Filled Pause Choice as a Sociolinguistic Variable

Josef Fruehwald[*]

1 Introduction

In this paper, I argue that filled pause selection (*um/uh*) is a sociolinguistic variable, conditioned by both internal and external factors. There appears to be a language change in progress towards selecting *um* more often than *uh.* In all respects, the (UHM) variable appears to pattern quantitatively just like all other sociolinguistic variables which have been examined, even though the locus of (UHM) variation would seem firmly in the speech planning domain. Combined with the quantitative systematicity of sociolinguistic variables across the full range of linguistic modules, I argue that the locus of variation may not be in the grammar, but rather constitutes a separate domain of knowledge, perhaps what Preston (2004) called the "sociocultural selection device."

2 Variation in Grammar

Labov (1969) demonstrated that the relationship between copula deletion and contraction in African American English exhibited quantitative systematicity, thus constituting a non-trivial component of speakers' knowledge of their language. He argued that this necessitated an expansion of the concept of "a rule of grammar" to include probabilistic weightings. At its best, the Variable Rule paradigm was exceptionally formal and explicit in the rules it proposed. For example, (1) is the proposed rule for TD Deletion in AAVE as formulated by Labov et al. (1968).

(1)   Variable TD Deletion Rule (Labov et al. 1968, p. 136)

$$\begin{bmatrix} - & \text{cont} \\ - & \text{grave} \\ - & \text{nas} \\ - & \text{comp} \end{bmatrix}_{\text{-t,d}} \rightarrow \langle \emptyset \rangle / \begin{bmatrix} \alpha & \text{consonantal} \\ \zeta & \text{obstruent} \end{bmatrix} \gamma(+)\delta(+) \begin{bmatrix} \varepsilon & \overline{\text{voice}} \end{bmatrix} \beta(\sim \text{V})$$

Despite some early concerns that inherent variability reflected a poverty of researcher imagination (Bickerton 1971) and that it may not be possible to identify two forms that "mean the same thing" outside of phonology (Lavandera 1978), the Variable Rule paradigm has persisted, and is still held up as the primary competitor to primarily usage-based approaches (Pierrehumbert 2006, Guy 2014). In fact, the variation-in-grammar approach has extended well beyond both the sociolinguistic and rule-based literature. Coetzee and Pater (2011) provide an overview of approaches to variation in the phonological literature including Partially Ordered Constraints, Stochastic OT, Noisy Harmonic Grammars and Maximum Entropy Grammars, the last three of which have defined learning algorithms.[1] What these approaches to variation all have in common is that they account for patterns of variation within the grammar usually by appealing to additional grammatical principles.

A slightly different approach to accounting for linguistic variation is the Competing Grammars model. Competing grammars, as formulated by Kroch (1989, 1994), characterizes speakers as having multiple grammars from which they can probabilistically select, resulting in output variants

---

[1]An underdiscussed issue regarding these constraint based models of variation is that they appeal to variation data as part of their motivation, but also "gradient well-formedness" (Hayes 2000, Keller and Asudeh 2002), meaning that these models predict that the less common variant is also less "grammatical," which is not a common assumption or result of research in language variation and change.

competing for use. This is formally identical to the proposal by MacSwan (1997) that bilingual code switchers have two lexicons from which they can draw. In a broad sense, the grammars define the possible selection space. For code-switching, these constraints are driven by whether the numeration which is drawn from the mixed lexicons can result in a valid derivation. However, in the case of variation, the probability of choosing, for example, a verb-raising grammar or a do-support grammar, is not itself encoded in the grammar nor governed by grammatical principles, but instead must constitute a separate kind of knowledge.

Proposing to separate at least some knowledge of variation from the grammatical modules is not an especially radical proposal. The fact that language-external factors, like gender, class and style, condition linguistic variation would seem to demand this sort of division of knowledge, although there are notable exceptions. Weinreich et al. (1968, ex (4)) incorporated style, class and age into the grammatical formulation of New York City (r), and Guy and Cutler (2011) argued that style is a post-lexical process. On the other hand, Preston (2004), in sketching out a psycholinguistic model of sociolinguistic variation, proposes a "sociocultural selection device," which moderates selection among options presented by speakers' grammar(s), but stands apart as a separate module. In a similar vein, Labov et al. (2011) investigated properties of the "sociolinguistic monitor" which they say "might be conceived of as a separate processing and storage module, or as the capacity to do a calculation on the fly at any time by an inspection of remembered tokens." Tamminga et al. (2016) take perhaps the most granular approach recently, subdividing the conditioning of variation into *s*(ociostylistic)-conditioning, *i*(nternal)-conditioning and *p*(sychological)-conditioning, specifically arguing that *i*-conditioning and *p*-conditioning are architecturally distinct. The logic of their argument rests on the fact that there are some factors that condition linguistic variables that never condition obligatory alternation. For example, an English auxiliary verb is more likely to contract the shorter the preceding NP is (MacKenzie 2013), but there are no attested languages with obligatory contraction if the preceding NP is less than 5 words, and obligatory non-contraction otherwise. Tamminga et al. (2016) argue that conditioning factors like these belong to the domain of *p*-conditioning.

This paper pursues a similar logic in arguing for separate domain of knowledge for linguistic variation. I will be arguing that a phenomenon typically considered paralinguistic, or properly part of the speech planning system, exhibits the same quantitative systematicity as any linguistic variable. Where Tamminga et al. (2016) argue on grounds that some conditioning factors are never involved in categorical alternation, and perhaps not expressible in the grammar internal formal system, I argue that the core variable itself is probably not part of the grammatical system.

3 Filled Pauses

"Filled pause" refers to the fillers *uh* and *um*, which are also commonly represented orthographically as *er* and *erm*. I will refer to these to variants as UH and UM to indicate the vowel final and nasal final variants, respectively, while completely glossing over any variation in vowel quality that may exist between them or between dialects. I will identify the variable itself as (UHM). (UHM) is commonly categorized as a speech disfluency along with repetitions, false starts and other repairs (e.g. Shriberg 1999). In terms of defining the envelope of variation, it would be reasonable to ask whether it is acceptable to include only UH and UM and not these other kinds of disfluencies and fillers. In this respect, I am following the lead in the psycholinguistic literature which often treats these fillers separately from others (e.g. Clark and Fox Tree 2002). Kendall (2013) also demonstrates that (UHM) has a quantitatively distinct effect on adjacent silent pauses, which are markedly longer than when adjacent to other fillers such as *like*, *well*, etc.

Clark and Fox Tree (2002) propose that UM and UH serve separable functions in speech. They argue that while both filled pauses serve as a signal of the speaker's cognitive load, or planning error, UM signals greater difficulty than UH. They find that silent pauses following UM are longer on average than silent pauses following UH, a result replicated by Kendall (2013, figure 6.11). They argue that this means UH and UM are "words like any other," with distinct meanings. This proposal has fueled some debate within the psycholinguistic literature regarding speaker awareness and

agency which is strikingly similar to issues surrounding socioindexical meaning in sociolinguistics (see Campbell-Kibler 2012). For example, in their review article on the psycholinguistic research on filled pauses, Corley and Stewart (2008) conclude that there is strong evidence that fillers occur in contexts of high cognitive load, "[b]ut this does not prove that they are used by speakers to signal, for example, that there will be a delay in the speech stream due to uncertainty, except in the sense that smoke signals fire."

In terms of sociolinguistic research on (UHM), Tottie (2014) found that the overall rate of filled pause use (specifically *not* analyzing UH and UM separately) differed between the Santa Barbara Corpus and the British National Corpus. There was also considerable inter-speaker and inter-context variation. She specifically connects a high (UHM) rate to task-related speech, and a lower (UHM) rate to small talk in both the SBC and the BNC. In a public speaking context, on the other hand, (UHM)-lessness is held up as the standard (Erard 2008).

4 UM Rising

A number of researchers have noted in English that women and younger speakers have a higher UM:UH ratio (Acton 2011, Tottie 2011, Laserna et al. 2014). Acton (2011) suggests that this may reflect a language change in progress. Wieling et al. (to appear) demonstrate a similar age and gender effect on the UM:UH ratio across a number of different Germanic languages, including North American English varieties, British English varieties, Dutch, German, Norwegian as well as Danish and Faroese.

The North American English analysis in Wieling et al. (to appear) included data from the Philadelphia Neighborhood Corpus (PNC) (Labov and Rosenfelder 2011). In total, 25,514 filled pauses were extracted from the PNC transcriptions and coded for their duration, following segment, following segment duration, and demographic information for each speaker. Figure 1 displays the proportion of UM usage out of all filled pauses for each speaker, with an average for each decade overlaid.
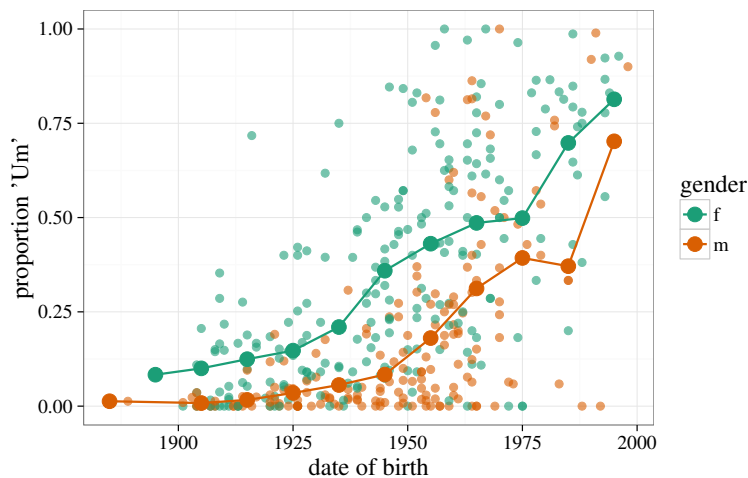

Figure 1: Proportion of UM in the PNC.

Under an apparent time analysis, the data in Figure 1 looks like a textbook language change in progress. The change follows a classic S-shaped curve, and men are approximately a generation behind women. Partially due to men's exceptionally low rate of UM usage early in the 20th century (close to 0%), fitting a logistic regression to the entire data set proved difficult. The most complex model which converged included gender, date of birth (centered at 1970 and divided by 10), their interaction, a categorical variable encoding whether the following segment was a silent pause, and a random intercept by speaker. Table 1 displays the fixed effects summary; all fixed effects are significant. The effect of being male puts speakers approximately 22 years behind in terms of the

	Estimate	Std. Error	z value	Pr($>$ |z|)
(Intercept)	-0.960	0.142	-6.769	<0.001
gender=male	-1.232	0.217	-5.675	<0.001
decade	0.549	0.042	13.061	<0.001
following pause	1.750	0.053	33.072	<0.001
gender=male×decade	0.194	0.072	2.699	0.007

Table 1: Fixed effects predicting the use of UM. Decade is centered at 1970. The reference level for gender is female.

rate of change, but it appears that the male rate of change is significantly faster. As expected based on Clark and Fox Tree 2002, a following silent pause has a large effect favoring UM.

However, it is worth considering whether these results truly reflect a change in filled pause selection itself, or whether this is an illusory result based on other unmodeled covariates. For example, a speakers' date of birth is highly correlated with their age in the PNC, so Figure 1 may reflect a process of age grading in which as speakers get older, they are more likely to select UH. Alternatively, there may be some other underlying change occurring, of which greater UM selection is merely symptomatic. D'Arcy (2012) proposes this kind of analysis for rise of the verb of quotation *be like* in New Zealand English. She argues that the rise of *be like* is partially driven by a shift in what kinds of "speech" (e.g. internal speech, mimetic speech, thoughts) are introduced into discourse at all, and that these kinds of speech favor use of *be like*. For (UHM), there may be a similarly new kind of discourse act coming into usage that disproportionately favors UM.

4.1 The effect of generational cohorts

In order to address whether the relationship between UM usage and date of birth reflects a lifecycle or generational effect, I fit a generalized additive mixed effects model predicting UM usage by a two dimensional tensor product smooth of date of birth and year of interview (see Wieling et al. 2011). Figure 2 plots highest posterior density confidence intervals for the predicted logit UM usage from the model for a number of representative date of birth cohorts. The way to interpret the strictly horizontal estimate for the 1940 cohort is that their UM usage rate is stable regardless of whether they were interviewed in 1973, 1993, or 2013. There is not, in fact, a reliable effect of the year of interview on predicted UM usage for any date of birth cohort.
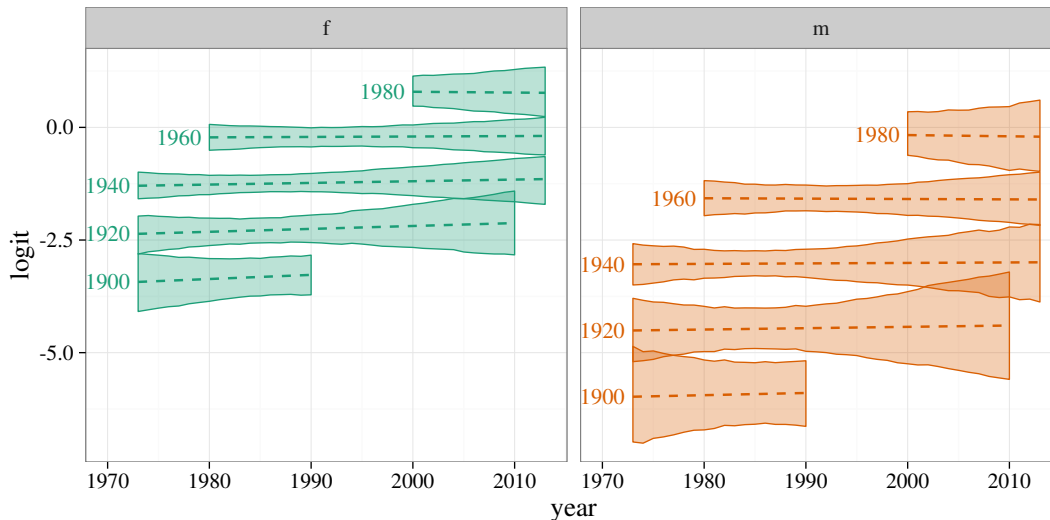

Figure 2: The interaction of date of birth cohort and year of interview on UM usage.

The conclusion to take away from these results is that this change, whatever it may be, is incrementing between generational cohorts, not within them.

4.2 Trading Frequencies

Another alternative is that a new discourse use of UM has entered into the language which boosts UM usage, rather than UM usage replacing UH usage in roughly equivalent contexts. An analogical example to this is the relationship between *computer* and *typewriter*. In the Google Books Corpus (Michel et al., 2011), *computer* appears to replace *typewriter* in a characteristic S-shaped curve across the 20th century. However, if we look at the frequency of *computer* and *typewriter* in isolation, rather than trading off in frequency, it is obvious that *computer* has rapidly increased in frequency due to other cultural shifts, while *typewriter* has remained relatively stable.
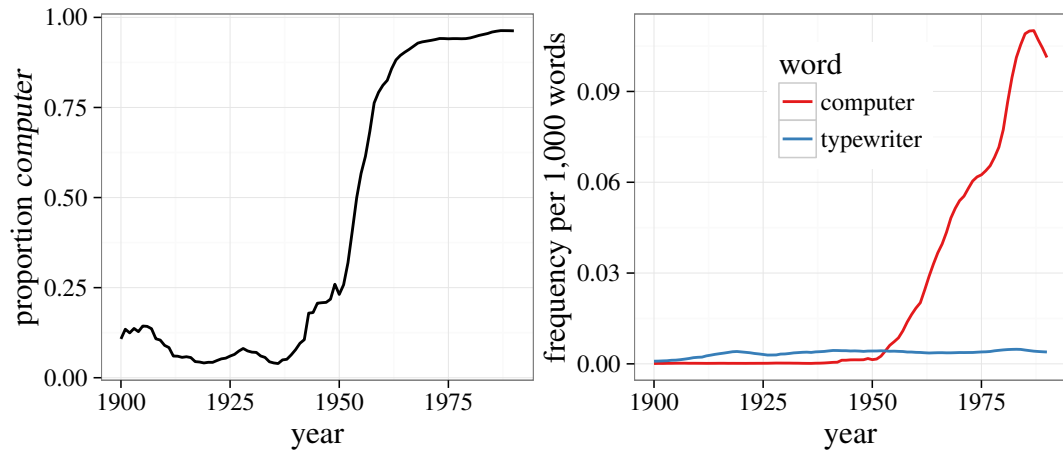

Figure 3: On the left, the proportion of *computer* out of *computer+typewriter*. On the right, isolated frequencies of *computer* and *typewriter*.
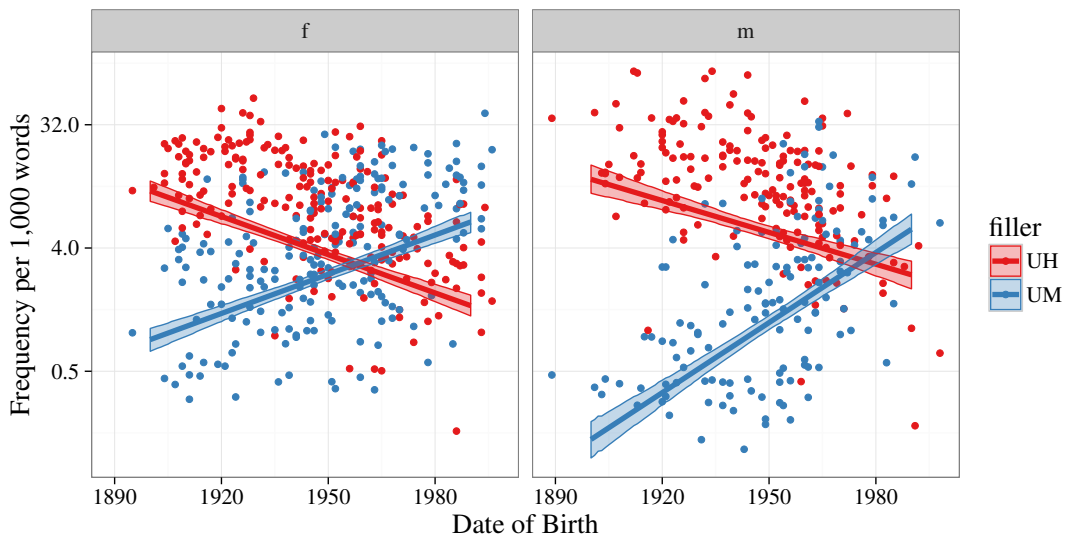

Figure 4: UM usage and UH usage trading in frequency.

If the increasing use of UM reflected a new discourse function coming into usage that favored UM, we would expect to see UM increasing in frequency with UH remaining relatively stable. Figure 4 plots the frequency of UM and UH per 1,000 words for each speaker's interview, with

the estimated frequency from a mixed effects poisson regression model overlaid. For women, the declining slope of UH is nearly equal and opposite that of UM. For men, the increasing slope of UM is notably steeper than the declining slope of UH, but this may be a statistical artifact driven by the fact that early in the 20th century men hardly used UM at all. Overall, it appears that UM and UH are trading off in frequency of use. This is more compatible with the analysis that UM is replacing UH in usage broadly in equivalent contexts, rather than UM increasing in frequency independent from UH usage.

4.3 No effect of being turn initial

There is a long standing hypothesis that filled pauses are used to hold the floor in conversation (e.g. Lallgee and Cook 1969), and some people have the intuition that initial UM carries connotations of being "snotty" (Zimmer 2014). Given these facts, it seems important to investigate the role of turn position on (UHM), especially since turn initial UM may be the best candidate for a new discourse function coming into use.

The PNC (UHM) data was coded according to who the speaker was immediately preceding the filled pause: the same speaker ("self"), or another speaker such as the interviewer or other third party ("other"). Figure 5 plots the proportion of UM in each turn context for each speaker, with the fits of a mixed effects logistic regression model overlaid. The models included predictors for date of birth (centered at 1970 and divided by 10), the previous speaker, their interaction, and random slopes for the previous speaker and random intercept by speaker. Separate models were fit for men and women to aid convergence. In these models, there was not a significant main effect of the previous speaker identity, nor was its interaction with date of birth significant. The size of the effects are observable in Figure 5.
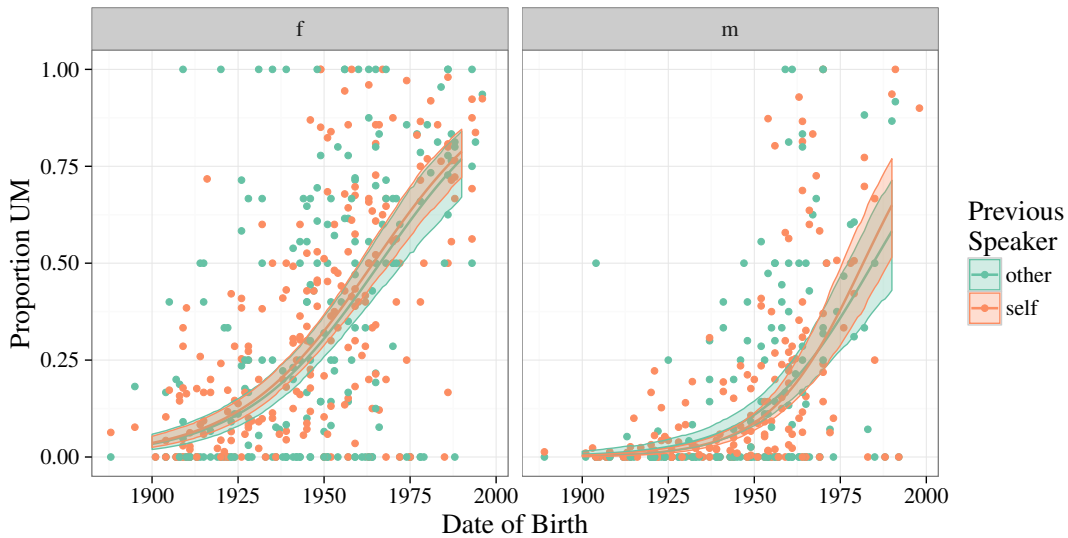

Figure 5: The effect of turn position on (UHM).

It does not appear that the most likely candidate for a new discourse function for UM (turn-initial UM) has an effect on this variable.

4.4 (UHM) Change Conclusion

It appears that (UHM) is undergoing a language change just like any other. The change is increment-ing between generational cohorts, and is being led by women. When analyzing the frequencies of UM and UH in isolation, it appears that they are trading in frequency. This result is most compatible with the usual interpretation of language change in progress, that UM is replacing UH in equivalent contexts. While there doesn't seem to be an effect of turn position on (UHM), it is conditioned by

whether it is followed by a silent pause, as well as the duration of the silent pause (see Wieling et al. to appear).

5 Placing (UHM) in the linguistic system

Based on the interaction of (UHM) with preceding determiners, it appears as if (UHM) insertion happens relatively late in the speech planning process. It has been noted that (UHM) behaves like any other vowel initial word by triggering the unreduced form of *the* [ðiː] (Arnold et al. 2003). However, it does not condition the phonologically conditioned allomorph *an* (Pak, 2016). In the PNC, when (UHM) follows the indefinite determiner, the determiner is realized either as [ə] or [eɪː], but never as [æn] or [ən]. This would suggest that (UHM) is not incorporated into utterances like most normal vowel initial words are. According to the speech planning system from Levelt 1989, there is a speech monitor that is "a device that is external to the production system" that inserts (UHM) into an utterance, presumably after the indefinite determiner allomorph has been selected. The elongation of the determiners [ðiː] and [eɪː] would seem to be a symptom of interrupting the speech stream, an effect which also impacts (UHM), as Shriberg (1999) demonstrated that the duration of the vowel in (UHM) is much longer than [ə] or [ʌ] are in any other lexical item.

6 Conclusion

There is a remarkable quantitative systematicity in variation across multiple domains of language, including:

(2)   a. Discourse (e.g. Tagliamonte 2008)
       b. Morpho-syntax (e.g. MacKenzie 2013)
       c. Phonology (e.g. Guy and Boberg 1997)

To this list we can now add the system of speech planning and monitoring that inserts (UHM) when a speaker experiences speech planning difficulty. The conclusion we can draw is either that the formal properties of these linguistic modules are nearly identical, such that they produce similar quantitative systematicity, or that the source of the quantitative systematicity is a property of a separate domain of linguistic knowledge that interacts with all of these systems. Developing analyses of the grammatical structure of linguistic variables undoubtedly improves our understanding of their quantitative patterning and interaction, since the grammar generates the pool of options speakers can choose from (for recent excellent examples, see Tamminga 2014). However, it may be a mistake to incorporate speakers' quantitative knowledge entirely into the grammar itself. Moving the quantitative knowledge that selects among the alternatives generated by the grammar into a sociocultural selection device, or a sociolinguistic monitor, may simultaneously simplify grammatical theories and account for the cross-domain systematicity of sociolinguistic variation. Moving quantitative knowledge outside of the narrow grammar does not mean that this knowledge is "extra-linguistic." It should remain clear that speakers' arbitrary knowledge of their language is not coextensive with the narrow grammar, and it is this knowledge of language that is the primary object of linguistic inquiry.

References

Acton, Eric K. 2011. On gender differences in the distribution of um and uh. *University of Pennsylvania Working Papers in Linguistics* 17.2.

Arnold, Jennifer E., Maria Fagnano, and Michael K. Tanenhaus. 2003. Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research* 32:25–36.

Bickerton, Derek. 1971. Inherent variability and variable rules. *Foundations of Language* 7:457–492.

Campbell-Kibler, Kathryn. 2012. The Implicit Association Test and sociolinguistic meaning. *Lingua* 122:753–763.

Clark, Herbert H., and Jean E. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition* 84:73–111.

Coetzee, Andries W, and Joe Pater. 2011. The place of variation in phonological theory. In *The Handbook of Phonological Theory*, ed. J. Goldsmith, J. Riggle, and A. Yu, 401–434. Blackwell, 2nd edition.

Corley, Martin, and Oliver W. Stewart. 2008. Hesitation disfluencies in spontaneous speech: The meaning of um. *Linguistics and Language Compass* 2:589–602.

D'Arcy, Alexandra. 2012. The diachrony of quotation: Evidence from New Zealand English. *Language Variation and Change* 24:343–369.

Erard, Michael. 2008. *Um...: Slips, Stumbles and Verbal Blunders, and What They Mean*. Anchor.

Guy, Gregory R. 2014. Linking usage and grammar: Generative phonology, exemplar theory, and variable rules. *Lingua* 142:57–65.

Guy, Gregory R., and Charles Boberg. 1997. Inherent variability and the obligatory contour principle. *Language Variation and Change* 9:149–164.

Guy, Gregory R., and Cecelia Cutler. 2011. Speech style and authenticity: Quantitative evidence for the performance of identity. *Language Variation and Change* 23:139–162.

Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. *Optimality Theory: Phonology, Syntax, and Acquisition* 88–120.

Keller, Frank, and Ash Asudeh. 2002. Probabilistic learning algorithms and Optimality Theory. *Linguistic Inquiry* 33:225–244.

Kendall, Tyler. 2013. *Speech rate, pause and sociolinguistic variation: Studies in corpus sociophonetics*. New York: Palgrave MacMillan.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K. Beals, 180–201.

Labov, William. 1969. Contraction, deletion, and the inherent variability of the English copula. *Language* 45:715–762.

Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski, and Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15:431–463.

Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. *The Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City: Phonological and Grammatical Analysis. Report on Co-operative Research Project 3288*. New York: Columbia University Press.

Labov, William, and Ingrid Rosenfelder. 2011. The Philadelphia Neighborhood Corpus.

Lallgee, Mansur G., and Mark Cook. 1969. An experimental investigation of the function of filled pauses in speech. *Language and speech* 12:24–28.

Laserna, Charlyn M., Yi-Tai Seih, and James W. Pennebaker. 2014. Um... Who like says you know: Filler word use as a function of age, gender and personality. *Journal of Language and Social Psychology* 33:328–338.

Lavandera, BR. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7:171–182.

Levelt, Willem J. M. 1989. *Speaking: From Intention to Articulation*. MIT Press.

MacKenzie, Laurel. 2013. Variation in English auxiliary realization: A new take on contraction. *Language Variation and Change* 25:17–41.

MacSwan, Jeffrey. 1997. A minimalist approach to intrasentential code switching: Spanish-Nahuatl bilingualism in Central Mexico. Doctoral dissertation, UCLA.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331:176–182.

Pak, Marjorie. 2016. How allomorphic is English article allomorphy? *Glossa* 1:20.

Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34:516–530.

Preston, Dennis R. 2004. Three kinds of sociolinguistics: A psycholinguistic perspective. In *Sociolinguistic Variation: Critical Reflections*, ed. Carmen Fought, 140–158. New York: Oxford University Press.

Shriberg, Elizabeth E. 1999. Phonetic Consequences of Speech Disfluency. *Proceedings of the International Congress of Phonetic Sciences* 1:619–622.

Tagliamonte, Sali A. 2008. So different and pretty cool! Recycling intensifiers in Toronto, Canada. *English Language and Linguistics* 12:361–394.

Tamminga, Meredith. 2014. Persistence in the Production of Linguistic Variation. Doctoral dissertation, University of Pennsylvania.

Tamminga, Meredith, Laurel MacKenzie, and David Embick. 2016. The dynamics of variation in individuals. *Linguistic Variation* 16.

Tottie, Gunnel. 2011. Uh and Um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics* 16:173–197.

Tottie, Gunnel. 2014. On the use of uh and um in American English. *Functions of Language* 21:6–29.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics*, ed. W. Lehmann and Y. Malkiel. U. of Texas Press.

Wieling, Martijn, Jack Grieve, Gosse Bouma, Josef Fruehwald, John Coleman, and Mark Liberman. To appear. Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change* .

Wieling, Martijn, John Nerbonne, and R. Harald Baayen. 2011. Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PloS one* 6:e23613.

Zimmer, Benjamin. 2014. Comment on "UM / UH geography". URL `http://languagelog.ldc.upenn.edu/nll/?p=14015#comment-1154111`.

University of Edinburgh
3 Charles Street
Edinburgh, EH8 9AD
United Kingdom
josef.frueh@ed.ac.uk