香港教育大學
The Education University of Hong Kong

Phonological Processing of Chinese Characters during Sentence Reading

Jinger Pan
Department of Psychology
The Education University of Hong Kong

Script-specific properties and word recognition

- Alphabetic languages
  - Keg: How to pronounce? What does it mean?

- Chinese
  - 畾： How to pronounce? What does it mean?

- How do the characteristics of a specific script affect our processing of words?

Phonological processing in different writing systems

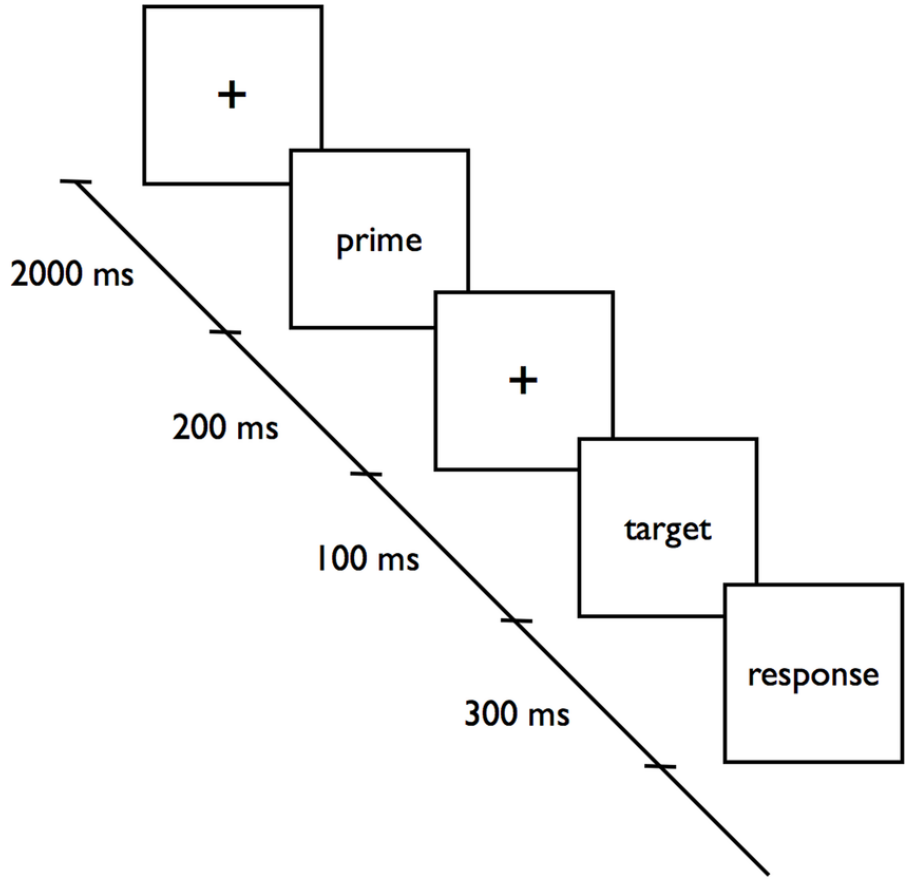
- Writing systems differ in their relative importance of phonological and semantic information.

- In alphabetic languages
  - orthography-phonology-semantics (Van Orden, 1987)

- In Chinese
  - Chinese characters represent morphemes and not directly represent the sounds of those morphemes (Hoosain, 1992)
  - Phonological mediation may be bypassed (Chen & Shu, 2001; Zhou & Marslen-Wilson, 2000)

Foveal priming

- Prime-Target
  - doctor—nurse
  - bread—nurse

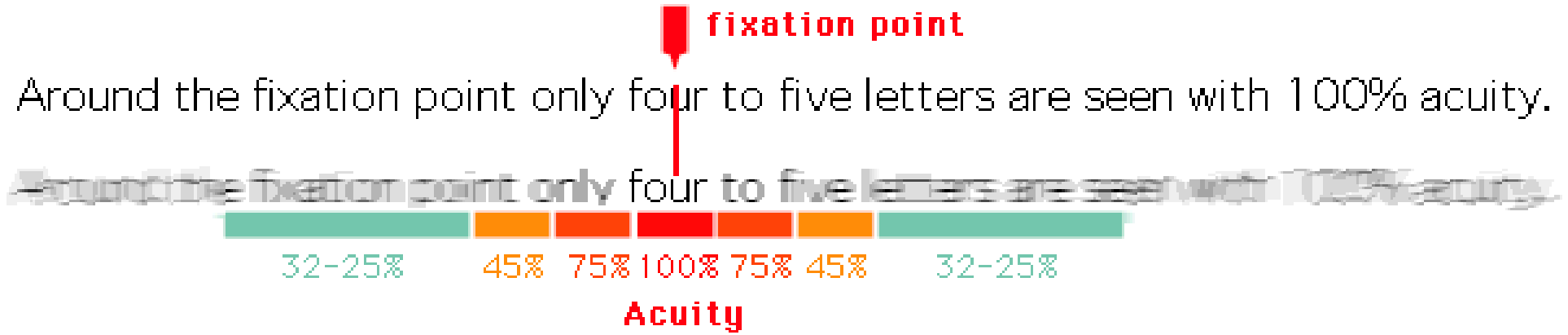


- Why do we actually need to move our eyes when we read?

The Retina of the Eye

- Fovea
- Parafovea
- Periphery

Visual acuity during reading



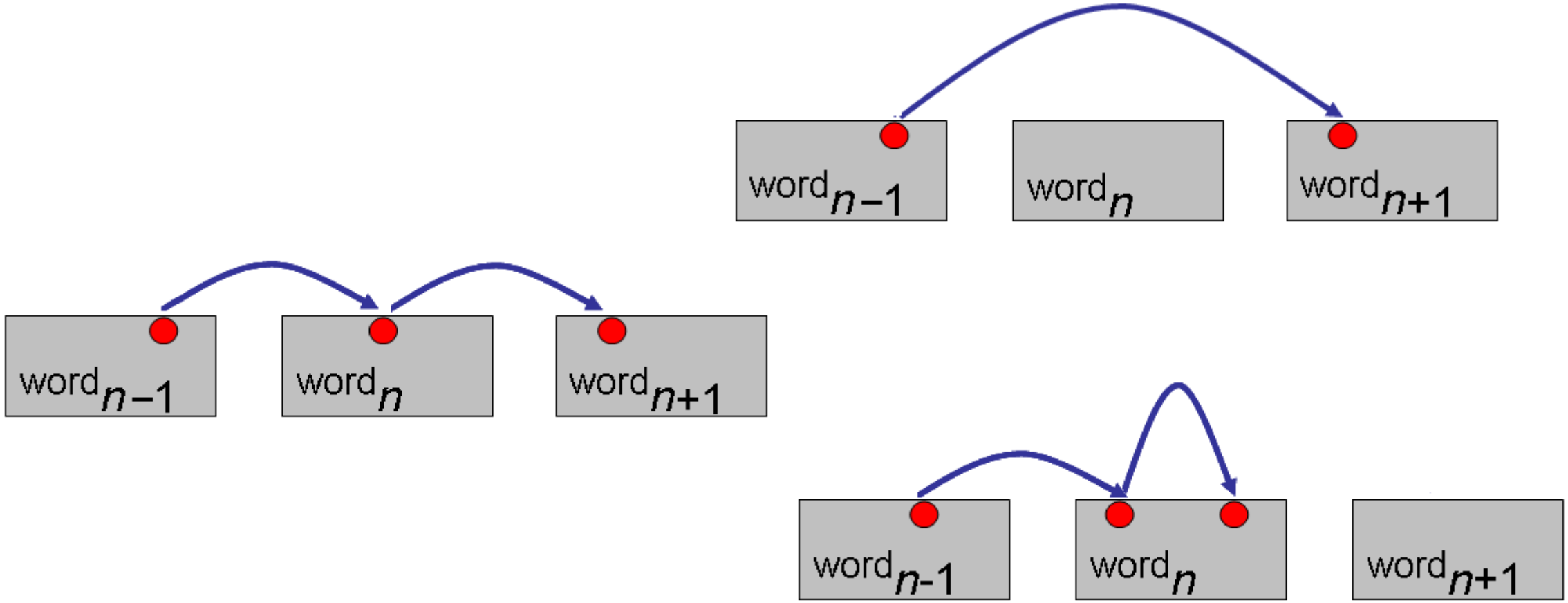
Wikipedia

EyeLink 1000/1000 Plus

Duration

Fixation : 150—300 ms

Saccade : 20—40 ms



Eye movements during silent reading

Error disruption paradigm (Daneman & Reingold, 1993)

Nice to meet you.

Nice to meet you.

Nice to meat you.

Nice to meot you.

Nice to haun you.

- Error disruption paradigm (foveal processing)
  - Wong and Chen (1999)
    - ✔ Orthographical substitution
    - ✗ Homophonic substitution
  - Feng, Miller, Shu, and Zhang (2001)
    - ✔ English phonological substitution
    - ✗ Chinese phonological substitution

研究发现，海豚有极为精密的声纳系统，它们能利用回声定位的方法识(食/织/考)别周围的环境，准确率之高令人惊叹。科学家们正努力探索海豚的这种回声定位的奥秘，以改进现有的船用声纳系统，从而提高捕鱼、扫雷、猎潜等多种工作的效率。此外，海豚还会“说话”，海豚那复杂的“语言”，至今未能被人类完全“破译”，它吸引着许多科学家的注意。

Although style manuals have existed for many years, showing established standards for creating footnotes, bibliographies, etc., for a wide range of document types in print and microform format, little has been available to help students and scholars *cite* [*sight/cute/gift*] electronic sources. Yet, it is equally important that these resources can be accurately credited to the author and, if desired, accessed by the reader.

The gaze-contingent boundary paradigm (Rayner, 1975)

… brilliantly composed a new xxxxg for the …

Condition	Word Fixated	
	n − 1	Target
Identical (song-song)	251(228)	246(214)
Semantically Related (tune-song)	250(228)	286(230)
Unrelated (door-song)	251(222)	290(234)
Visually Similar Nonword (sorp-song)	248(219)	251(215)

Phonological preview:

历史表明腐败是古代专制亡朝走向灭亡的原因。

*

Semantic preview:

历史表明腐败是古代专制帝朝走向灭亡的原因。

*

Unrelated preview:

历史表明腐败是古代专制西朝走向灭亡的原因。

*

Target sentence:

历史表明腐败是古代专制王朝走向灭亡的原因。

Late phonological preview benefit in simple and compound Chinese characters (Tsai et al., 2004; Tsai et al., 2013; Yan et al., 2009).

This talk

- Processing of different phonological units in reading and the factors affecting its processing.

- Study 1: Sandhi-tone processing
  - **Pan, J.**, Zhang, C., Huang, X., & Yan, M. (2021). Sandhi-tone words prolong fixation duration during silent sentence reading in Chinese. *Reading and Writing, 34,* 841-857.

- Study 2: The roles of rimes and tones in visual word processing
  - Yan, M., Tsang, Y.-K., & **Pan, J.** (2024). Phonological recovery during Chinese sentence reading: Effects of rime and tone. *Language, Cognition and Neuroscience, 39,* 501-512.

- Study 3: Lexical and sublexical phonological processing
  - **Pan, J.**, Yan, M., Laubrock, J., & Shu, H. (2019). Lexical and sublexical phonological effects in Chinese silent and oral reading. *Scientific Studies of Reading, 23,* 403-418.

- Study 4: Cross-language phonological processing of bilinguals
  - Yan, M., Luo, Y., & **Pan, J.** (2023). Monolingual and bilingual phonological activation in Cantonese. *Bilingualism: Language and Cognition, 26,* 751-761.

Study 1 Sandhi-tone processing

- One syllable in Chinese often maps onto a number of different characters and morphemes.

- There are over 50,000 Chinese characters in total, of which about 5,000 are commonly used, but there are only 400 meaningful syllables in Mandarin Chinese (Tsou, 1976).

- Chinese is a tonal language. One way to increase lexical distinction in speech is the lexical tone.

- Are detailed phonological features used during visual word identification?
  - Minimality hypothesis (Frost, 1998): simple and abstract, devoid of speech features.
  - Readers of alphabetic languages use relatively detailed phonological features, such as vowel duration, spoken syllable boundaries, and lexical stress, during visual word identification (e.g., Huestegge, 2010; Inhoff, Connine, & Radach, 2002).

- Eye-movement studies have shown that, during the silent reading of sentences, relatively detailed phonological features can be obtained.
  - Longer fixations on words with longer vowels (Huestegge, 2010)
  - Words with two stressed syllables took longer to read and received more fixations than those with one stressed syllable (Ashby & Clifton, 2005)

- Tone sandhi: T3 + T3 -> T2 + T3
  - 宝马 /bao**3** ma3/ becomes like /bao**2** ma3/

- Minimality hypothesis
  - No difference between sandhi-tone and base-tone (i.e., non-sandhi) words in viewing duration during the silent reading of Chinese sentences.

- If detailed phonological features are activated and utilized, we predict that fixation duration should be longer for sandhi-tone words than for base-tone words.

- Word frequency effect
  - Lexical access is faster for more commonly used words than for those seen less often (see Brysbaert, Mandera, & Keuleers, 2018, for a recent review).
  - When words are familiar to subjects, they tend to bypass phonology in favor of direct access, which leads to a general prediction that experimental effects due to phonological mediation should diminish or disappear (Coltheart, 1978; Van Orden, 1987).

- Method
  - Participants: 27 university students
  - Material and Design

	Sandhi tone	Base tone
Example word	赶紧	赶快
Meaning	Immediately	Immediately
Pronunciation	gan3-jing3	gan3-kuai4
Log-word freq	0.43 (0.71)	0.42 (0.72)
N. of strokes	16.6 (4.7)	16.5 (4.7)
Log-2nd char. freq	0.85 (0.76)	0.79 (0.75)
AoA	8.3 (1.2)	8.3 (1.2)
Imageability	4.0 (0.6)	3.9 (0.7)
Predictability	2.3% (8.4%)	1.6% (5.1%)

Sandhi-tone condition

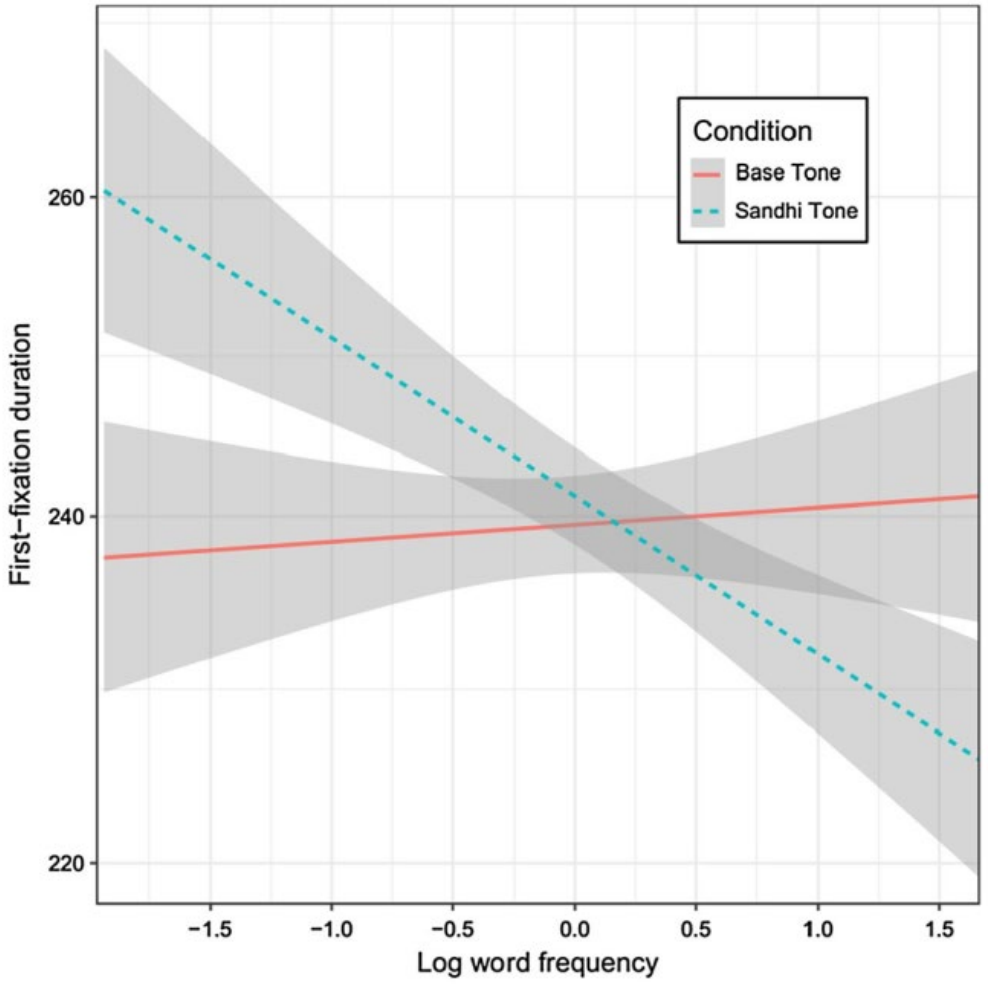
当地百姓非常希望赶紧结束这样颠沛流离的日子。

Base-tone condition

当地百姓非常希望赶快结束这样颠沛流离的日子。

- Results

	High frequency (N = 1865)		Low frequency (N = 1986)	
	Sandhi tone	Base tone	Sandhi tone	Base tone
SP	14.0 (15.9)	14.4 (13.6)	12.1 (11.2)	12.7 (12.7)
RP	16.8 (11.8)	18.3 (13.3)	22.1 (19.0)	17.4 (15.2)
FFD	249 (39)	250 (31)	259 (32)	249 (32)
GD	288 (48)	293 (43)	315 (59)	292 (51)
TRT	301 (52)	308 (54)	342 (76)	315 (64)



- Discussion
  - High frequency words
    - Rely on direct lexical access by passing phonological features; or
    - Phonological features were not yet activated by the time the lexical access of the frequent target words was completed
  - Low frequency words
    - Phonological activation: a prolongation in viewing time is a consequence of a conflict for the sandhi-tone words between their base tones represented at a morpheme/character level and surface realizations represented at a word level (i.e., sandhi tones)

Study 2
The roles of rimes and tones in visual word processing

- Previous findings (Luo et al., 2016; Yan et al., 2020) that lexical tone can influence visual word identification during natural silent reading of Chinese sentences were mostly based on some special tone variation phenomena in Chinese.

- Using the event-related potential (ERP) recording technique, Zou et al. (2020) found that tone violation was more difficult to detect than rime violation in spoken sentence comprehension.

- The present study investigated the relative importance of rimes and tones in character processing during reading.

- Method
  - Participants: 55 university students
  - Error disruption paradigm
  - Material and Design

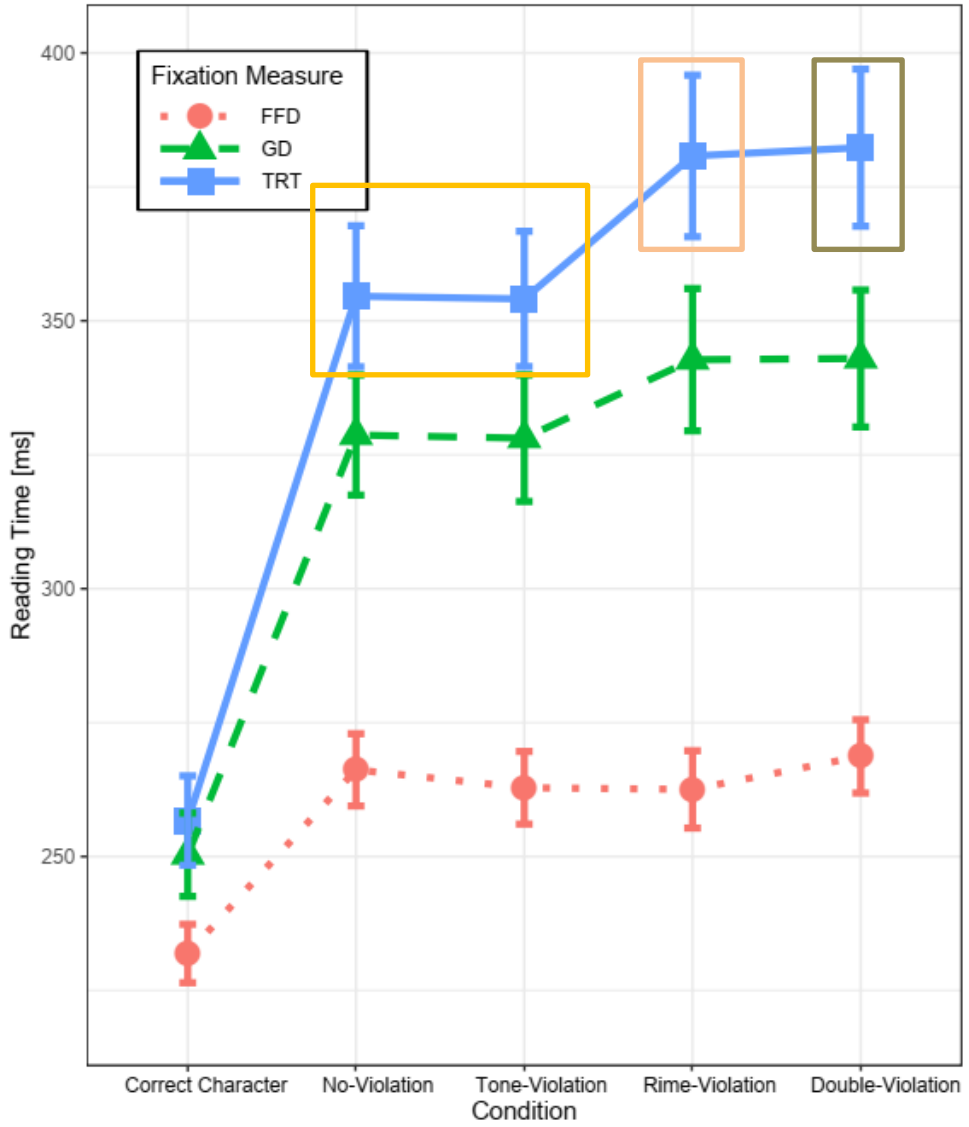
	Substitutions for Target Character (载 /zai3/)			
	No-Violation	Tone-Violation	Rime-Violation	Double-Violation
Example Character	宰	灾	枣	糟
Meaning	slaughter	disaster	date	dregs
Pronunciation	zai3	zai1	zao3	zao1
Printed Frequency	335 (659)	356 (605)	342 (615)	343 (463)
Phonological Frequency	1288 (1346)	1298 (1322)	1257 (1758)	1287 (1671)
Number of Strokes	9.5 (3.0)	9.5 (3.0)	9.5 (3.3)	9.5 (2.9)

Example sentence

这部史书详细记载了从上古到西汉约三千年的历史。	**Correct Character**
这部史书详细记宰了从上古到西汉约三千年的历史。	**No-Violation Character**
这部史书详细记灾了从上古到西汉约三千年的历史。	**Tone-Violation Character**
这部史书详细记枣了从上古到西汉约三千年的历史。	**Rime-Violation Character**
这部史书详细记糟了从上古到西汉约三千年的历史。	**Double-Violation Character**

- Results

	Condition				
	Correct Target	No-Violation	Tone-Violation	Rime-Violation	Double-Violation
SP	26.4 (17.1)	18.0 (16.4)	18.3 (17.9)	19.9 (17.0)	19.1 (17.0)
RP	9.2 (12.9)	24.9 (16.7)	24.8 (16.4)	28.0 (18.8)	25.9 (19.2)
FFD	233 (32)	269 (46)	266 (45)	267 (52)	271 (42)
GD	253 (45)	334 (68)	334 (71)	351 (83)	347 (79)
TRT	261 (51)	364 (84)	363 (86)	392 (107)	391 (95)



- Discussion
  - Late activation of phonological representation, even highly predictable target words did not lead to an earlier phonological activation.
  - The role of rime:
    - Participants were unable to recover the target word when the correct rime information was missing from the rime- and double-violation conditions; these findings suggest that rime has a decisive role in word recovery.
    - Given that target words were highly predicted, participants were expected to recover target words after finishing the sentence context. Consequently, although it can be said that rime affects the word recovery, it does not play a decisive role.

Study 3
Lexical and sublexical phonological processing

- Most previous studies on phonological processing have investigated only silent but not oral reading.

- Pan et al. (2016) found that, phonological information in the parafovea is activated early on during oral reading:
  - Simple characters (王-亡, wang)
  - Two-character words (轨迹-诡计, gui ji)

- 80% of the Chinese characters are phonograms composed of a semantic radical and a phonetic radical.

- Lexical and sublexical phonological representations
  - Regular　　牲sheng (生sheng)　　奶nai3 (乃nai)
  - Irregular　姓xing　(生sheng)　　仍reng (乃nai)

- We tested lexical and sublexical activation of foveal and parafoveal phonological information in silent and oral reading of Chinese sentences in two eye-tracking experiments.

- Lexical + Sublexical
  - Regular: both the phonetic radical and the whole substitution/preview character share the same pronunciation with the target character

- Sublexical
  - Irregular: only the phonetic radical of substitution/preview character shares the same pronunciation with the target character

Experiment 1: Foveal processing

- Method
  - Participants: 40 university students
  - Error disruption paradigm
  - Material and Design

	Identical	Lexical	Sublexical	Unrelated
Character	夕	惜	错	临
Translation	dusk	pity	fault	arrive
Lexical Pronunciation	xi1	xi1	cuo4	lin2
Phonetic Radical		昔	昔	
Radical Pronunciation		xi1	xi1	
Log. Frequency	1.8 (0.5)	1.6 (0.9)	1.5 (0.7)	1.5 (0.8)
No. Stroke	9.4 (3.0)	9.7 (2.6)	9.8 (2.5)	9.4 (2.3)
Semantic Relatedness Rating		1.5 (0.3)	1.4 (0.3)	1.4 (0.3)

Example sentence

Lexical Homophonic Substitution:
他的这幅油画意在描绘惜阳下事物的祥和美好。

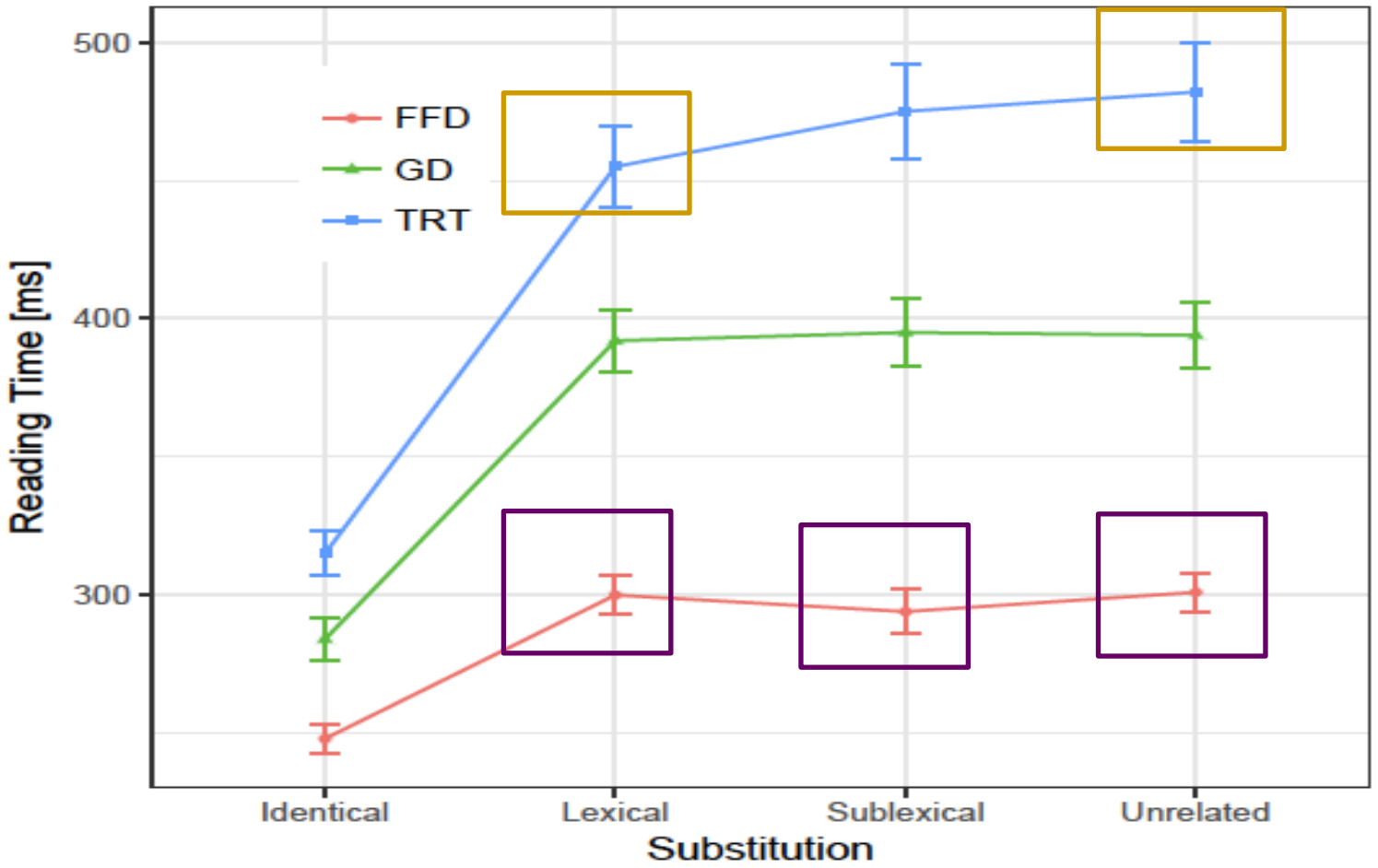
Sublexical Homophonic Substitution:
他的这幅油画意在描绘错阳下事物的祥和美好。

Unrelated Substitution:
他的这幅油画意在描绘临阳下事物的祥和美好。

Identical Substitution:
他的这幅油画意在描绘夕阳下事物的祥和美好。

Translation： *His oil painting is meant to describe the harmony of the things under the sunset.*

- Results & Discussion



- No evidence for early use of phonological information.
- Lexical phonological effects emerged in a late processing stage.

- Method
  - Participants : 52 university students
  - Gaze-contingent boundary paradigm
  - Oral vs. silent reading

Example sentence

他的这幅油画意在描绘夕阳下事物的祥和美好。 Identical Preview

*

他的这幅油画意在描绘惜阳下事物的祥和美好。 Lexical Homophonic Preview

*

他的这幅油画意在描绘错阳下事物的祥和美好。 Sublexical Homophonic Preview

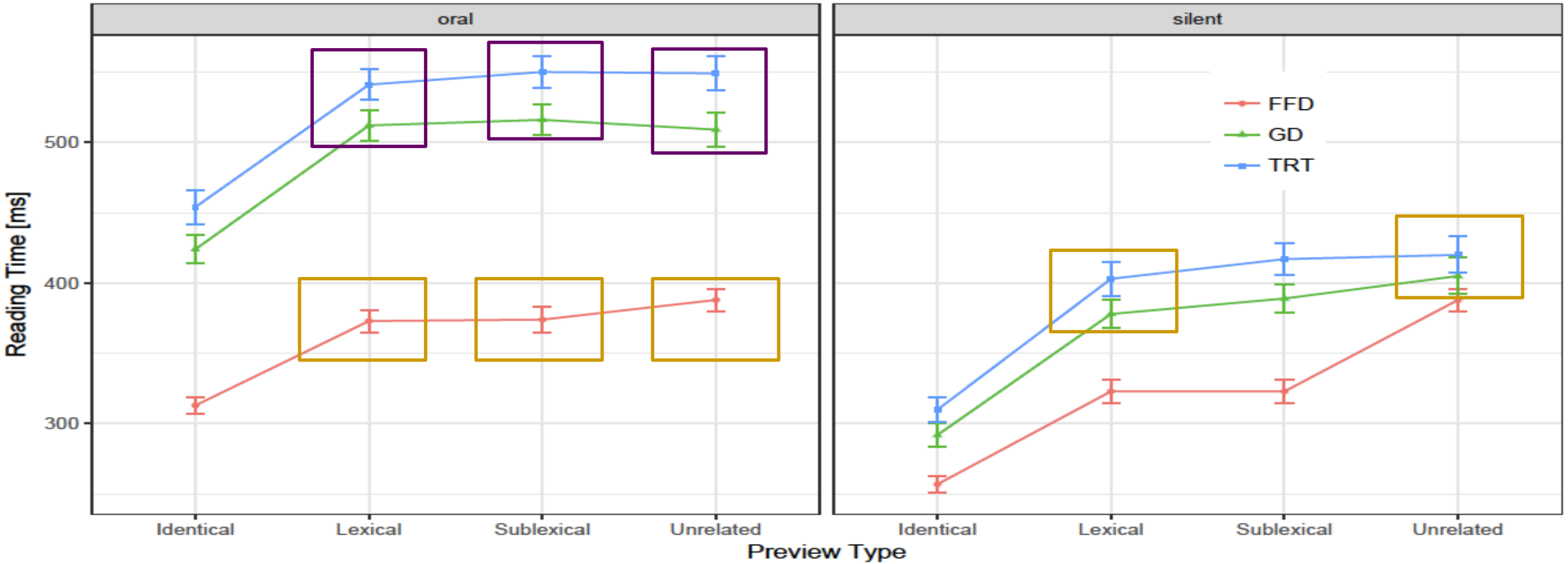
*

他的这幅油画意在描绘临阳下事物的祥和美好。 Unrelated Preview

*

他的这幅油画意在描绘夕阳下事物的祥和美好。 Target

- Results & Discussion



- Silent reading: reliable lexical homophone PBs appeared only in GD and TRT but not in FFD, with no evidence for sublexical homophone PB.
- Early lexical & sublexical phonological PB in oral reading
- Cancellation of phonological PB from early to late measures

Study 4
Cross-language phonological processing of bilinguals

- Phonological priming effects for cognates, even when the two languages are cross-scripted suggesting that lexical phonology is cross-linguistically integrated and represented for bilinguals.

- Inconsistent findings in Chinese:
  - For late Chinese–English bilinguals, Chinese words can phonologically prime English targets that are similar in pronunciation (e.g., Zhou, Chen, Yang & Dunlap, 2010).
  - Priming effects have been found not to differ between phonologically similar and dissimilar Chinese–Japanese cognate word pairs among late bilinguals (Liu et al., 2019; 2022).

- Confounds: Script familiarity
  - The identical written form of Chinese is mapped to several very different spoken systems, of which the most widely used are Cantonese and Mandarin. This offers a unique opportunity to study phonological representation among bilingual reader.

- Method
  - Participants: 65 Cantonese-Mandarin late bilinguals
  - Design: 2 (language) x 5 (substitution type) within-subject and within-item design
  - Materials

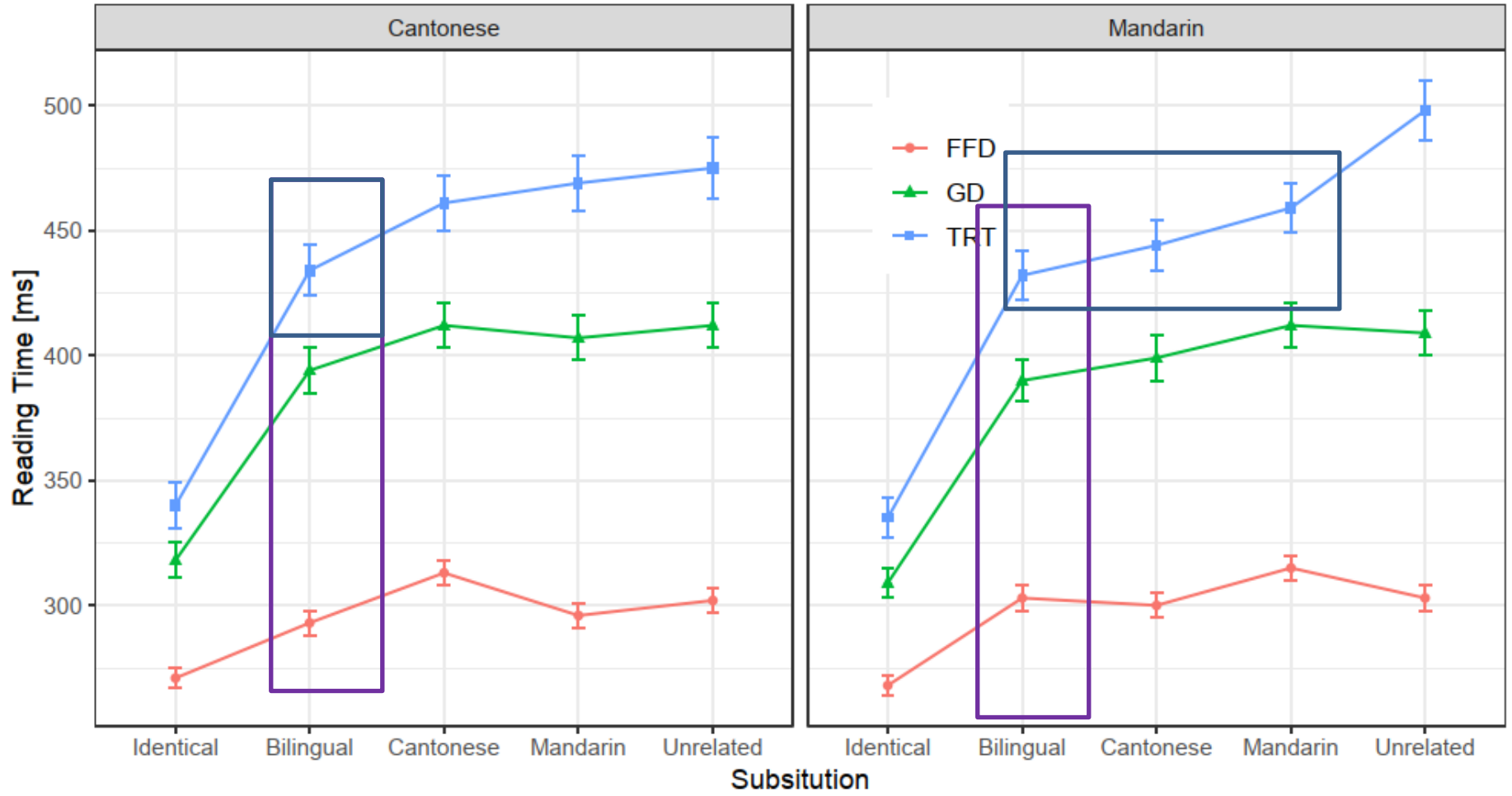
	Type of Substitution				
	Identical	Bilingual	Cantonese	Mandarin	Unrelated
Example	效	孝	拷	笑	眷
Can. Pronunciation (Jyutping)	haau	haau	haau	siu	gyun
Man. Pronunciation (Pinyin)	xiao	xiao	kao	xiao	juan
Log Frequency	3.16 (0.79)	3.01 (0.73)	2.94 (0.75)	2.90 (0.70)	3.02 (0.69)
N. strokes	10.8 (3.6)	11.4 (3.5)	11.5 (3.9)	11.1 (4.8)	10.9 (4.2)
Plausibility	4.01 (.56)	2.38 (.51)	2.39 (.46)	2.38 (.48)	2.30 (.40)
Predictability	.49 (1.24)	.03 (.23)	.00 (.00)	.00 (.00)	.01 (.08)

- Example sentence

他們需要在提高效益這個方面下工夫。	Identical Substitution
他們需要在提高孝益這個方面下工夫。	Bilingual Homophone (C+M+)
他們需要在提高拷益這個方面下工夫。	Cantonese-Only Homophone (C+M-)
他們需要在提高笑益這個方面下工夫。	Mandarin-Only Homophone (C−M+)
他們需要在提高眷益這個方面下工夫。	Unrelated Substitution (C-M-)

Translation: *They need to work on improving efficiency*.

- Results & Discussion



Summary

- In general, phonological features are activated in a later stage in Chinese word recognition during sentence reading.
- The time course of phonological activation can be influenced by different factors, such as task demand, word frequency, language proficiency.
- Our findings contrasted with the universal phonological principle (Perfetti et al., 1992), suggesting that language processing is affected by language-specific properties.

More studies on phonological processing

Pan, J., Laubrock, J., & Yan, M. (2016). Parafoveal processing in silent and oral reading: Reading mode influences the relative weighting of phonological and semantic information in Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1257-1273. doi: 10.1037/xlm0000242

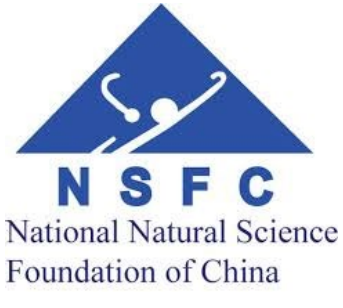
Pan, J., Laubrock, J., & Yan, M. (2021). Phonological consistency effects in Chinese sentence reading. *Scientific Studies of Reading, 25,* 335-350. doi: 10.1080/10888438.2020.1789146

Pan, J., Shu, H., Wang, Y., & Yan, M. (2015). Parafoveal activation of sign translation previews among deaf readers during the reading of Chinese sentences. *Memory & Cognition, 43,* 964-972. doi: 10.3758/s13421-015-0511-9

Thank You!


UGC
大學教育資助委員會
University Grants Committee


DFG
Deutsche
Forschungsgemeinschaft

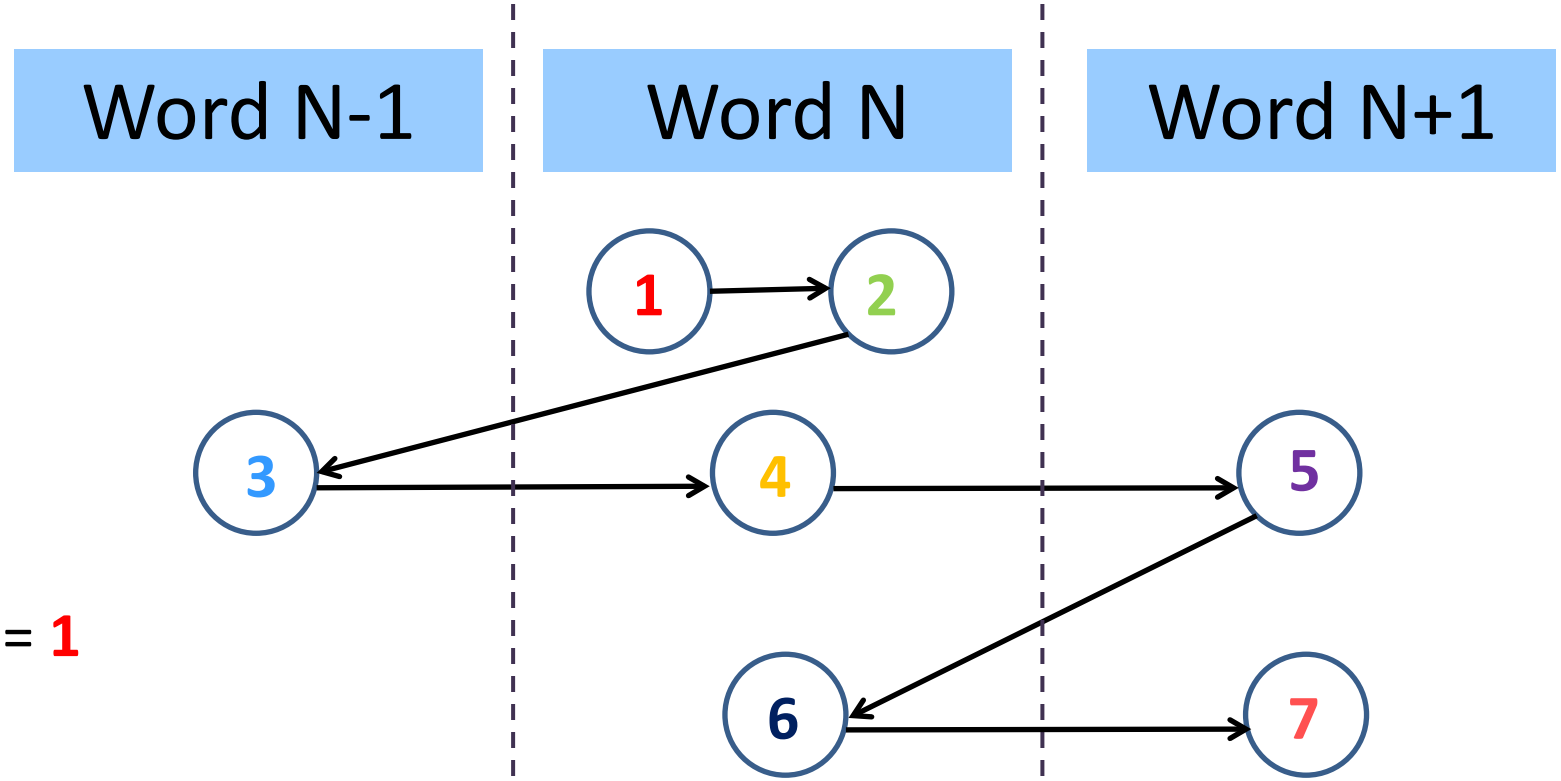

NSFC
National Natural Science
Foundation of China


UNIVERSIDADE DE MACAU
仁義禮知信
澳門大學


中国社会科学院
Chinese Academy of Social Sciences


科學技術發展基金
FDCT

Measures



First-Fixation Duration (FFD) = **1**

Gaze Duration (GD)= **1** + **2**

Go-past time (GPT) = **1** + **2** + **3** + **4**

Total reading time (TRT) = **1** + **2** + **4** + **6**