1pSC4. The Perception of Rate Induced Resyllabification in English

Kenneth de Jong, Byung-jin Lim, and Kyoko Nagao

(kdejong@indiana .edu, bylim@indiana.edu, knagao@indiana.edu)

Department of Linguistics, Indiana University

Main Points

*This paper replicates Stetson's perceptual observation of **rate-induced resyllabification:**

Codas repeated at fast rates -> Onsets

* **Naïvelisteners** perceive it

* The perception is robust to voicing contrasts and stimulus editing techniques

* **However …**it is not entirely categorical: fast rate items are perceived somewhere between the two syllabification types

* Acoustic correlates of syllable affiliation suggest that listeners rely heavily on indicators of juncture. Fast repetition largely removes these indicators

Introduction

Stetson (1951 and much earlier):
* Early articulatory work
* Repeated VC forms (such as 'eep') at fast rates
* Gets perceived as CV (such as 'pea')
* Replicated by Tuller & Kelso (1991)

de Jong (In press a, In press b):
* Acoustic production study
* Repeated codas become similar to onsets at fast rates

But
* Repeated onsets also become similar to codas at fast rates

And
* Fast rate onsets and codas are not neutralized. E.g., F2's are different between vowels with onsets and vowels with codas both at fast and slow rates
* Rate scaling affects onset temporal structure proportionally, while coda temporal structure is resistant to changes
* Phonemic voicing of the consonant also restricts how coda temporal structure is changed

In the current study, we …
* Examine naïve listeners' responses to repeated onsets and codas at a range of rates
* Generalize responses across editing techniques
* Generalize responses across voiced and voiceless stops
* Determine acoustic correlates of syllabification perception

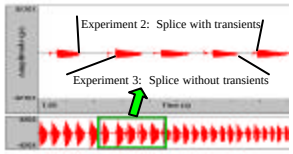
Stimuli

* Repetition rate controlled with a metronome, start slow (450 ms/ syll), and increase throughout trial (to 200 ms/ syll)
* Production rates range across stable production range from 2Hz to 5Hz as found in Nelson et al. (1984)

* Productions of four items

	CODAS	ONSETS
'voiced':	'eeb'	'bee'
'voicelesss':	'eep'	'pea

* Stimuli = 3-syllable pieces spliced from overall utterance:



Acknowledgements

Work supported by the NIDCD (grant #R03 DC04095-01A2) and by the NSF (grant #BCS-9910701)

Experiments

Experiment 1: Open-set labeling

Procedure
* Present extreme stimuli from beginning and end of trials
* 22 Listeners
* Open listening environment, subjects run as a group
* Asked to write down the repeated syllable

Results
* 72% of responses were one of the intended syllables
* 23% of responses split consonant into two consonants (e.g.,: 'eep' -> 'beep')
* Responses showed rate resyllabification of codas (3% onset response at slow rates -> 52% at fast rates)
* *d-prime* conversion indicates large bias toward onsets, and low detectability of both onsets and codas at fast rates

	SLOW RATE		FAST RATE	
Response	Onsets	Codas	Onsets	Codas
CV (Onset)	76%	3%	59%	52%
VC (Coda)	2%	88%	3%	6%
CVC (Both)	20%	5%	34%	35%
Other	2%	5%	4%	7%

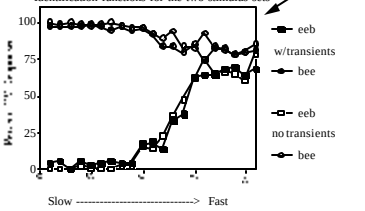
Experiment 2: Closed-set perception

Procedure
* Present 336 three-syllable stimuli singly over headphones
* 18 Listeners from IU population
* 4 choice identification with confidence rating as below

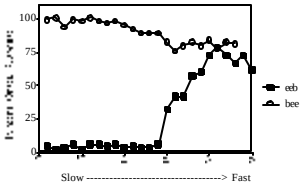
Results: Identification Shifts
* Horizontal scale indicates proportion of onset responses
* Listeners label intended codas at fast rates (to the right) as onsets
* Labels shift from coda (low onset responses to the left) to onset responses suddenly around stimulus number 15
* **However,** fast rate tokens are not identified 100% as onsets
* As for intended onsets (top function):
  Labeled 100% as onsets at slow rates (to the left),
  20% not labeled as onsets at slow rates (to the right)
* Shift in labeling also occurs around stimulus 15
* Both intended onsets and intended codas at fast rates elicit roughly the same proportion of CV and VC responses

Results: Voicing Generality
Below are Identification Functions for voiceless consonants
* The pattern is the same as for voiced consonants (above)

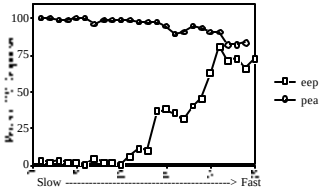
Results: Stimulus transients
Identification functions for the two stimulus sets



Results: Identification shifts
Average identification functions for voiced consonants



Results: Identification shifts
Average identification functions for voiceless consonants



Experiment 3: Stimulus Transients

Procedure
* The existence of a stop release at the final edge of coda stimuli might account for incomplete resyllabification
* To test this, new stimuli were spliced without transients
* Stimuli were presented to 18 listeners

Results
* Results indicate no appreciable difference
* Proportion of Onset Responses for each stimulus in first set (with transients) was regressed against matched stimulus in second set (without transients)
* Proportion of Onset Responses correlate linearly with r-squared = 0.953, m = 0.93

Predictors of Perceived Syllabification

Variables to be related

CONTROL VARIABLES
intended syllabification (onset or coda)
voicing ('voiced' or 'voiceless')
rate (numerical location in overall utterance)
ACOUSTIC MEASURES
(various measures of the following taken from the literature)
- duration of the syllable
- duration of sub-portions of the syllable
- duration of segments expressed as various proportions
PERCEPTUAL LABELLING
syllabification (onset or coda)

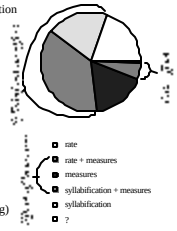
Procedure
* Regress
CONTROL VARIABLES -> PERCEPTUAL LABELS
CONTROL VARIABLES -> ACOUSTIC MEASURES
ACOUSTIC MEASURES -> PERCEPTUAL LABELS

Results
Variance in perception of syllabification throughout the corpus is accounted for as presented to the right:

* Intended syllabification (left-sloping shading) accounts for approximately 2/3's of variance in response

* Acoustic measures account for 67% of variance, 25% is independent of intended syllabification (dark shading) 5% is shared with rate (dark right-sloping shading)

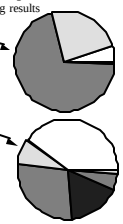
* Voicing accounts for nothing in syllabification perception



The Rate Effect:
Dividing the corpus in half and analyzing slow and fast halves separately yields the following results

Slow rates:
* 90% of response due to intended syllabification
* 65% of this is indexed by acoustic measures
* rate accounts for nothing

Fast rates:
* intended syllabification accounts for only 1/3 of response
* rate effect found in faster half of the corpus
* Acoustic measures account for 1/2 of variance, including most of the rate effect



Correlates of Syllable Affiliation

Useful Durational indicators of intended syllabification

% occlusion = consonant closure in proportion to v -to-v gap (c.f. Boucher, 1988)
voice latency = time between release and onset of vowel
glottalization = duration of creaky vowel onset
CV-ratio = consonant closure in proportion to vowel duration
%voicing = proportion of closure with voicing

* First three measures indicate the presence of juncture
* Last two are indicators of the structure of the syllable itself

Splitting corpus by rate reveals:
* The same measures account for perception at both rates
* Connection between intended syllabification and measures is obscured at fast rates
* Rate correlates with some measures only at the fast rates

Acoustic measures against intended syllables		Perception against acoustic measures
SLOW RATES		
Intended -> 63.1% of %occlusion	->	59.1% of Perceived
-> 58.2 % of voice latency	->	56.1% of
-> 20.2 % of CV -ratio	->	20.6% of
-> 20.2 % of %voicing	->	17.0% of
-> 20.2 % of glottalization	->	15.3% of
FAST RATES		
Intended -> 7.0% of %occlusion	->	18.9% of Perceived
-> 6.6% of voice latency	->	20.9% of
-> 10.8% of glottalization	->	29.8% of
-> 12.1% of CV -ratio	->	17.6% of
-> 19.8% of % voicing	->	19.9% of
FAST RATES: EFFECT OF RATE		
Rate -> 14.4% of voice latency	->	20.9% of Perceived

Discussion
* Measures which perform consistently across rates seem to be direct indicators of juncture
* These measures are affected by rate at fast rates
* Indicators of syllabic quality are less affected by rate at fa st rates, but seem not to be as heavily weighted
* Hence, fast repetition rates obscure juncture markers, contributing to perceptual neutralization of onsets and codas

Summary

* Naïve listeners perceive resyllabification at fast rates
* Perceptions are robust to splicing techniques, indicating that people are good at extracting syllable affiliation
* Consonant voicing doesn't affect syllabification perception
* Correlates of syllable affiliation involve both the quality of the syllable itself and juncture marking
* Rate resyllabification seems to be due to a loss of juncture - marking aspects of the signal at fast rates, and possibly a general bias toward identifying onsets

References Cited

BOUCHER, V.J. (1988). A parameter of syllabification for VstopV and relative timing invariance. *Journal of Phonetics*, **16**: 299 – 326

DE JONG, K.J. (in press a). Effects of Syllable Affiliation and Consonant Voicing on Temporal Adjustment in a Repetitive Speech Production Task. *Journal of Speech, Language, and Hearing Research*

DE JONG, K.J. (in press b). Rate -induced Resyllabification Revisited. *Language and Speech*, **44(2)**

NELSON, W.L., PERKELL, J.S., & WESTBURY, J.R. (1984). Mandible movements during increasingly rapid articulations of single syllables: Preliminary observations. *Journal of the Acoustical Society of America*, **75**: 945 – 951

STETSON, R.H. (1951). *Motor Phonetics: Nederlandaises de Phonetique Experimental*, **3**: 1-216. Amsterdam: North- Holland

TULLER, B., and KELSO, J.A.S. (1991). The production and percep tion of syllable structure. *Journal of Speech and Hearing Research*, **34**: 501508