LINGUISTICS 105:

Morphology

December 5, 2012:
The Mental Lexicon and Processing

Announcements

- Homework #8 due now.
- Will be handed back on Friday.
  - So you want to be here to get it.
- Online course evals are open until a week from tomorrow.
  - PLEASE FILL IT OUT.
- Example citation for data in papers:

e. Jien taj-t il-ktieb lil=ha u 'l Marija.
I gave-1.SG DEF-book **to=her** and DAT Mary
"I gave the book to her and Mary." (Sadler, 2012:1)

Background: Mental Lexicon

- *So far:* modeling language as a mathematical object.
  - No claims about the reality of this object in the mind/brain.
  - Ideally, this mathematical object would be what language-users actually have in their minds.
  - However, this isn't necessarily the case.
- *Up Next:* are there behavioral effects which help us understand what representations are actually being employed by speakers in real time production/ comprehension.
  - *Today:* What is the representation used in comprehension of complex words?

The Role of Frequency

- Processing/comprehension studies have often shown the importance of MORPHEME FREQUENCY.
- FREQUENCY $=_{def}$ how often a morpheme appears in a representative sample of natural language use.
- In general, more frequent words are comprehended more quickly than less frequent words.

go (1151045) composting (1090) (source: COCA)

- Many studies investigating other properties need to control for the CONFOUND of frequency.

Background: Lookup and Parsing

- **Question:** How do speakers process morphologically complex words?
  - LOOKUP: words stored as whole forms in memory; processing is retrieval from memory.

    [run] [runs] [running]
  - PARSING: words formed on-line by morphological rule application; processing is rule application.

    [run] [run-s] [runn-ing]
- DUAL-ROUTE models involve some component of each of the two options.
- Baayen, *et al.* (1997): dual-route is needed for Dutch.

Dominance and Frequency

- CUMULATIVE STEM FREQUENCY =def the frequency of a stem with *any* affix attached.
- Baayen, *et al*. claim that, for processing of nouns, whether the singular or plural is more frequent matters for processing:
  - SINGULAR-DOMINANT: pairs where the singular is more frequent than the plural.

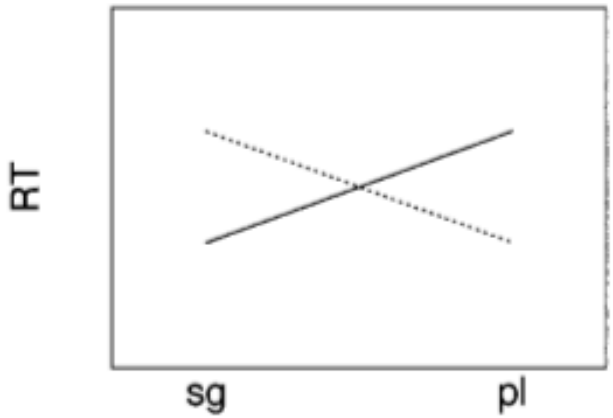
    boyfriend (9163) ~ boyfriends (1239)
  - PLURAL-DOMINANT: pairs where the plural is more frequent than the singular.

    eyes (146962) ~ eye (48659)

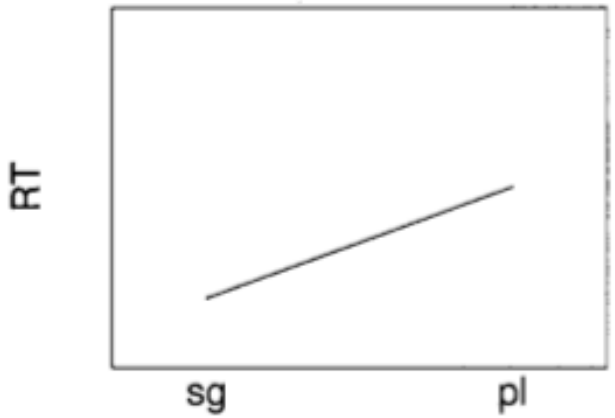
Lexical Decision and Priming

- All the experiments are LEXICAL DECISION tasks.
  - Participants shown a word and asked to decide if it is a word of Dutch or not.
  - DEPENDENT VARIABLE is reaction time (in ms).
  - Times not used if participants answered incorrectly.
- Faster reaction time = easier processing.
  - This is often correlated with whole-word access.
- Plurals are just harder to process, regardless of affixation.

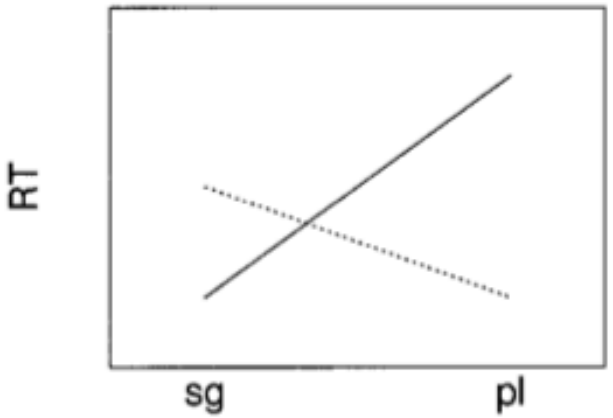
Predictions of Various Models



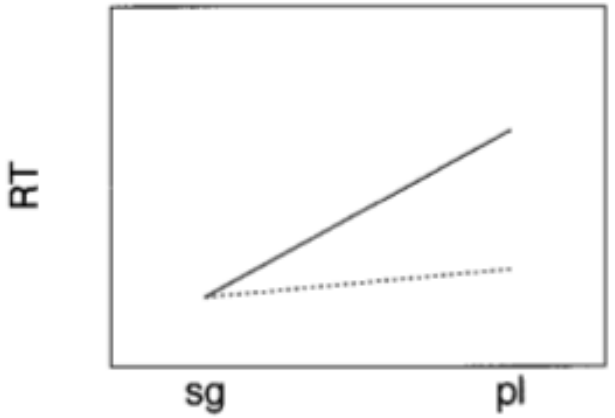
Full Listing Model



Full Parsing Model



AAM Model



Schreuder and Baayen (1995)

Experiment 1

Testing Dominance
in Lexical Decision

Materials

- Two kinds of stem frequency (= frequency of singular + plural + diminutive):
  - HIGH FREQUENCY: frequency is ~ 147/million
  - LOW FREQUENCY: frequency is ~6/million.
- Crossed with two kinds of frequency dominance.
  - SINGULAR-DOMINANT pairs
  - PLURAL-DOMINANT pairs.
- Result is 2X2 EXPERIMENTAL DESIGN:

	HIGH FREQ.	LOW FREQ.
SG DOM	SG-High	SG-Low
PL DOM	PL-High	PL-Low

Results: Experiment 1

TABLE 1

MEAN LATENCIES (IN MILLISECONDS) AND PERCENTAGES OF ERRORS FOR -EN PLURALS IN EXPERIMENT 1

Stem frequency	Dominance	Singular	Plural
High	SgDom	561 (2)	615 (6)
High	PlDom	551 (2)	558 (2)
Low	SgDom	612 (6)	708 (19)
Low	PlDom	606 (7)	645 (9)

Statistics Primer

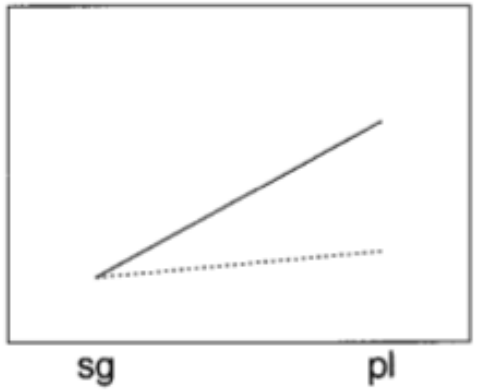
- ANOVA = ANALYSIS OF VARIANCE
  - Main effect: how much of the effect can be explained by changing one variable?
  - Interaction effect: how much of the effect can be explained by changing both variables at once?
- Results come with a *p-value*: a probability that the conclusion is incorrect and there is in fact no effect.
  - $p < 0.05$ is considered good enough for SIGNIFICANCE in linguistics.
  - Many results $p < 0.01$ and less.

Results: Stats

- Main effects in Experiment 1:
  - NUMBER: singulars faster than plurals ($p < .001$)
  - FREQUENCY: high-frequency faster than low frequency ($p < .001$).
  - DOMINANCE: SgDom faster than PlDom ($p < .001$)
- Interaction effects:
  - NUMBER X FREQUENCY ($p < .001$)
  - NUMBER X DOMINANCE ($p < .001$)

Conclusions - Exp. 1

- Results consistent with S&B's model
- Plurals with higher surface frequency than their singulars processed much faster (about as fast as singulars!).
- High frequency singulars in singular dominant pairs processed faster.
- Some plural forms are stored whole-word, especially when they are higher frequency than the singular.
- All of these numbers are a bit high compared to lexical decision results in other languages -- why?



Experiment 2

Keeping Singular Frequencies Constant

Materials

- All nouns matched for surface frequency of the singular, with three conditions:
  - Singular more frequent than plural (SGDOM)
  - Plural more frequent than singular (PLDOM)
  - Equal frequencies for singular and plural (NEUTRAL)
  - *Expectation*: summed frequency differs, so they expect different RTs for singulars.
- Result: 1X3 EXPERIMENTAL DESIGN

Results

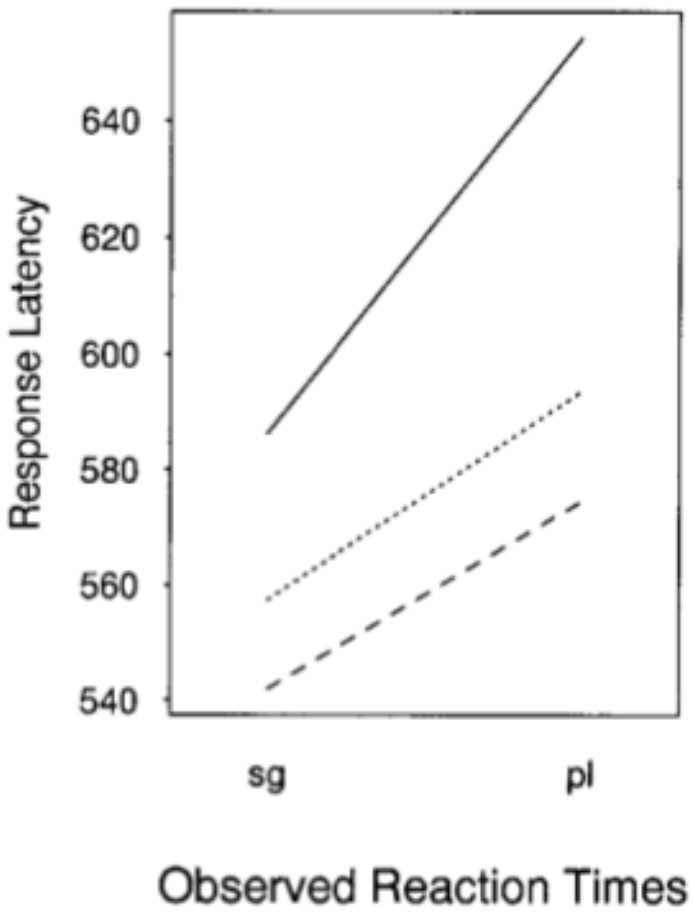
TABLE 2

MEAN LATENCIES (IN MILLISECONDS) AND PERCENTAGES OF ERRORS FOR -EN NOUN PLURALS AND CORRESPONDING SINGULARS IN EXPERIMENT 2

Dominance	Singular	Plural
PlDom	542 (4)	575 (3)
Neutral	557 (4)	593 (6)
SgDom	586 (7)	654 (16)

Conclusions - Exp. 2

- If one holds singular frequency constant, plural frequency is what matters.
- Lowest-frequency plurals had the slowest response times.
- Experiments 1&2: cumulative singular and plural frequency modulates RT to singular.
  - Also determines times for some plurals.



Experiment 3

Bringing Verbs into the Picture

Materials

- -*en* in Dutch also marks plural number on verbs and appears on infinitives. Could the POLYFUNCTIONALITY of -*en* cause slowed reaction time?
  - Most uses are verbal. Do we expect no longer RT with verbal -*en*?
- Experiment 3 crosses CATEGORY x NUMBER:
  - CATEGORY = {Verb, Noun}
  - NUMBER = {Singular, Plural}

Results

TABLE 3

MEAN LATENCIES (IN MILLISECONDS) AND PERCENTAGES OF ERRORS FOR -EN NOUN AND VERB SINGULARS AND PLURALS IN EXPERIMENT 3

Word category	Singular	Plural
Noun	545 (2)	611 (6)
Verb	603 (7)	612 (10)

Conclusions - Exp. 3

- Just like exp. 1: Plurals have much higher RTs than their corresponding singulars.
  - These are all singular dominant stems.
- Plural verbs, however, are as fast as singulars.
- *Idea:* there is one representation for *-en*, and it is default verbal.
- Nominal *-en* requires a stage to "check subcategorization," and that is why nominal *-en* is slower than verbal *-en*.