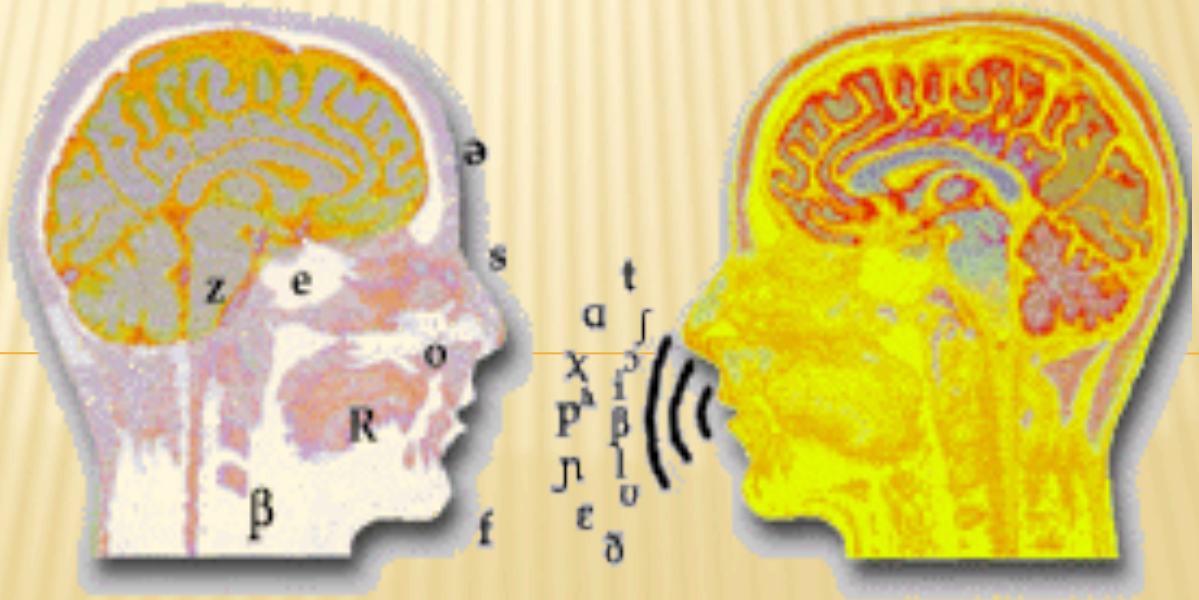
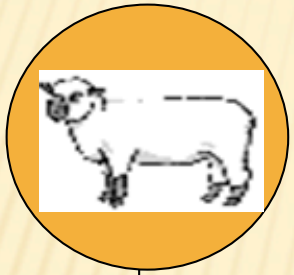
THE FREQUENCY EFFECT IN FIRST, SECOND, AND THIRD LANGUAGE VISUAL WORD RECOGNITION

Xin Wang, National University of Singapore
Ken Forster, University of Arizona



HOW IS A WORD REPRESENTED?



Conceptual

Animal, fuzzy, baaa…
Noun

Lemma: Semantics, syntax

ʃ ɪ p
SHEEP

Lexeme: Form (phonological or orthographical)

WHAT HAPPENS WHEN WE READ?

- Visual Word Recognition
  - + Automatic
  - + Fast
  - + Efficient
- An English word is "read" in 0.2 seconds by native speakers

WORD FREQUENCY EFFECT

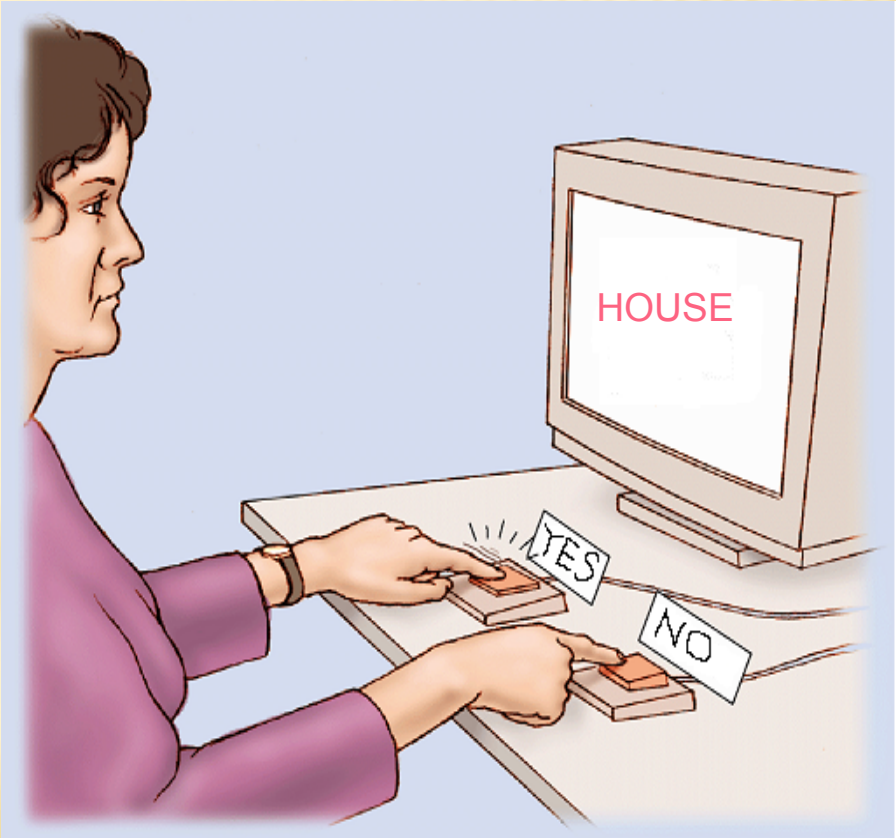
HF WORDS		LF WORDS	
bless		nasal	
relief		serge	
match		pedal	
effort		scant	
teach		oasis	
500ms	<<	550ms	RTs in Lexical Decision

THE LEXICAL DECISION TASK

Is it a word or not?

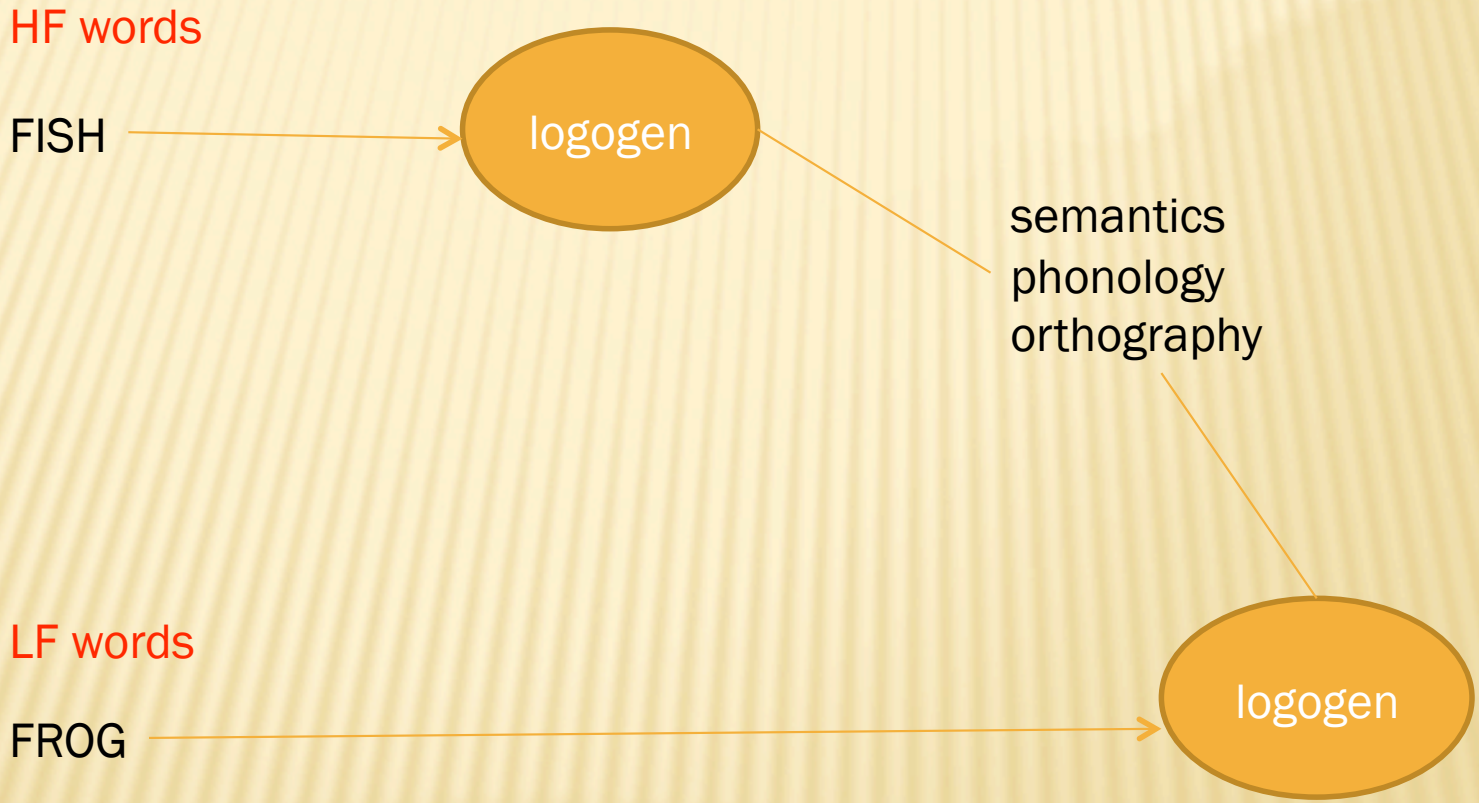
e.g.

HOUSE – Yes FLINK – No



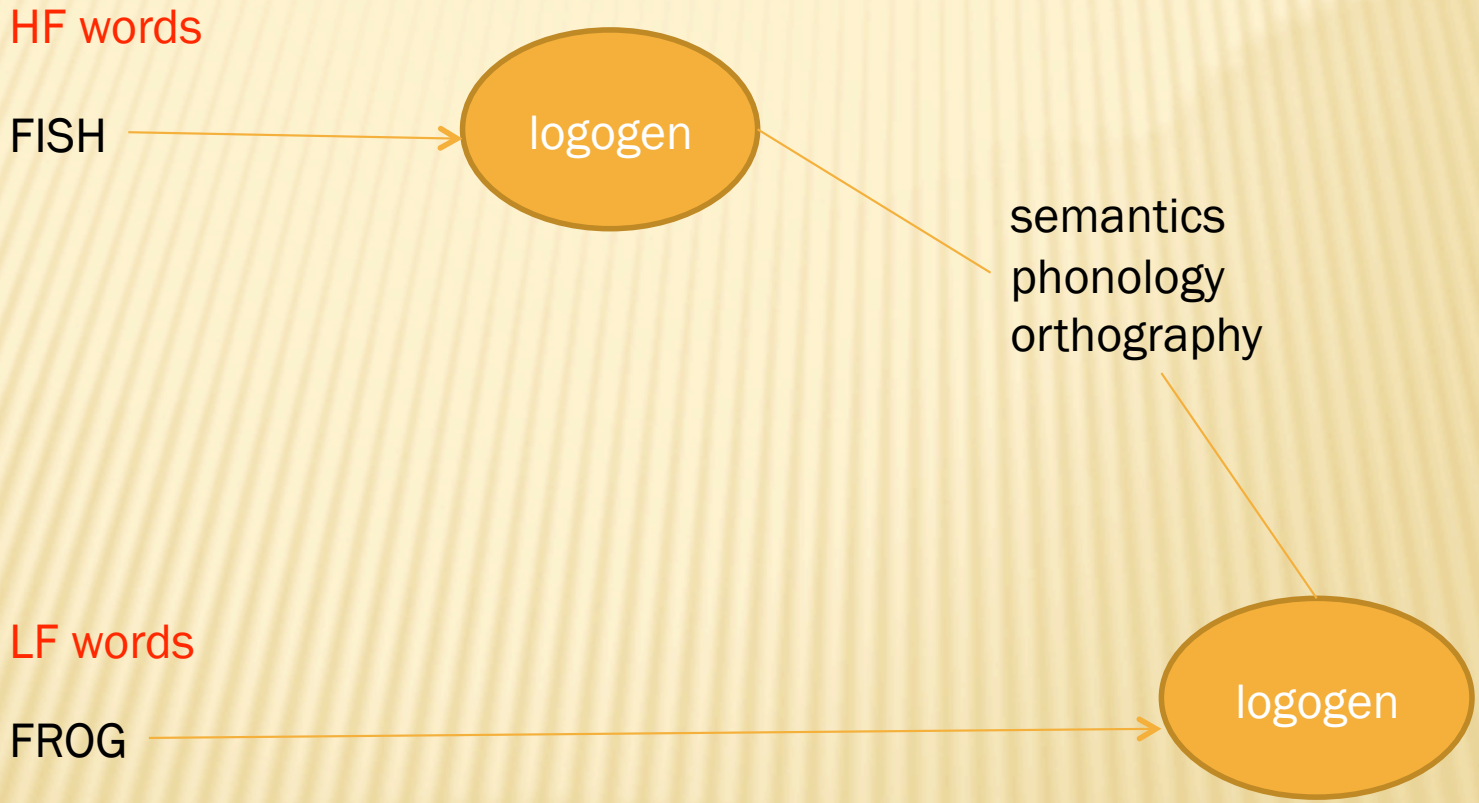
Subjects are required to make a response immediately after the presentation of a word target.

THE LOGOGEN MODEL (MORTON, 1969)



The logogen is activated by the sensory input, each encounter lowers the threshold of activation of the logogen.

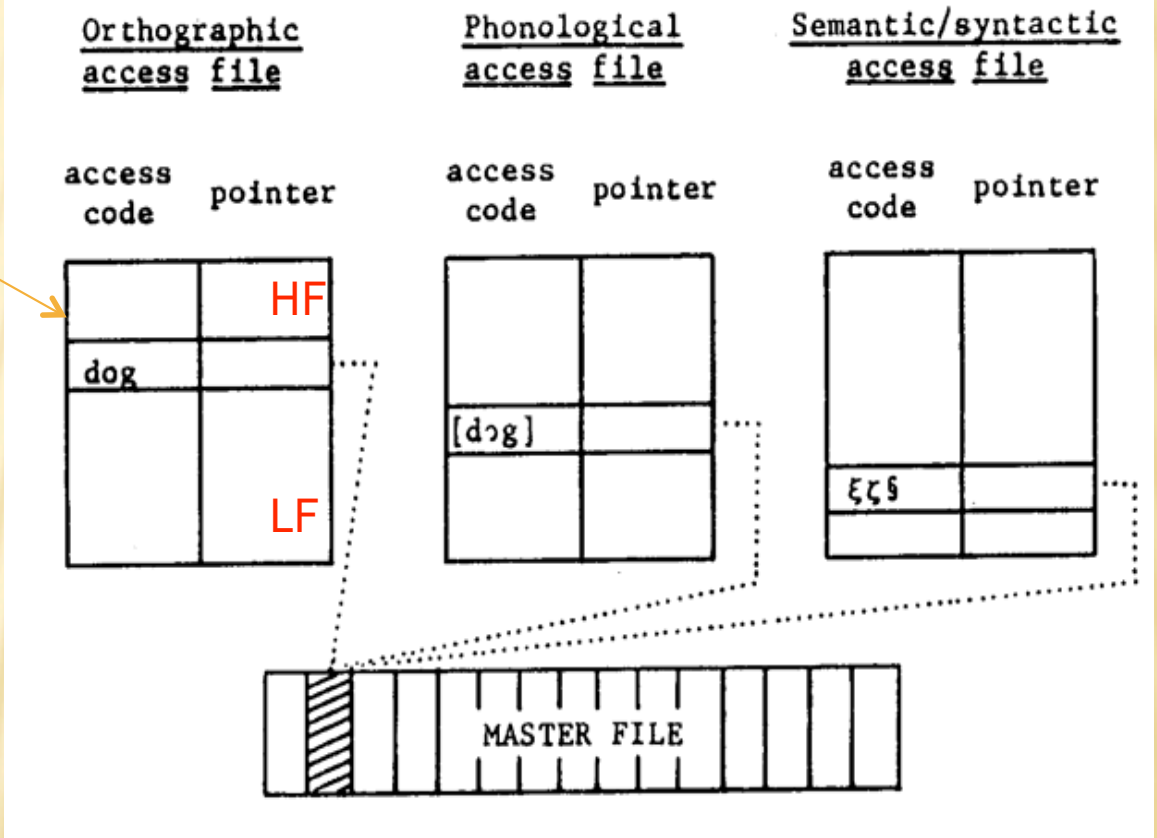
THE LOGOGEN MODEL (MORTON, 1969)



Frequency Effect is a result of repeated exposure of words --- Learning Account.

THE SEARCH MODEL (FORSTER, 1976, 1992)

Freq-based



The search terminates when the target is reached.

Frequency Effect is a result of search: HF words are searched earlier.

THE RANK HYPOTHESIS (MURRAY & FORSTER, 2004)

Rank: variable indicating the relative frequency of words

Better predictor of frequency
Effect: R 3- R1 = R5 –R3

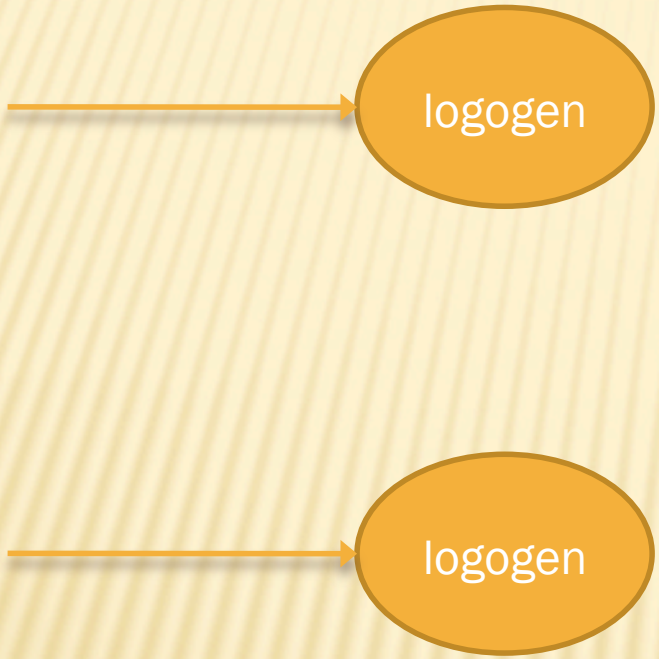
Bin A

Rank	*Freq.*	*Item*
1	10601	that
2	275	half
3	148	hair
4	104	role
5	80	lady
6	65	join
7	55	fort
8	47	ring
9	41	rare
10	36	crew
11	32	kids
etc.		

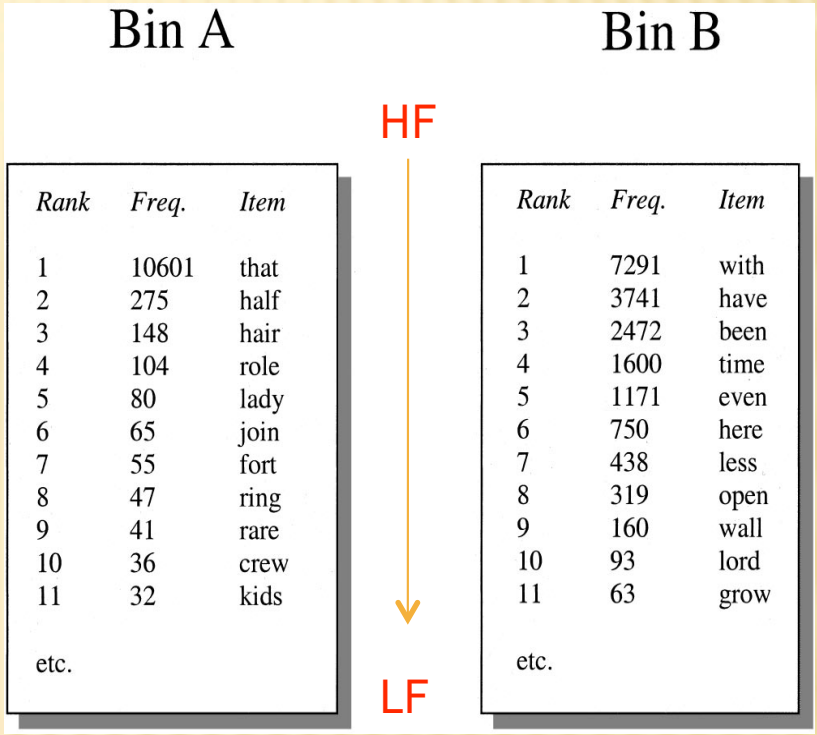
Bin B

Rank	*Freq.*	*Item*
1	7291	with
2	3741	have
3	2472	been
4	1600	time
5	1171	even
6	750	here
7	438	less
8	319	open
9	160	wall
10	93	lord
11	63	grow
etc.		

THE LOGOGEN MODEL AND THE RANK HYPOTHESIS



Parallel Processing



Rank	*Freq.*	*Item*
1	10601	that
2	275	half
3	148	hair
4	104	role
5	80	lady
6	65	join
7	55	fort
8	47	ring
9	41	rare
10	36	crew
11	32	kids
etc.		

Rank	*Freq.*	*Item*
1	7291	with
2	3741	have
3	2472	been
4	1600	time
5	1171	even
6	750	here
7	438	less
8	319	open
9	160	wall
10	93	lord
11	63	grow
etc.		

Serial Processing

The rank model is more specific in predicting the frequency effect.

THE LOGOGEN MODEL AND THE RANK HYPOTHESIS

- The learning account assumes that ‘exposure’ leads to FEs
- The Rank Model assumes that FEs should stay the same if the relative frequency does not change.
- Comparing L1 and L2, L3 FEs provides a window
  - test which account is more true
  - How L1, L2 and L3 are related in terms of FEs

RELEVANT DATA

- Production
  - Gollan, Montoya, Gera & Sandoval (2009): English dominant Spanish-English bilinguals showed a bigger FEs in picture naming in Spanish
  - Ivanova & Costa (2008): no such effect with Catalan-Spanish and Spanish-Catalan speakers
  - Could Gollan et al.'s results be due to AoA effects?

RELEVANT DATA

- Duyck et al. (2008)
  - Dutch-English bilinguals showed a larger FE in L2 (English) than L1 (Dutch)
  - Bilinguals' L1 FE is comparable to native English speakers (46ms)
  - Bilinguals' L2 EF is a lot bigger (103ms)
- The results support the learning account.
- The Rank Model needs modifications.

THE CURRENT STUDY

- Rationale
  - Replication of Duyck et al.'s results
  - The Rank Model:
    - If the frequency-based bins are language-specific: search speed differs in L1 and L2 (Chinese-English bilinguals)
    - If they are language-shared: L2 has lower rank within a bin. (Malay-English)

THE CURRENT STUDY

- The Rank Model in the bilingual/trilinguals situation:
  - Chinese-English bilinguals (orthographically different)
  - Malay-English bilinguals (orthographically same)
  - Chinese-English-Malay trilinguals
  - English native speakers
- The Learning Account of FEs
  - The more experiences with one language, the more likely the FEs is close to native speakers

EXPERIMENTS

- Materials
  - 60 Chinese words (2 characters) and 60 nonwords
  - 60 English words and 60 nonwords
  - 60 Malay words and 60 nonwords
  - Within each language, 30 HF and 30 LF words

EXPERIMENT

- Materials
  - English words were selected from ICE corpus (Singapore), nonwords from ARC database
  - Malay words were selected from the Malay Lexicon Project (the Malay Lexicon Project), nonwords were made by changing one letter to a real word
  - Chinese words were selected from Da (2004)-Chinese bigram frequency info, nonwords are illegal combination of two characters

EXPERIMENT

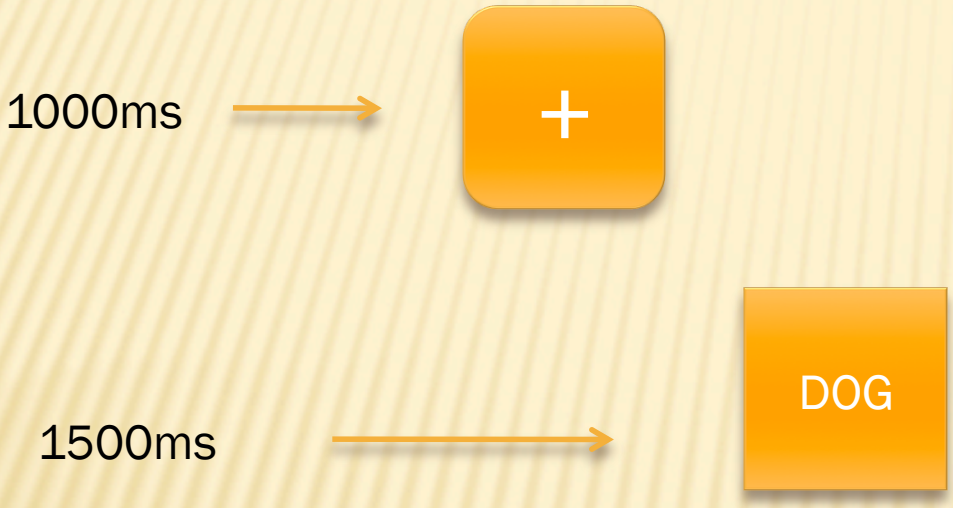
- Materials

Language	Freq	FreqPerMln	Length	N
Chinese	High	224.38	12.6 stroke C	0
	Low	9.36	12.6 stroke C	0
Malay	High	224.58	7.23	0
	Low	9.32	7.23	0
English	High	224.77	7.23	0
	Low	9.33	7.23	0

EXPERIMENTS

- Task
  - Lexical Decision
  - Counterbalanced testing
  - Each subject was tested with 2 or 3 languages
  - 10 practice trials + 120 test trials

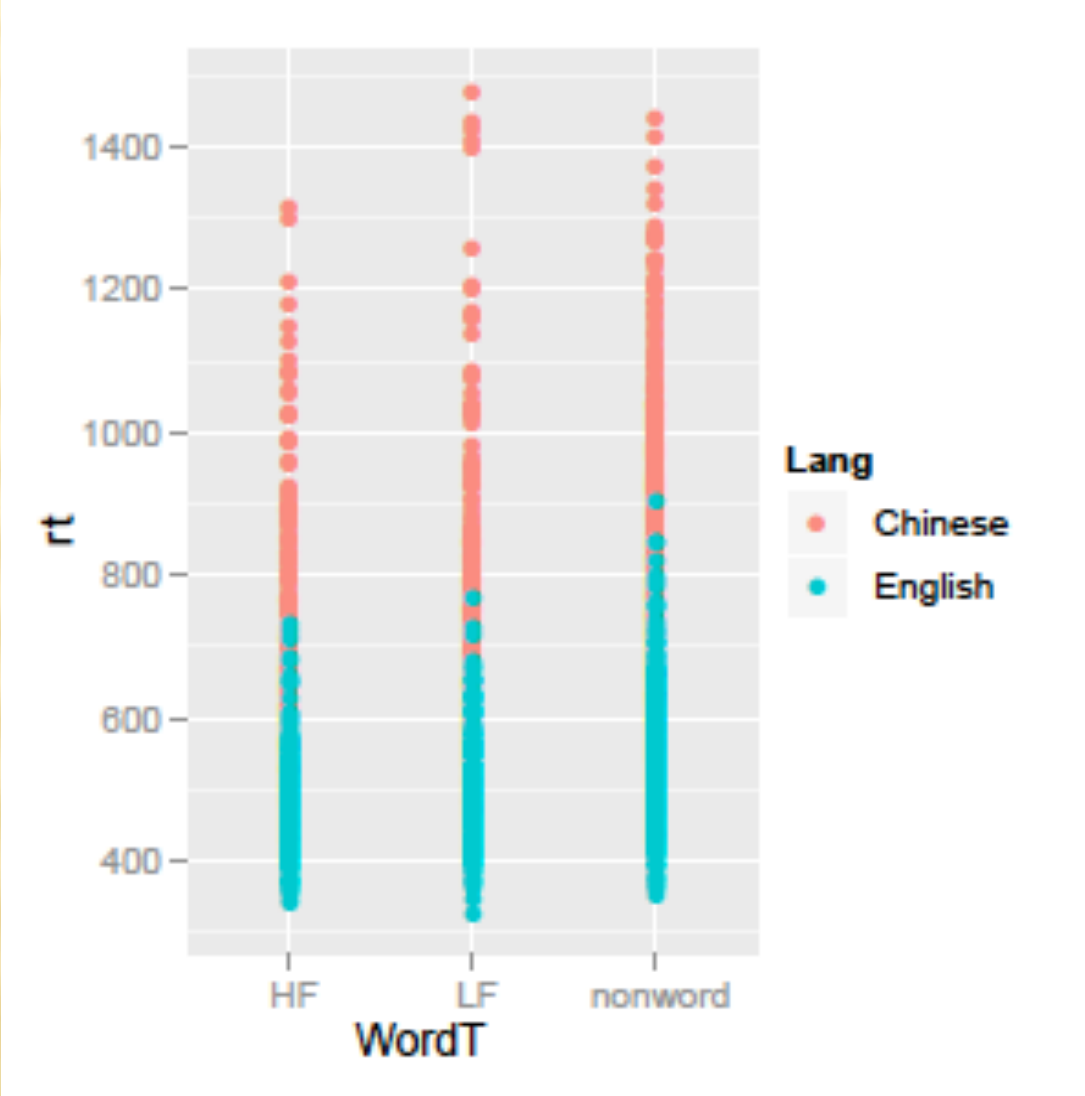
EXPERIMENT PROCEDURE



DATA TREATMENT

- Subject rejection
  - Error rates above 25%
- Item rejection
  - Cut offs: 2.5 SD
  - RTs above 1500ms, or lower than 300ms

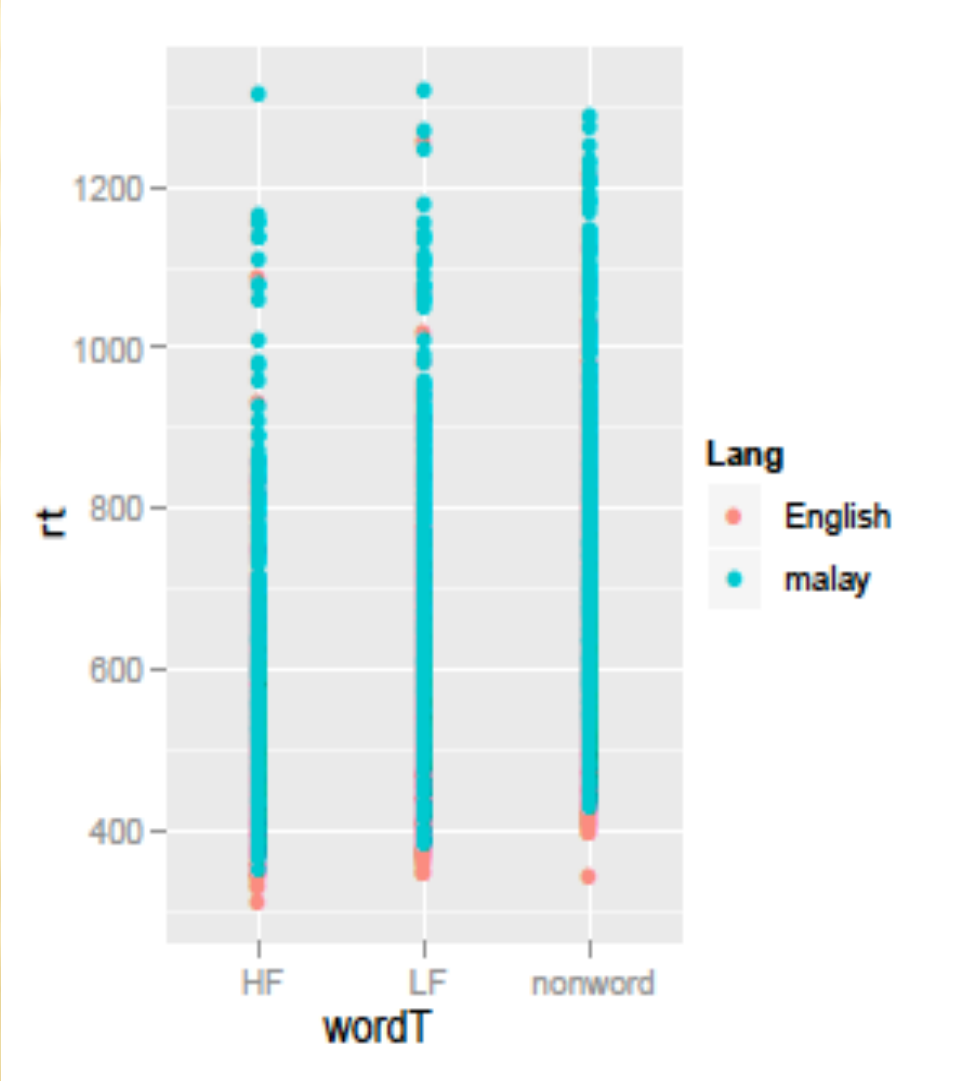
ENGLISH-CHINESE BILINGUALS



N=10
L2 Chinese: 62ms *
L1 English: 27ms*
Lang*FE

Native: 36ms*

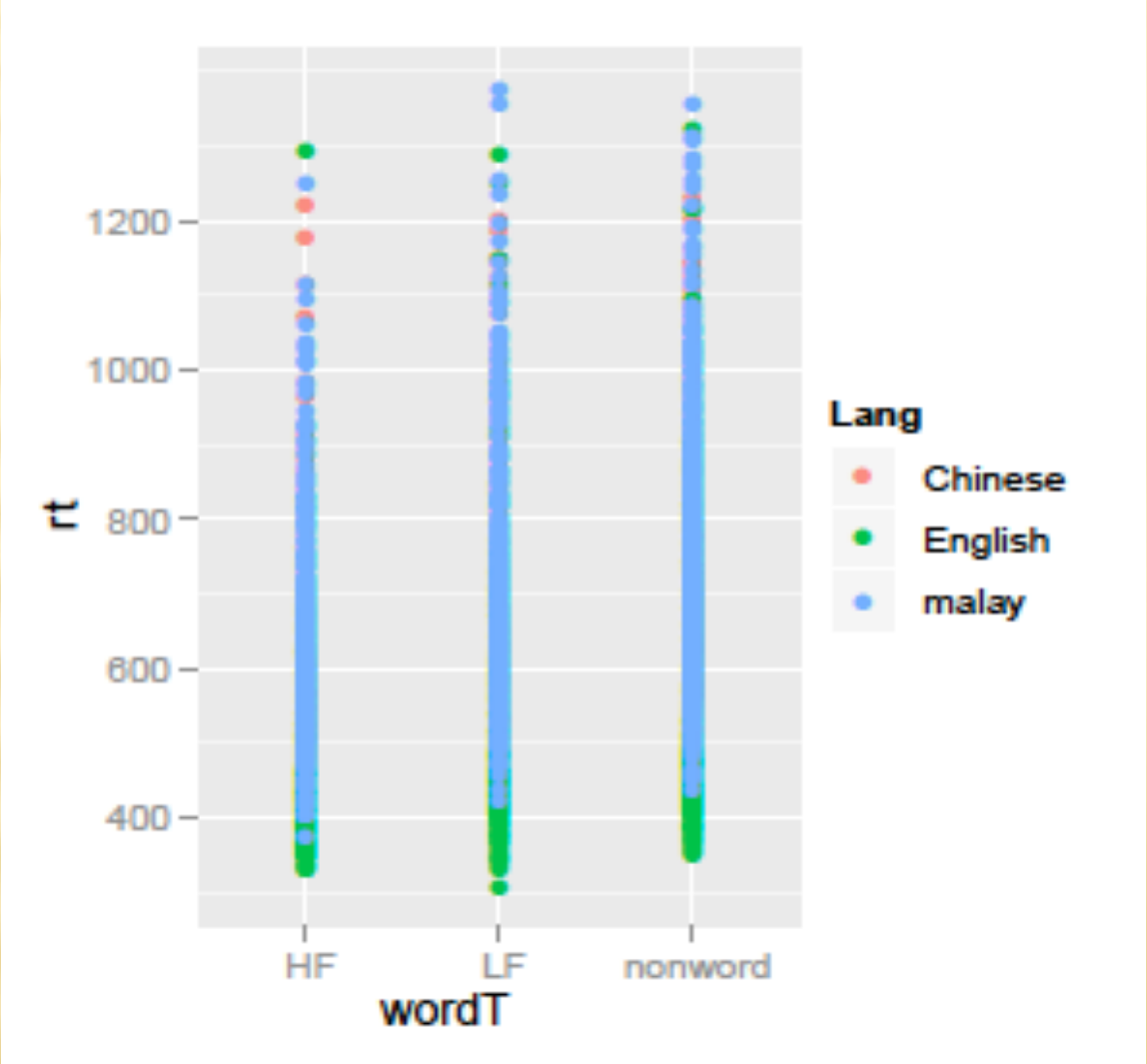
ENGLISH-MALAY BILINGUALS



N=14
L2 Malay: 82ms *
L1 English: 44ms *
Lang*FE

Data points are more overlapped than Chi-Eng bilinguals, script effect

ENGLISH-CHINESE-MALAY TRILINGUALS



N= 16
L1 Chinese: 54*
L2 English: 52*
L3 Malay: 81*

Malay* FE

DISCUSSION

- With English, bilinguals and tri-linguals showed comparable FEs to native speakers of English.
- The bilinguals data replicated Duyck et al.'s results: less dominant language showed bigger FEs.
- Trilinguals showed faster RTs in Chinese reading than bilinguals, comparable FEs between Chinese and English.
  - They read more? More proficient than bilinguals,

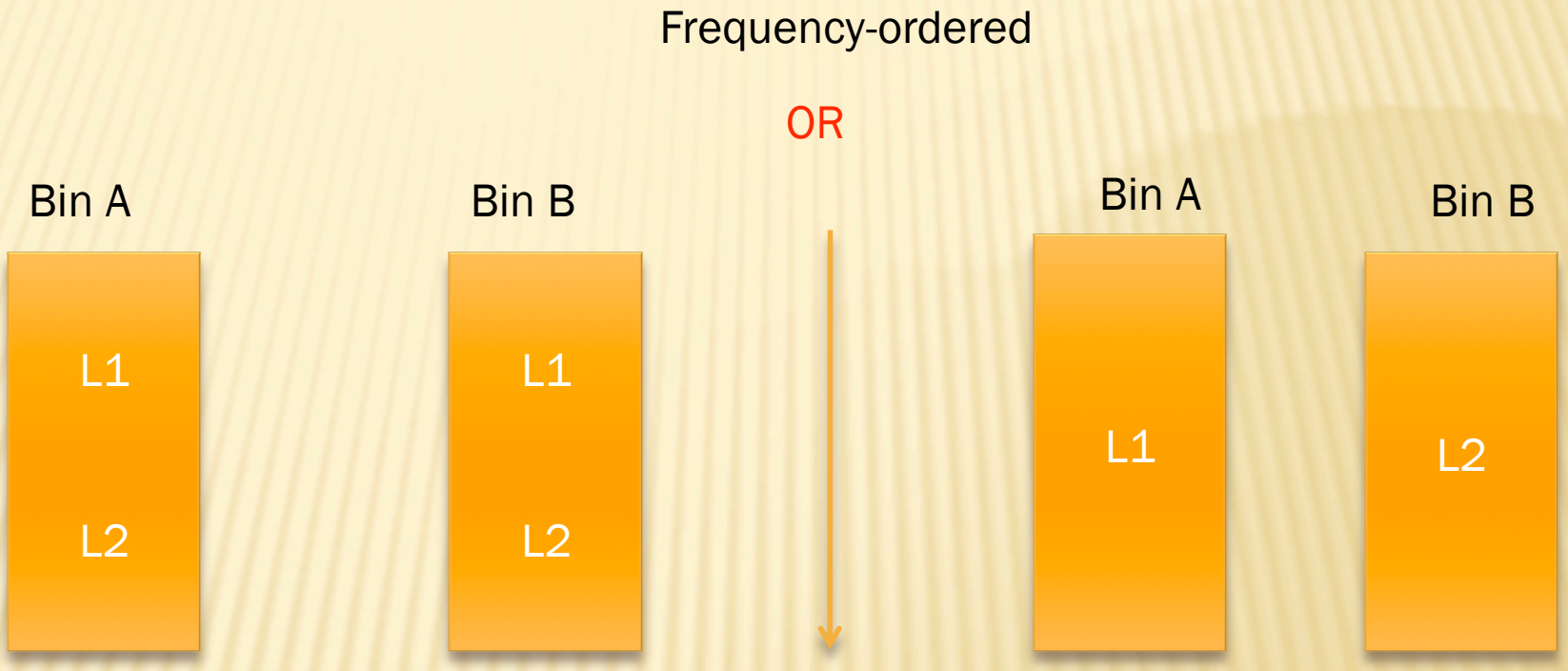
DISCUSSION

- The Learning Account (repeated exposure) is supported
  - Bilinguals/trilinguals are more experienced in reading the more dominant languages (English)
  - English subjective frequency is higher than Chinese or Malay subjective frequency for bilinguals
  - English and Chinese subjective frequency are equal, but higher than Malay for trilinguals.

DISCUSSION: THE RANK MODEL

- The bins are unlikely to be shared for Chinese for bilinguals or trilinguals.
- The mechanism for the Rank Model to be accountable for the bilingual lexicon is that L1 and L2 search speed significantly differs from each other.

THE BILINGUAL LEXICON



Malay-English Bilinguals

L1 is ranked higher than L2

Chinese-English Bilinguals

L1 is searched faster than L2

CONCLUSION

- There is a dissociation between the ‘objective’ frequency and ‘subjective’ frequency for bilinguals and trilinguals
  - L1 frequency effect is similar to native speakers
  - L2 frequency effect is confounded by the subjects
  - The trilingual situation is not only confounded by the subjects, but by their relative reading experiences in each language.

QUESTIONS?

- The End!

- Acknowledgement
  - Faculty Grant (ARF, National University of Singapore)
  - Participants