Rational phonological lengthening in spoken Dutch

Harry Tily

Brain and Cognitive Sciences,

MIT,

Cambridge MA 02139,

USA$^a$)

Victor Kuperman

Linguistics and Languages,

McMaster University,

Hamilton,

ON L8S4M2,

Canada
Abstract

Dutch allows optional schwa insertion between a liquid and obstruent in words like *film/*fɪlm* (“film”) and *dorp/*dɔrp* (“village”), lengthening the word by one syllable. This epenthesis is productive and widespread, and is understood to be a phonological rather than phonetic process. A corpus analysis shows that a speaker’s choice between the variant forms is influenced by probability: words that are less frequent or less probable given their immediate or discourse context are more likely to be lengthened. This may reflect a rational communication strategy in which language is manipulated to efficiently transmit information. As these results unambiguously show that *lengthening* is probabilistically influenced, they are informative to the understanding of the production mechanisms underlying pronunciation variation.

PACS numbers:

Keywords: communication, Dutch, epenthesis, frequency, phonology, predictability, production, rational analysis, uniform information density, variation
I. INTRODUCTION

A. Probability-sensitive variation

It is now widely accepted that the probability of linguistic material influences its realization in speech, with many different measures of probability having been demonstrated to affect word production. Higher frequency leads to faster speech onset, segment deletion/reduction, and shorter duration (see references in Kuperman et al., 2007). Words that are repeated, or predictable in context are more likely to be spoken with relatively shorter duration and reduced intensity or acoustic distinctiveness (for an overview see Bell et al., 2009).

Several researchers have argued that this tendency for more probable material to be accelerated, reduced, or omitted relative to less probable material reflects a rational online communication strategy. A rational strategy for speech (in the sense of Anderson, 1990:1-276) would suggest reducing or eliminating material which is more predictable, and therefore redundant enough to be reduced without impacting communicative success (for an overview of this and closely related proposals see Jaeger (2010)).

However, the production mechanisms underlying probabilistically conditioned variation are not fully understood. In fact, it is not even clear whether speakers are reducing more probable material or elaborating less probable material. Jurafsky et al. (2001) proposed a Probabilistic Reduction Hypothesis holding that speakers reduce wordforms which have higher probability (see Jaeger (2006) for a more general restatement). Such processes may reflect the architecture of the production system: Bybee and Hopper (2001) suggest that the articulatory representations of more frequently accessed words become reduced. Nevertheless, most findings are equally compatible with an alternative explanation, namely that speakers lengthen less predictable material. This could be a strategy to “buy time” when lexical retrieval is slowed by low accessibility of the words being spoken or planned (Bell et al., 2009).¹

It is hard to differentiate these explanations because most previous work has examined gradient

¹Electronic address: hjt@mit.edu
phonetic properties, like duration and amplitude, vowel centralization, degree of coarticulation, or spectral center of gravity. For none of these properties is there a basic or canonical representation which a given token could be compared against: e.g., there is no lexically fixed standard duration for a given word. Effects of probability on syntactic or word level choices typically involve the suppletion of two forms, such as the presence or absence of complementizer and relativizer that (e.g. Jaeger, 2006; Levy and Jaeger, 2007). Here too, it is unclear whether the form with or without the relativizer is basic. A case for reduction can be made for the copula alternation (e.g. *I am/I’m*) studied by Frank and Jaeger (2008) and the alternation between “full” and “reduced” phonological forms like */av/* vs */ə/* for *of* (Jurafsky et al., 2001). Still, as these involve a small (closed-class) set of forms, it remains possible that these cases involve suppletion between pairs of items rather than a productive reduction process.

The strongest candidate for an unambiguous reduction effect is the deletion of phonological segments. Bybee (published as Hooper, 1976) showed that schwa omission was more common in high-frequency words like *every* and *evening* than low-frequency *mammary* and *artillery*. Bybee (2000) also found higher rates of final t/d-deletion in high-frequency English past tense verbs like *told* relative to low-frequency *meant*. Similar findings obtain for Spanish final */r/* deletion (Díaz-Campos and Ruiz-Sánchez, 2008). Word predictability influences the deletion of medial and final */t/* or */d/* more generally (Jurafsky et al., 2001; Raymond et al., 2006). This deletion is well-studied (e.g. Guy, 1980, 1991; Tagliamonte and Temple, 2005), and widely taken to be a variable phonological process.²

To summarize, in variation more probable material often appears reduced relative to less probable material. Most variations cannot be definitively classified as reduction or lengthening, though several appear to be reduction. Although lengthening has been suggested to be an underlying production mechanism (e.g. Bell et al., 2009), there are no examples of unambiguous lengthening.

In this article, we demonstrate probability-sensitive variation in a domain which directly contrasts with Bybee’s schwa omission: the insertion of schwas in spoken Dutch. This is the finding which can only be explained as lengthening of low-probability material. We argue that this phenomenon can be considered a rational phonological process, in the sense that it is adaptive for
B. Dutch schwa epenthesis

Dutch allows variable insertion of a schwa between /l/ or /r/ and a following non-coronal consonant, as well as between /r/ and /n/. This variant is common, though more so when the affected consonant cluster is entirely within a syllable coda (1a) than when it crosses a syllable boundary (1b).

(1) a. melk /mɛlk/∼/mɛ-l@k/ ‘milk’
hulp /hυlp/∼/hυ-l@p/ ‘help’
berg /bɔrx/∼/bɔ-ʁ@x/ ‘mountain’
korf /koRF/∼/kɔ-ʁ@f/ ‘basket’
b. filmer /fil-m@r/∼/fi-l@-m@r/ ‘cameraman’
     ergens /ɛr-ɛns/∼/ɛ-r@-ɛns/ ‘somewhere’

(Adapted from Warner et al. 2001)

Warner et al. (2001) argue that this is a phonological alternation, ruling out the alternative possibility that the schwa is not present in the speaker’s phonological representation, but merely perceived due to retiming of the neighboring segments (“targetless schwa”). Their evidence comes from the realization of the preceding liquid: /l/ is articulated light (onset-like) before both epenthetic schwas and non-optional schwas in matched words, but dark (coda-like) in the same words produced without the optional schwa.

Schwa epenthesis manifests itself variably depending on many linguistic and sociolinguistic factors, including phonological environment, gender, age, and regional dialect (Swerts et al., 2001; Kloots et al., 2004). Here, we show that the probability of a word — the amount of information it carries — influences schwa realization, see (Hume and Bromberg, 2005).

Unlike the phenomena discussed in section I.A, Dutch schwa is unambiguously a lengthening process if the schwa-free form is basic. This is accepted by native speakers and linguists, on the following grounds. The schwa is not represented in standard spelling; epenthesis is generally less frequent; and participants tend to ignore the epenthized vowel in word manipulation tasks (van
Donselaar et al., 1999). Additionally, while segment reduction could be argued to be phonetic rather than phonological, insertion cannot be explained as gradient articulatory undershoot, so vowel epenthesis is unambiguously a phonological process (see Warner et al., 2001). Showing that a process of phonological lengthening is probabilistically conditioned will have implications for our understanding of online speech production.

II. CORPUS STUDY

From the manually transcribed sections of the Corpus of Spoken Dutch (Oostdijk et al., 2002), we extracted all words containing possible epenthesis environments. Since the transcription standards for the Flemish part of the corpus omit epenthetic schwas, we only report Netherlands Dutch data. We exclude speakers not born in the Netherlands. We used the provided manual phonological transcript to determine schwa presence. To verify the quality of this coding, the second author (VK, a non-native speaker) and a volunteer (MV, a native speaker) each listened to 132 randomly selected words out of context, blind coding each for schwa presence. Agreement between MV and the transcript was 89% (Cohen’s $\kappa = .78$), and between VK and the transcript 86% (Cohen’s $\kappa = .71$). These numbers indicate very good reliability of the transcript.

We consider four measures of word probability. The first, frequency, measures language-wide probability. The second and third measure probability given the local linguistic context: forward and backward bigram probability, the word’s probability conditioned on the previous/following word respectively. These three measures were estimated from the Twente Corpus of Dutch Newspapers (Ordelman, 2002), which contains over 300 million words. Finally, we define discourse-level probability, or thematicity as the ratio of the word’s frequency in the current document (conversation/text) to its log frequency in the language. This measure is based on the Term Frequency-Inverse Document Frequency measure commonly used in Information Retrieval (Salton and Buckley, 1988) measuring the word’s relatedness to the discourse topic, correcting for language-wide frequency (function words are frequent without being topical, for instance.) Following standard practice, we used the logarithm of all four probabilistic predictors (see Table 1 for pairwise corre-
4279 tokens were suitable environments for epenthesis. We excluded the 1373 (32%) cases which were pronunciation variants other than the basic and epenthesized forms. In many of these cases it would be impossible to tell whether a schwa was the epenthesized schwa or some other reduced vowel.

We coded the following control variables, to be entered as fixed effects in a regression:

- the *speech rate* in syllables per millisecond\(^4\)

- *across syllables:* whether syllabification leaves the liquid and following obstruent in the same (e.g. *berg*, /bɛrɔx/) or different syllables (e.g. *ergens*, /ɜrɛn̩s/)

- *morphological complexity:* single-morpheme vs multiple-morpheme

- *sentence finality:* whether the word is the last in a sentence

- *stress:* whether the environment falls in the rhyme of a lexically stressed syllable

- *liquid:* whether the previous liquid is /l/ (e.g. *melk*) or /r/ (e.g. *berg*)

- the *age* of the speaker in years

- the *sex* of the speaker

- the *spontaneity* of speech: conversation vs read speech

- the *occupation* type of the speaker: 3 skill-level classes plus student

Numerical predictors were centered and divided by two standard deviations to aid coefficient interpretability (Gelman, 2008). We removed the 1 case where any covariate was further than 3 standard deviations from its mean.

Finally, we included grouping factors to be entered into a multilevel regression model as random effects:

- word identity (52 types)
• speaker identity (387 individuals)

• speaker birthplace (77 locations)

We also coded for persistence, indicating whether the most recent potential epenthesis host was in fact realized with a schwa. This proved a significant predictor, with repeated epenthesis much more likely. However, it roughly halves the dataset since it is undefined for cases that follow the beginning of a document or are a pronunciation variation other than the dictionary form. In such a small subset, many of the control variables previously shown to affect the outcome do not reach significance whether persistence itself is included or not, so we exclude it from discussion here.

Using the lme4 package (Bates and Maechler, 2009) for R (R Development Core Team, 2009), we fitted a mixed-effects logit model (see Jaeger, 2008) to predict vowel epenthesis from the variables listed above. We modeled random slopes for the probabilistic predictors given speaker, and random intercepts only for the other two grouping factors. We added interactions between sex and age, since this seemed a plausible effect. To maximize the conservativity of any inferences about the probabilistic predictors of interest, we also included the interaction between each and speech rate.

This initial model was then trimmed to exclude predictors we could be confident had no or very little impact on epenthesis using the “drop1” procedure: we tested the removal of each variable in turn, and excluded the one with the least impact on goodness-of-fit. This was repeated until the resulting model was a better fit to the data than any nested model by $p_{\chi^2} > .1$ (see Table 2). We note that stepwise model selection procedures may lead to anti-conservative inferences in some cases, but we used this approach in order to yield a single interpretable model from the large set of predictors we have available. We did not consider the removal of any random effects. We then evaluated the significance and direction of effects in the final trimmed model. In the trimmed model, no correlation between two predictors exceeded $|0.25|$, indicating no obvious collinearity problem. The final model is shown in Tables 3 and 4.
III. RESULTS AND DISCUSSION

Like Swerts et al. (2001) we find that speakers’ choice between basic and epenthized forms varies along sociological divisions: older speakers epenthize more, as do those born in certain regions. Inspection of the best linear unbiased predictors of the model indicates that epenthesis rates are higher in the south, see Figure 1. Individuals also vary substantially from each other in their overall rates of epenthesis, as can be seen from the table of random effects. There is no apparent effect of sex or occupation type, and no difference between read speech and natural conversation. Epenthesis is more likely in single-morpheme words, syllable-internal environments, and with a preceding /r/. It is also more likely at the end of an utterance, which can be understood as phrase-final lengthening. In contrast, and similarly to Collins and Mees (1996:1-363), we find no main effect of speech rate. Lexical stress too appears to play no role, although this measure is somewhat confounded with the across-syllable and morphological complexity controls, so we cannot be certain which of these measures are independently relevant.

INSERT FIGURE 1 ABOUT HERE

FIG. 1. Model estimates for variation given birthplace.

Most interestingly, three of our four probability measures predict epenthesis. Although we found no evidence for an effect of predictability given the following word, there was a marginally significant tendency for an effect of frequency, and strong evidence for effects of predictability given the preceding word and for thematicity. All these effects act in the same direction: less probable words are more likely to include an epenthetic schwa. We note that frequency and thematicity are related by their definition, and in exploratory modelling we found the frequency effect to be significant when thematicity was excluded.

To check that our results are not confounded with the exclusion of words pronounced other than the full or epenthized form, we fitted a logistic regression model to predict the likelihood
of such pronunciation with the same fixed and random effects as above. Only one probabilistic measure showed reliable influence, i.e., a higher likelihood of another pronunciation is the following bigram is more probable. We conclude that factors affecting our exclusion criteria are largely different from the ones affecting the decision to epenthesize or not.

Given that all covariates were standardized by dividing by two standard deviations, the absolute coefficient values can be interpreted as relative effect sizes (Gelman, 2008), revealing that thematicity in particular has a relatively large effect on the outcome, comparable in size to the dispreference for epenthesis with the liquid /l/ previously reported by Swerts et al. (2001), for example. It is possible that the effect of thematicity is somewhat modulated by speech rate, having more influence in fast speech and less in slow speech, although this effect is only marginally significant. The random effect estimates show that speakers vary substantially in their sensitivity to thematicity, although there is no evidence for similar variation for frequency or predictability. Evidently, the production system can lengthen less predictable words as hypothesized by Bolinger (1963) and in keeping with Bell et al.’s (2009) suggestion that probabilistically conditioned variation allows speakers to buy more time when words being planned are hard to retrieve from memory or integrate into linguistic structure. Following Bell et al.’s hypothesis, one possible interpretation of these results is that they lend evidence to availability-based production (e.g. Ferreira and Dell, 2000).

A second explanation of our findings is that schwa epenthesis represents a (conscious or automatic) attempt on the part of the speaker to ensure that an unpredictable word is correctly perceived by the hearer. Lindblom’s (1990) hyper- and hypo-articulation (H&H) theory suggests that both phonetic reduction and elaboration can be understood as a compromise between speaker preference for minimize on one hand, and the necessity to convey enough explicit signal to enable the hearer to identify the word in context. A more general statement of this notion that extends beyond phonetic variation is Uniform Information Density (UID: Jaeger, 2006; Levy and Jaeger, 2007; Frank and Jaeger, 2008; Jaeger, 2010) and closely related theories (reviewed in Jaeger and Tily, 2011). UID uses information theory to formalize a hypothesis similar in spirit to Lindblom’s, deriving the prediction that speakers will mete out information at a roughly constant rate. There-
fore, if a word is less predictable (i.e. it conveys more information, by Shannon’s (1948) definition) it should be lengthened if possible. Information-theoretic explanations for linguistic phenomena and for phonological typology in particular are becoming increasingly influential (see Goldsmith, 2002; Hall, 2009). If our results do indeed indicate that an online phonological choice is made to maintain communicative efficiency, they align neatly with similar claims made for disfluency and gesture production, lexical choices, morphosyntactic choices, syntactic choice and even the content of full sentences in running texts (for an overview see Jaeger and Tily, 2011).

Since we argue that our production results reflect a rational trade-off between speaker and hearer pressures, it is useful to evaluate them in the light of related studies from comprehension. In word and phoneme detection tasks, van Donselaar and colleagues found that words with the inserted schwa were recognized faster than those without, despite their additional duration and lower frequency (van Donselaar et al., 1996, 1999). They argued that the schwa reduces gestural overlap between the consonants, making the form perceptually clearer. One might think that a good production strategy would be to always insert the schwa, thereby maximizing the probability that the comprehender correctly and easily processes the word. In fact, our results show that speakers insert it more frequently when the message is hard to predict, and hence when the comprehender benefits the most from any additional help: speakers chose the shortest wordform that will be comprehended with some level of reliability. Jaeger (2007) discusses a similar set of findings for that omission in relative clauses: although reading times are always faster when that is included, it is included less often when the relative clause is predictable. Linguistic alternations where length is inversely related to probability can be considered rational in the sense that they are adaptive for concise but error-bounded communication over a noisy acoustic channel. Longer forms are preferred where the speaker or hearer might require more time to process the form correctly, or where a form that is short and thus more easily lost in transmission would be difficult to infer from the context. This minimizes the expected level of communication error. In other situations, short forms are used, thus maximizing conciseness. In fact, Jaeger (2007) shows that reading times are accelerated by the inclusion of that by on average the same amount as they are slowed by unpredictability. Thus writers use the optional word sparingly but in a way that avoids
spikes in per-word comprehension difficulty.

There may also be speaker-internal motivations for epenthesis: Booij (1995:1-205) suggests that the schwa renders consonant clusters easier to articulate, even though it increases duration. Our findings could be explained as a compromise between purely speaker-internal pressures, if speakers opt to make articulation easier when retrieval/planning is harder. However, we know of no production theory in which that kind of trade-off is predicted. Additionally, the schwa epenthesis environment is far from the most complex coda permitted by Dutch phonotactics, and van Donselaar et al. (1999) note that abbreviations and nicknames show no tendency to avoid it (e.g. *direkteur* → *dirk*, “director”).

Schwa epenthesis in Dutch is often discussed alongside morpheme-final /n/ deletion in /ən/ contexts (Booij, 1995). Both phenomena are variable both within and between speakers, and the contexts which they favor appear to share at least some conditioning environments. Previous studies on both spontaneous (Van de Velde and van Hout, 1998) and read speech (Van de Velde and van Hout, 2000) have found /n/-deletion to be less common in monomorphemic than in polymorphemic words, in the Southern rather than the Northern dialect of Dutch, and to vary with sex and age. The reported morphological and dialectal biases for nonreduced forms in /ən/ converge perfectly with the preferences for fuller (epenthesized) forms that we observe here. Likewise, we replicate the sex by age interaction of Van de Velde and van Hout (2000) in predicting the epenthesis rate. Finally, both schwa epenthesis and final-/n/ deletion are sensitive to the rhythmic context of the word (Kuijpers and van Donselaar, 1998).

While strong parallels exist between these two variation phenomena, it is important to note that only epenthesis can be unequivocally labeled as lengthening. The /n/ deletion process appears to be a reduction of the full form: the presence of /n/ is codified in word spelling, and the full form is accepted as canonical by both linguists and native speakers.

Finally, we consider the implications of these results for dominant models of speech production. In models based on Levelt (1989:1-566), perceived variation may arise at an early *formulation* stage due to retrieval of a variant phonological form or modification of a phonological plan, or at an *articulation* stage due to gestural variation. Variations arising at different stages are predicted
to only show sensitivity to factors involved at that point in planning. For instance, Raymond et al. (2006) argue that word internal t/d-deletion primarily arises due to early phonological processes, and therefore is sensitive to phonological context, frequency, and stylistics, while word final t/d-deletion arises due to later gestural overlap and lenition, and so is sensitive to speech rate, fluency, and contextual predictability. Accepting that the Dutch schwa alternation is a phonological process (Warner et al., 2001), it must originate in the formulation stage, and therefore should only be sensitive to processes that act at that point. Accordingly, we and others have found lexical variation and an influence of stylistic and speaker-internal factors — all of which apply in phonological planning — and no influence of speech rate, an articulatory factor. However, we also found influences of predictability and thematicity, which is incompatible with the assumption that contextual probability can only influence articulatory level processes. Therefore, our results call into question the claim of Raymond et al. that the formulation and articulatory stages of speech production can be differentiated by their differential sensitivity to probabilistic factors. Rather, they add to a growing body of findings that probabilistic constraints on language influence production choice from the early planning stages, not merely in articulation (see discussion and references in Bell et al. 2009).

**Endnotes**

1. Note that if the duration of a word is affected by its own probability, this explanation supposes that articulation begins before words are fully accessed, or that longer duration and fuller articulation are the result of lower activation or less complete retrieval.

2. See however Browman and Goldstein (1991) and Bybee (2000) for the suggestion that perceived deletion may be the extreme end of a continuum of phonetic variation, leading to restructuring and categorical omission in some lexical items for some speakers.

3. Both speakers in fact rated 200 cases, but since 68 of these were pronunciation variants and therefore excluded from our analysis later, we ignore these for the purposes of calculating reliability.
4. Following Jurafsky et al. (2001) we approximately located the intonational phrase by taking the smallest region containing the word bounded by an utterance boundary or a pause of 500ms. The speech rate measure is the number of syllables in that region divided by its duration in milliseconds, excluding the critical word to avoid circularity with the dependent variable.

References


Bates, D., Maechler, M., 2009. lme4: Linear mixed-effects models using S4 classes. Vienna, Austria, r package version 0.999375-31.


URL http://www.ling.ohio-state.edu/ehume/papers/Epenth_info_Aix_final.pdf

Jaeger, T., Tily, H., 2011. On language utility: processing complexity and communicative effi-


URL http://www.bcs.rochester.edu/cls/abstracts/Jaeger_07_AMLAP.pdf

Jaeger, T. F., 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. Journal of Memory and Language 59, 434–446.


van Donselaar, W., Kuijpers, C., Cutler, A., 1996. How do Dutch listeners process words with


TABLE I. Correlations between probabilistic predictors

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Bwd Bigram</th>
<th>Fwd Bigram</th>
<th>Thematicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>1.00</td>
<td>0.19</td>
<td>0.31</td>
<td>-0.12</td>
</tr>
<tr>
<td>Bwd Bigram</td>
<td>0.19</td>
<td>1.00</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td>Fwd Bigram</td>
<td>0.31</td>
<td>0.13</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Thematicity</td>
<td>-0.12</td>
<td>0.05</td>
<td>0.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>
TABLE II. Predictors removed during model comparison, in order of removal

<table>
<thead>
<tr>
<th>Predictor</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$p_{\chi^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous</td>
<td>1</td>
<td>0.0432</td>
<td>0.835</td>
</tr>
<tr>
<td>Speech rate * Frequency</td>
<td>1</td>
<td>0.0813</td>
<td>0.775</td>
</tr>
<tr>
<td>Stress</td>
<td>1</td>
<td>0.365</td>
<td>0.546</td>
</tr>
<tr>
<td>Speech rate * Forwards bigram</td>
<td>1</td>
<td>1.35</td>
<td>0.245</td>
</tr>
<tr>
<td>Forwards bigram</td>
<td>1</td>
<td>0.563</td>
<td>0.453</td>
</tr>
<tr>
<td>Age * Sex</td>
<td>1</td>
<td>1.45</td>
<td>0.228</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>0.723</td>
<td>0.395</td>
</tr>
<tr>
<td>Occupation level</td>
<td>3</td>
<td>4.80</td>
<td>0.187</td>
</tr>
<tr>
<td>Speech rate * Backwards bigram</td>
<td>1</td>
<td>0.982</td>
<td>0.322</td>
</tr>
</tbody>
</table>
TABLE III. Final fixed effect estimates (positive outcome is epenthesis). Table shows coefficients (also plotted, with standard errors) and associated $p_z$ value for difference from 0. Likelihood ratio test statistics for improvement in fit are shown in the rightmost columns.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>$p_z$</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$p_{\chi^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.992</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid=/r/</td>
<td>1.09</td>
<td>&lt;.001</td>
<td></td>
<td>16.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Across syllables</td>
<td>-0.537</td>
<td>0.0193</td>
<td></td>
<td>5.04</td>
<td>0.0248</td>
</tr>
<tr>
<td>Morphologically complex</td>
<td>-0.929</td>
<td>&lt;.001</td>
<td></td>
<td>11.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age</td>
<td>1.08</td>
<td>&lt;.001</td>
<td></td>
<td>20.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Utterance final</td>
<td>0.751</td>
<td>&lt;.001</td>
<td></td>
<td>19.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Speech rate</td>
<td>-0.0367</td>
<td>0.801</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thematicity</td>
<td>-0.923</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backwards bigram</td>
<td>-0.385</td>
<td>0.00658</td>
<td>1 5.78</td>
<td>0.0162</td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>-0.383</td>
<td>0.0685</td>
<td>1 2.70</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>Speech rate $\times$ Thematicity</td>
<td>-0.667</td>
<td>0.0567</td>
<td>1 3.32</td>
<td>0.0686</td>
<td></td>
</tr>
</tbody>
</table>
TABLE IV. Final random effect estimates. Table shows standard deviations of group members (plotted as gaussian densities for visual comparison) and likelihood ratio test statistics associated with removal.
List of Figures

FIG. 1 Model estimates for variation given birthplace. .......................... 9