

FREQUENCY EFFECTS OF MULTIWORD SEQUENCES IN MANDARIN CHINESE

CHING CHU SUN, PETER HENDRIX, HARALD BAAZEN
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

INTRODUCTION

- Previous studies reported frequency effects of word n-grams in language comprehension (e.g. Arnon & Snider, 2010)
- Two studies that investigate frequency effects of multiword sequences in language production:
 - phrase reading experiment
 - spontaneous speech
- Language of interest: Mandarin Chinese

PHRASE READING EXPERIMENT

- Reading aloud of three-word phrases
- Randomly selected trigrams, rather than carefully controlled stimuli
- Data: 400 trigrams pronounced by 30 participants
- Dependent variables:
 - reaction times
 - acoustic durations

SPONTANEOUS SPEECH

- Taiwan Mandarin corpus of spontaneous speech (Fon, 2004)
- Data: 2700 trigrams pronounced by 55 native speakers of Mandarin Chinese
- Dependent variable:
 - acoustic durations

ADVERTISEMENT

- Chinese Lexical Database (CLD)
- Lexical database for Mandarin Chinese
- Lexical and lexical-distributional information for 48,000+ one to four character words



ANALYSIS

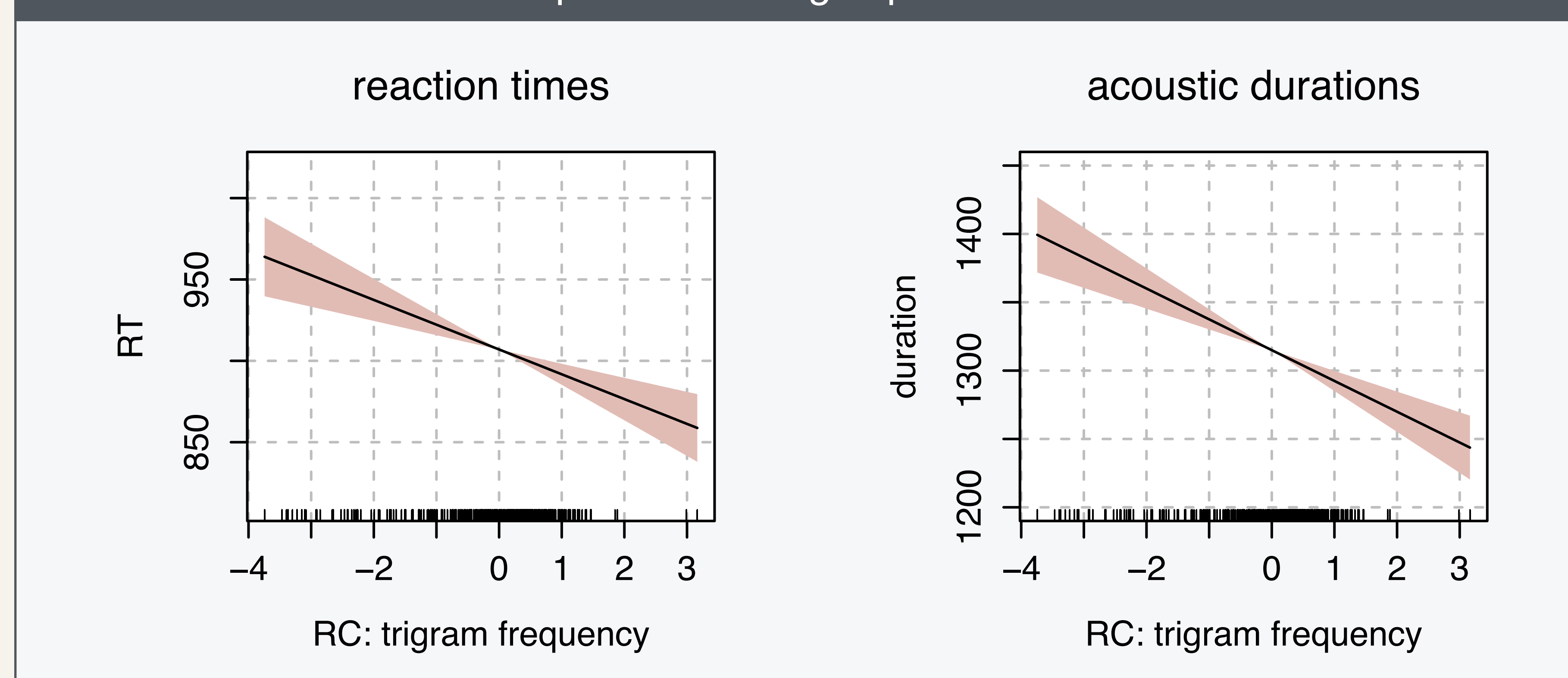
- Generalized additive mixed-effect models (GAMMs)
- Random effects for speaker and trigram
- Control for effects of visual complexity, parts of speech, and initial and final phoneme
- Principal components analysis with varimax rotation for:
 - word frequency
 - word bigram frequency
 - word trigram frequency

RESULTS

- Multiword frequency effects in both studies
- Phrase reading experiment:
 - shorter reaction times for more frequent trigrams
 - shorter acoustic durations for more frequent trigrams
- Spontaneous speech:
 - shorter acoustic durations for more frequent trigrams

VISUALIZATION OF RESULTS

phrase reading experiment



spontaneous speech

