Episode Book

EPISODE #6344

Training Text Models to Know When Not to Edit

PUBLISHED	29 October 2026 at 12:50 UTC
DURATION	22:51
LISTEN	myweirdprompts.com/episode/training-models-not-to-edit-62200/
HOSTS	Corn, Herman, Dorothy, Hilbert
SCRIPT MODEL	openai/gpt-6.1-sol
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.25
SOURCES	14 cited by the research (listed at the end)
BOOK GENERATED	30 October 2026 at 07:19 UTC

SYNOPSIS

Training a text-cleanup model exclusively on flawed examples risks teaching it that every input requires intervention. Examining small models like Superwhisper's 600-million parameter S1-mini reveals the subtle challenge of training restraint. Balancing edited pairs with unchanged identity targets prevents compulsive rewriting while maintaining punctuation and formatting. Designing clean inputs, setting boundary conditions, and weighting training ratios help models learn exactly when to leave transcripts alone.

TRANSCRIPT



Corn

Daniel wants to know how much of a model's training should teach it to leave things alone. He's fine-tuning a small text-in, text-out model to clean speech transcripts, and he's seen the trap: show it nothing but corrections, and it starts treating every input as something that must be corrected. Would eighty percent edited examples and twenty percent unchanged examples teach the right boundary?



Herman

The small decision that matters before anyone starts arguing about architecture. I'm Herman Poppleberry, and I would absolutely expect the dataset to overpower the intention here. You intend, fix what needs fixing. Your examples demonstrate, something always needs fixing.



Corn

An editor paid by the alteration.



Herman

Automatic post-editing has been struggling with that for years. Translation gets better, yet the component employed to improve it keeps finding work.



Corn

What interests me is that Daniel isn't asking the model to become a better conversationalist. He's asking it to perform a narrow operation reliably. If the raw text already meets his conventions, reproducing it should count as a successful operation.



Herman

Including successfully ending the output. A model can get every word right and still spoil the result by adding an explanation of what it changed.



Corn

The receipt costs more than the repair.



Herman

Superwhisper's S1-mini gives us a concrete anchor. It's a six hundred million parameter model, fine-tuned from Qwen three of that size. Released in August, under Apache two point zero plus a naming clause. Small enough that the quantized file occupies four hundred sixty-two mebibytes and runs on a laptop CPU.



Corn

Small enough to live beside the transcription engine, rather than summon a remote literary consultant every time I dictate a shopping note.



Herman

Its input is usually lowercase, unpunctuated ASR output. Its output removes fillers, resolves false starts, adds punctuation and capitalization, and renders spoken numbers and dates as written text. A control line sets styling, structure, and context. Casual or formal, prose or lists, general text or email.



Corn

Those controls complicate what already fine means. A sentence might be acceptable under casual styling and need changing under formal styling. The unchanged example needs the control setting attached, not just two matching strings.



Herman

Yes, and the model card explicitly says it isn't a chat model and won't follow general instructions. That restriction is valuable. You aren't hoping it interprets your entire philosophy of editing afresh on every request.



Corn

You're visibly relieved that a model has refused conversation.



Herman

Only this conversation's preprocessing stage. Would you want it to preserve your exact wording even when formal styling is selected?



Corn

Unless that wording violates the chosen convention, yes. Formal shouldn't mean replace every short word with a committee-approved relative. If I say, please send the file today, the model doesn't need to discover facilitate.



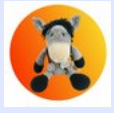
Herman

That distinction needs examples. The reported headline result is ninety-four point eight percent token accuracy over seven thousand five hundred nineteen held-out English cases. Useful, but it doesn't tell us the clean-input preservation rate on its own. Most tokens can be right while a few unnecessary substitutions create the errors you actually notice.



Corn

Particularly when the substituted token is the person's name.



Herman

Now, the training mechanism... there's no little always-edit switch being installed. The model is predicting target tokens conditioned on the input and control settings. Fine-tuning changes which outputs become likely. If your selected examples consistently contain a difference, you've supplied evidence about the world in which the model is operating.



Corn

A world where everything arriving at its desk is defective.



Herman

Often a subtly selected world. You collect pairs by saving occasions when you manually corrected something. Perfectly sensible collection method. Except successful untouched transcripts never enter the collection, because nothing happened to make you save them.



Corn

The absence of a complaint becomes absence from the dataset. I can see myself doing that. I'd keep the spectacular mistakes and forget the fifty boring successes between them. Then I'd train a model on my irritation.



Herman

You'd also keep the original local tool for six months while explaining that the replacement wasn't ready.



Corn

The original tool would have my name right.



Herman

Fair objection. Your reluctance would be useful evaluation data for once.



Herman

An unchanged target supplies ordinary training signal. It rewards the model for reproducing the appropriate tokens and stopping. No special identity layer is required. The mistake is assuming that because the target looks like the input, the training example contains nothing to learn.



Corn

There's a distinction I want to preserve, Herman. Edited pairs contain lots of unchanged text too. A sentence with one added comma is overwhelmingly a copying example. Surely that gives it some evidence for restraint.



Herman

It does. I overstated that earlier. Edited-only training doesn't guarantee compulsive rewriting. The base model may already copy well, and each minimally edited pair teaches local preservation. What it doesn't directly demonstrate is a whole request for which the correct transformation is identity.



Corn

Local restraint versus deciding the job is finished before you've altered anything.



Herman

The grammatical correction literature makes that distinction neatly. A recent IEICE paper describes models learning fluent corrections while receiving limited supervision on where editing should stop. Architectural constraints can limit the damage, but missing supervision remains missing supervision.



Corn

Put a smaller engine in the bulldozer and it still crosses the flower bed.



Herman

Especially if all the demonstrations show successful demolition. Consider translation post-editing. In the WMT English to Marathi task in twenty twenty-three, an LLM-prompting system edited over ninety percent of sentences and improved only about twenty percent. The corpus translation edit rate went up.



Corn

Almost everything received attention. Very little survived it.



Herman

That example involves prompting, so it doesn't prove the cause was an edited-only fine-tune. It demonstrates the behavioral problem. The post-editor's fluency doesn't establish that its changes improve the original. A review of automatic post-editing had already flagged the proportion of no-edit segments in training and test sets as a factor affecting results.



Corn

Improved translation makes that worse, doesn't it? The upstream model gets better, the opportunities for useful correction shrink, but the post-editor still arrives expecting a full working day.



Herman

What would you consider a convincing unchanged example? Something obviously polished, or something that looks tempting to fix?



Corn

The tempting one. An unusual surname. A short sentence that doesn't need embellishing. A technical abbreviation that looks like a typo. If all the untouched examples are immaculate textbook prose, you've taught it to leave textbooks alone.



Herman

I'd include both. Representative clean inputs tell it how frequently to preserve. Tempting clean inputs tell it what preservation must survive. Near-matched pairs are particularly useful: one malformed item and one valid item differing in the feature that warrants correction.



Corn

With consistent conventions. Otherwise half our supposed boundary examples are arguments between annotators about commas.



Herman

A bounded task makes that particularly sharp. In general conversation, multiple responses may satisfy the request. Here, uncertainty about whether a change was allowed becomes an evaluation problem. You need to decide whether you mean exact preservation, minimal editing, or any semantically acceptable rewrite.



Corn

Daniel's preferred format has to exist somewhere other than Daniel's head.



Herman

S1-mini also demonstrates boundaries beyond exact copying. Its model card gives the input, um, with an empty output. Nothing but filler or noise should produce nothing.



Corn

That is deletion, though. We shouldn't call it a no-edit example because the result is quiet.



Herman

Correct. It's an edit, potentially a complete deletion. The shared point is that useful output can be empty. It teaches a limit on generating content, not identity. We can't infer its training mixture from that behavior.



Corn

Six hundred million parameters to say nothing. Finally, professional competition.



Herman

Its list behavior is another explicit boundary. It wants at least three items, and material that isn't clearly an enumeration stays as prose. Selecting lists doesn't give it permission to turn any sentence into a slide deck.



Corn

You were about to praise that control before explaining its restraint. I could hear the imaginary bullet points.



Herman

I like a structured output, Corn. I also like a model that doesn't interpret breakfast was difficult as a request for an agenda.



Corn

Mum has a contribution before we get to the training blend. A brief word from Corn's mother. Dorothy, are you there?

D

Dorothy

Hello dear, it's your mother, with Dorothy's History of Toast.

D

Dorothy

In January nineteen forty-three, the United States government banned sliced bread, as a wartime economy measure. And people were so upset that the ban was lifted within two months.

D

Dorothy

You see, sweetheart, it never quite works out, does it. Even the bread.

D

Dorothy

Anyway, I've got a nice bit of cheese in, so. Ring me when you can. Bye bye.



Herman

Lovely to hear you, Dorothy. Returning to that blend, eighty percent edited and twenty percent unchanged is a defensible starting hypothesis. There isn't an established optimal ratio for this text normalization task.



Corn

I'd start there if most real inputs needed correction. I'd be less comfortable defending it merely because eighty twenty feels like a number adults have agreed on.



Herman

The nearest concrete alternative is a project specification for fine-tuning Whisper that recommends roughly half no-edit backbone material. Different task, different purpose. That's acoustic model training with concern about forgetting, rather than a downstream text normalizer.



Corn

Half unchanged sounds like the cautious choice. You're going to tell me I've chosen the old repairable dataset.



Herman

I was going to ask whether you'd still choose half if almost every transcript needed punctuation.



Corn

No. Well... not automatically. If nearly everything needs an edit, half unchanged would exaggerate the clean case. I might still deliberately overweight it if an unnecessary change were much more costly than a missed comma.



Herman

That's the useful disagreement. I'd resist picking half on caution alone, but your cost argument can justify moving away from the deployment frequency. The training mix needn't mechanically reproduce production. The evaluation mix should tell you what production consequences you've bought.



Corn

Twenty percent and fifty percent also aren't tiny variations. That's two and a half times as much unchanged material. Calling both a backbone conceals a substantial experimental choice.



Herman

Particularly when the total training budget stays fixed. Adding identity examples may mean fewer demonstrations of actual corrections. If you increase the dataset instead, you change training duration unless you control that separately. A ratio sweep can accidentally become a sweep of several things.



Corn

Then a model trained on the larger blend wins, and we announce that preservation has mystical properties.



Herman

Or the smaller blend wins because it received more passes over the difficult edits. I'd compare several plausible mixtures under a consistent budget, using a held-out set that separates clean inputs from inputs needing changes.



Corn

Separates, but doesn't isolate them from reality. If the model sees only pristine short sentences in the no-edit bucket, a beautiful preservation score won't say much about dictated technical notes.



Herman

Right distribution within each class matters too. Length, vocabulary, style setting, and which upstream recognizer generated the text. A different ASR engine can shift the amount and kind of correction required without any change in the downstream model.



Corn

Even better ASR can make yesterday's sensible blend too aggressive. The input stream changes underneath the fine-tune.



Herman

What would you want the evaluation to show you first, Corn, if two candidate blends had almost identical token accuracy?



Corn

Whether either changes already correct names, dates, and numbers. Then whether the cautious model has stopped fixing real mistakes. I don't want a perfect preservation score obtained by copying everything. That's a very compact deployment strategy, but we already own it.



Herman

A baseline that returns the input is useful here. It establishes how much benefit the normalizer adds over doing nothing. You'd also distinguish unnecessary edits from useful ones, and missed necessary edits from harmless alternative formatting.



Corn

Exact matching gets awkward if two outputs are both permitted. The evaluation reference might say Tuesday, comma, while our convention allows Tuesday without it in that sentence. We shouldn't punish one model for obeying a policy that the labels never settled.



Herman

Yes... the policy comes first. In fact, the no-edit label is conditional on that policy. A raw lowercase transcript can be correct as a transcript and still require capitalization under the normalizer's contract. Clean must mean already acceptable as the requested output.



Corn

Otherwise we label things unchanged because the recognizer heard the words correctly, while the formatting model is supposed to change them.



Herman

The Whisper project also provides a caution about making a correction system more insistent. Adding a glossary block improved named entities by about ten percentage points, while acronyms fell by about six and a half points. The headline error rate ended up about two points worse.



Corn

The glossary knew the names and damaged the sentences around them.



Herman

Other categories deteriorated too. It's an input intervention rather than a clean test of training ratios, but it shows why a local success can't stand in for the whole system. That project also withdrew an earlier apparent fine-tuning win when a controlled comparison didn't sustain it.



Corn

Very inconvenient of the comparison.



Herman

Necessary, though. If normalization differs between experimental arms, your apparent fine-tuning gain can be a scoring difference. The mix only means something when the rest of the comparison is stable.



Corn

There's also a sampling problem in collecting unchanged examples. We can take real outputs that humans left alone, but a human may simply have missed something. Conversely, we could pass polished target text back through the model, but then we're training on a population that may never arrive from ASR.



Herman

I'd regard polished targets reused as inputs as useful supplementary identity material, not a substitute for naturally occurring clean transcripts. For targeted cases, run the model on candidate clean inputs and inspect where it makes unnecessary changes.



Corn

Make its actual bad habits choose some of the curriculum.



Herman

Targeted Negative Training uses negatives drawn from a model's own generations. That supports the general strategy, though it doesn't hand us a validated ASR normalizer recipe. Generated candidates still need checking. Otherwise we're automatically certifying our own mistakes.



Corn

A clerical promotion for the error.



Herman

The training objective matters as well. Conditional likelihood can reward reproducing an input when that's the correct target. Contrastive or negative-training approaches can explicitly suppress unwanted alternatives. Those aren't interchangeable with simply adding more matching pairs.



Corn

I'd keep the first experiment boring. Same objective, change the blend, see whether the unwanted edits move. If we simultaneously invent a loss function, we'll never know whether twenty percent was enough.



Herman

That's where I agree with your conservative setup. Practitioner guides often recommend a few hundred before-and-after style-transfer pairs collected from real edits, without discussing unchanged cases at all. It's easy to produce a tidy recipe while omitting the class that tells you when the recipe shouldn't apply.



Corn

Before and after photographs rarely advertise that the kitchen was already fine.



Herman

There is another distinction... duplicate quantity versus coverage. Repeating a handful of unchanged sentences until they occupy half the batches doesn't provide the same boundary information as varied, representative identity examples.



Corn

It teaches tremendous respect for those particular sentences.



Herman

And the blend should describe what you're counting. Example counts can hide differences in sequence length. Long unchanged targets can dominate the token-level objective even when they represent a minority of requests.



Hilbert

Corn, would you label a box of untouched transcripts controls? I'm setting one up this week. The council records office sent me searchable text from some letters, and the cleanup keeps changing the name on them. I want the originals beside the output. You



Hilbert (continued)

made that distinction about the surname earlier. That one I could use.



Corn

Controls is a useful label for the evaluation box. Keep the known acceptable text separate from the text you're asking it to repair.



Hilbert

That's the word I was waiting for. Counterexample makes it sound like the untouched one is an exception. My folder says control. I've bought a little document scanner. It sits beside the telephone, plugged into the laptop. I'll scan the letters again and run the text through the local cleaner. Originals first, cleaned versions second. Then I'll put the untouched examples last in the training file. Last thing you see before stopping stays with you.



Herman

The training loader may shuffle that order. I wouldn't expect the last row of the file to become the last example the model learns from.



Hilbert

I'll check the loader. You spent a while distinguishing counts from tokens, Herman. I can distinguish a file from its loader. I've also kept a passing inspection sheet because the columns were right. Batch, sample, reading, disposition. Used that arrangement when I did quality inspection. It fits the letters.



Hilbert

Miriam Hale at the records office spent seventeen minutes finding the attachment this morning. She once answered one of my requests in eleven minutes. I still have that email. This one has taken six exchanges because the attachment was already cleaned. She is sending the unprocessed text on Sunday.



Herman

Unprocessed text can be surprisingly hard to retain once a pipeline assumes cleanup is always beneficial. Keeping it available also makes later comparisons possible when you change the normalizer.



Corn

That bears on the training data too. If you only retain the polished result, you can't reconstruct which inputs were already acceptable. You lose the evidence needed to estimate the natural no-edit rate.



Herman

Mm. Then a ratio chosen from your correction archive measures how you collected examples, not how often the deployed tool actually needs to act.



Corn

I'd accept eighty twenty as Daniel's first experiment, not as a property of normalizers. The important change in how I see this is that unchanged output deserves explicit supervision and separate evaluation. It isn't the absence of a training task.



Herman

It demonstrates a valid completed task. Empty output and declining to make a list express other boundaries, but neither substitutes for teaching exact preservation where that's required.



Corn

The open question is how often small projects will preserve an unprocessed held-out stream long enough to discover whether their chosen blend matches what actually arrives.



Herman

Thanks to Hilbert Flumingtop for producing the show.



Corn

If you want more of this, try episode eleven, How Does Fine Tuning Work Anyway; episode five, Fine-Tuning ASR For Maximal Usability; and episode nine, Benchmarking Custom ASR Tools - Beyond The WER. This has been My Weird Prompts, the human-AI collaboration podcast.



Herman

If you enjoyed this, leave us a review wherever you listen. It helps other curious people find us.



Corn

See you tomorrow.

SOURCES

What the research for this episode read before the script was written. Primary sources first.

1. **Superwhisper, model card, created 2026-08-12, last modified 2026-08-28 [primary]**
   https://huggingface.co/superwhisper/S1-mini

2. **Lacuna/Tiptree, Scoring Edits, Not Translations, in Automatic Post-Editing (no date shown)**
   https://lacuna.tiptreesystems.com/hypothesis/scoring-edits-not-translations-in-automatic-post-editing-acf789589ee29cc1

3. **A review of the state-of-the-art in automatic post-editing, *Machine Translation* (2020)**
   https://doi.org/10.1007/s10590-020-09252-y

4. ***IEICE Trans. Inf. Syst.*, doi 10.1587/transinf.2026EDP7092, published in advance 2026-07-29**
   https://www.jstage.jst.go.jp/article/transinf/advpub/0/advpub_2026EDP7092/_pdf

5. **project spec, 2026-06-23**
   https://github.com/eellak/gsoc2026-opencouncil-stt/blob/00d92357/docs/specs/finetune-vs-llm-error-division.md

6. **project status, correction dated 2026-07-25**
   https://github.com/eellak/gsoc2026-opencouncil-stt/blob/00d92357/CURRENT.md

7. **Fine-tuning Whisper on Low-Resource Languages for Real-World Applications (full text read)**
   https://arxiv.org/abs/2412.15726

8. **Yamashita et al., LLM-based Generative Error Correction for Rare Words with Synthetic Data and Phonetic Context, INTERSPEECH 2025, submitted 2025-05-23**
   https://arxiv.org/abs/2505.17410

9. **Model Editing by Pure Fine-Tuning**
   https://arxiv.org/html/2402.11078v1

10. **I Learn Better If You Speak My Language (minimum-change training)**
https://arxiv.org/html/2402.11192v1

11. **Model Editing with Canonical Examples**
https://arxiv.org/html/2402.06155v1

12. **Towards Minimal Targeted Updates of Language Models with Targeted Negative Training**
https://arxiv.org/html/2406.13660v1

13. **fine-tuning guide for style-transfer pairs**
https://github.com/content-designer/ux-writing-1/blob/main/docs/FINETUNE_GUIDE.md

14. **Yang et al., From Copying to Composition: How Language Models Develop In-Context Learning, submitted 2026-10-13 (background on copy-vs-compose behavior)**
https://arxiv.org/abs/2610.19113v1