

Frequency effects in the production of Southern Min reduplication

Jane Tsay & James Myers

Graduate Institute of Linguistics
National Chung Cheng University

Thanks to:
Ching-Yuan Lin, Chia-Fei Hung
NSC-89-2411-H-194-053

1

Overview

- Morphological production and reduplication
- Southern Min reduplication
- Estimating frequencies in Southern Min
- Frequency effects in the production of Southern Min reduplication

2

Morphological production

- Research on morphology in production is still in its infancy
 - Speech errors (Fromkin 1971, Levelt 1989)
 - Experiments (Roelofs 1996, Chen & Chen 2002)
- As with word recognition, one test for morphemes in production is to look for frequency effects (Roelofs 1996)

3

Reduplication

- Reduplication poses interesting questions for models of production
- How many morphemes, and what are they?

天天 tian-tian “every day”

- ① compound? 天天 \leftarrow 天 + 天
- ② reduplicant? 天天 \leftarrow 天 + [σ σ]
- ③ rule? 天天 \leftarrow COPY RULE \leftarrow 天

4

The production of reduplication

- There is virtually no research on the production of reduplication
- Stemberger & Lewis (1986)
 - Ewe reduplication
 - Experimentally induced phonological errors
 - Errors copied in both base and reduplicant
 - Results inconsistent with hypothesis ①
- Many issues remain to be addressed

5

Southern Min reduplication

- Southern Min (e.g. Taiwanese) has several kinds of reduplication (Yang 1991)
 - AA
 - AAA ABB AAB
 - AABB ABAB ABCC ABAC ABCB
- In our study we focused on the two “pure” types: double and triple reduplication of monosyllabic adjectives

6

Adjectival reduplication

- Like the similar pattern in Mandarin, Southern Min double reduplication moderates the meaning of the adjective
- Southern Min triple reduplication intensifies the meaning of the adjective

紅 ang “red”

紅紅 ang-ang “somewhat red”

紅紅紅 ang-ang-ang “very red”

7

Triple from double?

- The tone of the first syllable in triple forms can be derived automatically from that of double forms (Cheng 1973, Yip 1980), without irregular tone sandhi
- The first syllable in triple forms is also more variable than the other two (Du 1988)
- Are triple forms derived from double forms by adding a syllable?

8

Triple not from double

- Huang (1992) argues against this, since there are triple forms with no corresponding double forms

靈 ling “spiritually powerful”

*靈靈 *ling-ling

靈靈靈 ling-ling-ling “very spiritually powerful”

9

Reduplication and meaning

- Double and triple reduplication have different meanings
 - Word production is meaning-driven (Levelt 1989)
 - Similarity in form may be irrelevant
 - Double vs. triple reduplication may show different production patterns

10

Reduplication and the base

- Double and triple reduplicated forms share the same base
 - Activation of the base should play a role in production of reduplicated forms
 - But only if reduplicated forms are derived directly from the base (hypotheses ② or ③)

11

Frequency as diagnostic tool

- If triple forms are derived from double forms, double frequency should affect production of triple forms
- If reduplicated forms are derived from base forms, base form frequency should affect production of both types
- Caveat: Frequency effects in production are still poorly understood (Levelt, Roelofs, & Meyer 1999)

12

Goals of the present study

- Collect frequency data on Southern Min adjectives and reduplicated forms
- Examine correlations between frequencies of various forms
- Perform an on-line experiment to see how frequencies for various forms affect production latencies for various forms

13

Corpus-based frequencies

- Step 1: create the corpus
 - Step 1a: collect 80 hours of recordings
 - Step 1b: confirm 17.5 hours of recordings (230,724 words [詞詞])
 - Step 1b': transcribe 43 hours of recordings
 - Step 1b'': write spell-checker to ensure consistency in orthography (thanks, Galvin Chang)
- Step 2: calculate frequencies

14

Corpus-based frequencies

- Mean proportional frequency of monosyllabic adjectives: 0.0007
- Mean proportional frequency of double reduplicated forms: 0.00002
- Mean proportional frequency of triple reduplicated forms: 0.0000001
 - There are only 10 tokens (8 types) of triple reduplication in the confirmed corpus

15

Subjective frequency judgments

- Gernsbacher (1984) showed that subjective frequency judgments can be more accurate predictors of reaction times than objective corpus-based frequencies

16

Materials

- Began with 254 monosyllabic adjectives listed in Cheng (1981)
- Native-speaking researchers chose 89 of these that have both double and triple forms
- Ten naive speakers read all adjectives aloud; the 60 that were read without mistakes or hesitation formed final set

17

Materials

- Three lists were created
 - base forms (monosyllabic adjectives)
 - double reduplication forms
 - triple reduplication forms
- Five different versions of each list in different random orders

18

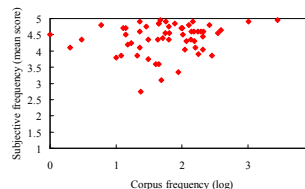
Participants and procedure

- 60 native speakers of Southern Min, judged fluent by a native-speaking assistant
- Each list given to group of 20 (4 per order)
- Judged each item on a 5-point scale
 - 1 = 最常聽到或最常講
(least often heard or said)
 - 5 = 最常被聽到或最常講
(most often heard or said)

19

Comparison with corpus

- Subjective frequencies for base forms showed small but positive correlation with log transform of corpus frequencies: $r = 0.20$



20

Comparisons among estimates

- Partial correlations among single, double, and triple form frequency estimates were all significant ($p < 0.05$)
 - base \times double: $r = 0.29$
 - base \times triple: $r = 0.26$
 - double \times triple: $r = 0.46$
- Linked processing, or sharing of “reduplicativity”?

21

Production experiment

- 30 native speakers of Southern Min, judged fluent by a native-speaking assistant
- Produced base, double, and triple forms
- Reaction times (production latencies) were measured

22

Materials and procedure

- Piloting reduced set of adjectives to 51
- All 51×3 possibilities presented in random order in single task
 - 「紅1」 \Rightarrow “ang”
 - 「紅2」 \Rightarrow “ang-ang”
 - 「紅3」 \Rightarrow “ang-ang-ang”
- Prompts thus used written characters (unavoidable, but possible confound?)

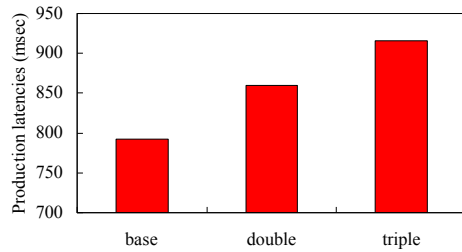
23

Analysis

- Bad responses (9% of total) were removed (too slow, incorrect pronunciations, voice-key errors)
- Three types of analysis
 - Comparison of mean RTs
 - Correlations between RTs and frequencies (base vs. double vs. triple)
 - Comparisons of RTs in controlled subsets

24

Comparison of mean RTs



25

Base frequency correlations

- For base condition, partial correlations with RTs were appropriate for both corpus frequencies and estimated frequencies

RT × corpus (log): $r = -0.40$

RT × estimates: $r = -0.23$

RT × character frequency (log): $r = 0.04$

26

RTs and estimated frequencies

- Partial correlations across all items

	RT1	RT2	RT3
Freq1	-0.28	0.20	-0.08
Freq2	0.07	-0.17	0.14
Freq3	-0.05	-0.07	-0.22

27

Controlled subsets

- Three sets of 10 adjective pairs, extracted from materials, in which only one estimated frequency varied (only base, only double, or only triple)
- Mean matched frequencies had p -values over 0.2 by two-tailed paired t -tests
- Mean nonmatching frequencies had p -values below 0.0001

28

RTs and estimated frequencies

- Comparisons in controlled subsets

	RT1	RT2	RT3
Freq1	Facilitate	(no effect)	(no effect)
Freq2	Inhibit	(no effect)	(no effect)
Freq3	Facilitate	Facilitate	Facilitate

29

Discussion

- Strongest effect was triple form frequency on triple form production: stored?
- Facilitatory effect of triple form frequency on base and double form productions: priming during task?
- Inhibitory relation between base and double forms: competition?

30

Conclusions

- Double and triple share “reduplicativity”
- Double and triple reduplication nevertheless are processed quite differently
- Reduplicated forms need not be derived from base forms in production