

Linguistic Characteristics of Code Review Feedback

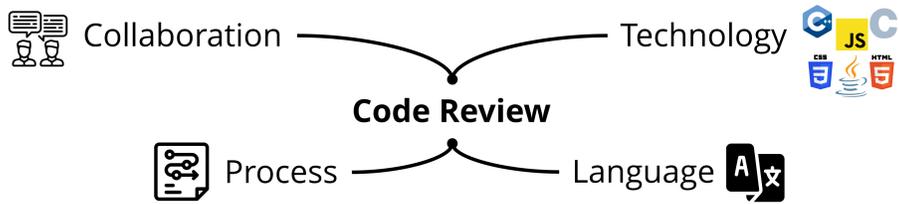
A Large Scale Study of Chromium

Nathan Munaiah¹, Benjamin S. Meyers¹, Cecilia O. Alm², Andrew Meneely¹, Pradeep K. Murukannaiah¹, Emily Prud'hommeaux^{2,3}, and Josephine Wolff²

¹Golisano College of Computing and Information Sciences and ²College of Liberal Arts, Rochester Institute of Technology ³Morrissey College of Arts and Sciences, Boston College

{nm6061,bsm9339,coag1a,axmvse,pkmvse,emilypx,jcwgpt}@rit.edu

Introduction



Code reviews help **collaboratively uncover quality concerns**
Developers perceive that code reviews **facilitate knowledge transfer and improve team awareness**

How Developers Can Provide Feedback

Fagan inspections afforded reviewers the ability to convey **nuanced meaning** using **paralinguistic cues** such as pitch, prosody, and facial expressions

In modern code review, reviewers may have to use **lexical cues** in creative ways to enrich the written feedback

Developer Reviewer

```
19 } /* validate */
20
21 + void heartbeat(message, size)
22 + {
23 +   return read_buffer(size);
```

Seriously? Not checking |size| is crazy. Fix it now.

NO! You must check the |size| before doing anything!

Please raise an error if |size| > len(|message|)

```
24 + } /* heartbeat */
25
26 void check(Credential *credential)
```

Is the way in which reviewers provide feedback relevant to it being acted upon?

Linguistic Characteristics

We collected and examined various linguistic characteristics

Syntax YNGVE and FRAZIER

Content PROPOSITIONAL DENSITY and CONTENT DENSITY

Tone FORMALITY, POLITENESS, SENTIMENT, and UNCERTAINTY

We Examined Acted-upon Feedback in Chromium

Phase I Automated using Done. Metric

```
Susan 2017/08/30 02:00:35
Can we initialize this to false, and only enable it where needed?
Reply Done Ack
```

```
18 static bool g_ooif_enabled;
```

```
Susan 2017/08/30 02:00:35
Can we initialize this to false, and only enable it where needed?
```

```
William 2017/08/31 21:00:25
On 2017/08/30 02:00:35, Susan wrote:
> Can we initialize this to false, and only enable it where needed?
Done.
```

```
19
```

Phase II Manual

```
79 while k < len(sub_dom_string):
```

```
Michelle 2017/09/27 17:15:13
Worth adding a unit test to validate the chunking logic?
```

```
80 sub_dom_string_part = sub_dom_string[k: k + step_size]
```

Acted-upon, not (known to be) acted-upon, and unknown

Data Set [1]

2,855,018	Comments	690,881	Acted-upon
1,591,431	Reviewer Comments	800	Not Acted-upon
		899,750	Unknown

Examples

You don't actually manage the deopt table's VirtualMemory, so you shouldn't act like you do.

Impolite
Informal
Negative

If you want to pursue this anyway, I'd probably be willing to at least look at the resulting patch.

Polite
Assertive
Uncertain

If this code is only reachable by mixed content (although I don't think that's the case), that's also fine, but then I'd like to see the methods renamed to reflect that, since otherwise it gives the impression that this magically handles all cases.

Polite
Informal
Positive
Assertive

We thought about this and still prefer to expose the operation as a poll (now called IsQueueEmpty), because that's what the client actually wants to do instead of reloading anything.

Impolite
Formal
Uncertain

Are you proposing that derived files, like components without an external file, would not have a mapping?

Polite
Informal
Negative

Empirical Analysis

How do linguistic characteristics of feedback relate to the feedback being acted upon?

Tests Mann-Whitney-Wilcoxon and Cliff's δ (Continuous) χ^2 (Logical)

Results Acted-upon comments had weak sentiment, were more formal and polite, and less uncertain

Do the linguistic characteristics of feedback vary with reviewer experience?

Tests Spearman's ρ (Continuous) and Cliff's δ (Logical)

Results No evidence of a relationship between the linguistic characteristics of feedback provided by a reviewer and corresponding review experience

Conclusion

The way developers speak matters!

The way developers provide feedback in code reviews is correlated with whether or not it is acted upon

The way in which reviewers provide feedback is likely to remain invariant with reviewer experience

Our work shows that **linguistic characteristics** can be (a) easily collected and (b) provide useful insights

BUT! Acted-uponness is just one validation—there may be other social implications of informal, impolite, etc. feedback

References

[1] B.S. Meyers, N. Munaiah, E. Prud'hommeaux, A. Meneely, J. Wolff, C.O. Alm, and P.K. Murukannaiah (to appear). A dataset for identifying actionable feedback in collaborative software development. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Short Papers).