *Xiandai hanzi xue [The study of modern Chinese characters]*. Higher Education Press.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*(1), 43–67. https://doi.org/10.1016/j.actpsy.2003.11.002

Grömping, U. (2006). Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of Statistical Software, 17*(1), 1–27.

Gutierrez-Sigut, E., Vergara-Martínez, M., & Perea, M. (2017). Early use of phonological codes in deaf readers: An ERP study. *Neuropsychologia, 106*, 261–279. https://doi.org/10.1016/j.neuropsychologia.2017.10.006

Gutierrez-Sigut, E., Vergara-Martinez, M., & Perea, M. (2019). Deaf readers benefit from lexical feedback during orthographic processing. *Scientific Reports, 9*(1), 1–13.

Gutierrez-Sigut, E., Vergara-Martínez, M., & Perea, M. (2022). The impact of visual cues during visual word recognition in deaf readers: An ERP study. *Cognition, 218*, 104938. https://doi.org/10.1016/j.cognition.2021.104938

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 3*, 662.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26*(4), 433–443.

Hocking, R. R. (2013). *Methods and applications of linear models: Regression and the analysis of variance*. John Wiley & Sons.

Hoffmeister, R. J., & Caldwell-Harris, C. L. (2014). Acquiring English as a second language via print: The task for deaf children. *Cognition, 132*(2), 229–242. https://doi.org/10.1016/j.cognition.2014.03.014

Hue, C.-W. (1992). Recognition Processes in Character Naming. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Advances in Psychology* (Vol. 90, pp. 93–107). North-Holland. https://doi.org/10.1016/S0166-4115(08)61888-9

Kelly, L. P., & Barac-Cikoja, D. (2007). *The comprehension of skilled deaf readers* (pp. 244–280). *Children's Comprehension Problems in Oral and Written Language*.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono-and disyllabic words and nonwords. *Frontiers in Psychology, 1*, 174.

Khanna, M. M., & Cortese, M. J. (2021). How well imageability, concreteness, perceptual strength, and action strength predict recognition memory, lexical decision, and reading aloud performance. *Memory, 29*(5), 622–636.

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*(1), 14.

Leong, C. K., Cheng, P.-W., & Mulcahy, R. (1987). Automatic processing of morphemic orthography by mature readers. *Language and Speech, 30*(2), 181–196.

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39*(2), 192–198.

Liu, Z. F., Chen, C. Y., Tong, W., & Su, Y. Q. (2021). Deafness enhances perceptual span size in Chinese reading: Evidence from a gaze-contingent moving-window paradigm. *PsyCh Journal, 10*(4), 508–520. https://doi.org/10.1002/pchj.442

Lytle, R. R., Johnson, K. E., & Hui, Y. J. (2006). Deaf Education in China: History, Current Issues, and Emerging Deaf Voices. *American Annals of the Deaf, 150*(5), 457–469. https://doi.org/10.1353/aad.2006.0009

Marschark, M. (2006). Intellectual functioning of deaf adults and children: Answers and questions. *European Journal of Cognitive Psychology, 18*(1), 70–89.

Mayberry, R. I., del Giudice, A. A., & Lieberman, A. M. (2011). Reading Achievement in Relation to Phonological Coding and Awareness in Deaf Readers: A Meta-analysis. *The Journal of Deaf Studies and Deaf Education, 16*(2), 164–188. https://doi.org/10.1093/deafed/enq049

McCullough, S., & Emmorey, K. (1997). Face Processing by Deaf ASL Signers: Evidence for Expertise in Distinguishing Local Features. *Journal of Deaf Studies and Deaf Education, 2*(4), 212–222.

Morford, J. P., Occhino-Kehoe, C., Piñar, P., Wilkinson, E., & Kroll, J. F. (2017). The time course of cross-language activation in deaf ASL–English bilinguals*. *Bilingualism: Language and Cognition, 20*(2), 337–350. https://doi.org/10.1017/S136672891500067X

Paul, P. V., Wang, Y., Trezek, B. J., & Luckner, J. L. (2009). Phonology Is Necessary, but Not Sufficient: A Rejoinder. *American Annals of the Deaf, 154*(4), 346–356. https://doi.org/10.1353/aad.0.0110

Peng, D., Deng, Y., & Chen, B. (2003). 汉语多义单字词的识别优势效应 [The polysemy effect in Chinese one-character word identification]. *Acta Psychologica Sinica, 35*(5), 569–575.

Perfetti, C. A. (2017). *Lexical quality revisited* (pp. 51–67). *Developmental Perspectives in Written Language and Literacy*.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of Functional Literacy, 11*, 67–86.

Perfetti, C. A., & Liu, Y. (2006). *Reading Chinese characters: Orthography, phonology, meaning, and the lexical constituency model*. na.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The Lexical Constituency Model: Some Implications of Research on Chinese for General

Theories of Reading. *Psychological Review, 112*(1), 43–59. https://doi.org/10.1037/0033-295X.112.1.43

Perfetti, C. A., & Sandak, R. (2000). Reading optimally builds on spoken language: Implications for deaf readers. *Journal of Deaf Studies and Deaf Education, 5*(1), 32–50.

Proksch, J., & Bavelier, D. (2002). Changes in the Spatial Distribution of Visual Attention after Early Deafness. *Journal of Cognitive Neuroscience, 14*(5), 687–701. https://doi.org/10.1162/089892902601385 91

R Core Team. (2023). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical. *Computing*. https://www.R-project.org/

Rowley, K. E. (2018). *Visual Word Recognition in Deaf Readers: The interplay between orthographic, semantic and phonological information* [PhD Thesis]. UCL (University College London).

Sehyr, Z. S., & Emmorey, K. (2022). Contribution of lexical quality and sign language variables to reading comprehension. *The Journal of Deaf Studies and Deaf Education, 27*(4), 355–372.

Sheather, S. (2009). *A Modern Approach to Regression with R*. Springer Science & Business Media.

Sze, W. P., Rickard Liow, S. J., & Yap, M. J. (2014). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2,500 Chinese characters. *Behavior Research Methods, 46*(1), 263–273. https://doi.org/10.3758/s13428-013-0355-9

Sze, W. P., Yap, M. J., & Rickard Liow, S. J. (2015). The role of lexical variables in the visual recognition of Chinese characters: A megastudy analysis. *Quarterly Journal of Experimental Psychology, 68*(8), 1541–1570.

Thierfelder, P., Durantin, G., & Wigglesworth, G. (2020). The Effect of Word Predictability on Phonological Activation in Cantonese Reading: A Study of Eye-Fixations and Pupillary Response. *Journal of Psycholinguistic Research.* https://doi.org/10.1007/s10936-020-09713-8

Thierfelder, P., Wigglesworth, G., & Tang, G. (2020a). Orthographic and phonological activation in Hong Kong deaf readers: An eye-tracking study. *Quarterly Journal of Experimental Psychology.* https://doi.org/10.1177/1747021820940223

Thierfelder, P., Wigglesworth, G., & Tang, G. (2020b). Sign phonological parameters modulate parafoveal preview effects in deaf readers. *Cognition, 201*. https://doi.org/10.1016/j.cognition.2020.104286

Traxler, C. B. (2000). The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students. *The Journal of Deaf Studies and Deaf Education, 5*(4), 337–348. https://doi.org/10.1093/deafed/5.4.337

Tse, C.-S., Yap, M. J., Chan, Y.-L., Sze, W. P., Shaoul, C., & Lin, D. (2017). The Chinese Lexicon Project: A megastudy of lexical decision performance for 25,000+ traditional Chinese two-character compound words. *Behavior Research Methods, 49*, 1503–1519.

Wang, Q., & Andrews, J. F. (2017). Literacy instruction in primary level deaf education in China. *Deafness & Education International, 19*(2), 63.

Wang, R., Huang, S., Zhou, Y., & Cai, Z. G. (2020). Chinese character handwriting: A large-scale behavioral study and a database. *Behavior Research Methods, 52*(1), 82–96. https://doi.org/10.3758/s13428-019-01206-4

Wang, Y., Trezek, B. J., Luckner, J. L., & Paul, P. V. (2008). The Role of Phonology and Phonologically Related Skills in Reading Instruction for Students Who Are Deaf or Hard of Hearing. *American Annals of the Deaf, 153*(4), 396–407. https://doi.org/10.1353/aad.0.0061

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In C. M. Wittrock (Ed.), *Handbook of Research in Teaching* (pp. 315–327). Macmillan.

Wen, Y., Qiu, Y., Leong, C. X. R., & van Heuven, W. J. B. (2023). LexCHI: A quick lexical test for estimating language proficiency in Chinese. *Behavior Research Methods.* https://doi.org/10.3758/s13428-023-02151-z

Wood, S., & Wood, M. S. (2015). Package 'mgcv.' *R Package Version, 1*(29), 729.

Yan, G., Lan, Z., Wang, Y., & Benson, V. (2020). *Phonological Coding during Sentence Reading in Chinese Deaf Readers: An Eye-Tracking Study*. *Scientific Studies of Reading*.

Yan, M., Pan, J., Bélanger, N. N., & Shu, H. (2015). Chinese Deaf Readers Have Early Access to Parafoveal Semantics. *Journal of Experimental Psychology. Learning. Memory & Cognition, 41*(1), 254–261. https://doi.org/10.1037/xlm0000035

Yang, J., McCandliss, B. D., Shu, H., & Zevin, J. D. (2009). Simulating language-specific and language-general effects in a statistical learning model of Chinese reading. *Journal of Memory and Language, 61*(2), 238–257.

Yang, S., Zhang, S., & Wang, Q. (2016). P2 and behavioral effects of stroke count in Chinese characters: Evidence for an analytic and attentional view. *Neuroscience Letters, 628*, 123–127. https://doi.org/10.1016/j.neulet.2016.06.006

Zhang, Q., Guo, C., Ding, J., & Wang, Z. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language, 96*(1), 59–68. https://doi.org/10.1016/j.bandl.2005.04.004

Ziegler, J. C., Tan, L. H., Perry, C., & Montant, M. (2000). Phonology matters: The phonological frequency effect in written Chinese. *Psychological Science, 11*(3), 234–238.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.