The effect of production planning locality on external sandhi: a study in /t/

Oriana Kilbourn-Ceron, Michael Wagner and Meghan Clayards
McGill University

1 Introduction

External sandhi processes, in which the target of an alternation is in a different word from the trigger of the alternation, differ from word-internal phonological processes in two important ways: they are subject to **locality** conditions that constrain which two word sequence the process can apply to, and they are more likely to be "optional" or **inherently variable**. Locality conditions on sandhi processes have received various analyses in the literature on prosodic phonology (Cooper and Paccia-Cooper, 1980; Selkirk, 1984; Kaisse, 1985; Odden, 1987; Nespor and Vogel, 1986; Pak, 2008; Selkirk, 2011; Šurkalović, 2016), but they do not link the locality conditions on sandhi to its variability, which is usually just noted and taken at face value. Recent phonological work interested in formalizing phonological variation (Pierrehumbert, 2001; Bybee and Scheibman, 1999; Coetzee and Pater, 2011; Coetzee and Kawahara, 2013, more relevant cites) has not tried to answer the question why cross-word processes specifically should be variable, and how this variability relates to their locality.

In this paper, we draw on speech production planning research in order to connect these two properties of external sandhi to a single underlying source: the locality of speech production planning (see also Wagner, 2011, 2012; Tanner et al., 2015). By developing and extending the predictions made by current models of speech production to external sandhi phenomena, we aim to provide a new perspective on the pattens of variability in connected speech, complementing existing theories of phonological alternations.

First, drawing on insights from current models of speech production, we present our **Production Planning Hypothesis** (PPH), and outline its predictions for external sandhi phenomena. Second, we present a case study of English coronal flapping in order to test its predictions: a production experiment which looks at the effects of syntactic junctures, and a corpus study looking at the effects of lexical frequency of the two words involved. In both studies, we control for the strength of the prosodic boundary separating the words in order to test whether these effects can be reduced to temporal compression and the resulting higher degree of gestural overlap.

1.1 Previous Analyses of Flapping

In English, the coronal stops /t,d/ can be realized as an alveolar flap when they appear between vowels.[1] When the VTV sequence is within a single word, the flap

[1] And also after certain sonorants, e.g. in /ntV/ and /ndV/ sequences, which we will not discus here.

realization is nearly categorical, unless the following vowel is stressed. But when a word boundary intervenes, whether a speaker will flap is variable, and depends on various factors such as speech rate and the strength of the boundary separating two words (e.g. Kahn, 1976, i.a.). This pattern of variability is common across many sandhi processes cross-linguistically – but *why* are segmental processes at word edges often more variable? Intuitively, the target and the trigger of an alternation must be "close enough" in order to interact, but different processes seem to have different locality restrictions governing what counts as close enough (cf. Kaisse, 1985).

Prosodic Phonology (e.g. Nespor and Vogel, 1986) captures locality conditions by pairing the structural description of a process with a domain of application defined by prosodic constituents, which are in turn derived from syntactic structure. Processes like flapping, however, do not seem to have an absolute syntactic upper bound that makes its application impossible, nor is there a clear split between a domain where it categorically applies or does not apply. It seems that flapping can in principle apply across any two words, even when separated by clause junctures, but is increasingly less likely to apply across bigger boundaries (Scott and Cutler, 1984; Fukaya and Byrd, 2005). This pattern is difficult to capture by restriction to a certain type of prosodic domain. Nespor and Vogel (1986) attribute this to variability in prosodic phrasing, in that often a single phonological utterance (the domain of flapping) is restructured into several such phonological constituents, blocking flapping from applying.

Another way that flapping has been analyzed is as a consequence of gestural overlap between the /t,d/ and adjacent vowels (cf. Fukaya and Byrd, 2005). This view is supported by observations that flapping is gradient rather than categorical (Fox and Terbeek, 1977), at least when looking at articulatory measures as opposed to acoustic ones (De Jong, 1998). It is also supported by findings that flapping does not neutralize the distinction between an underlying /t/ and /d/, which is still reflected in small but consistent phonetic differences in the length of the preceding vowel (Malécot and Lloyd, 1968; Herd et al., 2010; Braver, 2011). This is unexpected if flapping involves a categorical change (though see Bermudez-Otero, 2011). Finally, it is also supported by the obeservation that consonants other than /t,d/ are subject to similar temporal reductions in flapping environments (Turk, 1992).

Under the gestural overlap view, the temporal "closeness" of the consonant and vowel gestures is directly related to the likelihood of a flap, and flapping is simply a reflex of greater gestural overlap when the two words are temporally compressed and articulated more closely to each other. Under this view, syntactic and prosodic boundaries exert an effect on the flapping rate by way of inducing a temporal slow down a the boundary due the effect of final lengthening (Byrd and Saltzman, 2003). This view also accounts why other temporal modulations, such as changes in speech rate, affect the overall likelihood of flapping (cf. Browman and Goldstein, 1992).

The account of flapping in terms of gestural overlap hence does relate locality and variability, similar to the PPH. We will see evidence, however, that the locality of production planning has effects even when the degree of lengthening at a boundary is controlled for, suggesting that the temporal compression necessary to allow for gestural overlap may be a necessary but not a sufficient condition for flapping. We will argue that whether a stop is planned as a flap depends not just on the proximity of the gestures, but also on whether the upcoming vowel-initial word is sufficiently planned out at the time of encoding.

1.2 The Production Planning Hypothesis

The PPH can be motivated based on the research on speech production planning, which has shown that the window for detailed phonological encoding is quite narrow. For example, Sternberg et al. (1978) found utterance initiation time correlates with the overall number of words in and upcoming word list, suggesting that the number of items in a list is planned early on. Utterance initiation time also correlates with the number of syllables in the first word, while the number of syllables of later words was irrelevant. This suggests that higher level information is planned over a large planning window, while more detailed, phonological information about the content of a word is planned very locally. According to Levelt's influential model of speech production (Levelt et al., 1999), segmental information is encoded in roughly word-sized planning chunks, but of course the application of a sandhi process requires planning a chunk that encompasses the current word and at least the beginning of an upcoming word. The size of the window for phonological encoding has in fact been shown to vary. Wheeldon and Lahiri (1997, 2002) found that depending on the task, utterance initiation can be driven more by the number of upcoming prosodic words (with more planning time), or the internal complexity of the first upcoming (shorter planning time). We submit that the variability in the application of sandhi rules can be linked to this variablity in the size of the planning chunks.

The likehood of two words being phonologically encoded within the same planning window will be affected by any factor that influences the scope of production planning: syntactic constituency and semantic coherence (Wheeldon, 2013), lexical frequency (Konopka, 2012), cognitive load (Ferreira and Swets, 2002; Wagner et al., 2010), individual differences in planning scope (Swets et al., 2014), and potentially many other factors. For example, when encoding the articulatory plan for a verb, it seems plausible that the segmental content of the upcoming word is more likely to be planned simultaneously if it constitutes its complement (e.g. a direct object), compared to when it constitutes part of the subject of a following sentence. This could have semantic and syntactic motivations, and correlates with the fact that an upcoming subject would be more likely to be set off by a prosodic boundary than an upcoming object. The size of a prosodic boundary itself might affect the likelihood of the following word being planned, or it might be a reflex of whether it has been planned. Krivokapić (2007) showed that post-boundary pauses, an important cue to boundary strength, are affected by complexity of upcoming constituents both in terms of the length and branchingness of prosodic constituents, suggesting that both influence the course of phonological encoding.

The PPH predicts that any phonological alternation which relies on phonological information in upcoming words will show variability, since they cannot apply if the conditioning phonological environment in the next word has not yet been retrieved and encoded, and that the precise locality condition will depend on the kind information that is needed.

In this paper we test three specific factors that are predicted by the PPH to modulate variability:

1. **Strength of prosodic boundary**: a stronger boundary between the stop and the following vowel should have a negative correlation with the probability of flapping;
2. **Strength of syntactic break**: higher level syntactic boundaries should delay planning of the upcoming word and negatively correlate with flapping probability; and
3. **Frequency of the following (triggering) word**: higher frequency words should be retrieved more easily and hence planned earlier relative to the first word, and should show a higher flapping rate.

2 Boundary strength: production experiment

This production experiment examines the effect of phonological, prosodic and syntactic context on the realization of word-final /t/. For this paper, we restrict the analysis to the condition where vowels preceded and followed the target /t/, meeting the segmental conditions for flapping. The presence of a clause juncture is predicted to make it less likely that an upcoming word is planned, and hence decrease the rate of flapping. Similarly, a stronger prosodic boundary plausibly reduces the likelihood that an upcoming word is planned (or maybe strong boundaries are strong partly *because* the upcoming word was not planned yet–we will not try to tease this apart here), and hence the flapping rate is expected to decrease as prosodic boundary strength increases.

2.1 Methods

Participants Twenty-three participants were recruited from the McGill University community.

Materials The materials for the production study consisted of eight item sets in four conditions, varying two factors: *Phonology* and *Syntax*. Each item was a sentence with two clauses, as in (1), where the verb in the initial embedded clause was a nonce word that contained the target word-final /t/. The target /t/ was always preceded by a vowel, and followed by either a vowel- or consonant-initial proper noun depending on *Phonology* condition. The *Syntax* manipulation varied whether the nonce verb was followed by a clause boundary. In one condition, the following word was the object of the nonce verb (*No Clause Boundary* condition), forming a close synactic relationship, while in the other, the following word was the subject of the main clause (*Clause Boundary* condition), creating a large syntactic break after the target word.

Procedure Participants recorded each item in each condition once at normal speaking tempo, and once at a fast tempo. They were allowed to familiarize themselves with the sentence before the recording began.

Phonology	**Syntax**	
	Clause Boundary	*No Clause Boundary*
Consonant	If you **plit**, Alice will be mad.	If you **plit** Alice, John will be mad.
Vowel	If you **plit**, Penny will be mad.	If you **plit** Penny, John will be mad.

Table 1: A sample item from the production experiment, showing the four possible conditions.

Analysis The recorded utterances were force-aligned using the prosodylab-aligner (Gorman et al., 2011), and annotated by a research assistant for type of realization: released, unreleased, glottaled, deleted or flapped. Acoustic measures were also extracted: duration of the /t/, of the preceding vowel, and of the following segment.[2]

Since this paper's focus is on flapping, we restricted the data to only the condition in which a vowel follows the target word. Furthermore, we analyzed only the data elicited at the fast speech tempo, since flapping was very rare at a normal tempo. Finally, we excluded tokens that were followed by a pause, since none of these tokens were ever flapped. This left 368 tokens for investigation.

2.2 Results

The overall rate of flapping annotated in our data was 22.01%. This is in stark contrast to the flapping rate of 93.9% found by Patterson and Connine (2001) for word-medial /t/ in a corpus of conversational speech.

The rate of flapping was lower when a clause boundary followed the target word with a rate of 17.39%, compared to 26.63% when no clause boundary followed. As for the effect of boundary strength, which is operationalized here as lengthening of the preceding vowel, empirical examination suggests that there is a negative correlation with flapping rate. This is shown in Figure 1.

The empirical plots also show that the correlation between vowel duration and flapping rate may only hold if there is a clause boundary following the /t/.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.60	0.76	-3.41	0.00
Syntax.std	0.98	0.46	2.14	0.03
vDur.log.std	-0.92	0.69	-1.34	0.18
Syntax.std:vDur.log.std	0.97	0.82	1.18	0.24

Table 2: Fixed effects coefficients ($\hat{\beta}$, standard errors ($se(\hat{\beta})$), z-scores, and p-values for all model predictors.

The results were analyzed with a logistic regression. A mixed-effects model was fitted to the data, with *Syntax* and *Vowel Duration* as fixed effects. The interaction between these two was also included as a fixed effect, given the differing trends observed for the two conditions of *Syntax*, shown in Figure 1. The model also

[2] set up for why these measures in the experiment intro?

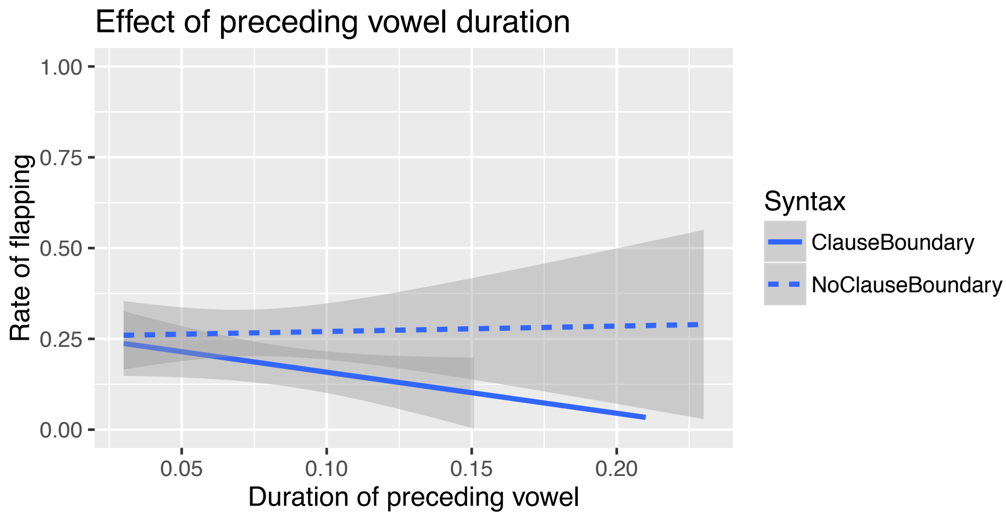

Figure 1: Empirical plots of the correlation between the rate of flapping for the target word and the duration of the vowel preceding the word-final /t/.

included full random effects structure by participant and by item, which controls for possible differences in baseline flapping rates and in effect size for each variable across individuals (Barr et al., 2013). This model is reported in Table 2.

Syntax significantly predicted flapping ($\beta = 0.98$, $p = 0.032$) and vowel duration was in the expected direction, negative ($\beta = -0.919$, $p = 0.18$). The interaction was not significant ($\beta = 0.96$, p = 0.23). The two factors were correlated (r = -0.16, p = 0.002) but model comparison suggested *Syntax* was the better predictor (*Syntax*: $\chi^2(3) = 6.97$, $p = 0.073$; *Vowel Duration*: $\chi^2(3) = 2.25$, $p = 0.52$).

2.3 Discussion

These results show that the presence of a clause boundary after a word-final /t/ has a significant influence on its likelihood of being flapped. Since the model controls for *Vowel Duration* (i.e. final lengthening) as an independent, continuous measure of prosodic boundary strength, we conclude that the effect of the syntactic manipulation is not completely reducible to durational effects associated with clause boundaries. Furthermore, the effect of *Syntax* is not categorical – a clause boundary does not completely block flapping, but rather decreases the probability in a gradient way. For example, holding *Vowel Duration* at its mean value, the probability of flapping is estimated to be 18% if a clause boundary follows, but increases to 36% if no clause boundary follows.

Capturing this effect either by direct reference to syntax or via reference to prosodic domains would require refining the blocking mechanism beyond a dichotomy of application/non-application. On the other hand, this type of subtle effect is completely consistent with the locality of production planning hypothesis. Previous studies have shown that greater syntactic complexity can delay production latencies, suggesting and increased planning load (Ferreira, 1991). Hence, an upcoming noun phrase (e.g. *Alice* in *plit Alice*) should incur less of a processing load if it is the final noun in the embedded clause, where it would be a small one-word constituent, rather than the first noun in a completely new clause, which would initiate planning of the entire new clause. Under the PPH, the delay incurred by the planning of a new clause has consequences for the low-level encoding of phonological segments, in that planning of the final segmental material in the first clause may have to proceed in the absence of information about upcoming segments. If this occurs, flapping will not apply since the licensing conditions for flapping are not met – no following vowel is present at that point in time, and hence a flap cannot be planned.

In sum, the results of this experiment revealed a consistent but gradient effect of syntactic clause boundaries on flapping. This is consistent with the predictions of the PPH, which predicts a probabilistic decrease in the availability of upcoming segmental information under increased planning load. Understanding the effect of syntax as mediated through production planning captures the non-categorical nature of this effect, which could not be explained by current accounts of external sandhi locality restrictions.

The PPH also makes a more general prediction that any factor which affects or delays production planning will have a similar inhibitory effects on flapping, and in the following we turn to frequency effects as observed in a corpus of spontaneous speech.

3 Lexical frequency: corpus study

The effect of lexical frequency on speech production has been shown in many studies, for example response latencies are lower in picture naming (Oldfield and Wingfield, 1965; Jescheniak and Levelt, 1994). A higher frequency of an upcoming word (the word containing the other vowel that the flap is released into) should therefore have the effect that it will be planned earlier relative to the first word (see Tanner et al. (2015) for discussion of potential effects of the target word's lexical frequency).

3.1 Data set

The data source for this study was the Buckeye Corpus of conversational speech (Pitt et al., 2007). We extracted[3] 11863 tokens of words which end in a vowel then /t/ or /d/ and were followed by a vowel-initial word (46.24% were transcribed as flaps). Of these, we excluded tokens where the following word was a disfluency marker[4] (18.26% of tokens), and where the following word was reduced to a syllabic consonant on the surface (0.07% of tokens). This left 11863 tokens for analysis.

Word frequencies were retrieved from SUBTLEX-US, a database of word frequencies based on film and television subtitles (Brysbaert and New, 2009). Boundary strength was included as a predictor, since results from our production experiment suggested that this also affects flapping rate. We operationalized bound-

[3]We gratefully acknowledge the assistance of Michael McAuliffe in extracting these data.

[4]These words were 'uh','um','okay','yes','yeah','oh','heh','yknow','um-huh','uh-uh','uh-huh','uh-hum','mm-hmm', and 'and', all of which were associated with flapping rates well below the mean by word-type.

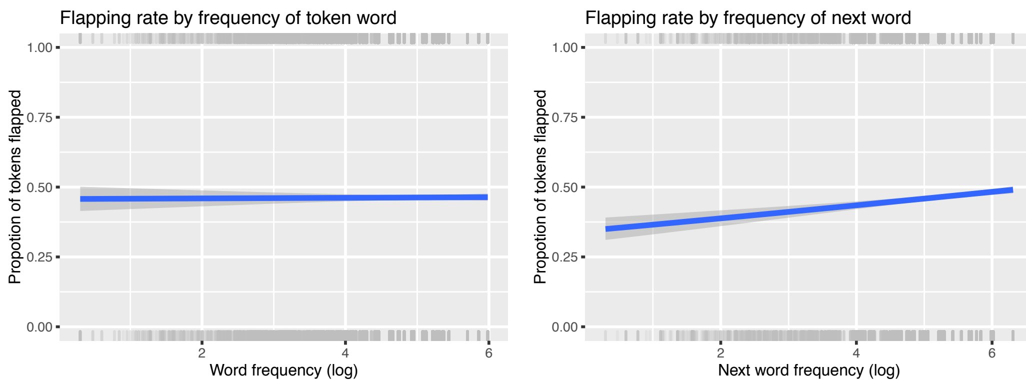

Figure 2: Correlation of token word and following word frequency with flapping in the current data set

ary strength as observed/expected word duration, where expected duration was the mean duration for that word in the entire corpus. This variable reflects the amount of pre-boundary lengthening, an effect which has been shown for English to be well-correlated with boundary strength (Wightman et al., 1992). This measure also relates to the predictions of gestural overlap, since a greater O/E ratio would reflect less temporal overlap between the adjacent segment gestures. Number of syllables was calculated for both the token and following word, with each syllabic segment in the Buckeye surface transcription counting as one syllable.

3.2 Model structure

A mixed-effects logistic regression was fit using the `glmer` function in the `lme4` package (Bates et al., 2014) package in R (Team et al., 2013). The dependent measure was whether or not the underlyingly /t,d/-final word was annotated as a flap in the surface transcription. The log-transformed lexical frequency of the token word and the following word were standardized and included as fixed effects. Control predictors included *presence of pause*, a binary variable, underlying voicing of the word-final segment (/t/ or /d/), *observed/expected (OE) word duration*, log-transformed and standardized, and binary variables tracking whether the token and following words were monosyllabic or not. Random effect structure included by-speaker and by-word intercepts, and by-speaker and by-word slopes for following word frequency.

3.3 Results

Table 3 shows the model estimates for the fixed effects coefficients. Each coefficient represents the estimated change in log-odds of flapping when other predictors are held at their mean observed values, except *pause* which is held at 0 (no pause).

The model finds a reliable difference between flapping rates for /d/ and /t/ once the effects of other variables are taken into account, with /t/-final words more likely to be flapped. The estimate for *pause* is negative and of very large magnitude compared to other effects ($\hat{\beta} = -4.8$, $p < 0.001$), confirming that flapping in the presence of a pause is very rare (just under 1% of tokens followed by pause in the subset under analysis are annotated as flaps). The effect of *O/E word duration* was not re-

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.654	0.115	5.681	0.000
underlying_final.std	0.338	0.117	2.888	0.004
log10wf.std	0.213	0.124	1.712	0.087
log10wf.next.std	0.290	0.096	3.010	0.003
oedur.log.std	0.047	0.133	0.356	0.722
sylls.std	-0.182	0.142	-1.284	0.199
sylls.next.std	-0.021	0.091	-0.231	0.818
pause	-4.796	0.239	-20.042	0.000
log10wf.std:log10wf.next.std	0.228	0.136	1.677	0.094

Table 3: Fixed effects coefficients ($\hat{\beta}$, standard errors ($se(\hat{\beta})$), z-scores, and p-values for all model predictors.

liably different from 0 ($\hat{\beta}$ = 0.05, p =0.722). Nor does the number of syllables in the target or following word have a statistically significant effect.

As for our crucial variable, the model confirms that the lexical frequency of the second, vowel-initial word in the sandhi pair has a reliable effect on the likelihood of flapping. Higher frequency vowel-initial words are more likely to trigger flapping on a preceding coronal stop ($\hat{\beta}$ = 0.29, p =0.003). The frequency of the coronal-final word itself showed a positive trend, but the effect was not statistically reliable in the full model ($\hat{\beta}$ = 0.21, p =0.087). There was also a positive interaction between these predictors: increasing the frequency of both words in the sandhi pair increases the likelihood of flapping even more than would be expected from the sum of the independent effects of each word's frequency ($\hat{\beta}$ = 0.23, p =0.094), though again the interaction was not reliable in the model with full random effect structure.

3.4 Discussion

The results of the corpus study show that there is a strong correlation between the *following word frequency* and the likelihood of flapping. This is consistent with the prediction of the PPH that flapping should be more likely when the following word is easier to plan. Lexical frequency is well-known to have a facilitatory effect on word form retrieval (Jescheniak and Levelt, 1994). This has the consequence that phonological encoding of the following vowel-initial word may begin sooner for more frequent words, thus making the vowel more available to trigger flapping on the target coronal-final word, according to the PPH.[5]

The measure of prosodic boundary strength that we included in this analysis, *O/E duration*, showed a negative trend in empirical analysis. However, its effect was not reliably different from 0 in the model once other factors were controlled and a by-word random slope was included. There was a high variance in estimated random slopes by word, suggesting that the *O/E duration* predictor may be reflecting different information in different words.

[5] We believe that a better measure of the predictability of an upcoming word would be an estimate of its conditional probability given the first word, but we have not yet computed this.

4 General discussion

These results show that both spontaneous and lab-elicited speech exhibit patterns of variabilty that are predicted by the locality of production planning hypothesis.

The production experiment results show that syntactic boundaries have a gradient blocking effect on flapping likelihood, and that this effect does not appear to be entirely due to the temporal slow-down effects at the clause boundary, since it was signficant when also controlling for pre-boundary lengthening. Interpreting this finding within the Prosodic Phonology framework is not straightforward – it is unclear which domain of application flapping could be restricted to in order to explain this pattern, since flapping does occur in both syntactic structures. One possibility would be to adopt Nespor and Vogel's (1986) restructuring mechanism, which can variably restucture two adjacent domains into a single one based on speaking rate and other factors. In the absence of independent evidence that the non-flapping cases have been prosodically restructured, an interpretation of the syntax effect under the PPH offers a more explicit mechanism for the effect, and one with many testable predictions. Although prosodic groupings are clearly an important part of phonological description, it may be possible to gain important insights into external sandhi segmental processes without explicitly tying them phonological domains by taking into account locality of speech production planning effects.

As discussed in the introduction, syntactic locality effects and variability in general could also receive a gestural overlap account, as in Fukaya and Byrd (2005). Edges of larger syntactic constituents are known to be associated with final lengthening (Wightman et al., 1992), which could interfere with the degree of overlap needed to cause a percept of flapping. However, the results of our production experiment suggest that the inhibitory effect of a clause boundary is above and beyond durational effects. Final lengthening, measured as the duration of the vowel preceding the target /t/, did not have a statistically significant effect in a model that also included syntax, and a likelihood ratio test showed that it did not significantly improve the model compared to a model with syntax as the only predictor. These results are not counterevidence against a gestural overlap account of flapping, but show that such accounts could be fruitfully paired with the PPH to account for syntactic and other non-temporally based effects that contribute to the variability of external standhi processes.

The corpus study revealed a positive effect of the following word's frequency. This confirms another prediction of the PPH: words that are planned more quickly are more likely to influence the phonological encoding of the word that precedes them, which in this case means triggering flapping. This type of effect is not part of any theories of locality, but has been investigated in the literature on probabilistic phonetic reduction, in the vein of Jurafsky et al. (2001). Pluymaekers et al. (2005) showed that for seven high frequency words in Dutch, mutual information with the following word was predictive of reduction, with fewer segments realized when mutual information was high. Torreira and Ernestus (2009) found an effect of bigram frequency with the following word on the acoustic realization of /t/. Ernestus et al. (2006) showed that a sandhi phenomenon in Dutch, voice assimilation, is more likely to occur within a compound when the two component words have a high cooccurence frequency.

On the interpretation of flapping as a reductive process, our results are in line with these previous findings, and compatible with the general idea that prosody in general and prosodic phrasing in particular is sensitive to predictability and information density (Aylett and Turk, 2004; Turk, 2010). The PPH offers an explanatory mechanism for such frequency and predictability effects.

The predictions of the PPH and, for example, the Probabilistic Reduction Hypothesis of Jurafsky et al. (2001), are fairly similar in the case of external sandhi processes that are reductive. One way their predictions could be teased apart is by examning sandhi processes that are *non*-reductive: the PPH predicts similar effects of the likelihood of an upcoming word being planned on the application rate of the sandhi process, say one in which a segment is inserted if an appropriate phonological environment follows, like liaison in French. The realization of liaison consonants, which depends on an upcoming word starting with a vowel, should increase with a greater predictabilty of an upcoming word. For such non-reductive processes, the Probabilistic Reduction Hypothesis would make no prediction, or maybe in fact predict a lower rate of liaison with greater predictability of the upcoming word, since predictability should correlate with more reduction.

5 Conclusion

This paper has developed the predictions of the production planning hypothesis, and tested three of them on a case study of English flapping. Our production experiment tested the effect of syntactic clause boundaries and prosodic boundary strength on the likelihood of flapping. Results showed that clause boundaries do make flapping less likely, but do not rule it out completely. Given findings that production planning is constrained by syntactic constituency (cf. Wheeldon et al., 2011), the PPH provides an explanatory mechanism for this effect.

The corpus study we conducted found that the frequency of the word following the coronal-final test word had a significant influence on the rate of flapping. The more frequent the following word, the more likely the coronal stop was to be flapped. This is again consistent with the PPH: given that higher lexical frequency facilitates a word's retrieval, the PPH predicts more frequent words to be phonologically encoded sooner relative to the first word, and therefore lead to a higher flapping rate.

Future work should investigate if the effects of the PPH also hold for non-reductive external sandhi processes, such as liaison in French.

References

Aylett, M. and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, 47(1):31–56.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278.

Bates, D., Maechler, M., Bolker, B., Walker, S., et al. (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version*, 1(7).

Bermudez-Otero, R. (2011). Cyclicity. In van Ostendorp, M., Ewen, C., Hume, E., and Rice, K., editors, *The Blackwell Companion to Phonology*. Wiley-Blackwell.

Braver, A. (2011). Incomplete neutralization in american english flapping: A production study. *University of Pennsylvania Working Papers in Linguistics*, 17(1):5.

Browman, C. P. and Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4):155–180.

Brysbaert, M. and New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for american english. *Behavior research methods*, 41(4):977–990.

Bybee, J. and Scheibman, J. (1999). The effect of usage on degrees of constituency: the reduction of don't in english. *Linguistics*, 37(4):575–596.

Byrd, D. and Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31:149–180.

Coetzee, A. W. and Kawahara, S. (2013). Frequency biases in phonological variation. *Natural Language & Linguistic Theory*, 31(1):47–89.

Coetzee, A. W. and Pater, J. (2011). The place of variation in phonological theory. In John A. Goldsmith, Jason Riggle, A. C. L. Y., editor, *The Handbook of Phonological Theory, Second Edition*, pages 401–434. Wiley-Blackwell.

Cooper, W. E. and Paccia-Cooper, J. (1980). *Syntax and speech*. Number 3. Harvard University Press.

De Jong, K. (1998). Stress-related variation in the articulation of coda alveolar stops: flapping revisited. *Journal of Phonetics*, 26(3):283–310.

Ernestus, M., Lahey, M., Verhees, F., and Baayen, R. H. (2006). Lexical frequency and voice assimilation. *The Journal of the Acoustical Society of America*, 120(2):1040–1051.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30(2):210–233.

Ferreira, F. and Swets, B. (2002). How incremental is language production? evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46:57–84.

Fox, R. and Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in american english. *Journal of Phonetics*, 5:27–34.

Fukaya, T. and Byrd, D. (2005). An articulatory examination of word-final flapping at phrase edges and interiors. *Journal of the international phonetic association*, 35(01):45–58.

Gorman, K., Howell, J., and Wagner, M. (2011). Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics*, 39(3):192–193.

Herd, W., Jongman, A., and Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in american english. *Journal of Phonetics*, 38(4):504–516.

Jescheniak, J. D. and Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4):824.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological studies in language*, 45:229–254.

Kahn, D. (1976). *Syllable-based generalizations in English phonology.* PhD thesis, Massachusetts Institute of Technology.

Kaisse, E. M. (1985). *Connected speech: the interaction of syntax and phonology*. Academic Press.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66(1):143–162.

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of phonetics*, 35(2):162–179.

Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(01):1–38.

Malécot, A. and Lloyd, P. (1968). The/t:/d/distinction in american alveolar flaps. *Lingua*, 19(3-4):264–272.

Nespor, M. and Vogel, I. (1986). *Prosodic phonology*. Walter de Gruyter.

Odden, D. (1987). Kimatuumbi phrasal phonology. *Phonology*, 4(01):13–36.
Oldfield, R. C. and Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17(4):273–281.
Pak, M. (2008). *The postsyntactic derivation and its phonological reflexes*. PhD thesis, University of Pennsylvania.
Patterson, D. and Connine, C. M. (2001). Variant frequency in flap production. *Phonetica*, 58(4):254–275.
Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. *Frequency and the Emergence of Linguistic Structure*, 45:137.
Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler-Lussier, E. (2007). Buckeye corpus of conversational speech (2nd release). *Columbus, OH: Department of Psychology, Ohio State University*.
Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken dutch. *The Journal of the Acoustical Society of America*, 118(4):2561–2569.
Scott, D. R. and Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior*, 23(4):450–466.
Selkirk, E. (2011). The syntax-phonology interface. *The handbook of phonological theory*, 2:435–483.
Selkirk, E. O. (1984). *Phonology and syntax*.
Sternberg, S., Monsell, S., Knoll, R., and Wright, C. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. *Information processing in motor control and learning*, pages 117–152.
Šurkalović, D. (2016). *The No-Reference Hypothesis: A Modular Approach to the Syntax-Phonology Interface*. PhD thesis, UiT Norges arktiske universitet.
Swets, B., Jacovina, M. E., and Gerrig, R. J. (2014). Individual differences in the scope of speech planning: evidence from eye-movements. *Language and Cognition*, 6(01):12–44.
Tanner, J., Sonderegger, M., and Wagner, M. (2015). Production planning and coronal stop deletion in spontaneous speech. In *Proceedings of the 18th International Congress of Phonetic Sciences*.
Team, R. C. et al. (2013). R: A language and environment for statistical computing.
Torreira, F. and Ernestus, M. (2009). Probabilistic effects on french [t] duration. In *10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, pages 448–451. Causal Productions Pty Ltd.
Turk, A. (1992). The american english flapping rule and the effect of stress on stop consonant durations. *Working papers of the Cornell phonetics laboratory*, 7:103–133.
Turk, A. (2010). Does prosodic constituency signal relative predictability? a smooth signal redundancy hypothesis. *Laboratory phonology*, 1(2):227–262.
Wagner, M. (2011). Production-planning constraints on allomorpy. *Proceedings of the Acoustics Week in Canada. Canadian Acoustics*, 39(3):160–161.
Wagner, M. (2012). Locality in phonology and production planning. In *Proceedings of Phonology in the 21 Century: Papers in Honour of Glyne Piggott. McGill Working Papers*.
Wagner, V., Jescheniak, J. D., and Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2):423.
Wheeldon, L. (2013). Producing spoken sentences: The scope of incremental planning. In Perrier, P. and Verlag, P. L., editors, *Cognitive and physical models of speech production, speech perception, and production-perception integration*.
Wheeldon, L. and Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, 37:356–381.
Wheeldon, L. and Lahiri, A. (2002). The minimal unit of phonological encoding: prosodic or lexical word. *Cognition*, 85(2):B31–B41.
Wheeldon, L., Smith, M., and Apperly, I. (2011). Repeating words in sentences: Effects of sentence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5):1051.
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of*

America, 92:1707–1717.