Allophonic variation
and the locality of production planning

Michael Wagner
McGill University

reporting on joint work with:
Meghan Clayards, Oriana Kilbourn-Ceron, Josiane Lachapelle, Morgan Sonderegger, James Tanner

Linguistics Colloquium, University of Maryland, September 13 2019
Outline

1. Locality and Variability in Phonological Processes
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
Sandhi

External Sandhi

Phonological processes in which (part of) the triggering context is not within the same word

(1) Liaison in French
   a. des vrai
   b. des vrai[z]a mis
   'real friends'

(2) Flapping in English
   a. A ca[t]meowed!
   b. A ca[R]ttacked!

Sandhi

External Sandhi

Phonological processes in which (part of) the triggering context is not within the same word

(1) Liaison in French
   a. des vrais copins
   b. des vrai[z] amis
   ‘real friends’
Sandhi

External Sandhi

Phonological processes in which (part of) the triggering context is not within the same word

(1) Liaison in French
   a. des vra\textit{i}s copins
   b. des vrai[\textit{z}] amis
   ‘real friends’

(2) Flapping in English
   a. A ca[\textit{t}] meowed!
   b. A ca[\textit{r}] attacked!
Locality of Sandhi Phenomena

Sandhi phenomena often only apply **locally**: The two words in question have to be in a certain locality relation to each other.
Locality of Sandhi Phenomena

Sandhi phenomena often only apply **locally**: The two words in question have to be in a certain locality relation to each other.

(Kilbourn-Ceron et al. 2016)
How is locality accounted for? Two common approaches:
How is locality accounted for? Two common approaches:

- **Syntactic domains** constrain phonological processes (Cooper and Paccia-Cooper, 1980; Kaisse, 1985; Chen, 1987; Pak, 2008, i.a.).
Locality

How is locality accounted for? Two common approaches:

- **Syntactic domains** constrain phonological processes (Cooper and Paccia-Cooper, 1980; Kaisse, 1985; Chen, 1987; Pak, 2008, i.a.).

- **Phonological domains** constrain phonological processes (and are influenced by syntax) (Selkirk, 1986; Kaisse, 1985; Nespor and Vogel, 1986; Odden, 1990; Selkirk, 2011, i.a.)
How is locality accounted for? Two common approaches:

- **Syntactic domains** constrain phonological processes (Cooper and Paccia-Cooper, 1980; Kaisse, 1985; Chen, 1987; Pak, 2008, i.a.).

- **Phonological domains** constrain phonological processes (and are influenced by syntax) (Selkirk, 1986; Kaisse, 1985; Nespor and Vogel, 1986; Odden, 1990; Selkirk, 2011, i.a.)

But **why** do particular processes apply within particular domains?
Variability

Sandhi processes are often **variable**. Two types of variability:

1. **Variability of Application**: Sandhi processes often only apply in a probabilistic way.
2. **Variability of Domain**: Sandhi processes often have a variable domain (e.g., locality window widens when speech rate increases, e.g. Kaisse 1985 on fast speech phenomena).
Variability

Sandhi processes are often **variable**. Two types of variability:

(i) **Variability of Application**: Sandhi processes often only apply in a probabilistic way.

(ii) **Variability of Domain**: Sandhi processes often have a variable domain (e.g., locality window widens when speech rate increases, e.g. Kaisse 1985 on fast speech phenomena)
Variability

Sandhi processes are often variable. Two types of variability:

(i) Variability of Application: Sandhi processes often only apply in a probabilistic way.

(ii) Variability of Domain: Sandhi processes often have a variable domain (e.g., locality window widens when speech rate increases, e.g. Kaisse 1985 on fast speech phenomena)
Current accounts in phonology usually assume the following:

(i) **Variability of Application**: Variable Rules/Variable Constraint Ranking (cf. Anttila, 2002; Coetzee and Kawahara, 2013)

(ii) **Variability of Domain**: Multiple prosodic constituents of a certain type optionally restructure into one constituent of that type or vice-versa. (e.g. Nespor and Vogel, 1986)
Variability

Current accounts in phonology usually assume the following:

(i) **Variability of Application**: Variable Rules/Variable Constraint Ranking (cf. Anttila, 2002; Coetzee and Kawahara, 2013)

(ii) **Variability of Domain**: Multiple prosodic constituents of a certain type optionally restructure into one constituent of that type or vice-versa. (e.g. Nespor and Vogel, 1986)

But **why** are sandhi processes often variable?
Can we make predictions about Locality?

Given the nature of a process, is there anything we can predict about the locality domain in which it is going to apply?
Can we make predictions about Locality?

Given the nature of a process, is there anything we can predict about the locality domain in which it is going to apply?

Can we make predictions about variability?

Given the nature of a process, is there anything we can predict about whether it is variable, and the structure of the variability?
The Locality of Production Planning

Can we make predictions about Locality?
Given the nature of a process, is there anything we can predict about the locality domain in which it is going to apply?

Can we make predictions about variability?
Given the nature of a process, is there anything we can predict about whether it is variable, and the structure of the variability?

The Basic Idea
We need to consider **locality of production planning**.
Evidence that phonological planning is very *local*:

- Sternberg 1978: Utterance-initiation-time is sensitive to # of upcoming words, but only to phonological detail (# of σ) of first word.
Evidence that phonological planning is very *local*:

- **Sternberg 1978**: Utterance-initiation-time is sensitive to \# of upcoming words, but only to phonological detail (\# of \σ\) of first word.
- **Levelt (1989)**: phonological detail is planned over a window roughly the size of a **single prosodic word**.
Evidence that the size of planning windows is variable:

- Lahiri & Wheeldon (1997, 2002) that prosodic size of planning window varies by task
Evidence that the size of planning windows is *variable*:

- Lahiri & Wheeldon (1997, 2002) that prosodic size of planning window varies by task
- E.g., the complexity of first prosodic word matters most when planning under time pressure...
The Locality of Production Planning

Evidence that the size of planning windows is *variable*:

- Lahiri & Wheeldon (1997, 2002) that prosodic size of planning window varies by task
- E.g., the complexity of first prosodic word matters most when planning under time pressure...
- ...while the # of upcoming prosodic words matters most when speakers have more time
Evidence that the size of planning windows is *variable*:

- Lahiri & Wheeldon (1997, 2002) that prosodic size of planning window varies by task.
- E.g., the complexity of first prosodic word matters most when planning under time pressure...
- ...while the # of upcoming prosodic words matters most when speakers have more time.
- Planning window also varies depending on cognitive load (Swets et al., 2013).
Production Planning Hypothesis (PPH)

Sandhi processes are local and variable because the phonological detail relevant to the process may not have been planned yet in time.
The Locality of Production Planning

The basic mechanism\(^1\):

\[ [t/d] \rightarrow r / \_\_ \, V \]

\(^1\)Note: This process is bled by aspiration!
The Locality of Production Planning

The basic mechanism\(^1\):

\[ [t/d] \rightarrow r / \_\_ V \]

- Why is tapping local? 

\(^1\)Note: This process is bled by aspiration!
The Locality of Production Planning

The basic mechanism\(^1\):

\[ [t/d] \rightarrow r / \_\_ V \]

- Why is tapping **local**?
  - Planning is local: Process can only applies if upcoming vowel available

\(^1\)Note: This process is bled by aspiration!
The Locality of Production Planning

The basic mechanism\(^1\):

\[ [t/d] \rightarrow r / \_ \_ V \]

- Why is tapping **local**?
  - Planning is local: Process can only applies if upcoming vowel available
- Why is tapping **variable**?

\(^1\)Note: This process is bled by aspiration!
The basic mechanism\(^1\):

\[ t/d \to r / \_ \_ \_ V \]

- Why is tapping \textit{local}?
  - Planning is local: Process can only applies if upcoming vowel available

- Why is tapping \textit{variable}?
  - Planning is variable: Scope of planning is affected by many factors

\(^1\)Note: This process is bled by aspiration!
Predictions of PPH for Phonological Processes

- Processes Sensitive to **upcoming phonological detail** (e.g. does next word start with vowel?):
  - necessarily local and variable
Predictions of PPH for Phonological Processes

- Processes Sensitive to **upcoming phonological detail** (e.g. does next word start with vowel?):
  - \(\rightarrow\) necessarily local and variable

- Processes sensitive to **higher-level information**, or **preceding phonological detail** (such as: is there another upcoming word? does preceding word end with vowel):
  - \(\rightarrow\) not necessarily local or variable
The Locality of Production Planning

Predictions of PPH for Phonological Processes

- Processes sensitive to **upcoming phonological detail** (e.g. does next word start with vowel?):
  \[ \rightarrow \text{necessarily local and variable} \]
- Processes sensitive to **higher-level information**, or **preceding phonological detail** (such as: is there another upcoming word? does preceding word end with vowel):
  \[ \rightarrow \text{not necessarily local or variable} \]

Evidence for PPH:
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
Effects of syntax and prosody

with Oriana Kilbourn-Ceron & Meghan Clayards

Effects of syntax and prosody

Tapping in American English (Kahn 76, Nespor & Vogel 1986):

*Monomorphemic words:*

**butter, later** → *pretty much always tapped*
Effects of syntax and prosody

Tapping in American English (Kahn 76, Nespor & Vogel 1986):

*Monomorphemic words:*

\textbf{butter, later} \quad \rightarrow \textit{pretty much always tapped}

*Words within a clause:*

If you \textbf{meet Ann, ...} \quad \rightarrow \textit{tapped in fast speech (cf. Kahn 76)}
Effects of syntax and prosody

Tapping in American English (Kahn 76, Nespor & Vogel 1986):

Monomorphemic words:
- butter, later → pretty much always tapped

Words within a clause:
- If you meet Ann, ... → tapped in fast speech (cf. Kahn 76)

Across Sentences:
- It’s late. I’m leaving. → (possible but rare: Kahn 76, Nespor & Vogel 86, ..)
Effects of syntax and prosody

Two factors affecting tapping:

1. **Strength of a prosodic boundary**: a stronger boundary between the stop and following vowel appear to reduce flapping rate
Effects of syntax and prosody

Two factors affecting tapping:

1. **Strength of a prosodic boundary**: a stronger boundary between the stop and following vowel appear to reduce flapping rate.

2. **Strength of syntactic break**: higher level syntactic boundaries appear to reduce flapping rate.
Two Types of Accounts

- **Prosodic phonology**: Syntax affects phrasing, phrasing in turn affects tapping, because tapping only applies within a particular prosodic domain (e.g. Nespor and Vogel, 1986)
Two Types of Accounts

- **Prosodic phonology**: Syntax affects phrasing, phrasing in turn affects tapping, because tapping only applies within a particular prosodic domain (e.g. Nespor and Vogel, 1986)

- **Articulatory phonology**: Tapping as the result of gestural undershoot/overlap, which is less likely across junctures
(3) Phonological tapping rule: 
\[ [t/d] \rightarrow r / \_\_ V \]
(3) Phonological tapping rule:
[t/d] \rightarrow r / _ _ V

- PPH: Rule will apply whenever environment is met
(3) Phonological tapping rule:
\[ [t/d] \rightarrow r / _V \]

- PPH: Rule will apply whenever environment is met.
- Whether the environment is available depends on whether the upcoming vowel has been planned at the time that the rule applies.
(3) Phonological tapping rule:
\[ [t/d] \rightarrow r / \_\_ \_ V \]

- PPH: Rule will apply whenever environment is met
- Whether the environment is available depends on whether the upcoming vowel has been planned at the time that the rule applies
- It will be less likely to have been planned (i) across word boundaries; (ii) across prosodic boundaries, (iii) across syntactic boundaries...
(3) Phonological tapping rule:
\[ \text{[t/d]} \rightarrow r / \_\_ V \]

- **PPH**: Rule will apply whenever environment is met
- Whether the environment is available depends on whether the upcoming vowel has been planned at the time that the rule applies
- It will be less likely to have been planned (i) across word boundaries; (ii) across prosodic boundaries, (iii) across syntactic boundaries...
- ...because we know independently that these factors affect planning scope
Tapping: Production Experiment

**Table:** A sample item set

<table>
<thead>
<tr>
<th>Phonology</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consonant</strong></td>
<td><strong>Clause Boundary</strong></td>
</tr>
<tr>
<td>If you <em>plit</em>, Alice will be mad.</td>
<td>If you <em>plit</em> Alice, John will be mad.</td>
</tr>
<tr>
<td>If you <em>plit</em>, Penny will be mad.</td>
<td>If you <em>plit</em> Penny, John will be mad.</td>
</tr>
<tr>
<td><strong>Vowel</strong></td>
<td></td>
</tr>
</tbody>
</table>

Additional manipulation: Speech rate
Tapping: Production Experiment

- 23 participants, 8 different item sets with the 4 conditions
- Participants could familiarize themselves with sentence before recording.
- They were recorded at two speech rates
- Utterances were annotated by RAs, and also forced-aligned
- Acoustic measures were extracted, in particular measures for the vowel preceding the [t] (‘final lengthening’, Price et al. 1991, and references therein) as a proxy for measuring prosodic boundary strength
Tapping: Production Experiment

**Figure**: Percent of tapped [t]s.
Tapping: Production Experiment

Effect of preceding vowel duration

Syntax
ClauseBoundary
NoClauseBoundary

Figure: Percent of tapped [t]s.

- Flapping rate lower when there is a syntactic boundary
- Flapping rate lower when there is a prosodic break...
- ...but only in intransitive case, when there is likely to be a boundary
Syntax affects tapping rate significantly even after controlling for prosodic boundary strength

- unexpected if purely driven by gestural overlap/undershoot, so AP account is insufficient

Significant interaction between syntax and prosody, and presence syntactic effect after prosody controlled for suggests its not reducible to prosodic effect

(Also: Effect with nonce-words hard to explain in terms of exemplars or storage of frequently co-occurring bigrams - Bybee 2001)
Syntax affects tapping rate significantly even after controlling for prosodic boundary strength
  ▶ unexpected if purely driven by gestural overlap/undershoot, so AP account is insufficient

Significant interaction between syntax and prosody, and presence syntactic effect after prosody controlled for suggests its not reducible to prosodic effect
  ▶ unexpected for Prosodic Phonology account
Syntax affects tapping rate significantly even after controlling for prosodic boundary strength
- unexpected if purely driven by gestural overlap/undershoot, so AP account is insufficient

Significant interaction between syntax and prosody, and presence syntactic effect after prosody controlled for suggests its not reducible to prosodic effect
- unexpected for Prosodic Phonology account

(Also: Effect with nonce-words hard to explain in terms of exemplars or storage of frequently co-occurring bigrams Bybee 2001)
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
Effects of Predictability 1: Tapping in Spontaneous speech


Kilbourn-Ceron, O., Clayards, M., Wagner M. (resubmitted). Predictability modulates pronunciation variants through speech planning effects: A case study on coronal stop realizations. Laboratory Linguistic
Tapping as probabilistic reduction?

- Tapping is a form of reduction

Tapping as probabilistic reduction?

- Tapping is a form of reduction
- Related to reduction of word duration
Tapping as probabilistic reduction?

- Tapping is a form of reduction
- Related to reduction of word duration
- Known factor influencing word duration: frequency and predictability
Tapping as probabilistic reduction?

- Tapping is a form of reduction
- Related to reduction of word duration
- Known factor influencing word duration: frequency and predictability
- Common approach to explain this: Information theoretic rationale (cf. Jurafsky et al., 2001; Pluymaekers et al., 2005; Jaeger, 2010, and many others): Less information $\rightarrow$ less oomph
Tapping as probabilistic reduction?

- PPH and probabilistic reduction often make similar predictions
Tapping as probabilistic reduction?

- PPH and probabilistic reduction often make similar predictions
- Glottalization does not require information about next word:
Tapping as probabilistic reduction?

- PPH and probabilistic reduction often make similar predictions.
- Glottalization does not require information about next word:
  - Tapping: [t/d] → r / _ V
- Should show similar pattern based on probabilistic reduction, but different pattern based on PPH.
Tapping as probabilistic reduction?

- PPH and probabilistic reduction often make similar predictions
- Glottalization does not require information about next word:
  - Tapping: \([t/d] \rightarrow r / \_ \_ V\)
  - Glottalization: \([t/d] \rightarrow ?/ \_ \_ \#\)
- Should show similar pattern based on probabilistic reduction, but different pattern, based on PPH
We look at Buckeye Corpus (Pitt et al., 2007) to look for effect predictability measures.

11863 tokens with word-findal /t/ or /d/ followed by a vowel-initial word (46.24% were transcribed as flaps).

Excluded: words followed by disfluency (18.26% of tokens).

Word frequencies were retrieved from SUBTLEX-US, a database of word frequencies based on film and television subtitles (Brysbaert and New, 2009).
Tapping: Corpus Data

**Figure:** Relationship between SUBTLEX-US word frequency (per million words) and proportion of tokens transcribed as flaps [\textit{dx}] (left panel, blue), glottal stops [\textit{tq}] (right panel, red) in the Buckeye corpus. Solid lines show trigger word frequency, dashed lines show target word frequency, with shading showing 95% confidence intervals of a linear smooth (GLM, logit-link). Rug plot on top and bottom margins represent distribution of tokens.
Tapping & Glottalization: Frequency Effects

- Higher **Target Word Frequency** → less tapping, more glottalization
Tapping & Glottalization: Frequency Effects

- Higher **Target Word Frequency** $\rightarrow$ less tapping, more glottalization
- Higher **Trigger Word Frequency** $\rightarrow$ more tapping, no effect on glottalization
Tapping & Glottalization: Frequency Effects

- Higher Target Word Frequency $\rightarrow$ less tapping, more glottalization
- Higher Trigger Word Frequency $\rightarrow$ more tapping, no effect on glottalization

The probabilistic reduction account does not explain the differences
Tapping & Glottalization: Conditional probability

The PPH predicts that the **conditional probability** of the second word given the first should be relevant for tapping:
Tapping & Glottalization: Conditional probability

The PPH predicts that the **conditional probability** of the second word given the first should be relevant for tapping:

---

**Tapping:**

---

**Glottalization:**

---

**Figure:** Relationship between Conditional Probability (of following word given target word) and proportion of tokens transcribed as flaps [dx] (blue, left panel) or glottal stops [tq] (red, right panel) in the Buckeye corpus. Solid lines and shading are linear smooths (GLM, logit-link) with 95% confidence intervals.
Why negative effect of Target Word Frequency on flapping rate?

There are conflicting results whether high Word1 frequency makes it more or less likely that Word2 is planned at the same time.
Effects of predictability 2: [t,d] Deletion in Clusters


[t,d] Deletion in Clusters (British English spontaneous speech)

fast ball \rightarrow fas’ ball
[t,d] Deletion in Clusters (British English spontaneous speech)

fast ball > fas’ ball

t/d-deletion: [t/d] → ∅ / C _# X
[t,d] Deletion in Clusters (British English spontaneous speech)

fast ball > fas’ ball

t/d-deletion: [t/d] → ∅ / C _-# X

**PPH:** Effect of X should be modulated by Prosody boundary strength and predictability of following word
[t,d] Deletion in Clusters

Effect of following segment is modulated by strength of boundary:

Figure: Deletion rate as a function of pause duration
[t,d] Deletion in Clusters

The higher the **conditional probability**, the bigger the effect of context:

**Figure**: Deletion rate as a function of conditional probability
Couldn’t these effects still just reflect gestural overlap/magnitude?

- Factors increasing planning scope (speech rate, predictability of words, cognitive load, ...) may also affect duration of gestures.
Couldn’t these effects still just reflect gestural overlap/magnitude?

- Factors increasing planning scope (speech rate, predictability of words, cognitive load, ...) may also affect duration of gestures
- Duration of gestures modulates **degree of overlap and magnitude** (Browman and Goldstein, 1990; Byrd and Saltzman, 2003; Krivokapić and Byrd, 2012; Temple, 2014)
Couldn’t these effects still just reflect gestural overlap/magnitude?

- Factors increasing planning scope (speech rate, predictability of words, cognitive load, ...) may also affect duration of gestures.
- Duration of gestures modulates **degree of overlap** and **magnitude** (Browman and Goldstein, 1990; Byrd and Saltzman, 2003; Krivokapić and Byrd, 2012; Temple, 2014).
- We tried to control for this by adding duration measures to model.
Couldn’t these effects still just reflect gestural overlap/magnitude?

- Factors increasing planning scope (speech rate, predictability of words, cognitive load, ...) may also affect duration of gestures.
- Duration of gestures modulates **degree of overlap** and **magnitude** (Browman and Goldstein, 1990; Byrd and Saltzman, 2003; Krivokapić and Byrd, 2012; Temple, 2014).
- We tried to control for this by adding duration measures to model.
Couldn’t these effects still just reflect gestural overlap/magnitude?

- Factors increasing planning scope (speech rate, predictability of words, cognitive load, ...) may also affect duration of gestures
- Duration of gestures modulates **degree of overlap and magnitude** (Browman and Goldstein, 1990; Byrd and Saltzman, 2003; Krivokapić and Byrd, 2012; Temple, 2014)
- We tried to control for this by adding duration measures to model
- ...but we could get more direct evidence by looking at **non-reductive** process
Couldn’t these effects still just probabilistic reduction?

- **Probabilistic reduction**: More predictable information is *reduced* for probabilistic/information-theoretic reasons (cf. Jurafsky et al., 2001; Pluymaekers et al., 2005; Jaeger, 2010, and many others)
Couldn’t these effects still just probabilistic reduction?

- **Probabilistic reduction**: More predictable information is reduced for probabilistic/information-theoretic reasons (cf. Jurafsky et al., 2001; Pluymaekers et al., 2005; Jaeger, 2010, and many others)
Couldn’t these effects still just probabilistic reduction?

- **Probabilistic reduction**: More predictable information is *reduced* for probabilistic/information-theoretic reasons (cf. Jurafsky et al., 2001; Pluymaekers et al., 2005; Jaeger, 2010, and many others)

- We found that reduction through glottalization works differently, which suggests that this explanation is not sufficient

- ...but again, by looking at **non-reductive processes** we could avoid similarity in predictions
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
Effects of predictability in non-reductive processes: Liaison

with Oriana Kilbourn-Ceron & Josiane Lachapelle

Corpus Study: Kilbourn-Ceron, Oriana (2016). Speech production planning affects variability in connected speech. Proceedings of AMP, USC

Experimental study: Kilbourn-Ceron, Oriana, Josiane Lachapelle, Michael Wagner (in prep)
**Liaison:** Latent consonant appears before vowel initial word

- **Un peti[t]** ami
  - cf. un peti[ ] chapeau

- **Des vrai[z]** amis
  - cf. des vrai[ ] chatons
**Liaison**

Different from tapping: Syntactic/Morphological Interactions (Côté, 2013, 157):

<table>
<thead>
<tr>
<th></th>
<th><strong>Singular</strong></th>
<th><strong>Plural</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj + N</td>
<td>le gros [z] enjeu</td>
<td>les gros [z] enjeux</td>
<td>‘the big stake(s)’</td>
</tr>
<tr>
<td>N + Adj</td>
<td>le pas *[z] enjoué</td>
<td>les pas [z] enjoués</td>
<td>‘the cheerful step(s)’</td>
</tr>
<tr>
<td>N + Verb</td>
<td>le pas *[z] endort</td>
<td>les pas *[z] endorment</td>
<td>‘the step(s) send(s) to sleep’</td>
</tr>
</tbody>
</table>
Predictions of PPH:

- Sensitive to upcoming phonological information → should be variable
Predictions of PPH:

- Sensitive to upcoming phonological information $\rightarrow$ should be variable
- Should be less likely with greater juncture (more liaison in adjective-noun vs. noun-adjective order)
Liaison

Predictions of PPH:

- Sensitive to upcoming phonological information $\rightarrow$ should be variable
- Should be less likely with greater juncture (more liaison in adjective-noun vs. noun-adjective order)
- For predictability effects, PPH makes same predictions as for reductive process
Liaison rate in Adjective-Noun (left) and Noun-Adjective (right) Cases:
Corpus evidence on liaison: Summary

- Effects for non-reductive process parallel those of reductive processes
Corpus evidence on liaison: Summary

- Effects for non-reductive process parallel those of reductive processes
- This is unexpected by probabilistic reduction account (it’s not reduction!), but expected by PPH

(Note that the observed frequency effects are also compatible with storage of larger-sized units, Bybee 2001; Côte 2013)
Corpus evidence on liaison: Summary

- Effects for non-reductive process parallel those of reductive processes
- This is unexpected by probabilistic reduction account (it’s not reduction!), but expected by PPH
- Modulation by syntax also as expected (see also Tamminga, M. (2018, Glossa) on [t/d]-deletion)
Corpus evidence on liaison: Summary

- Effects for non-reductive process parallel those of reductive processes
- This is unexpected by probabilistic reduction account (it’s not reduction!), but expected by PPH
- Modulation by syntax also as expected (see also Tamminga, M. (2018, *Glossa*) on [t/d]-deletion)
- (Note that the observed frequency effects are also compatible with storage of larger-sized units, Bybee 2001; Côté 2013)
Liaison: Production experiment

- Another factor affecting planning scope: word length
Liaison: Production experiment

- Another factor affecting planning scope: **word length**
- If word1 is long, then it is less likely that word2 will be planned at the same time (Miozzo and Caramazza, i.a.—but: Griffin)
Another factor affecting planning scope: **word length**

If word1 is long, then it is less likely that word2 will be planned at the same time (Miozzo and Caramazza, i.a.—but: Griffin)

Since only the *beginning* of word2 is relevant (*does it start with a vowel?*), its overall length might be less relevant
Another factor affecting planning scope: **word length**

If word1 is long, then it is less likely that word2 will be planned at the same time (Miozzo and Caramazza, i.a.—but: Griffin)

Since only the *beginning* of word2 is relevant (*does it start with a vowel?*), its overall length might be less relevant

Also manipulated: speech rate, repetition, word frequencies, conditional probability of upcoming word, syntax (adjective-noun vs. noun-adjective contexts)
Liaison: Production experiment

(4) Adjective-Noun (‘obligatory’ liaision context)

a. Low conditional probability;shortword1;shortword2:
   Elle discute avec les derniers élèves.
   she discusses with the last students
   ‘She is talking with the latest students.’ slow; fast

b. High conditional probability,longword1;shortword2:
   Vous regrettez vos dernières années.
   you regret your last years
   ‘You regret the previous years.’ slow; fast

(5) Noun-Adjective (‘optional’ liaision context’)

a. Low conditional probability;shortword1;longword2:
   Ils construisent des douches intérieures.
   they construct of showers interior
   ‘They are constructing interior showers.’ slow; fast

b. High conditional probability;shortword1;shortword2:
   Mathilde regarde ses dessins animés.
   Mathilde watches her drawing animated
   ‘Mathilde is reading comic books.’ slow; fast
Liaison: Production experiment

Plot of the effect of conditional probability, syntax, and length of word1
Liaison: Production experiment

Plot of the effect of conditional probability, syntax, and length of word2
The longer word1 was, the lower the liaison rate (this has not previously been shown for liaison, as far as we know)
• The longer word1 was, the lower the liaison rate (this has not previously been shown for liaison, as far as we know)
• The effect size for word2 length was about half compared to word1 length, and did not reach significance
The longer word1 was, the lower the liaison rate (this has not previously been shown for liaison, as far as we know)

The effect size for word2 length was about half compared to word1 length, and did not reach significance

The higher the conditional probability of the following word, the higher the liaison rate—just as for reductive processes
The longer word1 was, the lower the liaison rate (this has not previously been shown for liaison, as far as we know)

The effect size for word2 length was about half compared to word1 length, and did not reach significance

The higher the conditional probability of the following word, the higher the liaison rate—just as for reductive processes

No effect of Repetition
- The longer word1 was, the lower the liaison rate (this has not previously been shown for liaison, as far as we know)
- The effect size for word2 length was about half compared to word1 length, and did not reach significance
- The higher the conditional probability of the following word, the higher the liaison rate—just as for reductive processes
- No effect of Repetition
- No effect of Speech rate (already observed in Kaisse)
  - Why is it not like tapping in this respect? Does liaison operate at a different level of representation than tapping?
An alternative approach: Currie Hall et al. (2016)

Currie Hall et al. (2016); Turnbull et al. (2018) give information-theoretic rationale for phonological pattern:

- ‘conserve cost when message predictability is high’
- ‘additional material increasing signal specificity and redundancy is more likely to be invested when message predictability is low’

This predicts that liaison should be more likely if a following word is less predictable since pronouncing the liaison encodes information about it and hence increases redundancy (it encodes that the following word begins with a vowel).

The PPH (correctly) makes the opposite prediction.
An alternative approach: Currie Hall et al. (2016)

Currie Hall et al. (2016); Turnbull et al. (2018) give information-theoretic rationale for phonological pattern:

▶ ‘conserve cost when message predictability is high’

...and ‘additional material increasing signal specificity and redundancy is more likely to be invested when message predictability is low’

This predicts that liaison should be more likely if a following word is less predictable since pronouncing the liaison encodes information about it and hence increases redundancy (it encodes that the following word begins with a vowel)

The PPH (correctly) makes the opposite prediction
Currie Hall et al. (2016); Turnbull et al. (2018) give information-theoretic rationale for phonological pattern:

- ‘conserve cost when message predictability is high’
- ‘...and ‘additional material increasing signal specificity and redundancy is more likely to be invested when message predictability is low’
An alternative approach: Currie Hall et al. (2016)

Currie Hall et al. (2016); Turnbull et al. (2018) give information-theoretic rationale for phonological pattern:

- ‘conserve cost when message predictability is high’
- ...and ‘additional material increasing signal specificity and redundancy is more likely to be invested when message predictability is low’

This predicts that liaison should be more likely if a following word is less predictable

- ...since pronouncing the liaison encodes information about it and hence increases redundancy (it encodes that the following word begins with a vowel)
An alternative approach: Currie Hall et al. (2016)

Currie Hall et al. (2016); Turnbull et al. (2018) give information-theoretic rationale for phonological pattern:
- ‘conserve cost when message predictability is high’
- ‘...and ‘additional material increasing signal specificity and redundancy is more likely to be invested when message predictability is low’

This predicts that liaison should be more likely if a following word is less predictable
- ...since pronouncing the liaison encodes information about it and hence increases redundancy (it encodes that the following word begins with a vowel)

The PPH (correctly) makes the opposite prediction
An alternative approach: Currie Hall et al. (2016)

- However, Turnbull et al. (2018) find that the conditional probability of the following word makes nasal assimilation less likely.

This result contradicts the PPH! But there are potential confounds...
The PPH predicts predictability effect irrespective of their adaptiveness from an information-theoretic point of view (although which patterns 'survive' and are grammaticalized might still depend on their utility for message-retrieval).
An alternative approach: Currie Hall et al. (2016)

- However, Turnbull et al. (2018) find that the conditional probability of the following word makes nasal assimilation less likely.
- This result contradicts the PPH! But there are potential confounds...
An alternative approach: Currie Hall et al. (2016)

However, Turnbull et al. (2018) find that the conditional probability of the following word makes nasal assimilation less likely.

This result contradicts the PPH! But there are potential confounds...

The PPH predicts predictability effect irrespective of their adaptiveness from an information-theoretic point of view.
An alternative approach: Currie Hall et al. (2016)

- However, Turnbull et al. (2018) find that the conditional probability of the following word makes nasal assimilation less likely.
- This result contradicts the PPH! But there are potential confounds...
- The PPH predicts predictability effect irrespective of their adaptiveness from an information-theoretic point of view.
- (although which patterns ‘survive’ and are grammaticalized might still depend on their utility for message-retrieval)
Outline

1. Locality and Variability in Phonological Processes
2. Effects of syntax and prosody
3. Effects of Predictability
4. Effects of predictability in non-reductive processes
5. Conclusion and Outlook
What does this mean?

Four conceivable meta-responses to this hypothesis:

1a/1b: This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy) ...and that's a good thing/...and that's a bad thing

2a/2b: This is a retrograde generativist agenda going back to a more SPE-like theory where phonology doesn't see syntax or even prosody ...and that's a good thing/...and that's a bad thing

It may be a bit of both... but mostly it's just an empirical hypothesis...
What does this mean?

Four conceivable meta-responses to this hypothesis:

**(1a/1b):** This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy)

**(2a/2b):** This is a retrograde generativist agenda going back to a more SPE-like theory where phonology doesn’t see syntax or even prosody... and that’s a good thing/... and that’s a bad thing

It may be a bit of both... but mostly it’s just an empirical hypothesis...
What does this mean?

Four conceivable meta-responses to this hypothesis:

(1a/1b): This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy)
...and that’s a good thing/...and that’s a bad thing
What does this mean?

Four conceivable meta-responses to this hypothesis:

(1a/1b): This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy) ...and that’s a good thing/...and that’s a bad thing

(2a/2b): This is a retrograde generativist agenda going back to a more SPE-like theory where phonology doesn’t see syntax or even prosody
What does this mean?

Four conceivable meta-responses to this hypothesis:

(1a/1b): This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy) ...and that’s a good thing/...and that’s a bad thing

(2a/2b): This is a retrograde generativist agenda going back to a more SPE-like theory where phonology doesn’t see syntax or even prosody ...and that’s a good thing/...and that’s a bad thing
What does this mean?

Four conceivable meta-responses to this hypothesis:

**1a/1b**: This is a reductionist agenda trying to reduce the role of grammar (e.g. it removes some arguments for prosodic hierarchy)
...and that’s a good thing/...and that’s a bad thing

**2a/2b**: This is a retrograde generativist agenda going back to a more SPE-like theory where phonology doesn’t see syntax or even prosody
...and that’s a good thing/...and that’s a bad thing

It may be a bit of both... but mostly it’s just an empirical hypothesis...
Chen (1987, 2000): Locality of Tone sandhi in Taiwanese (also Xiamen) is constrained by syntax, but often contradicts prosody.
Toward a predictive theory of locality & variability in phonology

- Chen (1987, 2000): Locality of Tone sandhi in Taiwanese (also Xiamen) is constrained by syntax, but often contradicts prosody.
- ...other types of tone sandhi, e.g. Mandarin T3 sandhi, seem to be much more variable and constrained by surface prosody.
Toward a predictive theory of locality & variability in phonology

- Chen (1987, 2000): Locality of Tone sandhi in Taiwanese (also Xiamen) is constrained by syntax, but often contradicts prosody.
- ...other types of tone sandhi, e.g. Mandarin T3 sandhi, seem to be much more variable and constrained by surface prosody.

Why do these processes differ in their locality and variability?
Toward a predictive theory of locality & variability in phonology

- Taiwanese
  - every non-final word within a domain undergoes tone sandhi;
  - The following tone is irrelevant in determining which sandhi tone it shifts to.
  - Crucially, the only information relevant is whether a word is coming up within the same syntactic domain.
Toward a predictive theory of locality & variability in phonology

- **Taiwanese**
  - every non-final word within a domain undergoes tone sandhi;
  - The following tone is irrelevant in determining which sandhi tone it shifts to.
  - Crucially, the only information relevant is whether a word is coming up within the same syntactic domain.

- **Mandarin T3 sandhi**
  - Which sandhi tone you shift to depends on phonological identity of following tone
  - → the phonology of the following word has to have been planned out for T3 sandhi to apply
  - The PPH predicts the process to be local and variable.
Toward a predictive theory of locality & variability in phonology

- Taiwanese
  - every non-final word within a domain undergoes tone sandhi;
  - The following tone is irrelevant in determining which sandhi tone it shifts to.
  - Crucially, the only information relevant is whether a word is coming up within the same syntactic domain.

- Mandarin T3 sandhi
  - Which sandhi tone you shift to depends on phonological identity of following tone
  - → the phonology of the following word has to have been planned out for T3 sandhi to apply
  - The PPH predicts the process to be local and variable.

More predictions: Influence of prior vs. upcoming information in vowel coalescence (Lamontagne and Torreira, 2017)
Can we make predictions about Locality?

**Maybe yes:** When a process relies on phonological information about an upcoming word, it should necessarily be local; when it depends on phonological information about a previous word, or on higher level information, it does not need to be local.
Locality and Variability

Can we make predictions about Locality?

**Maybe yes:** When a process relies on phonological information about an upcoming word, it should necessarily be local; when it depends on phonological information about a previous word, or on higher level information, it does not need to be local.

Can we make predictions about variability?

**Maybe yes:** If a process relies on phonological information contained in an upcoming word, it necessarily has to be variable, but not if it relies on information from preceding word.
Thanks!

Co-authors:

Oriana Kilbourn-Ceron
Meghan Clayards
Josiane Lachapelle
James Tanner
Morgan Sonderegger

Thanks to SSHRC, NSERC, and to the members of the prosody.lab, speech.learning.lab, and mlml lab at McGill for their help.

Thanks also to Michael McAuliffe for helping us extract data from Buckey Corpus.


