



Um and Uh as Differential Delay Markers: The Role of Contextual Factors

Ralph ROSE <rose@waseda.jp>, Waseda University Faculty of Science and Engineering (Tokyo, Japan)

Disfluency in Spontaneous Speech
The University of Edinburgh
Scotland, United Kingdom (Aug 8-9)
DISS 2015

Abstract

The English filled pauses *uh* and *um* have been argued to correspond respectively to shorter and longer anticipated delays in speech production. This study looks at some contextual factors that might cause this difference by investigating filled pause instances in monologue and conversation speech corpora. Results are consistent with previously observed delay differences and further show that discourse-level processing may influence differential delay marking though monologue results are more conclusive than conversation results. However, no evidence was found that lexical factors (word type, frequency) correlate with filled pause choice. The findings suggest a limited view of how speakers use filled pauses as delay markers: Not all contextual factors may trigger differential delay marking.

Method: Sampling FPs from monologue and conversation corpora

Corpus of Oral Presentations in English (COPE) (Watanabe, unpublished)

- Unscripted monologue
- 10 mins preparation
- 10 mins speech
- 20 recordings
- 10-15 mins each
- Sample
 - First 2 mins of 10 recordings (20 mins, 10 speakers)
 - 163 FPs

my name is *****, and my title is most memorable moments in, *uh* my life. so definitely, *uh* one of my most memorable moments in life was, *um* when me and my family went on our panama canal cruise.

it's just something that's stayed with me *uh* for a while, and I still remember a lot of it vividly. *um* so it all began *um* when I was about sixteen years old and I realized I needed to start thinking about and start planning my eagle scout project in order to attain the rank of eagle in boy scouts.

last year I took a trip to new york with several friends of mine. we drove from orange, california to new york new york *uh* in about two weeks, *uh* stopping in michigan, utah, *uh* nevada, arizona, colorado, wyoming, bunch of different states.

Santa Barbara Corpus (SBC) (Du Bois et al 2000)

- Free conversation
- Various settings and speakers
- No investigative task
- 60 recordings
- 15-30 mins each
- Sample
 - 7 recordings (165 mins, 17 speakers)
 - 149 FPs

43.66 44.20 but I mean,
44.20 44.92 I'm not like,
44.92 47.60 (H) .. <@ I'm no-t uh= @>,
47.60 48.00 @
48.00 48.96 (H) I don't know how to say it.

604.35 605.55 [Have you heard] these figures.
605.55 606.05 that like=,
606.05 606.45 *um*,
606.45 609.22 ... it's something like forty percent of males, in .. the Bay Area, are supposed [to be infected]?

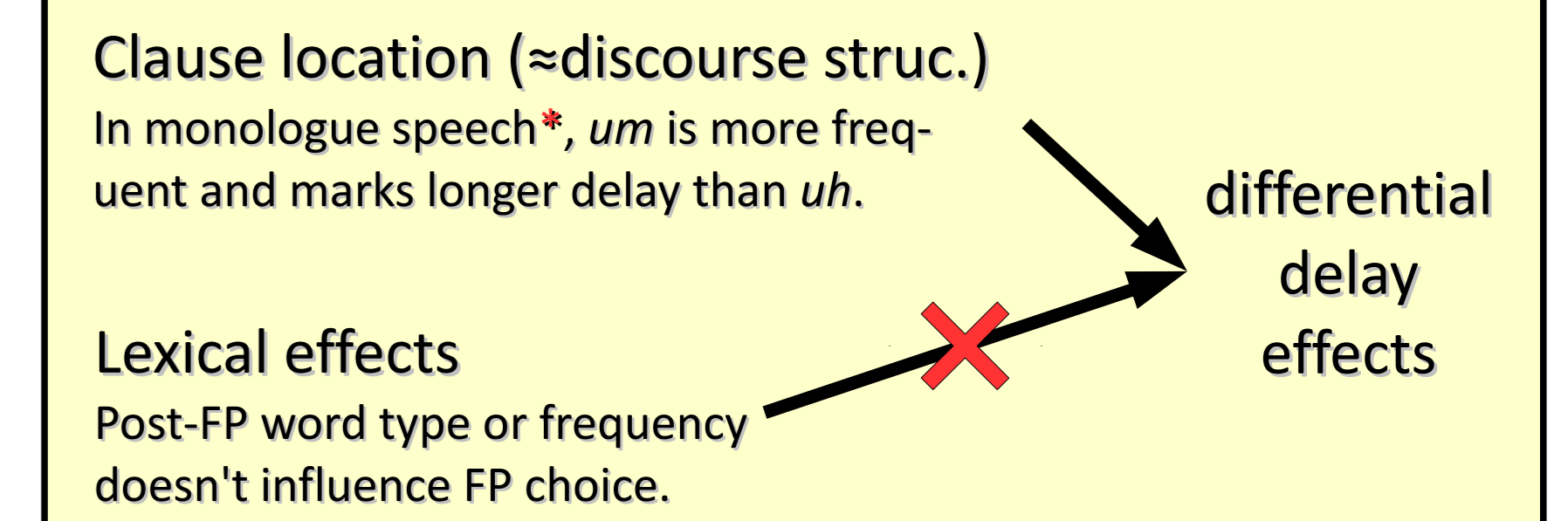
418.69 419.39 PETE: Where were they fishing.
419.39 420.14 .. Like in lakes,
420.14 420.39 or,
420.39 420.84 MARILYN: .. [Um=,
420.41 420.86 PETE: [rivers,
420.86 421.30 MARILYN: I think,

Measurements

- FP Duration
- Post-FP silent pause
 - Proportion
 - Duration
- Delay duration (FP dur. + SP dur.)
- Um proportion at clause locations (~discourse struc.)
- Post-FP content word proportion
- Post-FP word freq.

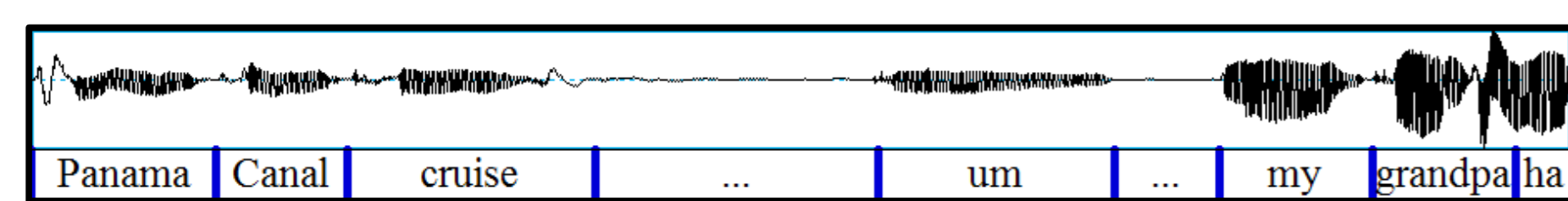
Discussion

Results show FPs are not generic markers of differential linguistic processing difficulty.



Future work could look at the gradience of other factors that may cause linguistic processing difficulties: e.g., articulation, (co)reference processing, syntactic or semantic effects.

Background: FPs as delay markers



Speakers are constantly monitoring their speech (i.e., perceptual loop theory: Levelt 1983, 1989) and when they detect a problem in their production, they (may) initiate a repair with a filled pause (*uh*, *um*). Thus, FPs mark a repair and constitute a delay in message transmission (Clark and Fox Tree 2002, 2014; Smith and Clark 1993; Kendall 2013; Rose 1998) with a differential in English, as follows.

uh → short delay *um* → long delay

What aspects of language production trigger this differential? Previous work shows FPs are more likely at major vs. minor discourse boundaries (Swerts 1998; Rose 1998) and FPs are more frequent before content than function words (Maclay and Osgood 1959) and before low-frequency than high-frequency words (Beattie and Butterworth 1979).

{A}Yesterday I was walking down the street when I saw a surprising thing. There was this guy selling toys {E} in a small {F} stall and everyone was watching him because he was so unique. {B} He would balance several toys at once in one hand {C} while demonstrating a new toy {D} with the other hand. All the kids couldn't help but watch and so many parents had no choice but to buy something!

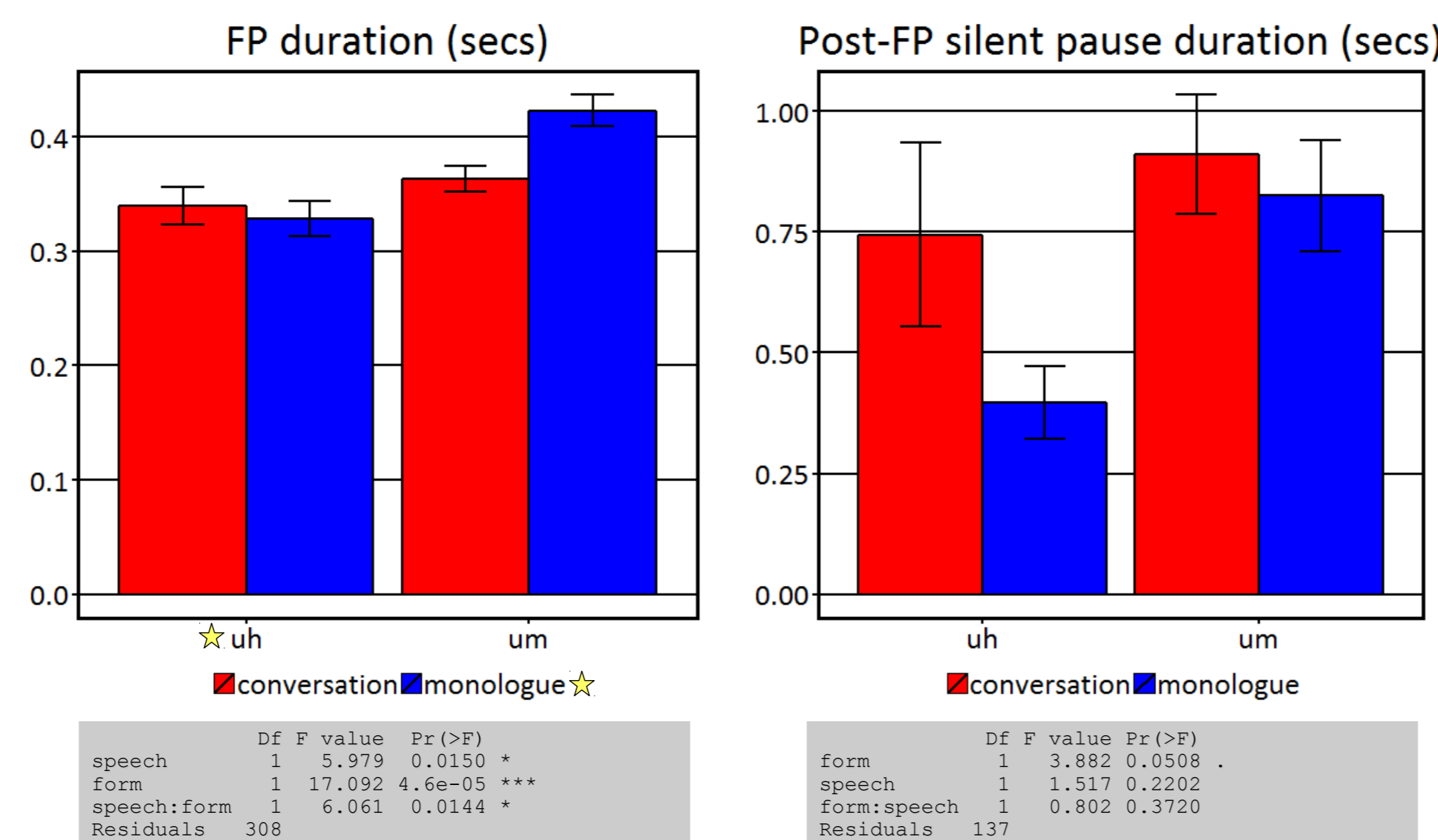
A > B > C > D F > E

Hypothesis: Gradient differences in discourse boundaries and following word status lead to major or minor repairs and hence greater or lesser choice of *uh* or *um*.

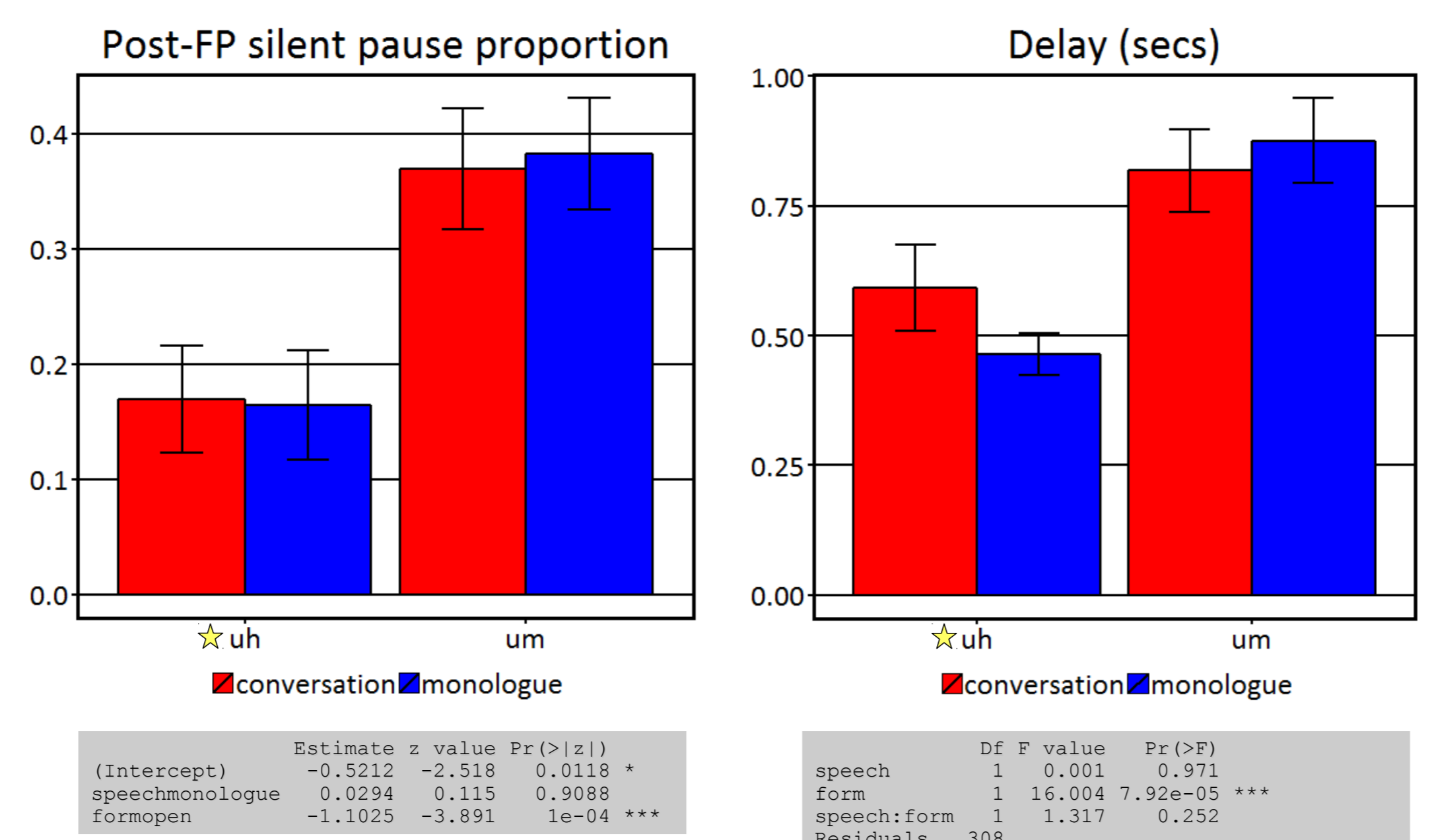
Acknowledgments

This research is supported by a Japan Society for the Promotion of Sciences (JSPS) Grant-in-Aid (#26284062): Kikuo Maekawa, Principal Investigator.

Results 1: Temporal parameters of FPs

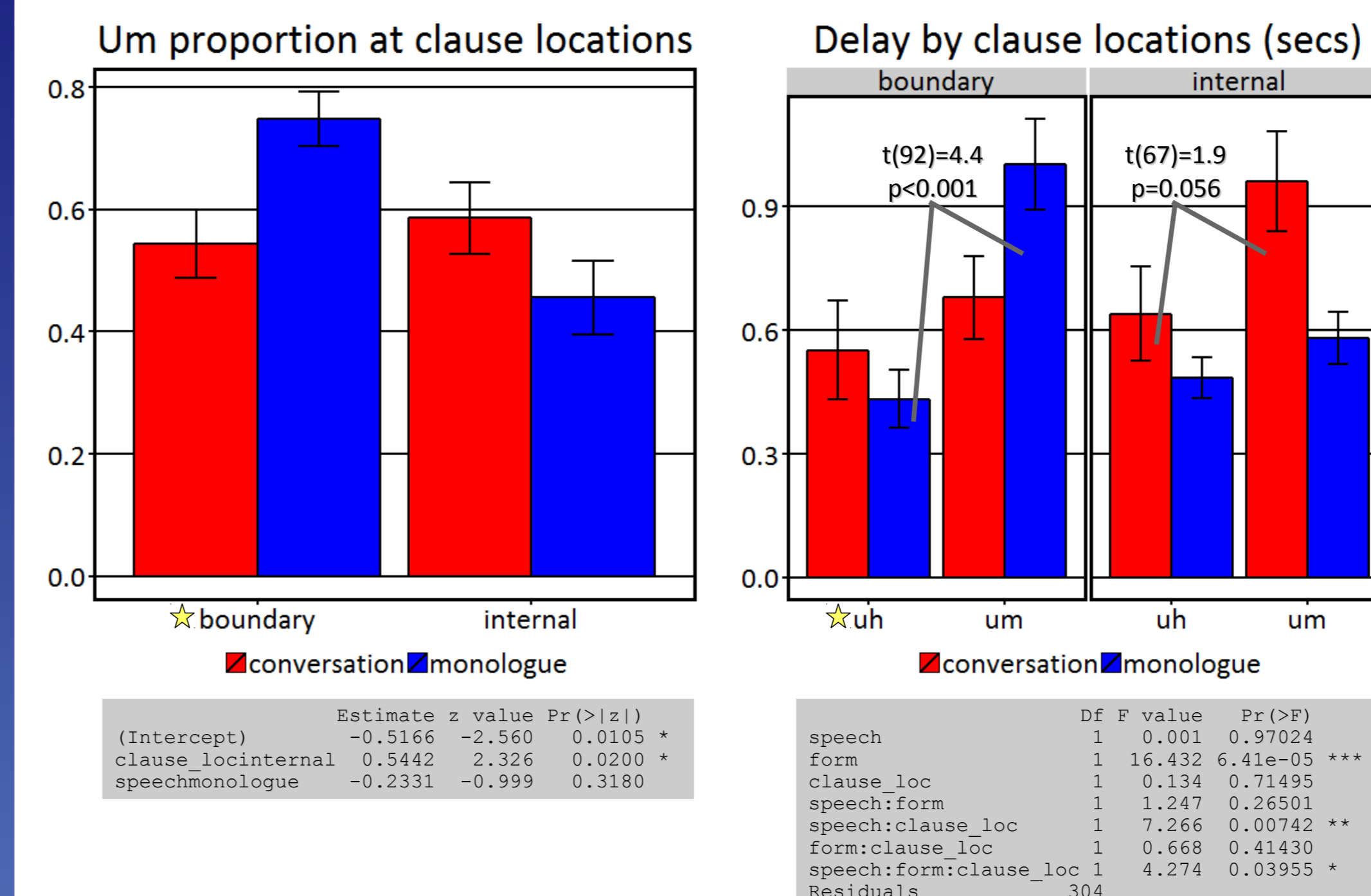


Articulatory duration of FPs differs between speech: *um* > *uh* in monologue, but *uh* = *um* in conversation.



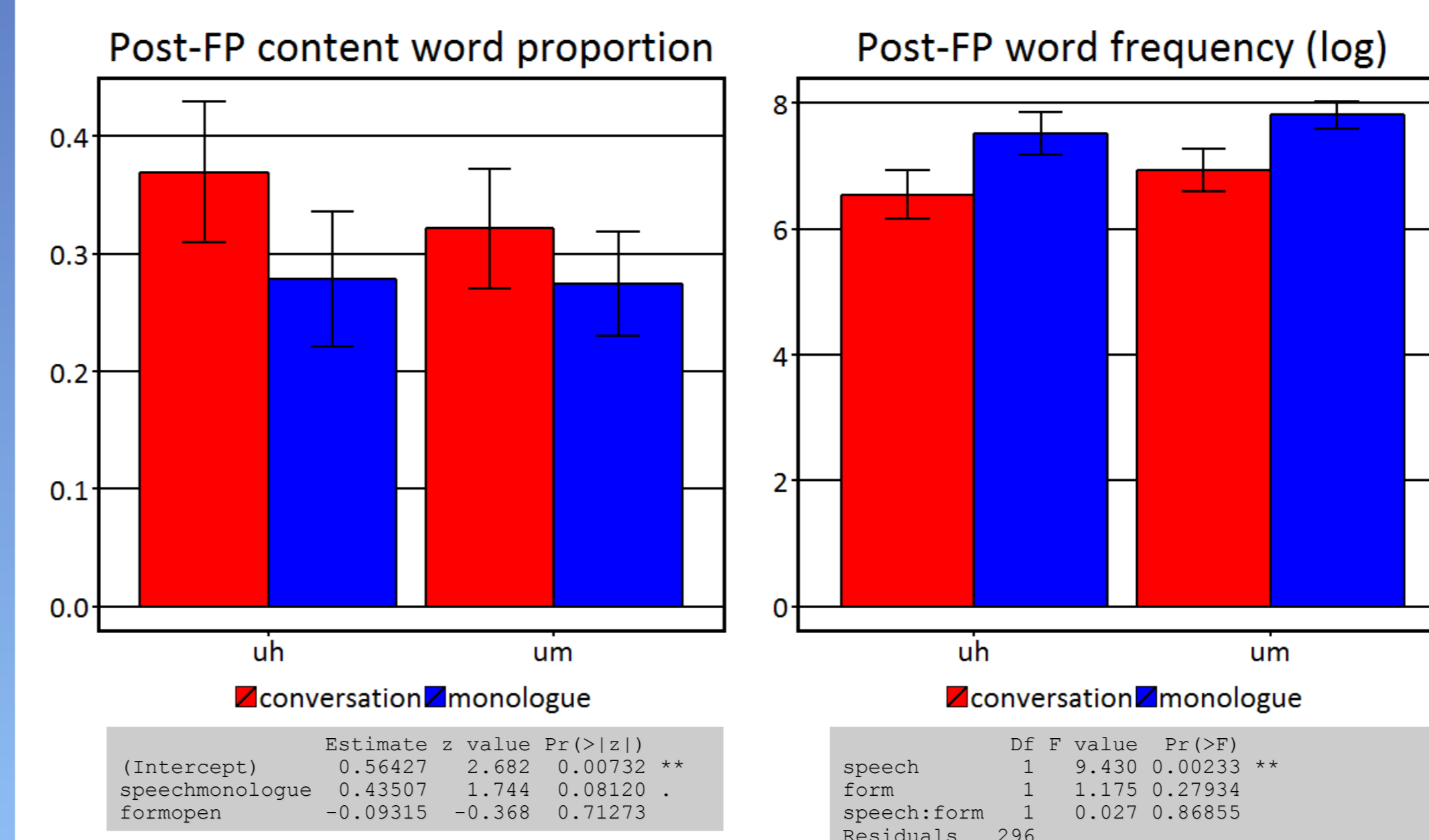
Um consistently marks a longer delay than *uh*. This replicates many previous studies (Clark and Fox Tree 2002, 2014, Smith and Clark 1993, Kendall 2013, Rose 1998).

Results 2: Clause location



In monologue, *um* > *uh* at boundaries, but *um* = *uh* internally. In conversation, *um* = *uh* consistently.

Results 3: Following word status



No clear differences are apparent based on following word status.

*Speculation

According to Clark and Fox Tree (2002, 2014), speakers use FPs differentially to communicate to interlocutors about their anticipated delay. Unpack this into two hypotheses:

- Differential delay hypothesis*: Different FPs in English correspond to different delay lengths. *well-attested*
- Differential conveyance hypothesis*: Speakers intend to convey their anticipation of a delay differentially. *not tested*

Intent is difficult to measure, but differences between the corpora may be suggestive.

Corpus	Time/Task constraints	Compulsion to communicate about anticipated delays
COPE (monologue)	Yes	Yes?
SBC (conversation)	No	No?

Monologue results support conveyance hypothesis as a default which is enabled under time or task constraints, though further work is necessary to confirm this.

References

Beattie, G.W., Butterworth, B.L. 1979. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech* 22, 201-211.

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

Clark, H., Fox Tree, J.E. 2014. On the-*uh* fillers *uh* and *um*. *Language Log*, November 11, <http://languagelog.ldc.upenn.edu/nl/?p=15718>

Du Bois, J.W., Chafe, W.L., Meyer, C., and Thompson, S.A. (2000) Santa Barbara corpus of spoken American English, Part 1. Philadelphia: Linguistic Data Consortium.

Kendall (2013) *Speech Rate, Pause and Sociolinguistic Variation: Studies in Corpus Sociophonetics*. London: Palgrave Macmillan.

Kucera, H., Francis, W.N. 1967. *Computational Analysis of Present-day American English*. Providence: Brown University Press.

Levelt, W.J.M. 1983. Monitoring and self-repair in speech. *Cognition* 14, 41-104.

Levelt, W.J.M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Maclay, H. and Osgood, C. (1959) Hestiation Phenomena in Spontaneous English Speech. In: *Word*, 15(19-44).

Rose, R.L. (1998) The Communicative Value of Filled Pauses in Spontaneous Speech. Unpublished Master's Thesis, University of Birmingham.

Smith, V. and Clark, H. (1993) On the course of answering questions. In: *Journal of Memory and Language*, 32:1(25-38).

Swerts, M. (1998) Filled pauses as markers of discourse structure. In: *Journal of Pragmatics*, 30:4(485-496).

Watanabe, M. Corpus of Oral Presentations in English (COPE). Unpublished corpus data.