Before the rise of um

Derek Denis and Timothy Gadanidis

September 17, 2019

1 Introduction

One of the most dramatic discourse-pragmatic changes in twentieth-century English has progressed under the radar of laypeople and (until recently) linguists: the rise of *um* as the predominant variant of the 'filled pause' variable (UHM) at the expense of *uh* (Fruehwald, 2016; Tottie, 2011; Wieling et al., 2016). Fruehwald (2016: 43) documents this "textbook" change over 100+ years of apparent time: *um* increases incrementally between generations and the rise is led by women. Why *um*? Why did this change occur? In this chapter, we investigate (UHM) at an early stage of change to determine what triggered the rise of *um*. We take up the hypothesis that the rise of *um* was connected to the development of a new discourse function for the variable (UHM), that *um* came to be favoured with. We remain agnostic about what the function is, but follow Tottie (2016) and Fruehwald (2016) who suggest a correlation between utterance position and function; and specifically Fruehwald (2016: 46) who suggests that "turn-initial *um* may be the best candidate for a new discourse function coming into use." We follow essentially a variationist approach, first treating *um* and *uh* as variants of a linguistic variable and using proportional analysis to assess the role of social and linguistic factors. We augment these results with a different quantitative perspective and examine the relative frequency of the variable itself in discourse to help in our interpretation.

2 (UHM) as a pragmatic marker

For the purposes of this study, we follow Fruehwald (2016), Wieling et al. (2016), and Tottie (2016), among others, in treating *uh* [əː] and *um* [əːm] (also written as *er* and *erm*) as variants of one variable, termed (UHM). It should be noted that this is not the only way that the variable context could be defined. For instance, Tottie (2018) includes (UHM) as one element of a set including *well*, *you know*, and *like*, on the basis that all of the elements are used to indicate speech planning. However, we argue that treating *um* and *uh* as an individual variable captures the two words' intuitive and structural similarity[1], both variants being phonologically and orthographically identical, modulo the coda. Both variants are also single-word constructions which, unlike *well*, *you know*, and *like*, do not appear to be derived from bleached lexical items, but from apparently non-lexical speech sounds. As Fruehwald (2016) notes, they have also traditionally been treated as a unique phenomenon in the psycholinguistic literature.

The exact nature of (UHM) as a linguistic feature is not a trivial question. A great deal of ink has been spilled over whether they are produced consciously or unconsciously, and what

[1]This is an (in our view justified) extension of the notion of "structural equivalence" (Pichler, 2010; Tagliamonte & Denis, 2010) to phonological/orthographic structure.

their purpose is. For example, Maclay and Osgood (1959: 41–42) characterize (UHM) as a floor-management device which speakers insert to indicate that they do not want to be interrupted when hesitating over what to say. Levelt (1983, 1989) describes (UHM) as an involuntary noise produced as a result of production problems: "*er* apparently signals that at the moment when trouble is detected, the source of the trouble is still actual or quite recent. But otherwise, *er* doesn't seem to mean anything. It is a symptom, not a sign" (Levelt, 1989: 484).

One problem with the involuntary "symptom" view is that, as Clark and Fox Tree (2002) point out, speakers have some control over whether or not they produce (UHM)—for example, it can be suppressed in a public speaking context (and indeed speakers are often counselled to do so). They argue that (UHM) is an "interjection" used to signify a delay, with *um* signalling longer delays than *uh*.

Recently, Tottie (2016) has put forward the argument that (UHM) is a pragmatic marker that, in speech, indicates planning. This is on the basis that (UHM) is used more frequently in contexts requiring more speaker planning, such as narratives and responses to questions. Tottie (2017) describes (UHM) as being on a "cline of lexicalization", where apparently-cliticized forms like *and-uh* and *but-uh* are not perceived as words, but *uh* and *um* alone are. Tottie (2017: 21–22) hypothesizes that the use of (UHM) between words and silent pauses, rather than in *and-uh* and its ilk, leads to the perception of (UHM) as a word in the lexicon, leading it to be available for conscious use. If this process was a factor in (UHM)'s diachronic development, we might expect to see an effect of cliticization at this early stage.

As we note above, the rise of *um* has now been described extensively in the variationist and corpus-linguistic literature, across a number of corpora and speech communities. The typical finding is that women have a higher *um–uh* ratio than men, and that younger speakers have a higher *um–uh* ratio than older ones. This pattern has been demonstrated in various speech communities and contexts in the United States (Acton, 2011; Fruehwald, 2016; Laserna, Seih, & Pennebaker, 2014; Wieling et al., 2016), as well as in England and Scotland (Tottie, 2011; Wieling et al., 2016), both in real and apparent time. Wieling et al. (2016) also show that this pattern extends beyond English to five other Germanic languages: Dutch, German, Norwegian, Danish, and Faroese.

While these accounts demonstrate definitively that a change is underway, an explanation remains elusive. What was the trigger for this "textbook" change? Fruehwald (2016) and Wieling et al. (2016) both suggest that a new meaning or function for *um* may have emerged in English[2]. Although Fruehwald (2016) found that *um* and *uh* appeared to be trading frequencies, casting doubt on a functional expansion explanation, it is possible that the emergence of a new function at some earlier point may have played a role nearer to the beginning of the change. Accordingly, in this chapter, we investigate data from before the rise of *um* with the goal of evaluating the functional expansion hypothesis.

3 Data and coding

The data for this study are from the *Farm Work and Farm Life Since 1890* oral history collection (Denis, 2016). The corpus consists of oral history interviews with 155 elderly farmers, recorded in 1984. The corpus covers five regions of Ontario, Canada: Temiskaming, Essex, Dufferin, Niagara Region, and Eastern Ontario; for this study, speakers from the latter two regions were considered.

[2]For Wieling et al. (2016), this is a possible explanation for the crosslinguistic nature of the change: a function could have emerged in English and then spread through contact to the other Germanic languages.

Speaker birth years range from 1891 to 1919, just before *um* began to take off per Fruehwald (2016).

We extracted each instance of *uh* and *um* from the transcripts, excluding unrelated instances such as *uh-oh*. Tokens from the two much-younger interviewers was also extracted, and analyzed separately. The transcription protocol emphasized faithful reproduction of *uh* and *um*.

The data were coded for the following social factors: year of birth, gender, and region (Niagara or Eastern Ontario). Year of birth and gender were used to operationalize the change-in-progress hypothesis. Table 1 presents a table of speakers by gender, region, and year of birth.

ID	Gender	Region	Year of birth
F-INT	F	N/A	unknown
M-INT	M	N/A	unknown
NO-1	M	Niagara	1906
NO-9	F	Niagara	1912
NO-11	M	Niagara	1917
NO-12	F	Niagara	1916
NO-20	F	Niagara	1911
NO-22	F	Niagara	1899
NO-23	M	Niagara	1898
NO-24	M	Niagara	1902
NO-27	M	Niagara	1911
NO-28	M	Niagara	1907
NO-32	F	Niagara	1904
NO-35	F	Niagara	1902
NO-36	F	Niagara	1903
SG-001	M	Eastern Ontario	1891
SG-004	F	Eastern Ontario	1907
SG-006	M	Eastern Ontario	1905
SG-012	M	Eastern Ontario	1910
SG-013	F	Eastern Ontario	1914
SG-014	F	Eastern Ontario	1899
SG-016	M	Eastern Ontario	1912
SG-019	M	Eastern Ontario	1904
SG-020	F	Eastern Ontario	1906
SG-022	F	Eastern Ontario	1915
SG-024	M	Eastern Ontario	1898
SG-028	F	Eastern Ontario	1919

Table 1: Table of speakers.

We also coded for two linguistic factors. To operationalize the functional expansion hypothesis, we coded for utterance position (initial or non-initial). (UHM) was defined as "initial" if it was the first element in an utterance, except in the case of *and-* or *but-* cliticization, where (UHM) was classed as "initial" if the containing utterance began with *and-uh* or *but-uh*. To test for a potential effect of cliticization (per Tottie's 2017 suggestion that this may have played a role in (UHM)'s lexicalization) we coded each token as "clitic" if it occurred immediately following *and* or *but,* and as "non-clitic" otherwise.

4 Results

4.1 Proportional frequency

Table 2 shows how our data compare with previous communities analyzed. The first block summarizes our data from Niagara and Eastern Ontario, as well as F-INT and M-INT, the two younger interviewers. The second block summarizes results from previous work on the Switchboard corpus (Godfrey, Holliman, & McDaniel, 1992), the Fisher corpus (Cieri, Miller, & Walker, 2004), the Philadelphia Neighborhood Corpus (PNC) (Labov & Rosenfelder, 2011), and the British National Corpus (BNC) (2007). The numbers for all of these other corpora are drawn from Wieling et al. (2016).

Community	Raw N *uh*	Raw N *um*	% *um*	Mean *uh* /1000	Mean *um*/1000	Mean UHM/1000
Niagara	1864	357	16.1	21.3	4.1	25.4
E. Ont.	1563	168	9.7	22.6	2.4	25.0
F-INT	321	318	49.8	12.4	12.3	24.7
M-INT	255	51	16.7	13.2	2.6	15.8
Switchboard	—	—	28.3	22.1	7.5	29.6
Fisher	—	—	64.1	6.8	9.9	16.7
PNC	—	—	27.6	13.2	4.5	17.7
BNC	—	—	46.1	4.5	4.3	8.8

Table 2: Cross-community comparison

As can be seen in the table, *um* is less frequent in our data compared to the more recent corpora; the female interviewer uses it around half the time, while the male interviewer's rate is comparable to the farmers'. Relative frequency of (UHM) taken as a whole is on par with other corpora, but we are cautious about making such a comparison because each corpus was collected and transcribed differently (for related discussion, see Pichler, 2010).

Looking at individual speakers' rates, we can see that all speakers use both *uh* and *um*, but there is no clear pattern by age (Figure 1) or gender (Figure 2).

Figure 3 shows the proportion of *um* in apparent time. In the plot to the left, year of birth is binned into five-year increments, whereas in the plot to the right, year of birth is continuous. In both cases, there is a modest trend upward over time. To determine the possible predictors underlying this trend, in the following figures we split the data by gender, position and cliticization. Figure 4 shows the pattern when splitting speakers by gender. Starting around 1905, women use *um* slightly less often than men do, with both genders' *um* rates trending slightly upward over time. Figure 5 shows the pattern when splitting tokens by position (initial vs. non-initial). Starting around 1905, *um* is used more frequently in initial position than in non-initial position. Figure 6 shows the pattern when splitting tokens by cliticization with *and* or *but* and position. *Um*'s proportional increase appears to be limited to non-cliticized initial tokens.

Figure 7 shows a conditional inference tree for all farmers. The model confirms several of the patterns indicated in Figures 3–6. The tree splits first at cliticization, with cliticized (UHM) having the lowest overall *um* rate. Within the cliticized tokens, there is a slight difference between noninitial and initial (UHM), with initial tokens having a higher *um* rate than noninitial ones. Looking at the noncliticized side of the tree, we see another positional split, again with noninitial

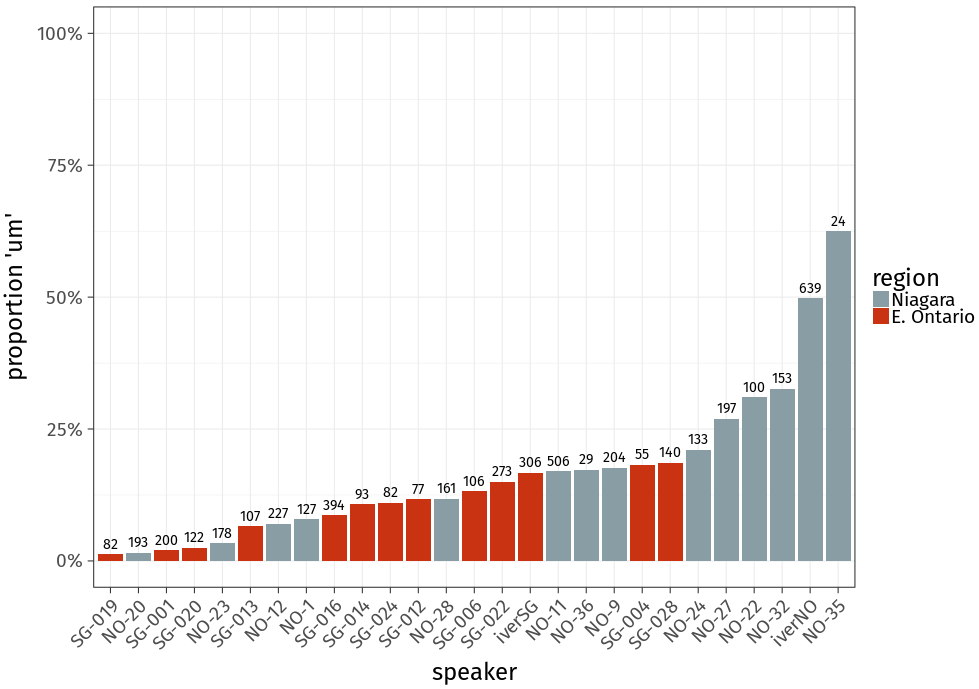

Figure 1: Proportion *um* per speaker by age.

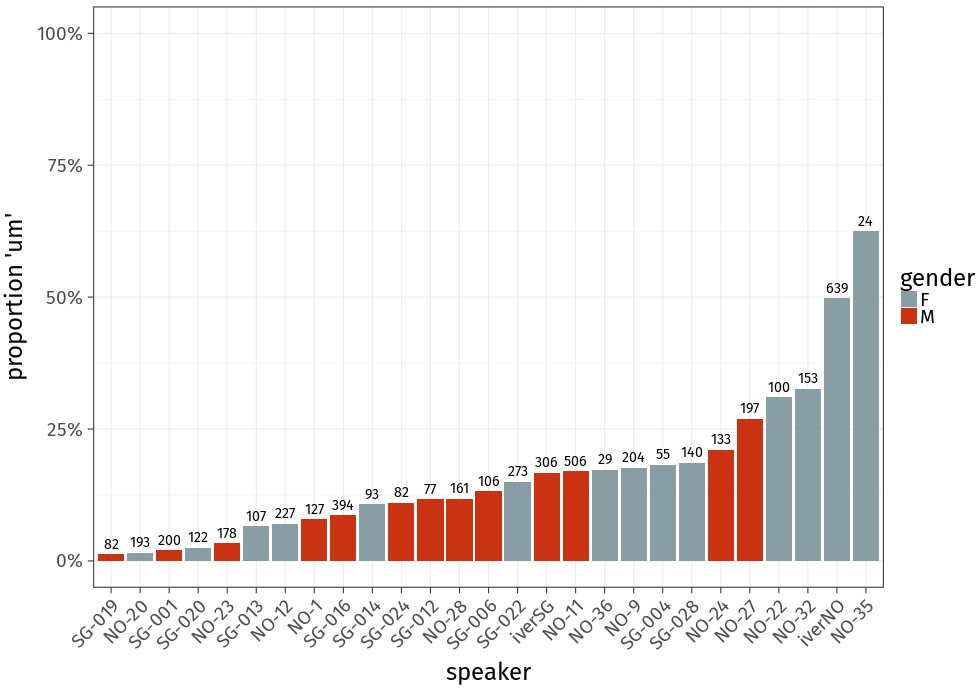

Figure 2: Proportion *um* per speaker by gender.

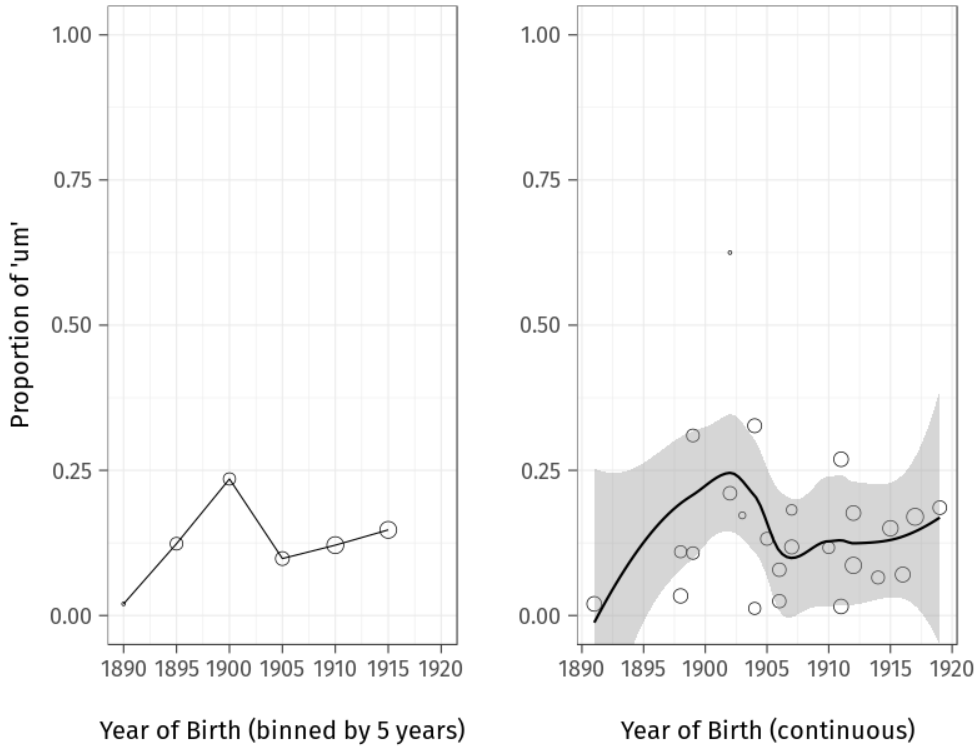

Figure 3: Proportion *um* in apparent time.

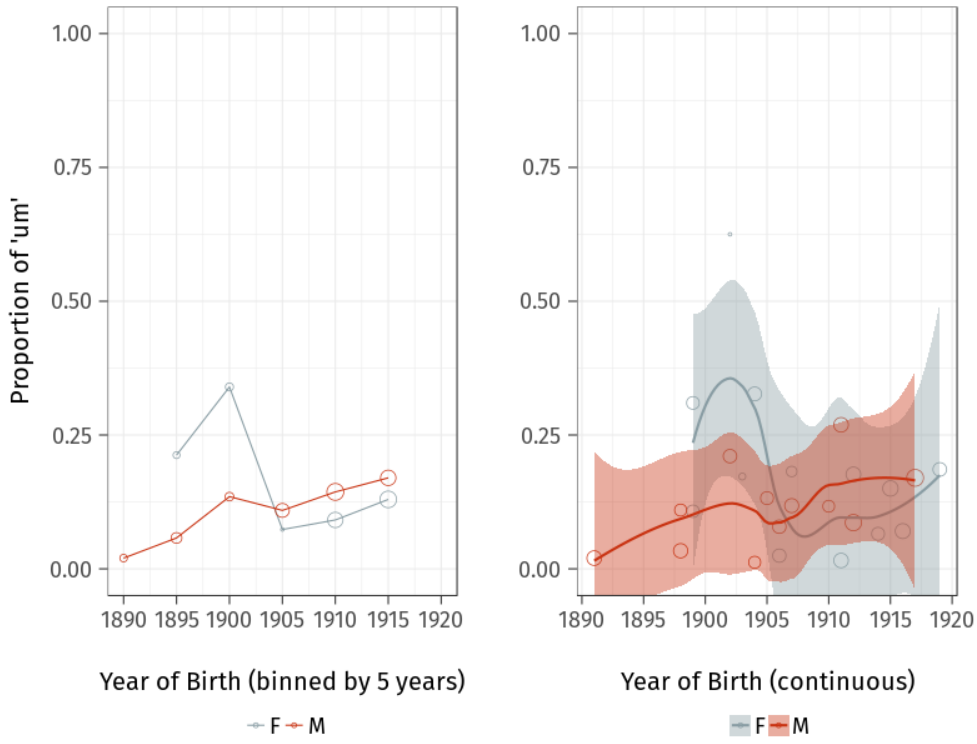

Figure 4: Proportion *um* in apparent time, by gender.

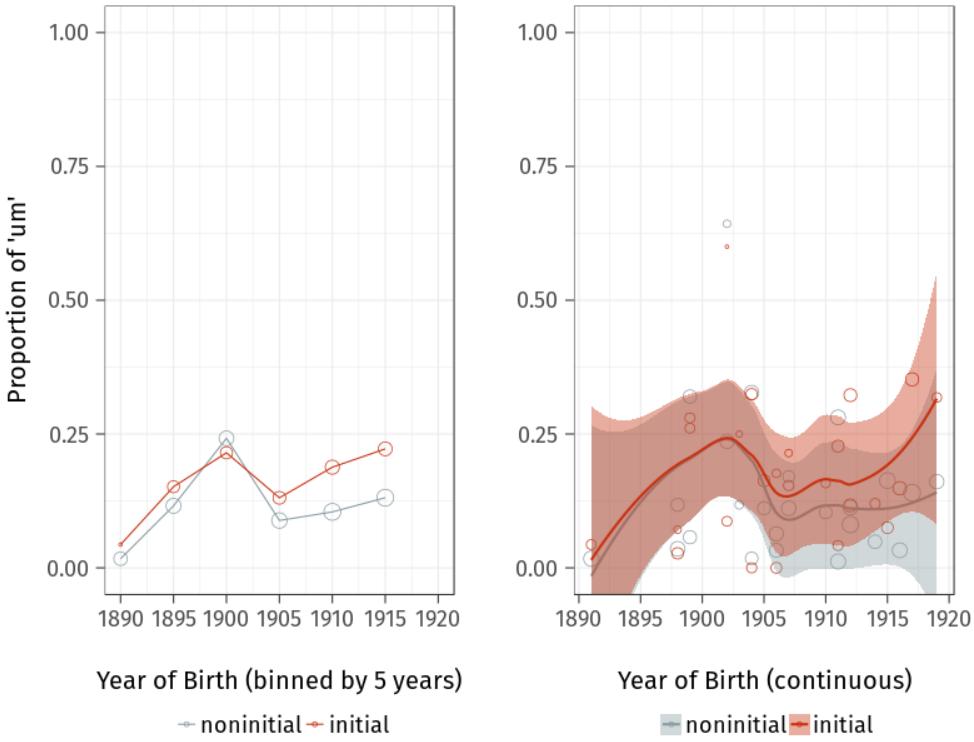

Figure 5: Proportion *um* in apparent time, by position.

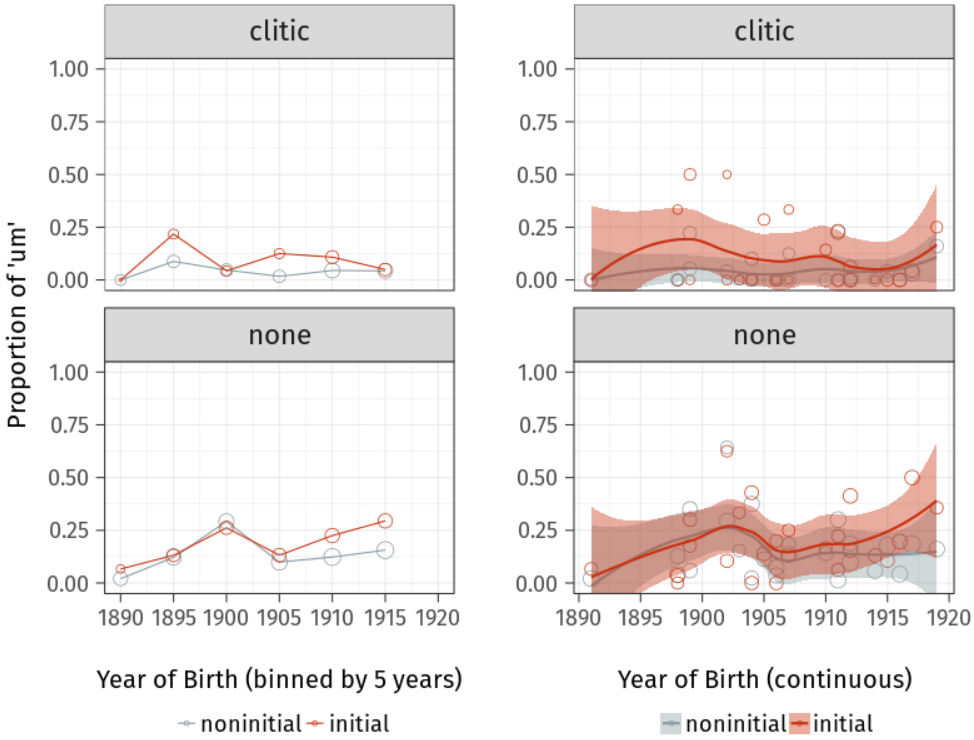

Figure 6: Proportion *um* in apparent time, by position and cliticization.

tokens having a lower rate than initial ones. Within the initial tokens, there is also an effect of year of birth: speakers born after 1898 have a much higher *um* rate in noncliticized tokens, and this is especially true in initial position.

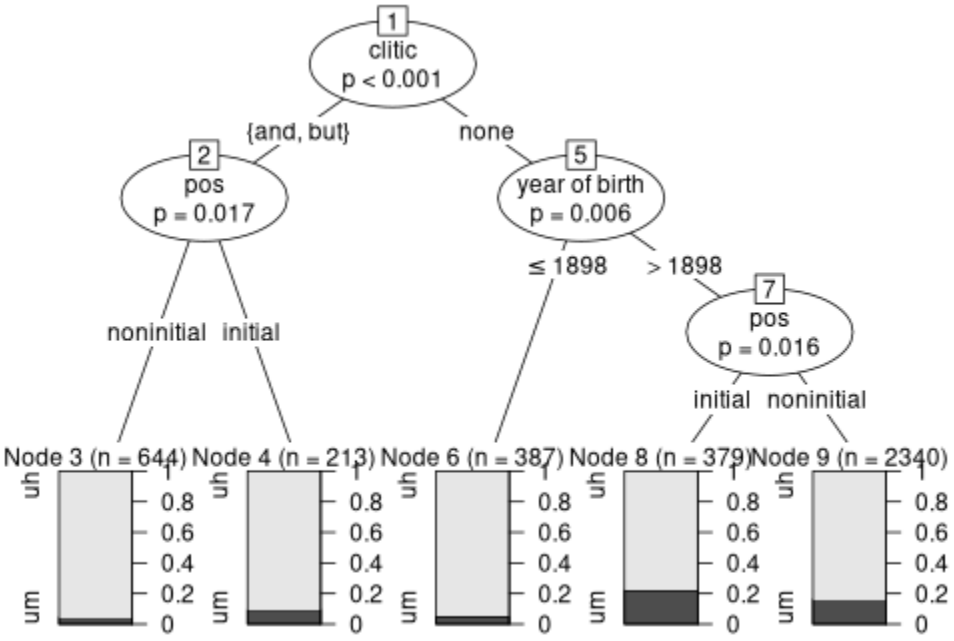

Figure 7: Conditional inference tree for farmers.

Figure 8 shows a conditional inference tree for the two interviewers. As shown in the tree, the internal constraints are much the same, but the baseline *um* rate is much higher (due in large part to the female interviewer). *Uh* is common in the cliticized forms, while *um* is more common in the cliticized ones—especially in initial position.

Taken together, the results presented in this section appear to show the beginning of the change toward *um* that has been observed by other researchers. While other work has shown that women lead this change, in our data, older women actually use more *um* than the younger women.

Looking at internal factors, we can see that cliticized forms, like *and-uh*, favour *uh*. There is some evidence for positional divergence, possibly consistent with a new utterance-initial discourse function that favours *um* (cf. Fruehwald, 2016, who found no turn-positional difference). Conditional inference trees confirm that the internal constraints persist with the younger speakers, while their baseline *um* rate is higher.

4.2 Relative frequency

Fruehwald (2016) tests the hypothesis that functional expansion triggered the rise of um by considering changes to the relative frequency of variants over time (e.g., frequency of *um* or *uh* per 10 000 words). When a new discourse-pragmatic function emerges, we expect that these functions would add to the relative frequency of the feature, and if the new function is restricted to one variant, the relative frequency of that variant should rise, with little change to the relative frequency of the other variant. In other words, we expect a fishtail pattern as with *computer* and *typewriter* over time: once *computer* gained its contemporary meaning, its relative frequency took

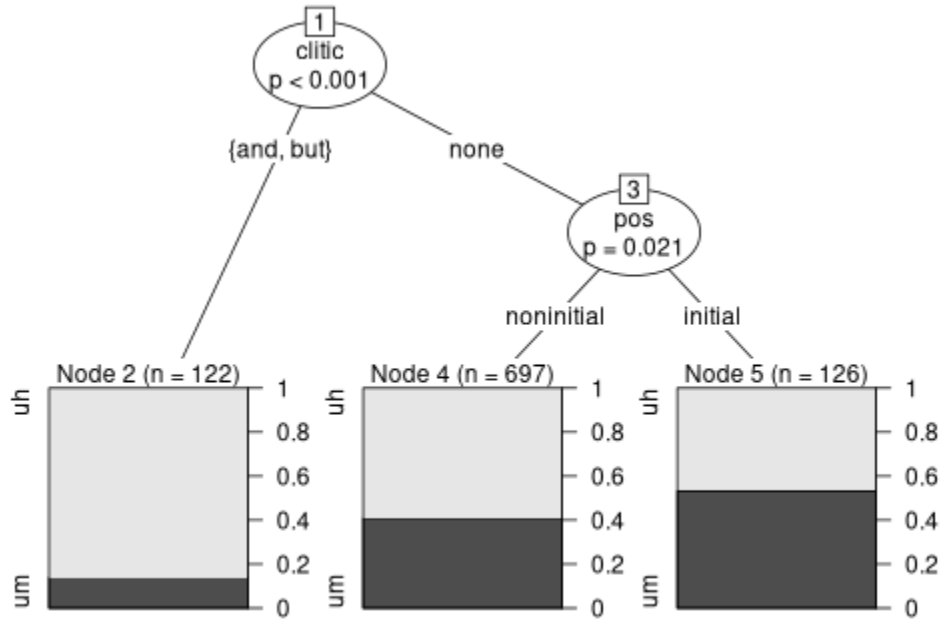

Figure 8: Conditional inference tree for interviewers.

off as that meaning became more frequent. This is illustrated in Figure 9 (Figure 3 from Fruehwald, 2016): looking at the proportion of *computer* over *typewriter* (left graph), *computer* appears to replace *typewriter* over time; but looking at the relative frequency of each word (right graph), it's clear that *typewriter* remained stable as *computer* took off.

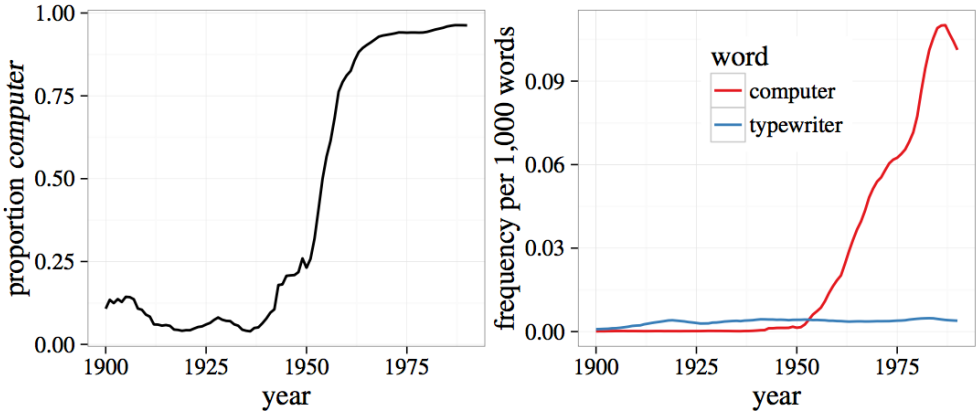

Figure 9: Proportional frequency and relative frequency of *computer* and *typewriter* (Figure 3 from Fruehwald, 2016).

If a new discourse function is what led to the rise of *um*, we should expect to see a similar fishtail pattern, with *um* rising and *uh* remaining stable. Conversely, if *um* were straightforwardly replacing *uh*, we should expect *uh* to fall concurrently with *um*'s rise.

Figure 10 shows the frequency of *um* and *uh* per 1000 words for each of the farmers. There is some evidence of a fishtail pattern, but in the opposite direction as expected: *uh* is increasing as

um remains relatively stable. The pattern is more extreme when we split the data by position, as in Figure 11. In initial position, both *um* and *uh* are largely stable, whereas in noninitial position, *uh* alone is increasing. Splitting the data again by gender, we can see that the increase can be attributed to the female speakers—there is no apparent increase over apparent time for male speakers, but the older female speakers have a relatively lower *uh* rate, rising to match the male speakers by the 1910s.

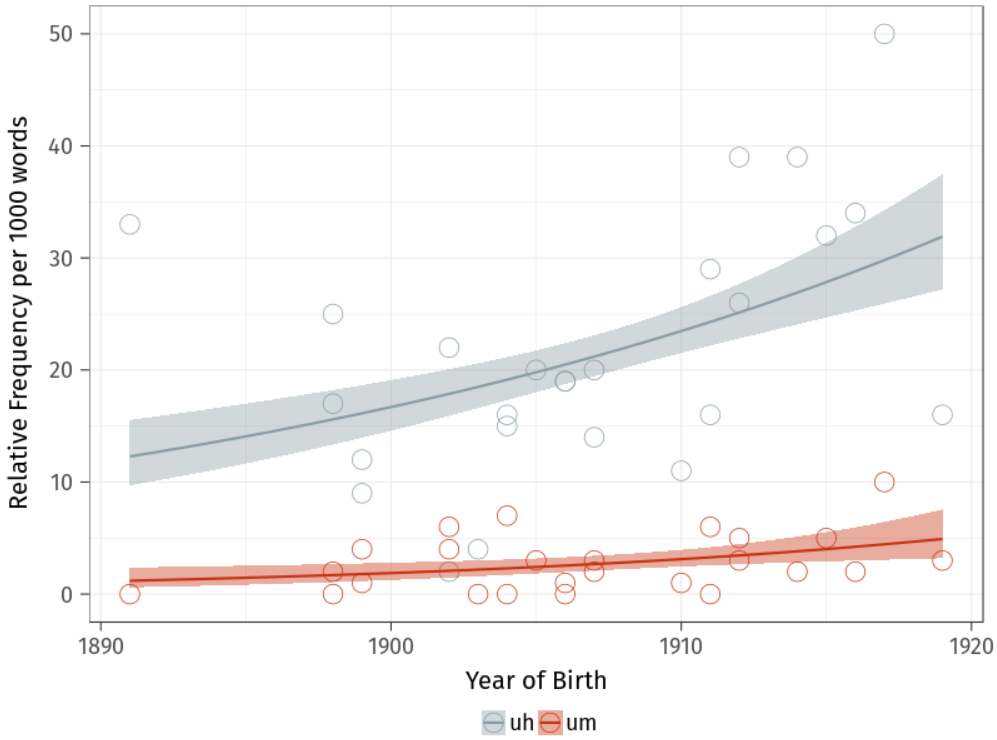

Figure 10: Frequency of *uh* and *um* per 1000 words

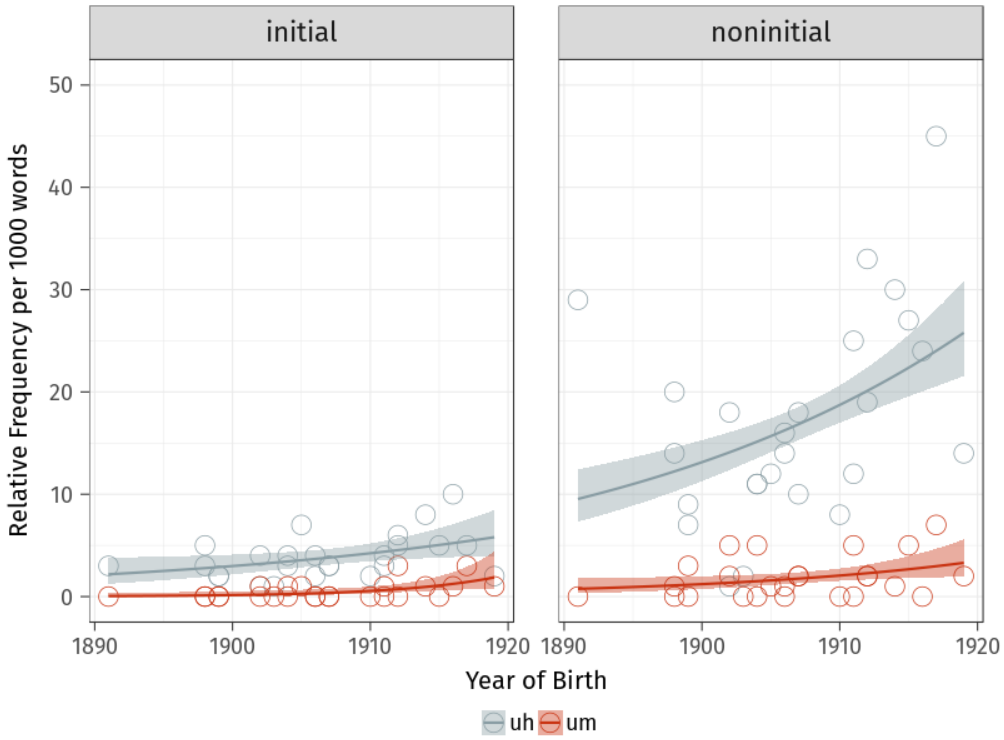

Figure 11: Frequency of *uh* and *um* per 1000 words, by position

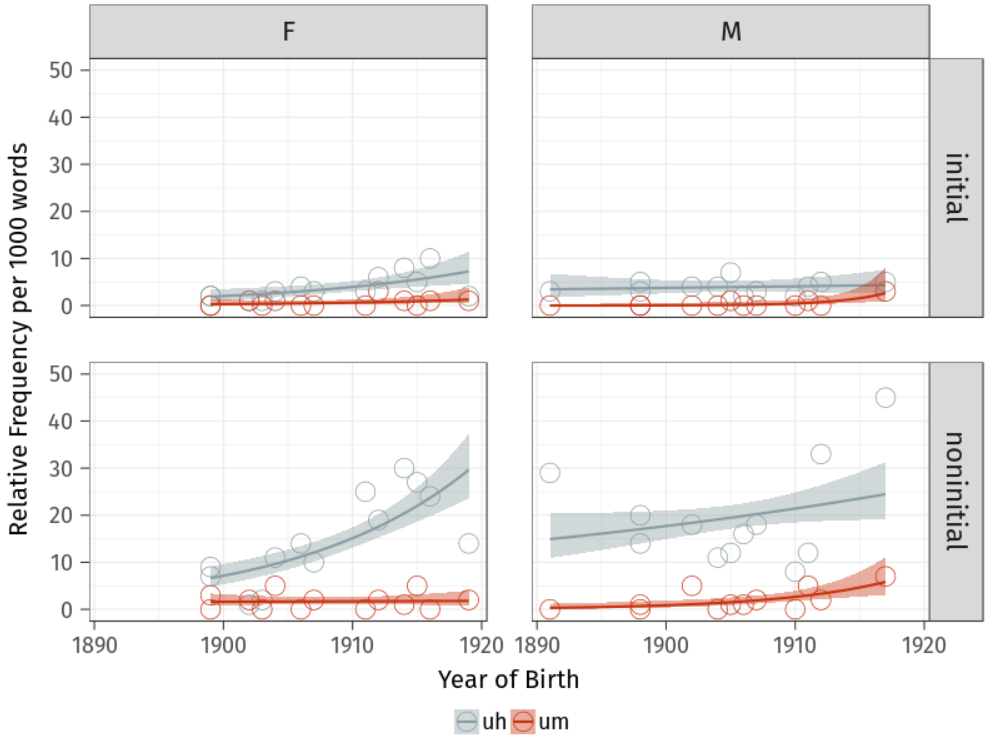

Figure 12: Frequency of *uh* and *um* per 1000 words, by position and gender

5 Discussion

Exploring data from before the rise of *um* has not yielded a definitive explanation for the change. Looking at the proportional frequency, we do find that for the younger farmers, *um* appears more frequent than *uh* in initial position, and the same is the case for the two (much younger) interviewers. Alone, this could be taken as suggestive evidence in favour of a new, initial-position function for *um*. Looking at the relative frequency, however, we see that the pattern does not appear to be driven by an increase of *um* in initial position—like Fruehwald (2016), we do not find strong evidence that a new, utterance-initial function for *um* is behind the rise of *um*. Instead, we find evidence of a different change: younger speakers fill more pauses than their older counterparts, and when they do, they are largely using uncliticized *uh* in non-initial position.

This difference is illustrated by Extracts 1 and 2: two passages of about the same length from NO-11, a younger woman (born 1917), and NO-36, an older woman (born 1903). In the transcriptions, (UHM) is bolded, and unfilled pauses are indicated using (.) or (…), depending on the length of the pause. In her extract, NO-11 uses (UHM) eight times—all but one of which are *uh*. In sharp contrast, NO-36 does not use (UHM) once, opting instead for lengthy, unfilled pauses. With respect to (UHM), the two speakers employ fundamentally different discourse strategies.

While our data are too early to shed much light on the rise of *um*, and it is important to be careful when generalizing across corpora and speech communities, it is possible that the *uh*-led shift from unfilled to filled pauses played a role in the competition between *um* and *uh* in the years to come. For example, if *uh* became specialized to non-initial position, which often appears to indicate word-search (Tottie, 2016, 2017), it may have become a less desirable variant (frequent word-search possibly giving the impression of disfluency). However, we have to stress that more work would be needed for us to be able to go beyond this kind of speculation.

INT: And what types of fruit (.) did you grow?
NO-11: Well the **uh** (.) originally **uh** when they came- **uh** grandfather bought the property in nineteen hundred and **uh** (.) **um** (.) to begin with there was very- there were very few fruit trees on it and they planted (.) **uh** (.) our orchard of **uh** (.) peaches. And **uh** waiting- while they waited for the peaches to come into bearing, they planted raspberries between the rows, so it started out as principally a raspberry farm I suppose but (.) it evolved into a farm that **uh** principally grew peaches and cherries, mainly sweet cherries.

Extract 1: High (UHM) user

These results highlight the importance of viewing discourse-pragmatic variation from multiple angles: the two perspectives we employ here, a proportional analysis and a relative frequency analysis, provide conflicting information about the functional expansion hypothesis.

INT: Okay. And how much (.) older was the very oldest?
NO-36: The oldest was born (…) in eighteen ninety two (…) and then my sister Lianne, eighteen ninety four (…) Greg, eighteen ninety eight (.) Sally nineteen hundred and one (…) I was born nineteen hundred and three (.) and that's it.
INT: Okay, and how old was your dad when you were born? At-
NO-36: (…) I- (…) how old was my dad when I was born? Oh.
INT: I think we had figured out that he was probably somewhere around forty five.
NO-36: Oh yes.
INT: And your mom was?
NO-36: Thirty (.) five?
INT: Thirty- oh-
NO-36: Is that it?
INT: Yup. Good.

Extract 2: Low (UHM) user

References

Acton, E. K. (2011). On gender differences in the distribution of um and uh. *University of Pennsylvania Working Papers in Linguistics*, *17*(2), 2.

Cieri, C., Miller, D., & Walker, K. (2004). The fisher corpus: A resource for the next generations of speech-to-text. In *Lrec* (Vol. 4, pp. 69–71).

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111.

Denis, D. (2016). Oral histories as a window to sociolinguistic history and language history: Exploring earlier Ontario English with the Farm Work and Farm Life Since 1890 oral history collection. *American Speech*, *91*(4), 513–516.

Fruehwald, J. (2016). Filled pause choice as a sociolinguistic variable. *University of Pennsylvania Working Papers in Linguistics*, *22*(2), 6.

Godfrey, J. J., Holliman, E. C., & McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. In *[proceedings] icassp-92: 1992 ieee international conference on acoustics, speech, and signal processing* (Vol. 1, pp. 517–520). IEEE.

Labov, W., & Rosenfelder, I. (2011). The Philadelphia Neighborhood Corpus.

Laserna, C. M., Seih, Y.-T., & Pennebaker, J. W. (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*(3), 328–338.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. MIT Press Series in Natural-Language Processing. MIT Press, Cambridge, Massachusetts.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*(1), 19–44.

Pichler, H. (2010). Methods in discourse variation analysis: Reflections on the way forward. *Journal of Sociolinguistics*, *14*(5), 581–608.

Tagliamonte, S. A., & Denis, D. (2010). The stuff of change: General extenders in toronto, canada. *Journal of English Linguistics*, *38*(4), 335–368.

The British National Corpus. (2007). Version 3. Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium.
Tottie, G. (2011). Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics*, *16*(2), 173–197.
Tottie, G. (2016). Planning what to say: Uh and um among the pragmatic markers. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the clause: Form and function of extra-clausal constituents* (pp. 97–122). John Benjamins Publishing Company.
Tottie, G. (2017). From pause to word: *uh*, *um* and *er* in written American English. *English Language & Linguistics*, 1–26.
Tottie, G. (2018). Variation and change among pragmatic markers as planners in American English, Paper presented at DiPVaC4, Helsinki, Finland.
Wieling, M., Grieve, J., Bouma, G., Fruehwald, J., Coleman, J., & Liberman, M. (2016). Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change*, *6*(2), 199–234.