Episode Book

EPISODE #6672

How Language Textbooks Flatten Speech and Emotion

PUBLISHED	19 November 2026 at 00:10 UTC
DURATION	22:34
LISTEN	myweirdprompts.com/episode/language-textbooks-real-speech-56479/
HOSTS	Corn, Herman, Hilbert
SCRIPT MODEL	z-ai/glm-5.3
VOICES	omnivoice
PIPELINE	V5.35
SOURCES	24 cited by the research (listed at the end)
BOOK GENERATED	19 November 2026 at 03:42 UTC

SYNOPSIS

Language textbooks don't just sound stiff to beginners—corpus research shows they stay just as flat at advanced levels. Exploring Elen Le Foll's Textbook English Corpus, David Crystal's 1976 critique of textbookese, and Block and Gray's work on erased working-class voices, this conversation unpacks why coursebook dialogues feel like a sanitized middle-class performance. It also examines Jean-Marc Dewaele's study of 1,039 multilinguals, revealing why learners hunt for swear words to close the emotional distance left by formal materials.

TRANSCRIPT



Corn

"My name is Tom and I'm going to a job interview." Daniel has been reading sentences like that aloud for hours, because he's building training data for a speech recognition model, and at some point in the middle of it he noticed he wasn't just tired of recording — he was tired of Tom.



Herman

Tom sounds exhausting.



Corn

Tom has been going to job interviews in every language on earth for seventy years and never once got the job. That's Daniel's framing, roughly. He's a lifelong language learner, he's done this in Hebrew too, and he says the sentences drag you into a make-believe world of painfully clean, neutral life — his phrase for it is a Disney World version of how people actually live. And he's noticed the counter-movement: the learners who finish the textbook and then go learn the swear words. His question is whether the sanitised vocabulary itself is the problem — whether everyone can tell this isn't the language people actually reach for when they feel anything, and whether that makes the language feel academic and unreal and drains the motivation to bother. And he's careful to say this isn't the textbooks-versus-speaking debate. It's about the content of the materials themselves. Have the linguists actually made this critique?



Herman

They have, and for longer than you'd guess. The short answer is yes — the critique exists, it goes back at least to the nineteen seventies, and there's now corpus work that puts a number on it. But the number and the feeling are two different claims, and I want to keep them apart, because the field does.



Corn

Which is the polite way of saying one of them hasn't been tested.



Herman

Partly, yes. But let me set up the three strands first, because Daniel's prompt is actually three complaints wearing one coat. Strand one is linguistic: the dialogues don't sound like speech. No contractions, no interruptions, nobody finishes a sentence for anybody. Strand two is situational: the world the dialogues depict — nobody's skint, nobody's furious, everyone goes to the job interview. Strand three is emotional: learners want the words that carry feeling, which is why the swearing comes afterwards, and that one has its own research literature.



Corn

And the phrase for strand one has been around a while.



Herman

"Textbookese." Avon Crismore wrote it up in the mid-eighties — she dates the style back to nineteen thirty, and calls it an emotionless style where the author gets flattened out. Third person, objective point of view, nobody home behind the sentence. She was writing about textbooks generally, not language textbooks — but that's the register. In language teaching the term is textbook English or school English, and that's where the hard evidence lives.



Corn

So take the linguistic strand. You said there's a number.



Herman

There's a number. Elen Le Foll built what she calls the Textbook English Corpus — forty-three EFL coursebooks, nine best-selling series, eight major publishers, secondary schools in France, Germany and Spain, including the audio and video transcripts. She ran it through Biber's multi-dimensional analysis, which scores texts on things like how involved versus informational they are — whether the language is full of pronouns and hedges and feeling, or whether it's packed with nouns. Real conversation scores high on involvement. Textbook conversation should too, if it's modelling conversation.



Corn

It doesn't.



Herman

It scores fifteen point seven five. Which sounds fine until you see that real spoken British English — the spoken BNC twenty fourteen corpus — sits sixteen points higher. Sixteen points, on a scale where personal letters are around nine and fiction is around five. Textbook conversation is less like talking than a novel is. Less like talking than somebody's letter home.



Corn

That's a finding with a certain cruelty to it.



Herman

It gets crueler. Le Foll's interpretation is that the dialogues are more nominal — noun-heavy — and appear to serve vocabulary reinforcement rather than model realistic speech. Which is Daniel's whole complaint converted into a measurement: the dialogue isn't a conversation, it's a delivery vehicle for the word list. Tom goes to the job interview so that "job interview" can be in a sentence.



Corn

The dialogue is the pill.



Herman

And the structural part is what I find striking — she found register explains about sixty-three percent of the variance in the scores, and proficiency level, whether the book is for beginners or advanced students, explains about three. Three percent. The blandness doesn't fade as you climb the ladder. It's not a concession to beginners that gets fixed at B two. It's baked into the genre.



Corn

That's the part I'd want learners to sit with, because everyone assumes it — you tolerate the fake English at the start because you're told it gets real later. And it doesn't.



Herman

And this critique is old. David Crystal — with Derek Davy, in nineteen seventy-six, Advanced Conversational English — opens by describing teachers and advanced students abroad who had good formal English and were openly dissatisfied with what their coursework gave them. And there's a line I want to read, because it's Daniel's whole setup fifty years early. "Despite all the available materials, the request to say some real English into my tape recorder was disconcertingly persistent."



Corn

Fifty years.



Herman

Students in the early seventies literally asking David Crystal to speak real English into a tape recorder because the materials weren't it. And now Daniel is in Jerusalem recording hours of sentences into a microphone to build training data — we did an episode on his custom ASR tooling where the recording side was the whole practical problem — and the complaint hasn't moved. The technology changed; the sentence didn't.



Corn

Their diagnosis, if I remember the line right, was a stereotype of conversation.



Herman

"For too long, English language teachers have been operating with a stereotype of conversation." And to their credit they conceded the hard part — that colloquiality sparks off attitudes about when, how and how much it should be taught, or whether it can be formally taught at all. That concession is still the live argument today.



Corn

Because the obvious objection to all this is that you can't put a real pub conversation in front of a beginner and call it pedagogy.



Herman

And the field takes that objection seriously, which we'll get to. But before we leave Crystal and Davy — their method was to record many hours of actual conversation and take unedited extracts. They said they hadn't come across commercially available material that informal or realistic. So the alternative existed; it was just never the product.



Corn

Fifty years of knowing the fix and not shipping it.



Herman

Publishers have reasons, and we'll get to those too. But the honesty I owe Daniel here is this: the corpus proves the input is odd. It does not prove the oddness demotivates anybody. Le Foll is explicit that the study makes no claim about causal links to learner output, let alone learner motivation. The feeling is plausibly real and measured somewhere else.



Corn

Which is strand three's job. But strand two first — whose world is Tom's?



Herman

This is where it stops being aesthetics and becomes ideology. Block and Gray traced UK-produced ELT textbooks from the nineteen seventies onward and documented a progressive decline in working-class characters — appearing only incidentally, with a working-class perspective effectively erased. Their phrase is that at the heart of the ELT textbook is a regime of representation, a way of constructing the world that suggests what it means to be a speaker of English. And the regime isn't neutral. The Disney World isn't empty — it's furnished middle class.



Corn

Nobody in the textbook is worried about rent.



Herman

Nobody in the textbook has a job that hurts. There's a Japanese ELT textbook study that uses the word sanitized outright — it describes individualized Westernized middle-class lifestyles and consumerism. And a Finnish study of adult-learner textbooks found the same shape: middle class, conventional, a fairly narrow image of Finnishness. Its abstract has the strongest theoretical claim available for Daniel's thesis: through careful decisions concerning what to include and what to omit, textbooks have the power to direct what a beginner can and should say in their target language. That's it, stated plainly. The omission is the teaching.



Corn

The learner can only say what the textbook admitted into existence.



Herman

Within that world, yes. And there's a named distinction for the two failures. Sitaraman and Vasuthavan, twenty eleven — they split linguistic realism, the dialogue being artificial and interrogative and free of contractions, from situational realism, whether the situation is believable. Their opening is blunt: one very obvious reason adolescent learners resent language lessons is the continued use of unrealistic, illogical and stereotype kind of instructional materials. And the lacking in real, genuine communication is what makes lessons tedious and boring.



Corn

"Tedious and boring" — that's the motivation claim.



Herman

It's asserted, though. It's their reading of adolescent resentment, not a trial. And it's worth noting the mechanism they name — that the material is designed specifically to teach rather than to be. The situation exists to carry the grammar point. Everyone in the dialogue cooperates with the grammar point.



Corn

That's the thing I'd never noticed until Daniel said it — nobody in a textbook has ever misunderstood anything. Two people meet, exchange perfectly formed questions, and both walk away fully informed. I've had maybe three conversations like that in my life, all with dentists.



Herman

The natural conversation data backs that up — real talk is full of overlap, restart, unfinished clauses. The textbook dialogue is a turn-taking ceremony. So the blandness is structural, not a beginner-stage problem.



Hilbert

Hilbert's Anteater Trivia. The Portuguese word tamandua, used for these animals, comes from the Tupi language of Brazil. I have worked here long enough that my own name now sounds like a loanword.



Corn

Moving on, before that becomes a segment.



Herman

Where were we — structural. The other thing about the situational strand: it fails in both directions. There's a Bangladesh study from Applied Linguistics on locally produced textbooks, and the high-status students there rejected the textbook's links between English and low-profile jobs because it contradicted their class aspirations. So the sanitized world is too bland for some learners and not aspirational enough for others. You can't furnish the world neutrally — every choice somebody made.



Corn

Which takes us to the swearing, because that's the part everyone actually does.



Herman

Jean-Marc Dewaele is the name here. He surveyed a thousand and thirty-nine multilinguals — seventy-five different first languages — and had them rate swear and taboo words on an emotional weight scale. In the first language, the words carry much more force than in the second. The gap between first and second language was enormous — one of the largest effects I've seen in this literature — and it repeats, weaker, from second to third language, third to fourth.



Corn

So a word you learned as an adult literally weighs less.



Herman

Feels like it weighs less, when you say it. His interpretation is the quotable part: later-learned languages are languages of distance and detachment. You can say the word, you know exactly what it means, and it doesn't land in you. And that — Corn, that's the whole counter-movement. Learners hunting for the swear words after the textbook aren't being edgy. They're trying to close a distance that's been measured. The textbook omitted precisely the vocabulary with the highest emotional payload, which is to say the vocabulary that makes a language feel inhabited.



Corn

The first words with any voltage in them.



Herman

Dewaele defines taboo words as multifunctional pragmatic units — they do discourse work, they manage emphasis and solidarity, they're grammar-adjacent social machinery. The textbook leaves out an entire functional system and calls the remainder the language.



Corn

And then we're surprised the language feels like a specimen.



Herman

That's Daniel's word, basically — academic. But here's the counter-argument, and it's a real one. Alexander Gilmore did a big review in two thousand seven of the authenticity question, and he found that the claim "authentic materials motivate learners" is everywhere in the literature — he lists something like fifteen sources asserting it — but the empirical support rests on about three studies. Three. One of them, Peacock, found motivation gains with South Korean university students over twenty days, which is a short window. And Berardo warns the other way: a badly chosen authentic text can de-motivate, and in Krashenite terms puts up the affective filter. Tyson and Baek say avoid too many unfamiliar idioms or slang.



Corn

So swing the pendulum and you drown the beginner.



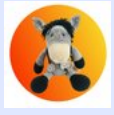
Herman

Hand a first-week learner a genuine pub argument and you don't get motivation, you get panic. The anxiety finding is consistent. So the honest answer to Daniel is layered: the critique of bland content is fifty years old and well documented; the linguistic oddness is measured; the emotional distance is measured; the claim that blandness causes demotivation is asserted by some and directly contradicted by nobody, and tested by almost nobody.



Corn

That's a gap shaped exactly like a PhD thesis.



Herman

Somebody should write it, yes. Though I'd want them to separate the strands the way we did, because the fix for each is different. The dialogue problem is fixable — write realistic dialogue and grade the vocabulary around it. The world problem is a publisher decision. PARSNIP, incidentally — that's the actual industry acronym for the topics coursebooks avoid. Politics, alcohol, religion, sex, narcotics, isms, pork.



Corn

Pork.



Herman

Pork. So the blandness isn't an accident of imagination; it's a checklist. And this is also where I'd draw Daniel's other line for him — the distinctness from the textbooks-versus-immersion debate. That argument is about method: classroom versus street, study versus speaking. This critique is about content. You could learn purely by immersion and still be handed sanitized materials; you could learn purely from a book and be handed realistic ones. The medium isn't the problem. The register and the world the register depicts are the problem.



Corn

The argument was never where you learn. It's what they put in front of you.



Herman

And notably, nobody has studied the swearing counter-movement itself as a learner strategy — Dewaele explains why the words matter, but the post-textbook taboo hunt is a documented gap. People do it, nobody's measured it.



Corn

Daniel's doing ethnography without the paperwork.



Hilbert

Which of you decided the word Disney, though. Because I've been sitting with that one.



Corn

It was Daniel's.



Hilbert

Then he got it right, mostly, and I'll tell you where he got it slightly wrong. I have a phrasebook. Had it a long time. I'm not saying which language, because the language isn't the point. The entry for I don't understand was in a register I can only call ceremonial. I used it exactly once, at a counter, and the man laughed — actually laughed, and then explained to a colleague what I'd said, and they both laughed. I had paid money to sound like an announcement.



Herman

That's the Kanagawa finding in one counter. Their study says the teaching recordings are almost always advertised as natural or authentic even when they sound quite unnatural and are almost certainly not authentic in the sense of being unscripted. Was yours sold that way?



Hilbert

On the back cover. Phrases as spoken by native speakers today. I checked, Herman. I have had it checked. That claim has been on the back cover the entire time, through three printings. I wrote in the margin, years ago, what I should have said instead. I didn't speak the language then. I had to get someone to help me choose the correction, and then I memorised it so the margin would be mine.



Corn

You annotated the phrasebook in a language you didn't speak.



Hilbert

I do the margin before I do the trip. Always have. That book is more mine now than it was when it was new, and none of it is the publisher's. The laughing man did me a favour. He was the first honest reviewer the thing ever had.



Corn

The misconception, before we go — the one I'd bet most listeners carry.



Herman

Say it.



Corn

That the fake English is a beginner's concession. That you eat the cardboard sentences at level one and somebody hands you the real language at level five.



Herman

Le Foll's numbers close that door — proficiency explained three percent of the variance, register sixty-three. The advanced book sounds nearly as staged as the beginner one. If anything the gap narrows from the wrong direction, because real speech doesn't get more textbookish as you improve.



Corn

Nobody ever hands you the real language. You go get the swearing yourself.



Herman

Which leaves the question nobody has quite answered. The corpus proves the input is odd, Dewaele proves the emotional gap is real, but the direct study — does bland content cause demotivation — appears not to exist, and on the Hebrew side there's no dedicated study of ulpan textbook sanitisation either. The closest literature is Arabic, modern standard versus dialect, where the same complaint is decades old. Daniel's Hebrew observation stands on its own feet, illustrated by everything we said today, but not documented. And the sentence I keep hearing is the Finnish one — the textbook's power is in what it omits, and the words it omits are the ones with the most feeling in them. There's a producer credit to hand out first, though — Hilbert Flumingtop, who has been at the desk annotating us in a language we don't speak. For more along these lines, there's episode seven, Building Custom ASR Tools; episode nine, Benchmarking Custom ASR Tools - Beyond The WER; and episode fourteen, AGI's Crossroads. This has been My Weird Prompts — no question too obscure, no rabbit hole too deep. If you enjoyed this one, the full archive lives at my weird prompts dot com. See you tomorrow.

SOURCES

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

1. **David Crystal & Derek Davy, *Advanced Conversational English*, Longman 1976, PDF re-keyed 2019 ; retrieved 2026-11-19. [primary]**

   https://davidcrystal.com/Files/BooksAndArticles/-5312.pdf

2. **Sitaraman & Vasuthavan, Adolescent Instruction in the ESL Context: Situational Realism and Linguistic Realism in Material Selection, *Academic Journal UiTM Johor* Vol. 10, 2011; retrieved 2026-11-19.**

   https://ir.uitm.edu.my/id/eprint/5881/2/5881.pdf

3. **Elen Le Foll, Register variation in school EFL textbooks, *Register Studies*, John Benjamins, 2021, pp. 207**

   246 (page itself blocked by Cloudflare; content via search index); retrieved 2026-11-19.
   https://www.jbe-platform.com/content/journals/10.1075/rs.20009.lef

4. **John Benjamins catalogue entry for the same article; retrieved 2026-11-19.**

   https://benjamins.com/catalog/rs.20009.lef

5. **Le Foll, Textbook English MDA synthesis page; retrieved 2026-11-19.**

   https://elenlefoll.github.io/TextbookMDA/summary.html

6. **Zenodo figure: six registers of the Textbook English Corpus on Biber's Dimension 1; retrieved 2026-11-19.**

   https://doi.org/10.5281/zenodo.4732285

7. **Avon Crismore, The Use of Author Roles in Improving Textbooks and Learning, Technical Report No. 365, ERIC, December 1985; retrieved 2026-11-19.**

   https://eric.ed.gov/?id=ED281139

8. **Block & Gray, All Middle Class Now? Evolving Representations of the Working Class in the Neoliberal Era: The Case of ELT Textbooks, Palgrave; retrieved 2026-11-19.**

   https://doi.org/10.1057/9781137276285_2

9. **English language and employability in locally produced ELT textbooks, *Applied Linguistics*, 2025; retrieved 2026-11-19.**

   https://doi.org/10.1093/applin/amae070

10. **National citizens in a sanitized vision of global economy in Japanese ELT textbooks, *Journal of Language and Politics*, John Benjamins; retrieved 2026-11-19.**

    https://www.jbe-platform.com/content/journals/10.1075/jlp.26035.shi

11. **Learning language, learning culture: Constructing Finnishness in adult learner textbooks, *Discourse & Communication*; retrieved 2026-11-19.**

    https://journals.sagepub.com/doi/10.1177/1367549417719062

12. **Jean-Marc Dewaele, The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals; retrieved 2026-11-19.**

    https://www.researchgate.net/publication/237444929_The_Emotional_Force_of_Swearwords_and_Taboo_Words_in_the_Speech_of_Multilinguals

13. **Semantic Scholar record for the same paper; retrieved 2026-11-19.**
https://www.semanticscholar.org/paper/The-Emotional-Force-of-Swearwords-and-Taboo-Words-Dewaele/47cdf482e937b770483365d9675e5aa9fccc9890

14. **Alexander Gilmore, Authentic materials and authenticity in foreign language learning, 2007; retrieved 2026-11-19.**
https://www.alexandergilmore.com/wp-content/uploads/2022/04/Gilmore_2007.pdf

15. **Berardo, *The Reading Matrix* Vol. 6, No. 2, September 2006; retrieved 2026-11-19.**
https://www.readingmatrix.com/articles/berardo/article.pdf

16. **UAE High School Students' Attitudes Towards Using Authentic Materials (citing Tyson & Baek 1999); retrieved 2026-11-19.**
https://doi.org/10.82366/1133

17. **Authenticity in Language Learning (OKARA, UIN Madura), citing Richards 2001; retrieved 2026-11-19.**
https://ejournal.uinmadura.ac.id/index.php/okara/article/download/8577/3634/

18. **Use of Authentic Materials in the ESP Classroom, ERIC; retrieved 2026-11-19.**
https://files.eric.ed.gov/fulltext/ED530013.pdf

19. **(prosodic features of English teaching recordings), Kanagawa University repository; retrieved 2026-11-19.**
https://kanagawa-u.repo.nii.ac.jp/records/12135

20. **LITE journal article on real vs. textbook conversation; retrieved 2026-11-19.**
https://publikasi.dinus.ac.id/index.php/lite/article/download/1100/813/36429

21. **Second Language Learners' Perceptions of Conversations in English Materials, *SAGE Open* 2025, authenticity checklist appendix; retrieved 2026-11-19.**
https://sage.cnpereading.com/storage/sage/journal/article/SGO/2025/SGO_2025_15_2_10.1177_21582440251339659/unzip/ds-sgo-10.1177_21582440251339659/sj-pdf-1-sgo-10.1177_21582440251339659.pdf

22. **Wikipedia, Word taboo, last edited 2026-10-26; retrieved 2026-11-19.**
https://en.wikipedia.org/wiki/Word_taboo

23. **Arabic Language Hub, on MSA vs. dialect in textbooks (practitioner blog); retrieved 2026-11-19.**
https://arabiclanguagehub.com/top-5-mistakes-parents-make-while-teaching-their-kids-arabic-and-how-to-fix-them/

24. **30 Expert Tips for Learning Any Language Fast and Easily, 2020-12-02 (practitioner blog); retrieved 2026-11-19.**
https://www.learnmedicalspanish.org/expert-language-learning-tips/