(In)dependence of lexical and syntactic production: that-reduction and omission in spontaneous speech
T. Florian Jaeger, Neal Snider, Laura Staum, and Dan Jurafsky (Stanford University)

Contact: tfl@stanford.edu (Florian)

Do syntactic and phonetic reduction differ in how they are affected by:
- probabilistic factors?
- experienced and/or anticipated production difficulty?
- phonological factors?

The Question
What is the role of probability and production difficulty in human language production? The case of reduction.

The Experiment

- Collected examples from the 800,000 word syntactically annotated Switchboard corpus of American English telephone conversations.
- We extracted 6,648 complement clauses and 3,465 relative clauses from the corpus.
- 1,265 complement clauses (19%) and 1,410 relative clauses (40%) contained complementer or relativizer that.
- Durations were automatically extracted from time-aligned orthographic transcriptions (Stanford-Edinburgh Paraphrase Link Project).
- Time aligned transcripts at a 10ms resolution (Bobkova et al., 1998).
- Earlier work on phonetic reduction averaged over different uses (lemmas) of that, you i.e. Switchboard or >40,000 words; Bell et al., 2003.
- Potentially different lemmas (complementizer and relativizer that) are investigated separately.

Methods

- Following/Following Phone
  - A following consonant increases production difficulty
  - A following stressed syllable affects both phenomena in the same way
- Adjacent disfluency affects both phenomena in the same way: in disfluent contexts, that is more likely to be produced more frequently and longer
- Stress affects phonetic reduction and syntactic reduction in the same way. Segmental context doesn’t
- Omission and reduction are similar (and may partly be driven by similar mechanisms)
- Omission is NOT an extreme form of reduction

Results/ Conclusions

- Recent (but separate) work on these reduction phenomena suggests that they are both probabilistic:
  - that is phonetically reduced when predictable (Jurafsky et al., 2003, Bell et al., 2003)
  - that produced less frequently if the relative/complement clause is more predictable (Wasow et al., in press, Jaeger et al., 2005 & Jaeger and Levy in progress)
  - Perhaps omission (syntactic reduction) is just an extreme form of phonetic reduction (see Luba, 1966)

- Do syntactic and phonetic reduction differ in how they are affected by:
  - probabilistic factors?
  - experienced and/or anticipated production difficulty?
  - phonological factors?

Factors

- Predictability factors
  - Conditional probability of RC/CC given the preceding word
  - Conditional probability of RC/CC given its first word
  - Raw probability of the preceding word
  - Raw probability of the following word

Disfluency factors

- Edit Phrase
  - Well I guess I know I guess I can see it from several different perspectives.
  - If you think I really like [for chicken enchiladas]
  - I think [that the gang violence is scary everyone to deal too]

Factors

- Predictability factors
  - Conditional probability of RC/CC given the preceding word
  - Conditional probability of RC/CC given its first word
  - Raw probability of the preceding word
  - Raw probability of the following word

Omission of that

- Complement Clauses
  - Predicability Factors in Omission
  - Fluency Factors in Omission
  - Phonological Factors in Omission

Reduction of that

- Complement Clauses
  - Predicability Factors in Reduction
  - Fluency Factors in Reduction
  - Phonological Factors in Reduction

Results: Comparison

- Omitted: Shorter/Longer
- Present: Shorter/Longer

Conclusions

- Syntactic reduction is largely sensitive to probabilities given preceding items, while phonetic reduction is primarily influenced by following items
- Adjacent disfluency affects both phenomena in the same way: in disfluent contexts, that is more likely to be produced more frequently and longer
- Stress affects phonetic reduction and syntactic reduction in the same way. Segmental context doesn’t
- Omission and reduction are similar (and may partly be driven by similar mechanisms)
- Omission is NOT an extreme form of reduction
References


For further references on *that*:

[www.stanford.edu/~lstaum/thatbib.htm](http://www.stanford.edu/~lstaum/thatbib.htm)