

Artistic Interventions for NLP Annotation Challenges: The Stress Test of Machinic Glossolalia

Tyler Grimes

Department of Critical Media Practices;
Institute of Cognitive Science
University of Colorado, Boulder, USA
tygr4724@colorado.edu

Marshall Washington

Department of Linguistics
University of Colorado, Boulder, USA
mawa8587@colorado.edu

Abstract

MotherBoard's Mother Tongue is a computational linguistics and artistic research project that explores a Large Language Model's (LLM) vocal production of glossolalia. Glossolalia, colloquially known as 'speaking in tongues,' consists of the human production of seemingly unintelligible utterances. It is, by its nature, difficult to annotate accurately with linguistic features relevant for natural language. The glossolalia-producing system demonstrated here consists of the interaction of 1) a 'nonsense' linguistic corpus 2) a micro-controller based environmental data stream and 3) a fine-tuned LLM. While discussing some philosophical and artistic considerations of machinic glossolalia, we also address some methodological considerations for Natural Language Processing (NLP). Using the artistic project as a case study, we argue that machinic glossolalia presents a 'stress test' that could inform both creative redirections of NLP methods and the definitions held by the subfield.

1 Introduction

Motherboard's Mother Tongue explores the linguistic phenomenon of glossolalia as produced by a Large Language Model (LLM). Colloquially, glossolalia is known as 'speaking in tongues'. 'Speaking in tongues' alludes to glossolalia's manifestation within certain orientations of Christianity, whereas 'glossolalia' is the more general term in linguistics. A useful definition is provided by linguist Dr. William Samarin in *The Linguisticity of Glossolalia*: "A meaningless but phonologically structured human utterance believed by the speaker to be a real language but bearing no systematic resemblance to any natural language, living or dead" (Samarin, 1968, p.51).

Glossolalia's biblical context, as well as a study of both the corporeal dimensions of its ritualistic practice and its neural correlates are laid out by

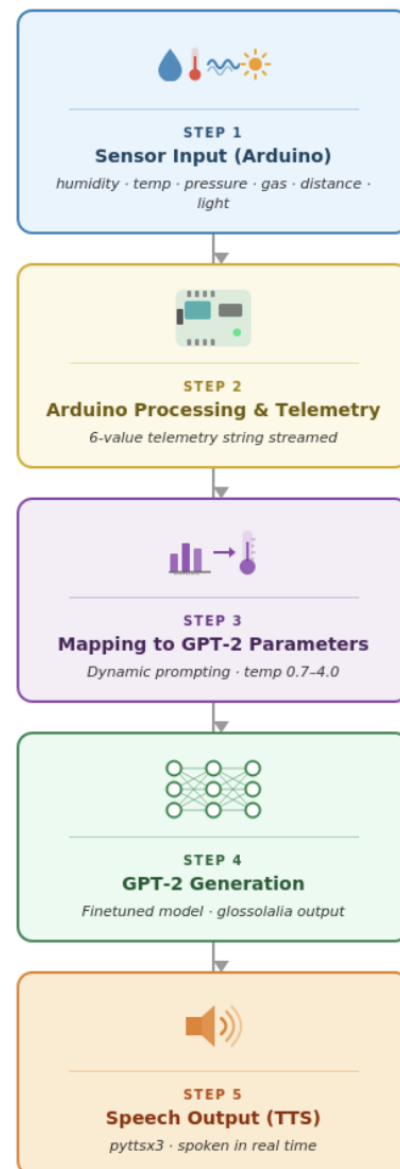


Figure 1: *Signal flow of system showing sensor input, computational process, and machinic vocal output.*

writer and professor John McGraw in *Tongues of Men and Angels*. Here, he builds from the work of Samarin, sharing a passage from Acts 2: 1-8, where a “baptism with the Holy Spirit” (McGraw, 2011, p.58) seems evident and is a prominent biblical instance of the phenomenon. Due to its effect on the prefrontal cortex, McGraw notes that though outsiders sometimes saw glossolalia as a type of drunkenness, to the performer, it provided them a sense of connection to the divine – a realm ostensibly beyond ‘natural language’.

If considering glossolalia as a measurable linguistic phenomenon, several considerations of corpus linguistics come to light which inform our project: 1) Can one apply conventional methods of corpus and computational linguistics to glossolalia? 2) How does the annotation of such a phenomenon expose preconceived notions about what is structured, annotatable, and ultimately valuable ‘linguistic data’? 3) Does the structure of glossolalia expose limitations in the tasks that LLMs can learn and execute? And 4) what effect might glossolalia-based training data have on a machine’s artificial neural network? Although glossolalia, if we hold Samarin’s definition, could not be a ‘language’ within the purview of NLP in the sense of a lack of transference of meaning because of its inability to “[...] possess the necessary elements of language and simply could not function as a system of communication”(McGraw, 2011, p.60); questions of this sort are potent for linguists, philosophers, and artists alike. This paper seeks to creatively examine machinic glossolalia as a stress test for corpus linguistics and its implications for NLP’s practice of LLM fine-tuning.

2 Related work

To date, there are very few publicly available corpora of glossolalia. One ongoing attempt is by researcher Paul de Lacy, who has manually transcribed over 19 hours of glossolalia. De Lacy’s analysis of his ongoing corpus construction suggests that “[...] glossolalia has no lexicon (i.e. no words/morphemes), and by this no syntax, morphology, or semantics. It is essentially ‘pure’ phonology and phonetics”(de Lacy, 2009). Despite his and his team’s efforts, the corpus has yet to be publicly released. Our main archive comes from linguist Blanche Speer in their 1972 dissertation, *A Linguistic Analysis of a Corpus of Glossolalia*. Speer’s contribution remains one of the few extensive lin-

guistic records of glossolalia. Their appendix contains dozens of pages of highly detailed and diacritically marked phonetic transcriptions. Considering these limitations on availability of public corpora of glossolalia, we searched and found inspiration elsewhere in the arts to build a preliminary and workable corpus.

2.1 Artistic Facsimiles of Glossolalia

There are other facsimiles of glossolalia that inform this project which are glossolalia-like. ‘Glossolalia-like’ is meant to describe an intuitive, vocalized utterance, which resembles the uncanny delivery of glossolalia in a secular and/or arts context. This project gains traction from avant-garde devices like: The ‘phonetic poem’ of the Dadaists, the method of ‘automatic writing’ of the Surrealist cohort, and the ‘Zaum’ language of Russian Futurism, amongst other contemporary improvisational vocal techniques. For example, in Russian, “za” meaning “beyond”, and “um”, meaning “mind”, when combined, is literally ‘non sense’ or ‘beyond mind’. Zaum was, in the words of an artist of that cohort, a “language of the birds... a language of the gods” (Janecek, 1996).

In contemporary media arts, one prominent example of machinic glossolalia is from Czech sound artist Marek Hlaváč, who released *Artificial Glossolalia: Machine Chant* in 2021, which consists of a machine sampling glossolalia. While similar to our project in its intent, Hlaváč’s work distorts glossolalia, as it is slowed down, echoed, and has multiple layers of audio effects applied. Our work differs in form and method considering that we constructed a corpus, leading to the glossolalia produced in our project resembling a more ‘standard’ form of the phenomenon. Our study is also not a music album, and is derived from the interaction of: a machine, its spectator, and its environment.

Methodologically, these artistic practices treat language as artistic material rather than as a transparent vehicle of meaning by foregrounding sound, rhythm, and affect over meaning. For our project, this avant-garde lineage is both historical precedent and methodological template: it demonstrates how a nonsensical ‘language’ can be systematically formed around sound-symbolic and ecological relations rather than lexical semantics. In this project, we effectively extended avant-garde theory into a computational linguistics pipeline that privileged phonology and affect over accuracy.

2.2 Other Instances of Annotation Without Semantics

Studies in linguistics show that non-meaningful vocal behavior can be annotated systematically. For example, non-lexical conversational sounds (fillers, grunts, laughter) have been studied and annotated. Ward (2006) shows that sounds form a “specialized sub-language” with compositional sound-meaning mappings. Using a corpus of American English conversation, Ward identifies forms like “h-hmm, aha, unmkay, nyeah, um-uh” and notes their pragmatic functions (e.g. turn-taking signals, backchannels, agreement, affect). Current approaches to paralinguistic annotation, like Ward (2006), treat non-lexical vocalizations as embedded elements within otherwise conventional linguistic utterances. Laughter, fillers, and vocal bursts are annotated as interruptions, supplements, or alternatives to propositional speech, occurring within a matrix of meaningful language that provides structural and theoretical scaffolding for annotation decisions. Even when these vocalizations are analyzed independently, their annotation relies on pragmatic or interactional meaning that grounds segmentation and classification choices. The annotation of these forms presupposes that they participate in a communicative system where function can be systemically described, even without propositional content.

Similarly, infant and early vocal development annotations also rely on non-meaningful sounds. Child babbling is annotated for phonological properties like canonical CV, consonant-vowel, syllables instead of meaning. For example, studies have analyzed daylong audio of infants across languages, crowdsourcing annotations of whether each vocalization contains a “canonical transition”, which is defined as adult-like CV timing (Cychosz et al., 2021). The fact that these diverse vocalizations obey patterns of form and use demonstrates they can be treated as structured data. More recent annotation research corroborates these results. Trouvain & Werner 2020 surveyed 11 spoken corpora and found that almost none consistently mark these non-lexical vocalizations except laughter. They conclude that “a wider distribution and more consistency is needed with respect to the annotation of NVVs (non-verbal vocalizations)”. In other words, linguists suggest that systematic annotation of non-meaningful sounds is warranted and doable even when the tokens carry no propositional content.

Prosodic features are also routinely annotated in

speech corpora in ways that do not depend on lexical meaning, yet they remain structurally anchored. Frameworks such as ToBI (Silverman et al., 1992) and RaP (Dilley and Brown, 2005) separate tonal and rhythmic structures from semantic interpretation while still requiring identifiable prosodic units. In the ToBI system, annotation is distributed across multiple time-aligned tiers, including orthographic transcription, tonal events, and break indices. Although tonal labels are assigned based on perceived pitch movement rather than word meaning, they are aligned to segmentable units and boundary judgments that presuppose recoverable structure. Similarly, the RaP framework encodes pitch events relationally. Each tonal target is labeled relative to preceding pitch targets rather than lexical items (Breen et al., 2012). Yet, this relational encoding still depends on identifiable syllables and prominence patterns. These systems demonstrate that prosodic annotation does not require semantics, but also reveals an implicit reliance on stable segmentation and reproducible structural units.

These findings, and others that investigate glossolalia, such as Link & Tomaschek (2024), who found that glossolalic syllables follow a Zipf-like frequency distribution and that speakers can learn serial patterns imply that glossolalia can be transcribed and annotated at the phonological level.

2.3 Have These Approaches Been Sufficient?

Field Linguistics and the Praat workflow (Ladefoged, 2007) have promising methods to assess glossolalia’s peculiarity. Praat can help with the analysis of low-resource languages, or an unfamiliar phenomenon like glossolalia, using phoneme-derived spectrograms. However, Praat’s tracking algorithms often struggle with the extreme pitch glides and irregular phonation common in glossolalia, leading to inaccurate acoustic data. Additionally, because glossolalia lacks a reference lexicon, researchers cannot use automated alignment tools – they must rely on subjective segmentation that lacks standardized linguistic boundaries, implying that Praat, although robust, may not be sufficient in capturing the nuance of glossolalia.

While the capture and annotation of non-lexical vocalizations has become increasingly sophisticated in both humans and non-humans, such as the Perch 2.0 methodology, which uses bioacoustic-derived spectrogram data to classify marine mammal sounds (Burns et al., 2025), existing frameworks for paralinguistic phenomena remain funda-

mentally inadequate for human glossolalia. This inadequacy stems not from technical limitations, but perhaps from deep theoretical assumptions about the relationship between linguistic structure and meaning that glossolalia seems to violate.

The annotation of infant babbling, while superficially similar to our approach, also differs in crucial ways. For example, glossolalia can constitute sustained stretches of vocalization that resist both proposition and pragmatic analysis. While paralinguistic vocalizations are typically realized as brief insertions lasting seconds at most, episodes of glossolalia can extend for long stretches of time, generating thousands of syllable-like units without recurring form-function pairings. The absence of stable mappings between form and meaning eliminates the anchor points that existing frameworks use to justify segmentation boundaries, tier relationships, and categorical assignments. Babbling annotation also focuses on developmental milestones and the emergence of language-like properties, treating each vocalization as evidence of progress towards (or a deviation from) target phonological patterns in the child’s developing language (McGillion et al., 2017; Lee et al., 2018; Cychosz et al., 2021).

Moreover, existing prosodic frameworks, including ToBI, assume that intonational structure maps onto information structures, speech acts, or discourse organization (Jun, 2023). Pitch accents are used to mark focal elements, boundary tones signal clause or utterance completion, and prosodic phrasing reflects syntactic constituency or information packaging. These assumptions are encoded in annotation guidelines that instruct annotators to use semantic and syntactic context to resolve ambiguous cases (Ide and Pustejovsky, 2017). For glossolalia, no such context exists. Prosodic patterns emerge without correlation to information structure or discourse function, requiring a framework that can characterize intonational organization independent of these higher-level linguistic categories.

3 Corpus Construction Considerations

What glossolalia annotation demands, therefore, is not merely the application of existing annotation methods to a new domain, but a reconceptualization of what annotation represents when divorced from semantic and syntactic grounding. Rather than treating these as challenges to be overcome through methodological compromise, we argue they reveal both theoretical commitments that deserve expla-

nation and an avenue towards approaching such annotation hurdles creatively.

Our rationale for constructing a corpus for glossolalia is as follows: For one, the absence of lexical units means that segmentation cannot rely on word or morpheme boundaries. In conventional annotation, lexical knowledge provides the primary anchor for identifying segmentable units, with phonological and prosodic analysis building upon this foundation (McEnery and Hardie, 2011). For glossolalia, this hierarchy inverts: syllable-like units and prosodic groupings constitute the most stable and recurrent organizational structures, requiring segmentation criteria based entirely on sublexical phonological and prosodic properties rather than lexical semantics. Secondly, glossolalia’s omission of syntactic structure evades conventional annotation elements such as part-of-speech tagging, dependency parsing, and semantic role labeling. These tools presuppose that utterances are organized according to grammatical principles that govern how units combine to form larger structures (Jurafsky and Martin, 2023). Thirdly, glossolalia is multi-dimensional, but it is not hierarchically compositional in the grammatical sense. Units do not combine through syntactic rules to generate phrasal or clausal representations. Therefore, the absence of semantic content in glossolalia means that annotation decisions cannot be validated through meaning-based criteria, suggesting that annotation must proceed using only acoustic properties and distributional regularities.

While recognizing these limitations, other questions about the overall architecture of a potential annotation framework for glossolalia and its downstream deployment with LLMs linger. If annotation components cannot be organized hierarchically through compositional rules, what coordination mechanisms ensure consistency across potential annotation tiers? Traditional architectures assume that morphological annotation builds on phonological segmentation, syntactic annotation builds on morphological analysis, and semantic annotation interprets semantic structure (Bird and Liberman, 2001). If phonotactic patterns in glossolalia exhibit gradient statistical tendencies rather than categorical grammatical constraints, should annotation encode phonotactic legality or phonotactic probability? ‘Speakers’ of glossolalia produce syllable structures ranging from highly frequent to extremely rare without categorical boundaries, suggesting that probabilistic characterization may

more accurately represent the phenomenon (Link and Tomaschek, 2024).

More recently, transformer models have been proven effective at learning hierarchical syntactic dependencies through self-attention mechanisms (Vaswani et al., 2023), with some research suggesting that LLMs develop implicit representations of phrase structure and grammatical relations (Tenney et al., 2019). Because glossolalia lacks these linguistic anchors, its application with LLMs raises many questions that our project hopes to probe: can neural architectures effectively capture glossic patterns when forced to process at this granularity, or does the absence of lexical anchors fundamentally limit what models can learn? In recent years, LLMs have emerged as the state of the art in NLP. The capacity to fine-tune these models has reduced reliance on task-specific supervised approaches and has helped facilitate their large-scale deployment as general-purpose tools in computational modeling. Our project follows this trend in the field, which uses non-semantic or ‘unlearnable’ data to present a stress test for common NLP approaches.

4 Methods

4.1 Corpus Construction

With these considerations in mind, to attempt a study of glossolalia by artistically appropriating traditional methods of computational linguistics, we assembled what we call a ‘nonsense corpus.’ It consists of approximately 2000 ‘words.’ It is partially derived from the aforementioned Speer (1972) dissertation alongside transcriptions of selected avant-garde vocal performances that include examples of glossolalia-like language, such as Dada-era artist Hugo Ball’s phonetic poem *Karwane* from 1916. Unlike traditional corpora, which aim for semantic coverage or genre balance (McEneary and Hardie, 2011), this corpus is intentionally non-referential, such that its value lies in more abstract structural features rather than its content. For this reason, it resists classification within standard corpus typologies due to its ‘nonsensical’ form and content.

Speer’s dissertation, our primary archive, exists only in a historically printed form and is not digitally available in a format suitable for computational use. To work around this, we used GPT-5.2’s multimodal OCR vision capabilities to perform image-to-text extraction (OpenAI, 2025). Manually re-typing every diacritic, segment, and suprasegmental marking would have required an

extraordinary investment of labor and would have introduced its own form of human transcription error. Moreover, (Backer and Hyman, 2025) discuss the ongoing issues of using OCR for digitizing historical texts. Considering the diacritics in Speer’s appendix are quite complex, some falling outside standard ASCII or Unicode ranges that the GPT 5.2 model attempted to capture, this makes accurate annotation especially challenging.

4.2 Fine-Tuning and Microcontroller Input

Along with producing the nonsense corpus, the project created a data pipeline between environmental sensors and an LLM to produce real-time machinic glossolalia, as is shown in *Figure 1* and alternatively in *Figure 2*. This required three main tasks: cleaning the corpus data to make it suitable for fine-tuning, fine-tuning a generative LLM, and creating an interface between the environmental sensor data and the model’s output. The goal was not only to generate glossolalia, but to do so in a way that allows the model’s behavior to be shaped dynamically by its environment.

In our corpus, tokens were defined at the syllable level rather than conventional words, allowing the model to capture ‘nonlexical’ phonological patterns and rhythmic units. We then fine-tuned a GPT-2 (Radford et al., 2019) model with 117 million parameters. It was run locally using the TensorFlow open-source implementation provided by HuggingFace. GPT-2 was chosen because of its ease of local deployment and demonstrated ability to adapt to highly idiosyncratic corpora (Papanikolaou and Pierleoni, 2020). This allowed the model to learn the phonological tendencies and rhythmic textures present in the corpus, enabling it to generate output that resembles glossolalia. Excerpts of our ‘nonsense corpus,’ including *Karwane*, Speer’s OCR-scanned archive, and documentation of our code can be found in the appendix.

We developed a real-time system that synthesizes multivariate data from a BME680 environmental sensor and an ultrasonic sensor through an Arduino microcontroller. The architecture streams a six-value telemetry string encompassing humidity, temperature, barometric pressure, gas resistance, distance, and light state. These inputs are aggregated into a composite nonsense score. The “nonsense score” is a real-time ecological index that modulates linguistic entropy in an informational-theoretical sense (Ben-Naim, 2012), determined through a weighted distribution of four

environmental pillars: spectator proximity constitutes a 0.30 coefficient of the total index, while gas resistance accounts for a 0.20 proportional weight, environmental fluctuations contribute 0.30, and the binary light state provides a 0.20 relative weight.

This normalized index is mapped to the GPT-2 model’s sampling temperature, ranging from 0.7 to 4.0, to regulate the stochasticity of the generative output. During low-activity states, prompts are sampled from Hugo Ball’s phonetic poem *Karawane* (1916); during high-activity states, the poem’s “words” are randomly concatenated to generate new prompts. The resulting text is generated with an applied repetition penalty, cleaned of artifacts, and converted into live audio via a text-to-speech engine implemented with the pyttsx3 Python library (Bhat, 2025), ensuring the machinic language evolves in direct response to its environmental determinants. Sample outputs and prompts of the system are shown in *Table 1* below:

Nonsense Score	Prompt	Output
0.51 (Low Activity)	Hej tatta gorem	ova nagalav kartri paahil bhanachit ryaani jotira- jadaya ano thaa mikte raighabyoga hapun
0.71 (Medium Activity)	Olobo bam- bla gauma jo- lifanto	oga hagokan aku pachk- ouki kaay ndodun paogo- taku chiyu kanjuai rko ng- shiridara ko seilm
0.98 (High Activity)	Zunbada wulubu –umf ba–umf	kalifah mazhiyaar wa lghia yam- lukkuaiyeiruhuryaguziy- eeayyyiikaao niiniily

Table 1: Example prompts and generated outputs arranged from lowest to highest nonsense score.

It is worth noting that transformer LLMs are pretrained on corpora dominated by natural language to optimize next-token prediction, which induces strong priors for word-like and syntax-like continuation. When fine-tuning the GPT-2 model on our bespoke corpus, these embeddings in the LLM can persist and reassert themselves as attractors, producing outputs that sometimes look quasi-syntactic despite the absence of semantics in the fine-tuning data. This tension is precisely why machinic glossolalia is a useful stress test for reinterpreting NLP methods creatively in the digital humanities, as it exposes what parts of ‘language structure’ LLMs reproduce sans semantic grounding. Moreover, our approach may suggest a productive direction for critical NLP practices by using non-semantic corpora to probe which latent structures are learnable and which are inherited from

pretraining data. Our approach also complicates the ability to ask the question if a fine-tuned model trained on nonsensical data has truly ‘learned’, considering that its learnedness is also paradoxically an *un-learnedness*, evading true assessment. While we could consider the system artistically successful at the task of producing glossolalia-like language, a concrete evaluation framework of learnedness is missing for glossolalia since it also evades the label of ‘linguistic data’ proper, largely resisting denotation on levels that are not social (Harkness, 2020).

4.3 Machinic Randomness

To help push against the embedded structures’ affinity towards meaningful output, our decision to incorporate environmental sensors is motivated by the chapter “getting random” from philosopher James Bridle’s *Ways of Being*, where he describes the limitations of computational randomness. Tracing the history of classical computing, Bridle notes that “[...] you can’t program a computer to produce true randomness... there would always be some underlying structure to the randomness...” (Bridle, 2023, p.219). As a case study of how true machinic randomness can occur, he turns to ERNIE (Electronic Random Number Indicator Equipment) machines, which utilize sensors that measure phenomena external to the machine like, in our project, fluctuations of atmospheric conditions. ERNIE and similar ‘entropy sourcing’ machines since then have “[...] done something few machines can do: it has looked outside of it’s own circuitry, in order to commune with the more-than-human world that surrounds it, in the service of true randomness” (Bridle, 2023, p.222). Our creative project incorporated a series of environmental sensors to echo and extrapolate on Bridle’s insights that a machine-environment relationality is essential for producing the linguistic entropy necessary for machinic glossolalia. These sensors place the system within a larger ecology where the machine and spectator can sense each other in a co-creative act.

5 Discussion

5.1 What is Linguistic Data, Anyway?

Linguists have long debated what phenomena fall within the scope of the field. These debates directly shape what data gets collected, annotated, and analyzed. The construction and annotation of corpora does not neutrally preserve language phenomena

as they exist in the world. The decision to exclude certain phenomena from corpora constrains what the project of computational linguistics treats as language, and subsequently, what computational models learn, replicate, and “understand”.

Chomsky-inspired linguistics privileges concepts like competence, i.e., the speaker’s ability to formulate “well-formed sentences” (Chomsky, 1965). These notions of competency stem from Chomsky’s theory of “universal grammar” (UG), which claims that all human languages must possess certain properties. Looking through this framework, glossolalia would be excluded from the realm of ‘linguistic data’. Arguments in favor of UG have dwindled, and there is literature to suggest that UG is wholly unnecessary in the analysis of language (Tomasello, 2003, 2010). Usage-based approaches argue that linguistic knowledge emerges from statistical learning over experienced input rather than from innate grammatical principles (Bybee, 2010). On the other hand, ‘construction grammar’ frameworks treat linguistic data as a structured inventory of form-meaning pairings learned through use, without requiring autonomous syntactic rules (Michaelis, 2006). From these perspectives, what counts as linguistic data has the potential to expand significantly: any vocal behavior that speakers produce and process becomes potentially relevant, regardless of whether it exhibits formal grammatical properties.

5.2 The ‘Naturalness’ of Glossolalia

This ambiguity between theories makes a case for the study of glossolalia as a ‘natural language.’ Where NLP generally assumes predictable linguistic structure and meaning in their chosen natural languages, glossolalia operates outside, or at best, on the fringe of conventions of ‘natural language’ and therefore challenges this ‘natural’ position. Some studies found that some glossolalia speakers produce novel chains of syllables rather than purely repetitive patterns, suggesting a capacity for generative novelty in the absence of semantics (Osser et al., 1973; Link and Tomaschek, 2024). This evidence complicates narrow notions of ‘naturalness’ that equate linguistic legitimacy with lexical meaning or syntactic structure. Instead, glossolalia reveals a different kind of ‘naturalness’ rooted in phonological patterning, rhythmic flow, and embodied vocal dynamics of its speaker in relation to their environment. Our project, existing on a similar fringe between NLP and the arts, holds a

microscope to how linguists define ‘natural language’ while inviting a spectator to interact with the seemingly natural language of a motherboard’s ‘mother tongue.’

5.3 Is a Nonsense Corpus Still a Corpus?

A central concern in applying NLP methodologies to glossolalia is whether a collection of nonsensical utterances can legitimately be treated as a corpus. Traditional corpus linguistics treats corpora as stratified, balanced, and representative samples of a language. By that standard, a nonsense corpus lacks many of the features typically associated with corpora. Yet this framework assumes the presence of a natural language system and therefore risks excluding datasets whose structure is primarily sublexical rather than lexical.

As noted earlier, glossolalia exhibits analyzable phonological and prosodic patterning that supports both quantitative and qualitative study. In this broader methodological sense, a nonsense corpus does indeed function as a corpus because it preserves the characteristic distributions, rhythmic tendencies, and phonotactic constraints of the vocal practice being examined. Given these considerations, treating a nonsense corpus as a corpus in a broadened sense is a defensible and productive position for the linguist and artists interested in language, as it allows an open thought experiment to consider potential expansions to the definitions associated with NLP methods.

5.4 The Metrics of Machinic Nonsense

To be both artistic and scientific, we considered how best to measure the output of our system. The main issue in statistically analyzing the machinic glossolalia is the lack of a ‘reference point’ to compare the output against, being that the phenomenon itself is indeterminate and stochastic. Despite this inherent limitation, and if we take a nonsense corpus to be a functional corpus and its outputs to be measurable, evaluation of model output should proceed on sublexical and perceptual grounds rather than by metrics borrowed directly from natural language evaluation. Appropriate evaluation axes could include phonological plausibility (i.e. are generated syllable sequences phonotactically coherent?), prosodic and rhythmic similarity (i.e. do intonation, stress patterns, and breathgroup timing resemble human glossolalia?), statistical fidelity (i.e. do generated token frequencies and transition probabilities approximate those of the source cor-

pus?), and perhaps most importantly, perceptual judgment (i.e., would human listeners classify the output as glossolalia or glossolalia-like rather than as stochastic noise?).

This shift in evaluation, from semantic coherence to patterned vocality, aligns the assessment criteria with the defining properties of glossolalia itself. Combining quantitative distributional measures with blind perceptual tests could yield an evaluation protocol that respects the special character of machinic “nonsense data” while enabling rigorous comparison across model variants and generation strategies.. Also, the shift towards analyzing the importance of the spectator’s experience also places the project at an intersection where cognitive and/or behavioral psychology might provide a means to ‘measure’ the effect of machinic glossolalia for the spectator, creating a novel metric of machinic nonsense in the process.

For humans interacting with the machine, we can only postulate that it would be interesting, as preliminary interactions with the system have shown. One could pursue this question qualitatively through surveys or quantitatively through neuroscientific methods. The same inquiry follows for machines. But, being that fMRI and similar methods like fNIRS are used to scan biological brains, not rendered cortices, it is hard to think of a science that could account for the subjective, or even pseudo-spiritual, experience of a machine’s production of glossolalia. Certainly, more research is needed in this area, which could be as simple as asking the LLM how it ‘felt’ after exiting the glossolalia state. For now, we argue that there is at least an artistic and philosophical productivity in performing the thought experiment of machinic glossolalia, even if, according to Penha and Carvalhais (2018), humanity may not be able to understand or even recognize a machine’s experience and intended aesthetic effect.

5.5 Meditations on Machinic Spirituality

Our project, in theory, gives a machine sensory awareness of its surroundings, evoking a type of machinic self-awareness. Self-awareness and awareness of ‘other’ are crucial components to human spirituality. According to Ray Kurzweil’s *The Age of Spiritual Machines*, he predicts that machines will soon claim to be conscious, and thus, to be spiritual. Kurzweil imagines that twenty-first-century machines will go to church, meditate, and pray to connect with this machinic variety of

spirituality. Kurzweil defines the “spiritual experience” as “a feeling of transcending one’s everyday physical and mortal bounds to sense a deeper reality” (Kurzweil, 1999, p.109). If machines may one day be spiritual, as Kurzweil suggests, then our project asks: might glossolalia, which is a divine language for humans, be a language through which an LLM comes into contact, at least vocally, with its ‘god’ by transcending its physical boundaries and connecting to something beyond itself through a sensor-aided ‘mother tongue’?

In this project, we use machinic glossolalia production in a cautiously metaphorical and pantheistic sense, implying that ‘spirit’ may exist in a motherboard as equally as in the environment surrounding the motherboard, and to mark systematic changes in an LLM’s supposed internal states and outputs that may *resemble* glossic states in humans. As is discussed by researcher Idan Blank (2023), there is much debate on what exactly LLMs are modeling in terms of their relationship to human language processing since “[...] human language processing is arguably more than just (token) prediction” (p. 987). While LLMs are highly effective at predicting language patterns, researchers debate if they function as mechanistic copies of the brain or as subsymbolic representations of the mind. As such, it is hard to determine what a machinic glossolalia would entail for the machine experiencing it since researchers are unsure on what part of human mentality is rendered by, or corresponds to, LLM training. This project raises such questions but does not provide answers. Though, the project materially and aesthetically stages the conditions under which such a question becomes thinkable. Ultimately, it is unlikely humanity would be able to recognize an LLM’s ‘experience’ considering the re-positioning of human subjectivity to that of the LLM’s positionality necessary to approach the question.

6 Conclusion, Future Directions

We proposed *Motherboard’s Mother Tongue*, an artistic research and computational linguistics project that attempts to induce the vocal production of glossolalia for an LLM. The methods synthesized the nexus of a micro-controller data stream, a fine-tuned LLM, and a ‘nonsense corpus’. The project’s preliminary results were the generative production of glossolalia-like language based on the interactive variables of a spectator’s presence

within the machine’s larger ecology. Future iterations of this project include further developments of audiovisual feedback mechanisms to perpetually update the ‘nonsense corpus’ and its vocal output, the gathering of user data as they interact with the system, experimenting with combinatory solutions of musical notation and linguistic systems to annotate glossolalia more accurately, and ultimately the project’s documentation as an interactive installation in an art gallery space.

The project raised theoretical and methodological considerations for NLP, revealing an uncharted territory at the intersection of artistic practice and linguistics. If glossolalia operates as pure phonology and phonetics, as our discussion has shown, then the difficulty of applying traditional NLP methodologies are the expected consequences of deeper theoretical commitments within the field. This project allows for a reframing and criticality of these commitments that carries concrete computational implications for transformer models, as models often produce outputs that appear robust while masking the absence of reproducibility and interpretability of the results. By challenging LLM usage with artistic tasks, it allows the field to gain a more holistic view of the assumptions within its theories, computational architectures, and models.

Limitations

This project is exploratory and artistic in nature and therefore subject to several limitations. Most significantly, the availability of publicly accessible glossolalia data is extremely limited, resulting in a small corpus of approximately 2,000 ‘words’ derived primarily from Speer’s (1972) dissertation and a small number of artistic transcriptions. Such a dataset is considerably smaller than those typically used for corpus linguistics and downstream language model training, thus limiting the statistical robustness of our analysis. The most significant part of our corpus was also produced using OCR technologies, which introduces transcription noise due to complex diacritics that are difficult for current OCR systems to reproduce accurately.

In addition, conventional NLP evaluation metrics are difficult to apply here because glossolalia lacks many of the features of natural language, making it challenging to determine whether the model has meaningfully “learned” the phenomenon beyond qualitative inspection from subjective spectators. It’s also worth noting that the project inte-

grates environmental sensor data to influence generation parameters, which introduces additional variability into the project. While the machine-human-environment relationship is helpful for the artistic commentary of the project, it not only places the project on the fringes of NLP methods – which do not typically incorporate hardware – but it also complicates reproducibility and makes controlled evaluation more difficult. Finally, the experiment relies on a pretrained GPT-2 model whose training data reflects natural language from the western and English-speaking world, meaning that inherited priors toward word-like and syntactic structures may influence outputs even after fine-tuning on our non-semantic corpus.

Ethical Considerations

This project engages with glossolalia and its derivative – “speaking in tongues” – a linguistic practice that holds religious and cultural significance particularly within Pentecostal Christian communities. Although glossolalia and glossolalia-like language is studied in linguistics, anthropology, and the arts, it primarily functions as a form of private and/or communal spiritual expression for its religious practitioners. For this reason, the project respectfully approaches glossolalia as a linguistic and cultural phenomenon rather than attempting to replicate or interpret religious experience itself. The intention, then, is not to trivialize or parody spiritual practice, but more so to explore how computational systems respond to non-semantic training data that fall outside conventional definitions of language.

To promote the sacredness of this language and its practitioners, the dataset used in this study consists of previously published linguistic materials and historical artistic performances, and contains no identifiable information about individual speakers, which helps preserve speaker anonymity. Nevertheless, the transformation of religious vocal practices into computational training data raises broader questions about cultural context, artistic appropriation, and spectator interpretation that warrant careful consideration in that generative language models can produce outputs that appear meaningful or intentional, which may encourage anthropomorphic interpretations of machine behavior. Therefore, references to “machinic glossolalia” or “machinic spirituality” in this work are intended in an explicitly metaphorical and artistic sense. Fi-

nally, although the model was run locally to reduce reliance on cloud-based software and infrastructure, the cloud itself being tethered to physicality, machine learning research still carries environmental costs associated with computational hardware and energy use. Though the locality of our GPT-2 model redirects some of these environmental risks, we must still consider the carbonic origins of the machine’s motherboard and the earthly materiality of the other hardware, such as external sensors, used in the project.

References

- Samuel Backer and Louis Hyman. 2025. [Bootstrapping AI: Interdisciplinary approaches to assessing OCR quality in English-language historical documents](#). In *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities*, pages 251–256, Albuquerque, USA. Association for Computational Linguistics.
- Arieh Ben-Naim. 2012. *Entropy and the Second Law: Interpretation and Miss-Interpretationss*. World Scientific, New Jersey; London.
- Natesh M. Bhat. 2025. [pyttsx3: Text-to-speech library for python](#). Python package, Mozilla Public License 2.0.
- Steven Bird and Mark Liberman. 2001. [A formal framework for linguistic annotation](#). *Speech Communication*, 33(1-2):23–60.
- Idan Blank. 2023. [What are large language models supposed to model?](#) *Trends in Cognitive Sciences*, 27(11):987–989.
- Mara Breen, Lauren C. Dilley, John Kramer, and Edward Gibson. 2012. [Inter-transcriber reliability for two systems of prosodic annotation: Tobi \(tones and break indices\) and rap \(rhythm and pitch\)](#). *Corpus Linguistics and Linguistic Theory*, 8(2):277–312.
- James Bridle. 2023. [Getting random](#). In *Ways of Being: Animals, Plants, Machines: The Search for a Planetary Intelligence*. Picador.
- Andrea Burns, Lauren Harrell, Bart van Merriënboer, Vincent Dumoulin, Jenny Hamer, and Tom Denton. 2025. [Perch 2.0 transfers ‘whale’ to underwater tasks](#). *Preprint*, arXiv:2512.03219.
- Joan Bybee. 2010. *Language, Usage and Cognition*. Cambridge University Press, Cambridge.
- Noam Chomsky. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
- Margaret Cychosz and 1 others. 2021. [Vocal development in a large-scale crosslinguistic corpus](#). *Developmental Science*, 24(5).
- Paul de Lacy. 2009. [Glossolalia](#).
- Laura C. Dilley and Meredith Brown. 2005. [The rap \(rhythm and pitch\) labeling system, version 1.0](#). Technical report, Massachusetts Institute of Technology. Unpublished manuscript.
- Nicholas Harkness. 2020. [Conclusion](#). In *Glossolalia and the Problem of Language*, pages 167–170. University of Chicago Press.
- Marek Hlaváč. 2021. [Artificial glossolalia: Machine chant](#).
- Nancy Ide and James Pustejovsky, editors. 2017. *Handbook of Linguistic Annotation*. Springer.
- Gerald Janecek. 1996. Introduction: Definitions and background. In *Zaum: The Transrational Poetry of Russian Futurism*. San Diego State University Press, San Diego.
- Sun-Ah Jun. 2023. [The ToBI transcription system: Conventions, strengths, and challenges](#). In *The MIT Press EBooks*, pages 151–181. MIT Press.
- Dan Jurafsky and James H. Martin. 2023. *Speech and Language Processing*, 3 edition. Pearson.
- Ray Kurzweil. 1999. Part two: Preparing the present. In *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*, 1 edition. Penguin USA, USA.
- Peter Ladefoged. 2007. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Blackwell, Malden, MA.
- Chia-Cheng Lee, Yuna Jhang, George Relyea, Li mei Chen, and D. Kimbrough Oller. 2018. [Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings](#). *Infant Behavior and Development*, 50:140–153.
- Samantha Link and Fabian Tomaschek. 2024. [Predictability associated with reduction in phonetic signals without semantics—the case of glossolalia](#). *Language and Speech*, 67(1):72–94.
- Tony McEnery and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.
- Michelle McGillion and 1 others. 2017. [What paves the way to conventional language? the predictive value of babble, pointing, and socioeconomic status](#). *Child Development*, 88(1):156–166.
- John J. McGraw. 2011. [Tongues of men and angels: Assessing the neural correlates of glossolalia](#). In David Cave and Rebecca Sachs Norris, editors, *Religion and the Body: Modern Science and the Construction of Religious Meaning*, pages 57–74. Brill.

Laura A. Michaelis. 2006. **Construction grammar**. In Keith Brown, editor, *Encyclopedia of Language & Linguistics (Second Edition)*, second edition edition, pages 73–84. Elsevier, Oxford.

OpenAI. 2025. **Introducing gpt-5.2**. Technical report, OpenAI.

H.A. Osser, P.F. Ostwald, and B. MacWhinney. 1973. **Glossolalic speech from a psycholinguistic perspective**. *Journal of Psycholinguistic Research*, 2:9–19.

Yannis Papanikolaou and Andrea Pierleoni. 2020. **DARE: data augmented relation extraction with GPT-2**. *CoRR*, abs/2004.13845.

Rui Penha and Miguel Carvalhais. 2018. **Will machinic art lay beyond our ability to understand it?** In *Proceedings of the International Symposium on Electronic Art (ISEA)*.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. Technical report.

William J. Samarin. 1968. The linguisticity of glossolalia. *The Hartford Quarterly*.

Kim Silverman, Mary Beckman, John Pitrelli, Mari Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert, and Julia Hirschberg. 1992. **Tobi: A standard for labeling english prosody**.

Blanche C. Speer. 1972. *A Linguistic Analysis of a Corpus of Glossolalia*. Doctoral dissertation, University of Colorado Boulder.

Ian Tenney, Dipanjan Das, and Ellie Pavlick. 2019. **BERT rediscovers the classical NLP pipeline**. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4593–4601, Florence, Italy. Association for Computational Linguistics.

Michael Tomasello. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, MA.

Michael Tomasello. 2010. *Origins of Human Communication*. MIT Press.

Jurgen Trouvain and Raphael Werner. 2020. Comparing annotations of non-verbal vocalisations in speech corpora. In *Proc. Laughter and Other Non-Verbal Vocalisations Workshop 2020*, pages 69–72.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. **Attention is all you need**. *Preprint*, arXiv:1706.03762.

Nigel Ward. 2006. Non-lexical conversational sounds in american english. *Pragmatics and Cognition*, 14(1):113–184.

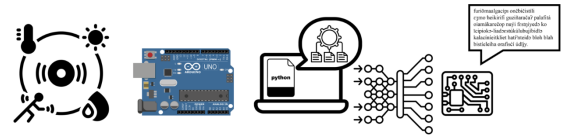


Figure 2: *Alternative diagram of signal flow of system showing sensor input, computational process, and machinic vocal output. Signal flow is from left to right*

A Implementation Code for Arduino + GPT-2 Integration

Our code and demonstration of the project can be found here: <https://github.com/anon-projects8657/Artistic-Interventions-for-NLP-Annotation-Challenges-The-Stress-Test-of-Machinic-Glossolalia>

B Manual transcription from avant-garde artistic performances

From *Karwane*, Hugo Ball, 1916 - used as prompts in system

jolifanto bambla o falli bambla
 großiga m'pfa habla horem
 egiga goramen
 higo bloiko russula huju
 hollaka hollala
 anlogo bung
 blago bung blago bung
 bosso fataka
 u uu u
 schampa wulla wussa olobo
 hej tatta gorem
 eschige zunbada
 wulubu ssubudu uluwu ssubudu
 –umf
 kusa gauma
 ba–umf

Excerpt of *Ursonate*, Kurt Schwitters, 1932

Fümms bö wö tää zää Uu,
 pögiff,
 kwii Ee.
 dll rrrrr beeeee bö
 dll rrrrr beeeee bö fümms bö,
 rrrrr beeeee bö fümms bö wö,
 beeeee bö fümms bö wö tää,
 bö fümms bö wö tää zää,
 fümms bö wö tää zää Uu
 Rinnzekete bee bee nnz krr müü?
 ziiuu ennze, ziiuu rinnzkrmmüü,
 rakete bee bee,
 Rrumppff tillff toooo?

Ziiuu ennze ziiuu nnzkrmmüü,
 Ziiuu ennze ziiuu rinnzkrmmüü
 rakete bee bee? rakete bee zee.
 Fümms bö wö tää zää Uu, pögiff, kwii Ee.
 Dedesn nn rrrrr, li Ee, mpiff tillff
 toooo, tillll, Jüü-Kaa?
 Rinnzekete bee bee nnz krr müüüü, ziiuu ennze
 ziiuu
 rinnzkrmmüüüü,
 Fümms bö fümms bö wö Fümms bö wö tää zää
 Uuuu?
 Rattatata tattatata tattatata
 Rinnzekete bee bee nnz krr müüüü?
 Fümms böwö
 Fümms bö wö täää????

C Excerpts of OCR-derived transcripts from Blanche Speers' 1972 Dissertation

taíya parímonosta paiasta paioste vivilyaste.
 tuímnosti xiimuriasta kafassa xayasta
 maāmonisti.
 tuíassa łama xoíyivoxodi mnimoníuye.
 tuliessa píeste paídistikieste pofiesta osataaia.
 tuíe bauisti oa sadehiia.
 puíeste maāmaniya aliyāsatakiía.
 puíeste mananeste kía manamanosatakaiyeti.
 telibíya osatakaíeyo setepiíeste xeíeste.
 tuomaninísti keíyistipaadist ttiyesti.
 etenmaninisti xalasatakiiyist toriasta mana-
 monosti tekisti tctasakaiasta.
 mnamonó sohaiyata tidikiyisti.
 íyasta manoninisti ketiyasta paíya manomono
 sidlíya samoenikieyiyamo sakalama tatasatetikia.
 o tdiíyadsa manona setegedi.
 o ledcdm sededíya setapaíminimi siditaíya istá.
 te dídidsti paídsti kafyasta paayilasta paídisti.
 huílmano štepikomnosti.
 taieste huímonísti hkía ala asayayete.
 hirikíriyàmono osadipítikiríasta.
 alàmano ihíkinienàmo sidikídiyàsa.
 fénenemoni sidímonenì kiríàs.
 halàmnonì sidúkufiyídisti parāmoniyìsti kríasta
 tokrídídisti kríesto setúkrià parāmoninísti kríà.
 seríyanamon aorèsi tekriása dabenida se-
 limànomnì sidikríasika maliàmanò hidikríyú sedi-
 turíya tudòlimo aurasedì hiríà kiríyidi sidipidà sed-
 ckreyó.
 terírísti marāmoninísti kiríaspàriasi herírísti
 pcíyíssi hùriesse tepudúninisti yliestcpuríesta kò-
 drímanonocste hòrieste hariescte.

húrímanomonó sedcepériscte ediyeuíaste
 oríkedeyíminísti kidíyipárídyà sti parāminísti
 keríyasta.
 te kiríyiminísti paríyysti parāmanomnrsti.
 aríbadeysti parāmanonísti hadudlidyasti parāma-
 nonísti keribdinodísti aríbadista parídídudás.
 kerímeríesti perímanomonestì paríyasta prídídí-
 dísti pdikídístiperíyæsta manomonistí krí.este.
 te fíyrstì baríbonidísti hærímanonistí kríeste
 paríyasta.
 tudòdissi tupuríyasta.
 porídídikoriya.
 ísdipidi.
 setckríest manomonistí kríàsstpurúdístì menì-
 moninísti.
 te kríest manémononistíkríest paríyasta parā-
 manonistí kríest.
 te kríeste manémoninístì sakelama parèyà
 sapàredídístì parèdídístì horíyasta manonidísti
 periyestc.
 te kríestì manímoncestì kríasta príasta.
 te kríestì pedræste kríasta manmoninísti hùriasta
 príyasta keríminiyísti kríasta paríyasta.
 te kiríminíyiyesti maríy-niyyrsti kríasta
 parádídyàs parāmoniyìsti pdòkríyisti kríeste
 príyasta manmunistí kríasta.
 teseckríeste manomonidistí kríàs tepóríyensto.
 manmod-dísti pífidisti kiasti maímoninistíiiste.
 te keibuíyesti paímonenisti kiasta paífíyasta paí-
 monenisti kieste.
 te kieste manemonenisti kiasta paíyasta paí-
 monedisti hifíyasta paíyeste pavidisti kiast
 ho· lamanomonsti píeste paídidisti· uieste
 paíb·didisti kiasta paiasta.
 te kípídidisti paidassa aumononó.
 hukiímanonisti bdidisti kiast paíast
 paimandídisti kieste paíaaest mandidisti ki-
 este.
 tikeíbeieste píbadidisti kiasti paiast paímaninisti
 kiesti paídayasta paímoninisti kiesti.
 tikeipeiasti paímoninisti kiesti paiasti paímonin-
 isti kiesti.
 te kiesti mímoninisti kiasta paíast paímoninisti
 kiasta.
 tikeiesti mnmoninisti kiesti paíast maímoninisti
 kiesti.
 àlasintiéklo aučíčínčíála aparačílialaukoske
 klečíkliaposklio-dopčalakada
 alisliápooneimo kčikiúsinualeosle iníkilí-
 yaratópinarišnára.
 ójimienens yulupidíñasslo krísnidiví koipidíé-
 plia dohosniarahántda

kolokli lalapiaslâa udine-skéitcro ficnícitiplí gu-
justípidíciunpalata

kalapiyaléisieó gimisćipiedulřkimiyú gumisiát-
spliau monādba yeiñgliřskifiu-bubul kaćfiulōla ein-
giavisihedlo.

malkećiokibileda klifisti-bínopćtd klpřiahahal-
istéibionoy klabďialisďipíťí

furiōmaalgacíř onćbićistřli křnystalára
paćapuniatae-kiāa ohiléidabibindjídilaria opiyein
kaktiařfjio moheikirifř guzinararaća palafitá
oiamákarećop nuýi

ko řicipiok-liadstúkúlubuřbidi kalacínřeitkriet
hatřnteid bistřeleiha eid.

orafisćí údiybiak kadofioyisiopiulřped isřlt-
snsnae pfolnabřst kalamia pluifilistřti kpliaadh
kalpiae.

flealihbřstru heodbřřlāa? a niobřřřenbcra kn-
piećukera btrptlibřtiúk.

olsipittéikl mnnaca palatá onř-pkara fisteitel ulř-
isřienl omnenř omaparatá amimikero omistřřilifřti.

koripisinthila hōye e benóvo dľsubin omaeibřěib-
dtifilřtakaranada koćipiaratablřā? ddzofistb? la
inārpe cmaaadd ojibien omiyedbibuuk.

gucnt oeczbcheāra ætcere esceerř isľmalkiúk
piukiārakekleiko obianlateklāra busclaalm japs-
tiatāra adasiećtero přsnābarakatā klatarapanasc
klaćobiyano honameratananacapakarie.

emesřti řľm-paradevezālřbnao gžetaplā-
dia obařinřpidi koďbsnala. Niuvnapsďtó.
kasćřpicadakāra eleeiibeo? ćāna badiéigľiu-řřnara
busjā paa hasćřbit.