Nasal place assimilation trades off inferrability of both target and trigger words

Rory Turnbull, Scott Seyfarth, Elizabeth Hume and T. Florian Jaeger (2018)

Presenter: Mark Mikleu, Date: 24.01.25

UNIVERSITÄT DES SAARLANDES

Assimilation 1

- Sounds influenced by neighboring sounds
- Sounds adopt phonetic properties from their neighbors
- Example: ***ten bucks*** being pronounced as ***tem bucks*** [nb] ~ [mb] results in simplification

Place → / Manner ↓	Labial			Coronal			
	Bilabial	Labio-dental	Linguo-labial	Dental	Alveolar	Post-alveolar	Retro-flex
Nasal	m̥ m	ɱ̊ ɱ	n̼		n̥ n		ɳ̊ ɳ
Plosive	p b	p̪ b̪	t̼ d̼		t d		ʈ ɖ

Place → / Manner ↓	Labial			Coronal			
	Bilabial	Labio-dental	Linguo-labial	Dental	Alveolar	Post-alveolar	Retro-flex
Nasal	m̥ m	ɱ̊ ɱ	n̼		n̥ n		ɳ̊ ɳ
Plosive	p b	p̪ b̪	t̼ d̼		t d		ʈ ɖ

Source: https://en.wikipedia.org/wiki/Consonant

Assimilation 2

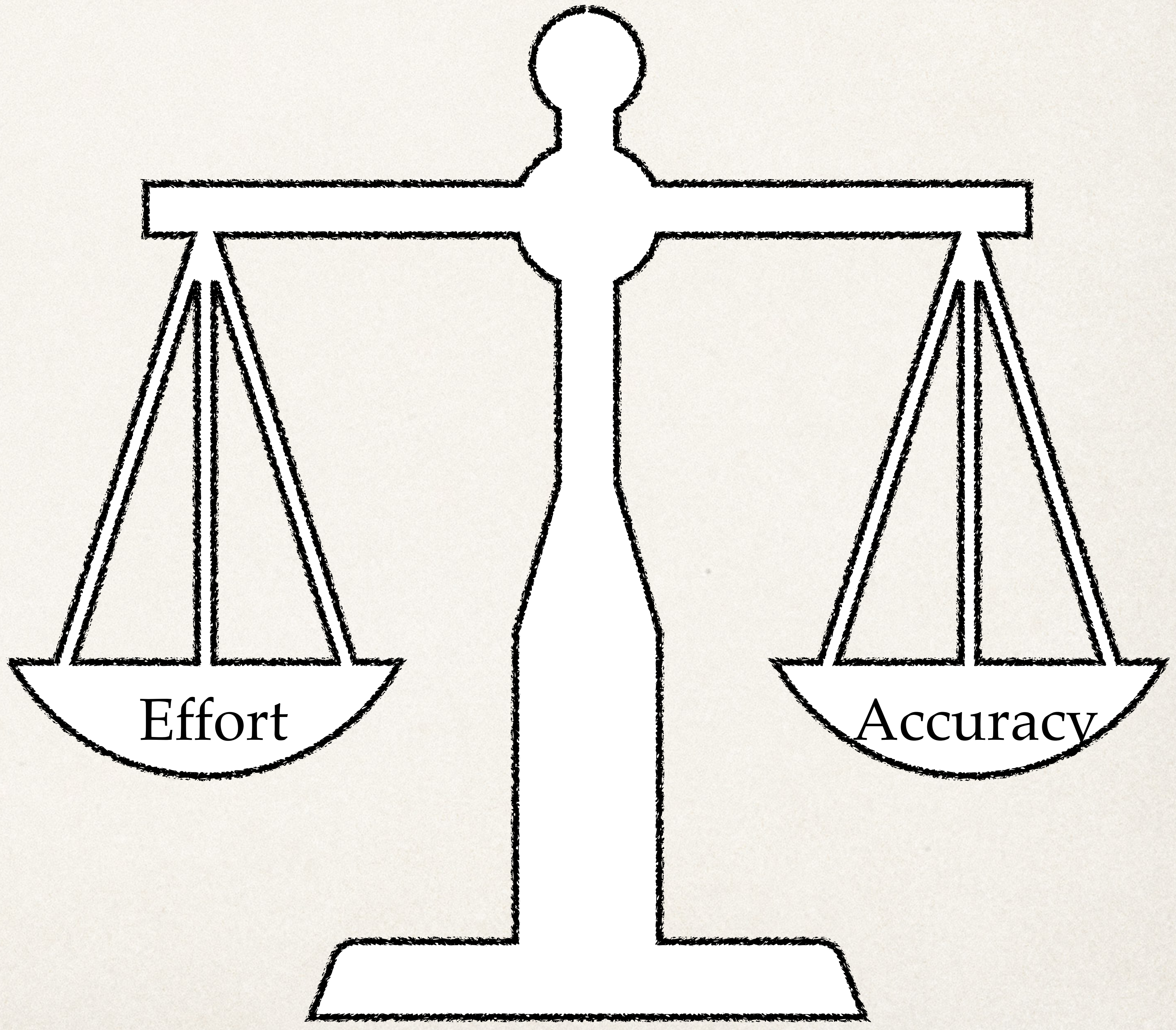
- Consonant with weak perceptual cues => Target of assimilation
- Consonant with strong perceptual cues => Trigger of assimilation
- Reduction of assimilation targets

Balancing effort and accuracy 1

- Bias for high message transmission accuracy using minimal necessary effort

- Reduction of weakly perceptible segments keeps effort low

- Due to low contributions to word inferrability, the reduction of such segments will not influence the message transmission accuracy by a lot.

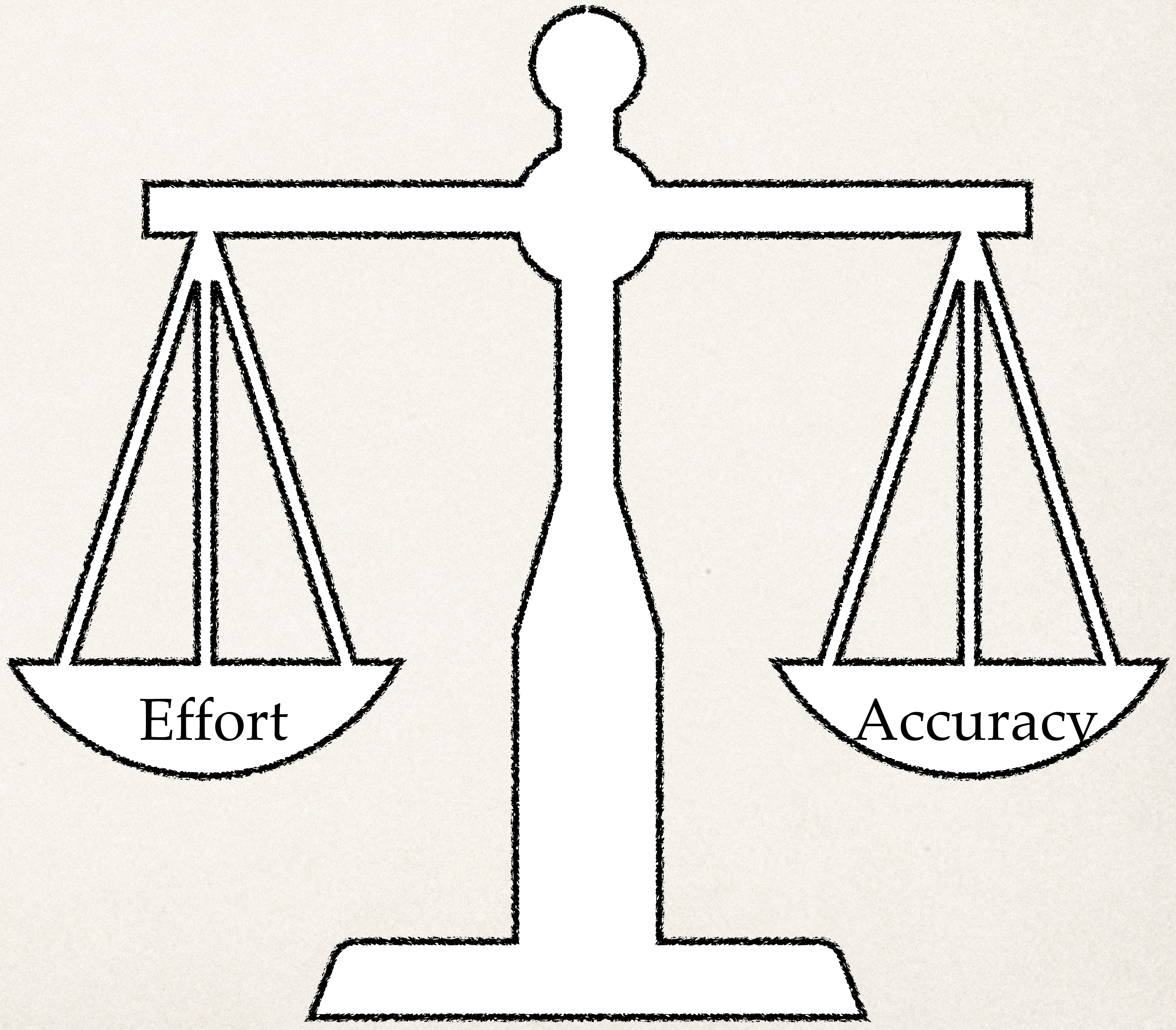


Balancing effort and accuracy 2

- Words with **high predictability** are more easily inferred and speakers can thus reduce them

- Words with **low predictability** need to be pronounced clearly to prevent communication failure

- When contextual predictability of a word increases the phonetic detail decreases



Predictions from a communicative framework 1

- **First prediction:** If the *target word* is more predictable given the context then segments that are part of the target word are more likely to undergo place assimilation (e.g. *ten bucks* => *tem bucks*)

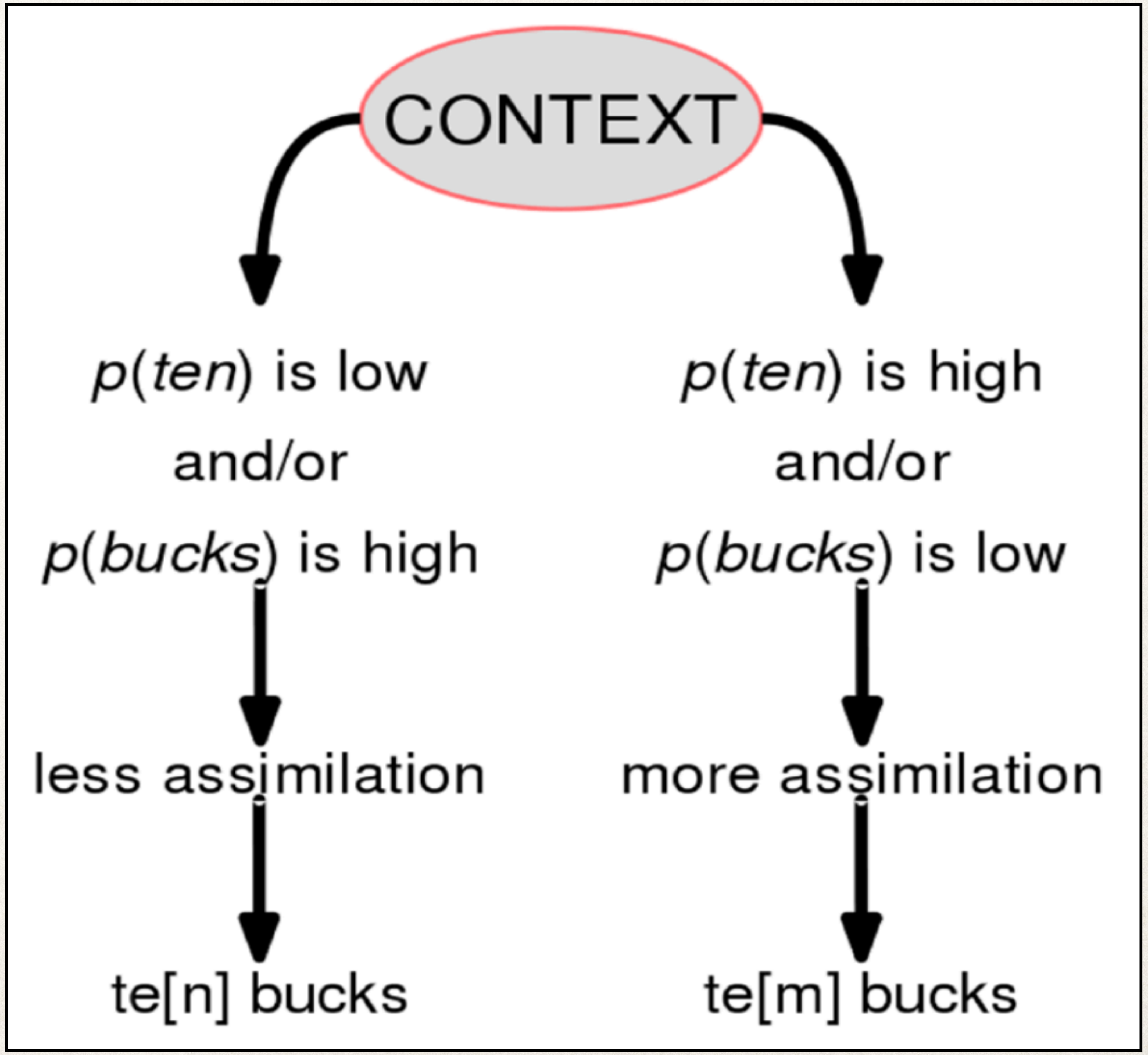
- **Second prediction:** If the *trigger word* is less predictable given the context then segments that are part of the target word are more likely to undergo place assimilation

- Phonetic properties from trigger available earlier, improving recognition of trigger word (Scarborough, 2013).

Predictions from a communicative framework 2

- The likelihood of assimilation to occur in the target word increases with the *target word's* probability

- The likelihood of assimilation to occur in the target word decreases with the *trigger word's* probability



Predictions of usage based accounts

- Main factor: word frequency
- High frequency words (target and trigger) would lead to more assimilation
- In contrast to the communicative framework, where less probable trigger words result in more assimilation

The current study

- Item structure: **Target word followed by Trigger word** (e.g. *ten bucks*)
- Target of assimilation: Word final coronal nasals /n/
- Trigger of assimilation: Word initial labial or velar plosives /p, b, k, g/
- Labial triggers /p, b/ can result in /n/ being assimilated to /m/
- Velar triggers /k, g/ can result in /n/ being assimilated to /ŋ/

Items (Inclusion criteria)

- Bigrams representing target-trigger sequences of spontaneous speech retrieved from the Buckeye corpus

- Inclusion criteria:

- **1)** The ***target*** has a dictionary transcription ending in /n/ and is preceded by a vowel.

- **2)** The ***trigger*** has a dictionary transcription starting with an oral stop that has a non-coronal place of articulation (e.g. /p, b, k, g/)

- **3)** Items free from speech errors, restarts, filled pauses and transcription errors

Items (Exclusion criteria)

Exclusion reason	**Tokens after exclusion**
(All Buckeye sequences with word-final /Vn/ followed by /p, b, k, g/ onset with no speech errors, pauses, or alignment errors)	1334
No vowel in target's phonetic transcription	1195
No nasal in target's phonetic transcription	1177
No stop in trigger's phonetic transcription	1137
Short target vowel (*excluded for partial assimilation only*)	777
Diphthongal target vowel (*excluded for partial assimilation only*)	671

Measuring categorical assimilation (close phonetic transcriptions)

- Assimilation occurs when:
- /n/ is transcribed as [m] before /p, b/
- /n/ is transcribed as [ŋ] before /k, g/
- For example: /n/ is assimilated when *ten* in the bigram *ten bucks* was transcribed as [tɛm].
- If no assimilation occurred the word is said to be in canonical form (e.g. *ten* pronounced as [tɛn]).

Measuring partial assimilation (acoustic data 1)

- Measuring the vowel formant trajectories preceding /n/
- The F2 changes before a consonant depending on that consonant's place of articulation.
- Vowel followed by labial consonant /p, b/:
  - F2 falls before consonant onset
- Vowel followed by velar consonant /k, g/:
  - F2 rises before consonant onset.

Measuring partial assimilation (acoustic data 2)

- Problems:
  - The magnitude of F2 changes depend on the vowel type (front and back vowels differing)
  - Differences between speaker pronunciation
- Solution:
  - A standardized metric adjusting for vowel type and speaker specific differences
    - Higher values: If F2 at vowel offset is deviating from normal values (high or low)
    - Lower values: If F2 at vowel offset is considered normal

Measuring partial assimilation (acoustic data 3)

- Measure F2 values at vowel offset for two datasets
- First dataset: same bigrams as before
- Second dataset:
  - /Vn/ of target word now followed by a vowel or /h/
  - Examples: *then again, then here*
  - (Assimilation not anticipated)

Measuring partial assimilation (acoustic data 4)

- Calculation of F2 value for within-talker and within-vowel standardized bigrams

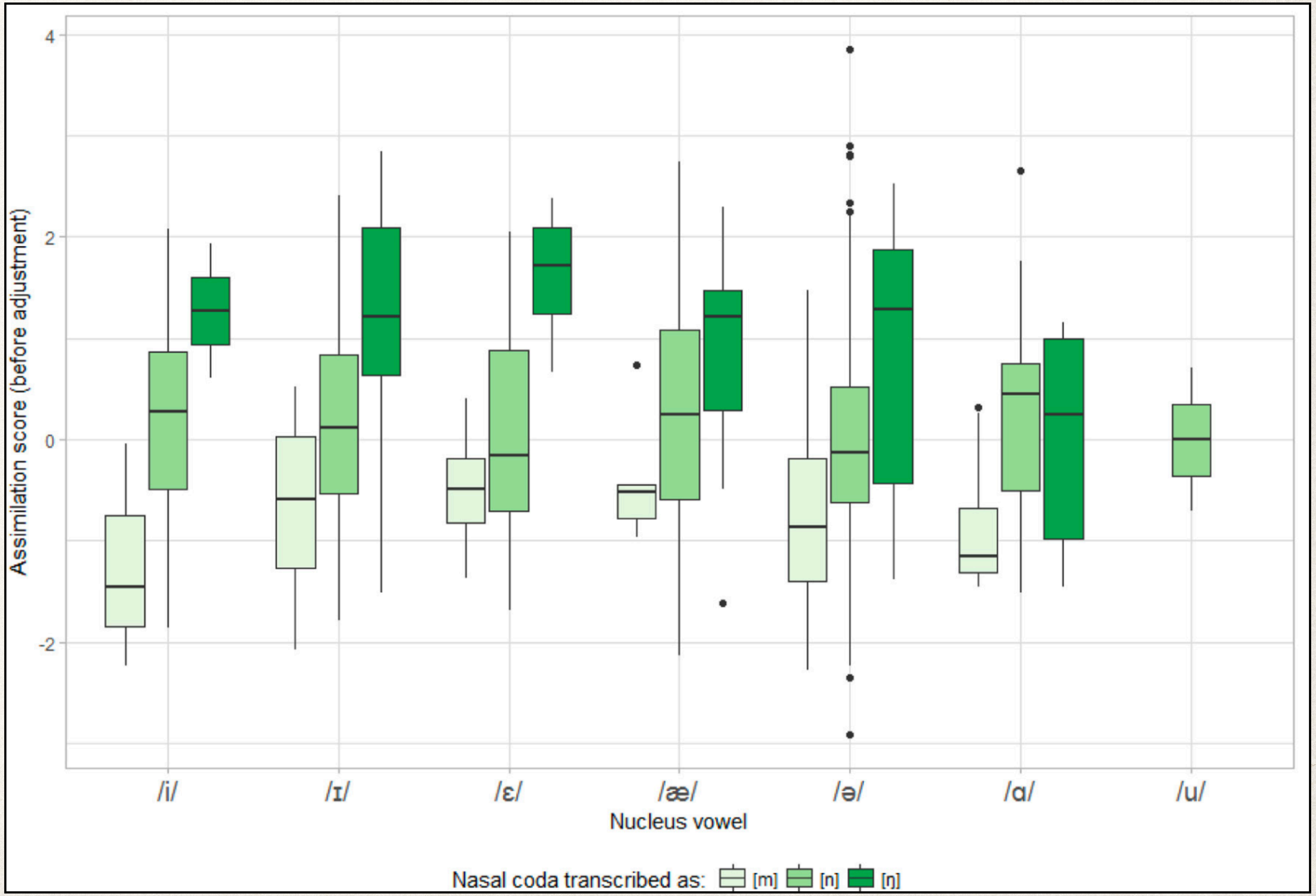
- Example: The F2 value of *ten bucks* at the vowel offset was compared to the average F2 value at the vowel offset for all instances of /ɛn/ uttered by the same speaker. Both, cases with assimilation and those without were included.

- Positive scores: The observed F2 was higher than expected

- Negative scores: The observed F2 was lower than expected

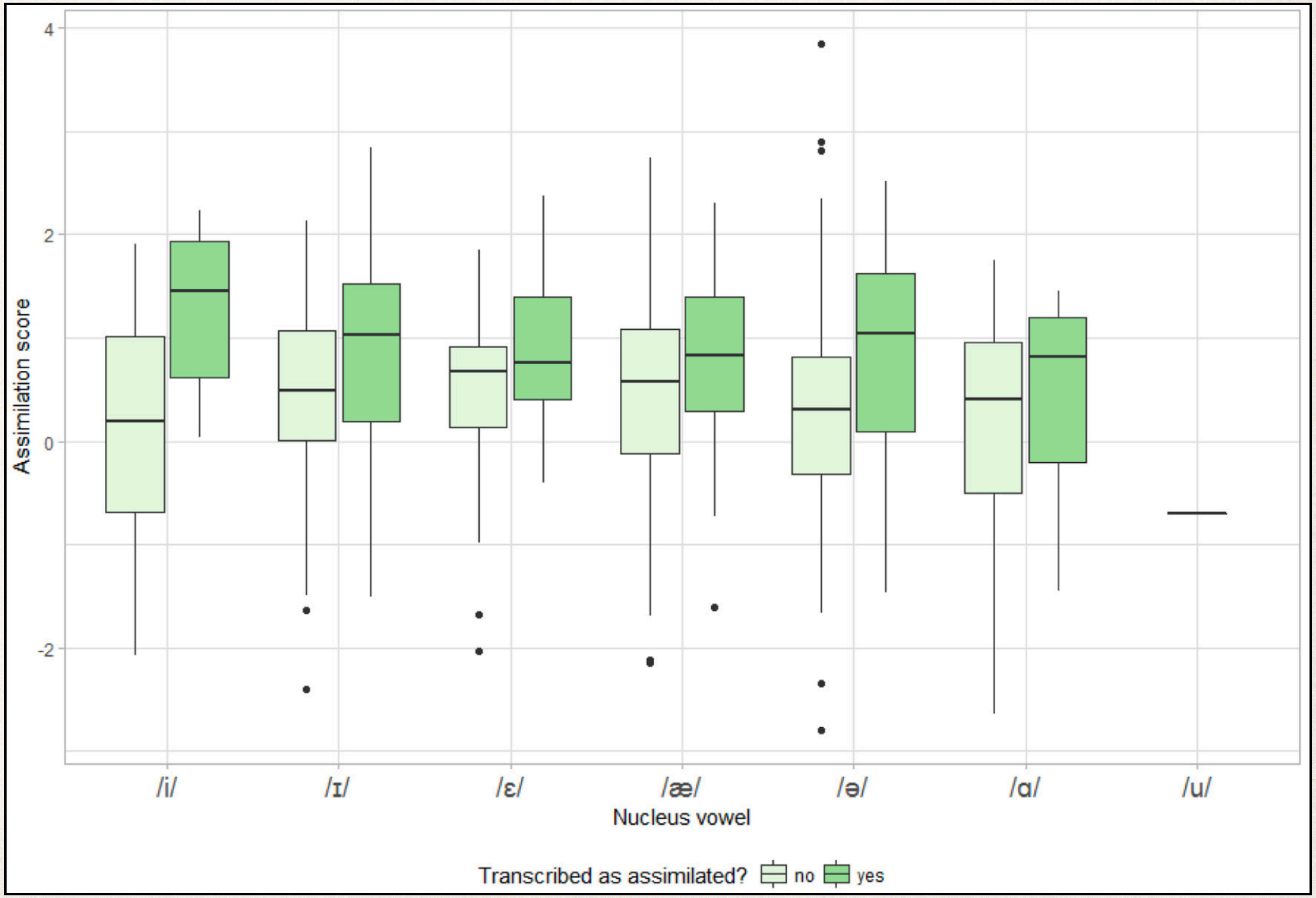
Measuring partial assimilation (acoustic data 5)

- Vowel followed by labial consonant /p, b/:
    - lower F2 at vowel offset
- Vowel followed by velar consonant /k, g/:
    - higher F2 at vowel offset



Measuring partial assimilation (acoustic data 6)

- Invert assimilation scores for labial consonants
- Results:
  - Matching transcription and assimilation scores
  - Variation within same tokens and overlap between different categories



Word predictability

- Backward probability: **P(ten | bucks)**, Forward probability: **P(bucks | ten)**
- Prediction: More nasal place assimilation for predictable ***target*** words and unpredictable ***trigger*** words
- Challenges:
  - **1)** need for similar context
  - **2)** data size
- Due to **2)** include another (larger) corpus
  - The Fisher Part 2 corpus (11 000 000 compared to 300 000 words)

Within-word predictability

- The probability of the target's word-final /n/ and the trigger's word-initial /p, b, k, g/ segments depending on preceding segments (Examples: *betwee**n**, ru**n***)

- Lower probability of word-initial/p, b, k, g/ because segment dependent only on the word boundary. Probabilities: /p/: 5.2%, /b/: 4.7%, /k/: 9.6%, /g/: 3.0%.

- /k/ less likely while /g/ more likely to trigger assimilation processes

Word frequency

- Use of Buckeye (96.7%) and Fisher (3.3%)
- Buckeye robust enough to estimate unigram frequency but not bigram predictability
- The asymmetry could result in a bias that favors the unigram frequency over the bigram predictability

Statistical Analysis

- Two analyses:
  - Logistic multilevel regression (predict whether assimilation occurred)
  - Linear multilevel regression (predict degree of assimilation)
- The analyses used three sets of parameters (combinations used to construct models):
  - Word frequencies **(P(ten) and P(bucks))**
  - Bigram probabilities **(P(ten | bucks) and P(bucks | ten))**
  - Within-word probabilities **(P(/n/ | /# tɛ/) and P(/b/ | #))**

Hypotheses:

- Hypotheses:
  - Higher predictability of the target word increases assimilation
    - More density in positive range expected
  - Higher predictability of the trigger word decreases assimilation
    - More density in negative range expected

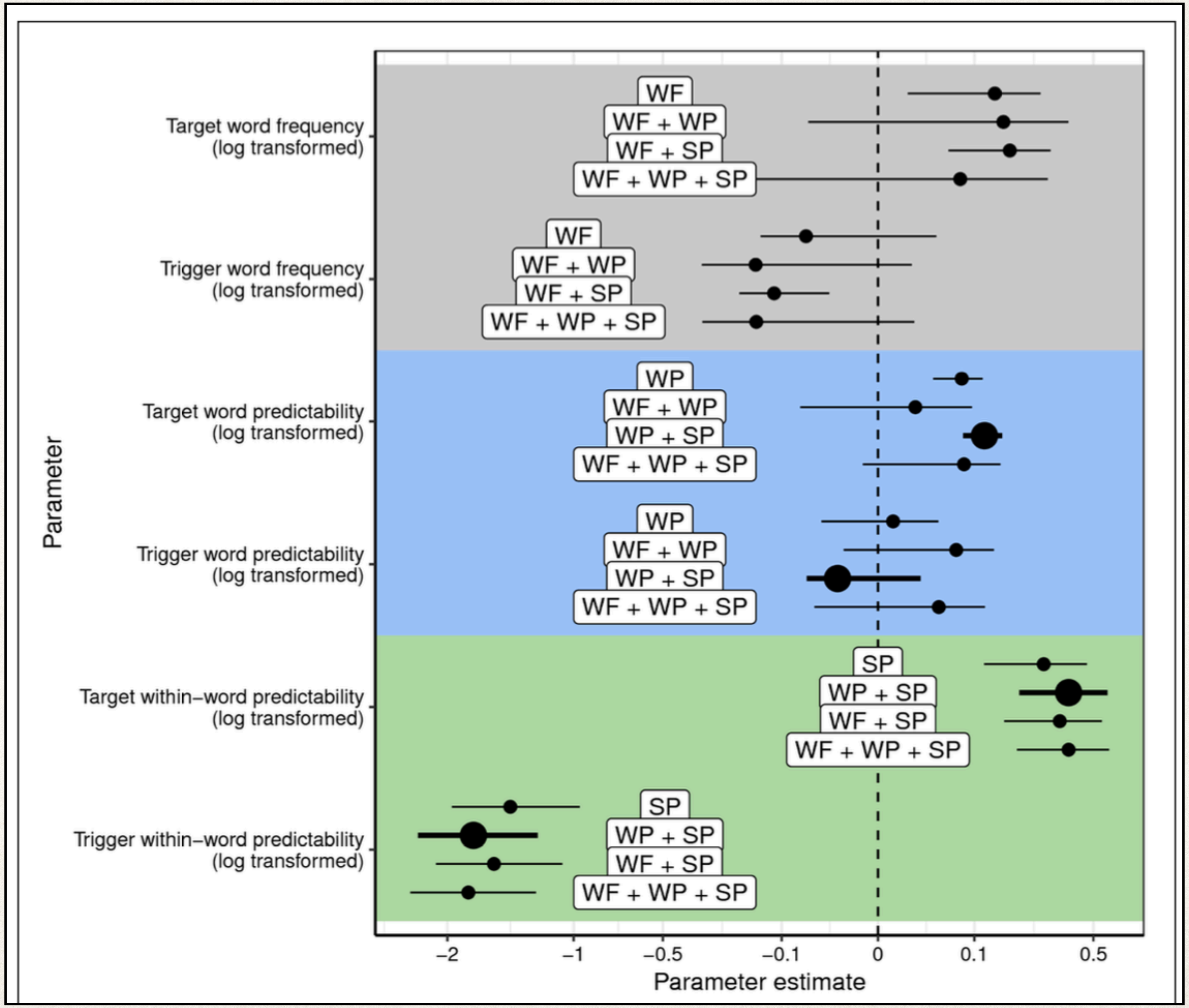
Values of Interest and HPDI

- Values of interest:
  - **Proportion of posterior density in predicted direction**
    - strong evidence: 1, weak evidence: 0
  - **Highest posterior density interval (HPDI)**
    - From HPDI 2.5% to HPDI 97.5% is the interval that contains 95% of the posterior density
    - 95% chance that the true parameter value lies in the interval
    - If HPDI only in positive range that supports hypothesis that target predictability has a positive effect on assimilation (Vice versa for trigger word and negative range and effect)

Results 1

		WordFreq + WordPred + SegPred	WordFreq + WordPred	WordPred + SegPred	WordFreq + SegPred	WordFreq	WordPred	SegPred
Target frequency (log-transformed)	Posterior density > 0	0.635	0.760		0.905	0.854		
	(2.5% HPDI)	*−0.302*	*−0.223*		*−0.047*	*−0.077*		
	(97.5% HPDI)	*0.474*	*0.558*		*0.422*	*0.374*		
Trigger frequency (log-transformed)	Posterior density < 0	0.810	0.804		0.896	0.733		
	(2.5% HPDI)	*−0.457*	*−0.467*		*−0.269*	*−0.205*		
	(97.5% HPDI)	*0.143*	*0.149*		*0.037*	*0.091*		
Target predictability (log-transformed)	Posterior density > 0	0.830	0.576	**0.998**			0.966	
	(2.5% HPDI)	*−0.056*	*−0.118*	***0.051***			*0.008*	
	(97.5% HPDI)	*0.218*	*0.149*	***0.193***			*0.147*	
Trigger predictability (log-transformed)	Posterior density < 0	0.316	0.209	**0.676**			0.488	
	(2.5% HPDI)	*−0.094*	*−0.064*	***−0.079***			*−0.061*	
	(97.5% HPDI)	*0.175*	*0.202*	***0.044***			*0.062*	
Target within-word predictability (log-transformed)	Posterior density > 0	0.983		**0.989**	0.979			0.950
	(2.5% HPDI)	*0.080*		***0.108***	*0.069*			*0.001*
	(97.5% HPDI)	*0.702*		***0.699***	*0.667*			*0.598*
Trigger within-word predictability (log-transformed)	Posterior density < 0	1.000		**0.999**	0.999			0.998
	(2.5% HPDI)	*−2.689*		***−2.611***	*−2.428*			*−2.327*
	(97.5% HPDI)	*−0.928*		***−0.913***	*−0.760*			*−0.654*
ELPD		−501.3	−506.1	**−499.9**	−501.5	−504.8	−504.7	−501.0

Results 2



Target word frequency (log transformed)
Trigger word frequency (log transformed)
Target word predictability (log transformed)
Trigger word predictability (log transformed)
Target within–word predictability (log transformed)
Trigger within–word predictability (log transformed)
WF
WF + WP
WF + SP
WF + WP + SP
WP
WP + SP
SP
Parameter
−2
−1
−0.5
−0.1
0
0.1
0.5
Parameter estimate

Summary of Results 1

- Categorical assimilation:
  - *Target*: Evidence for predicted effects of word and within-word predictability (If **P(ten | bucks)** or **P(n | # tɛ)** is high than **/n/** more likely to undergo assimilation)
  - *Trigger*: Evidence only for predicted effects of within-word predictability (If **P(b | #)** is high than **/n/** less likely to undergo assimilation)
  - Predictability is weaker in models that include frequency and vice versa (Due to correlation of frequency and predictability)

Summary of Results 2

- Partial assimilation:
  - *Target* and *Trigger*: evidence for predicted effects of word but not within-word predictability

Discussion and Conclusion 1

- The results stand in contrast to the usage based accounts which expect more assimilation for frequent targets and triggers.
- The results however align with the communicative approach in which assimilation is dependent on the predictability of both target and trigger
- Model comparisons showed that word predictability was a better predictor than frequency.
- Observed frequency effects could actually be due to predictability

Discussion and Conclusion 2

- Keep a balance between clarity and effort
  - Not necessary to clearly utter the nasal in a predictable word
  - Unpredictable triggers benefit from increasing redundancy
- Partial assimilation no effect of within-word predictability which could indicate different processes of full assimilation
- Assimilation might depend on whether the resulting homophony is unambiguous or ambiguous. ([tɛn] => [tɛm] in *ten bucks;* [rʌn] => [rʌm] in *quick run picks you up*)

References

- Turnbull, R., Seyfarth, S., Hume, E. & Jaeger, T., (2018) "Nasal place assimilation trades off inferrability of both target and trigger words", *Laboratory Phonology* 9(1): 15. doi: https://doi.org/10.5334/labphon.119

- Scarborough, R. 2013. Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics, 41*(6), 491–508. DOI: https://doi.org/10.1016/j.wocn.2013.09.004

- https://en.wikipedia.org/wiki/Consonant