Composing with Constraints: AI-Assisted Music Composition in Toki Pona and 31-EDO Tuning

Miika Hyytiäinen
Visiting Researcher
University of Arts, Helsinki
Helsinki, Finland
e-mail: miika.hyytiainen@uniarts.fi

***Abstract*—This paper documents and analyses the AI-aided compositional process behind *kalama musi lili*, a vocal work for soprano and fixed-media audio composed using two radical constraints: text written in Toki Pona, a constructed language containing approximately 123 words, and a tuning system of 31 equal divisions of the octave (31-EDO or equal division of the octave). Both constraints were chosen for their capacity to generate productive friction with AI-assisted tools, since neither Toki Pona nor 31-EDO is present in the training data of large language models or music generation systems. The paper describes a multi-stage compositional process in which Artificial Intelligence (AI) tools (Perplexity, Claude, Tempolor, Mubert, Soundverse, Loudly, and Beatoven) were employed at different junctures, alternating with human intuitive decisions. The findings suggest that AI tools consistently gravitate toward aesthetic conservatism, producing pleasant but artistically unambitious results, and that their creative utility depends on the degree to which the composer actively exploits their limitations, surprising outputs, and chaotic juxtapositions rather than accepting their default suggestions. The paper contributes to the growing discourse on human-AI collaboration in music creation, with particular attention to how constraints in language and tuning shape and disrupt the AI-generated material.**

Keywords-sound and music computing; music; music retrieval and generation; constraints; language translation

I. Introduction

The question of how artificial intelligence may serve as a creator or creative collaborator in music composition has attracted increasing scholarly attention from diverse perspectives [1]–[4]. Nevertheless, the Artificial Intelligence (AI) models as we know them now are limited by their training data [5]. Since the AIs used in this research are mainly trained by freely available musical examples, the results could be described as the conservative aesthetic average. Additionally, for the AIs specialised in creating music, the preference is for the aesthetic pleasing to as many customers as possible. This tends towards conventionally chosen musical parameters: common-practice harmonic languages, standard Western tuning, and texts drawn from old rhythmic patterns with classical subjects.

In this paper, I examine the creative possibilities that emerge when those parameters are deliberately disrupted, and the result becomes the creative starting point. I ask, how can the composer use the friction between constraint and AI's tendencies as a source of new artistic language? In this case study, I am not trying to identify the challenges of the specific, rapidly changing language models, nor to create data of statistical significance on composers using them. Rather, my goal is to use the methodologies of the artistic research to document and analyse one possible creative process to conceptualise and give insight into the composer's practice.

The work under discussion, *kalama musi lili*, is a piece for soprano and fixed-media audio file, henceforth called "the tape". The title of the piece and its text are in Toki Pona, a language that does not use capital letters, so I also use this practice for the title in this article. Toki Pona is a minimalist language constructed by Sonja Lang and published in 2014 in a form containing approximately 123 words [6]. This radical restriction gives rise to a dense web of semantic ambiguity, poetic compression, and unexpected resonance. For example, most words can be used as nouns, verbs and adjectives [7]. As another constraint, *kalama musi lili* employs a 31-EDO (equal division of the octave) tuning system, a perceptually challenging framework that divides the octave into 31 equal steps rather than the conventional 12. Because of the complexity of playing in this tuning, I opted to create a fixed-media stereo file for the performance.

I chose constraints in language and tuning partly for aesthetic reasons and partly as a strategic means of probing the creative limits of contemporary AI systems, since music in Toki Pona and 31-EDO does not occur in training data. The resulting responses are therefore more revealing of the tools' aesthetic biases and, maybe even, their underlying architectures than responses to more familiar prompts would be.

The central argument of this paper is that AI tools, in their current form, are structurally inclined toward aesthetic conservatism: they tend to produce outputs that are pleasant and coherent, and in the process sacrifice the particularity and risk that distinguish artistically ambitious work [8]. Under sufficiently radical constraints, this limitation can be a creative power. This paper thus proposes a model of human-AI collaboration oriented not toward AI as a source of novelty but toward AI as a provider of partial, biased, and aesthetically legible material that the human composer can and must actively transform.

The remainder of this paper is structured as follows: Section II situates the work within existing literature on AI-assisted composition, microtonal music, and constructed languages in artistic practice and describes the compositional methods employed. Section III presents the compositional process, focusing on specific AI outputs and the decisions made in response to them, whereas Section IV describes the results

of this process. Section V discusses the broader implications for human-AI collaboration, including an evaluation of the individual AI tools used. Section V offers conclusions and directions for future research.

II. Related work

Research into AI-assisted music composition has developed along two primary trajectories. The first concerns generative systems that produce musical material autonomously, from early rule-based systems [9]. The second group are the contemporary deep-learning models designed to collaborate with human composers, but still include many aesthetic decisions made before the consumer gets their hands on the program. The present work belongs firmly to the second category, though it uses elements of the first category's AI music-continuation platform.

Within the literature on human-AI collaboration, a recurrent theme is the tension between AI following the given data and AI creating novel combinations. Already, early systems demonstrated that AI could produce stylistically convincing imitations of canonical composers [10]. These results did not catch the attention of the big audience, and also among composers and other professionals, it was mostly seen as an interesting experiment. Although there has been a lot of study in the AI's possibilities and limitations [1]–[4], they are developed at such a speed that researching a specific model at a given time does not always provide very long-lasting results. I also aim to provide the artist's point of view and vocabulary that is special for the artistic research [11] and which is often lost in the discussion of generative AI.

As for the question of constraints in the artistic creation, there are numerous artists and artist groups I could refer to. The reference most close to my heart is the French literary group OuLiPo. Their philosophy is captured beautifully by an OuLiPan author, Paul Fournel [12]:

> The first role of constraint is to act as a stimulus to literary creation. By setting bounds on the imagination, it paradoxically makes writers more aware of the range of their freedom. [...] Secondly, constraints challenge the formal conventions that we have accepted by conscious or unconscious collective submission, or out of the historical inertia. In that role, it is a tool for questioning the relationship between form and meaning.

Previously, I have used similar constraints in my work *Žižek Says!* where I composed anti-acousmatically, a process I also documented by the means of auto-ethnography [13]. The use of language models in specifically the constraints of *kalama musi lili* has only limited references. The AI translation of Toki Pona is not only aesthetic, but computational [14]. Unfortunately, I have not been able to find any literature on AI-aided music creation in 31-EDO.

III. Creating a song cycle in five easy steps

In this section, I describe the creation of *kalama musi lili* first from the point of view of understanding the constraints and then from the practical aspect of the five compositional stages.

A. The constraints

Toki Pona's extreme minimal vocabulary of approximately 123 words must express the full range of human communicative needs through compounding, context, and semantic flexibility, which created productive constraints for text setting. The translation of Toki Pona resists one-to-one lexical equivalence with natural languages, so considerable interpretation is required. AI translation tools, trained primarily on natural-language corpora, tend to over-specify meanings.

31-EDO is a tuning system with a substantial history in theoretical and experimental music. Its advocacy by Adriaan Fokker from the 1940s onward [15], and its implementation in instruments such as the Fokker organ, have inspired some composers interested in extended harmonic resources. It also seems that the online interest towards it is increasing [16]. Nevertheless, learning to play an instrument with 31 keys per octave is not an easy task. Ere Lievonen is a specialist in using the only existing Fokker organ in the world, and even he often opts for a modified 12-EDO midi keyboard. More recently, composers including Easley Blackwood and Marc Sabat have explored its possibilities, and software implementations have become increasingly accessible [17][18]. No standardised notation system for 31-EDO is universally adopted, a circumstance that proved significant in the AI-assisted harmonic generation phase.

B. The creative process

The compositional process is listed in five distinct phases below. At each stage, the composer retained editorial and aesthetic control, using AI outputs not as authoritative suggestions but as raw material to be evaluated, modified, or deliberately misused.

1) I produced a phonetically organised word list by prompting Perplexity to arrange all Toki Pona words in an order of graduated phonetic similarity, and a text composed by the composer exploiting the sonic qualities of Toki Pona phonemes independently of semantic content. Subsequently translated between English and Toki pona.
2) I prompted Perplexity and Claude to generate chords and chord progressions in 31-EDO, requiring the tools to devise their own notation systems for communicating microtonally specified harmonies. Even after many prompts, the results remained conservative, to say the least, and the material can only be used in short fanfare-like structures.
3) I composed a sketch without AI assistance, establishing rhythmic structure, approximate melodic contour, and tentative harmonic gestures.
4) I finished the composition while assembling the backing track in Dorico using these materials, combined with spoken voice and field recordings of rain.
5) I submitted the existing audio to five different AI music-continuation platforms: Tempolor, Mubert, Soundverse, Loudly, and Beatoven [19]–[23]. I layered the resulting

outputs and edited them to create a final movement over which the soprano improvises in live performance.

The roles of composer and AI in these phases are described in Table 1.

TABLE I. STAGES OF THE CREATION

Phase	*Human*	*AI*
1. Text	Creating	Translating, Ordering
2. Chords	Prompting	Creating
3. Writing the score	Creating	-
4. Creating the tape	Creating	-
5. The last song	Organising	Creating

The composition process mentioned in phases 4 and 5 can not be discussed in full detail here. I was heavily inspired by the rhythms and phoneme combinations emerging from the language. In addition, the musical language emphasises the intervals available in this tuning, including natural minor thirds, septimal harmonics, and a range of microtonally inflected consonances and dissonances [15]. Out of the scope of this paper is also the phase that could be seen as most important: *kalama musi lili* is premiered on the 6th of June 2026 by soprano Iida Hirvola in the concert created around computer-aided music at the Arabia Kaikuu festival.

The final composition comprises five songs and runs roughly 15 minutes. The first song features a live vocalese section that overlaps recorded meditative material listing all the words of Toki pona. The third song has a similar list, but the live song sings simple text construction based on onomatopoeia rather than language. The second song is inspired by the Toki pona word "mu" that is used for all animal communication. The fourth song uses the translations of Toki pona most explicitly. It started as some English phrases that I translated into Toki pona, then back into English and back again, with varying results. Here is the text in Toki pona and English.

mejuna nanpa pilin linja alan
Messy threads of feelings.
lukin kiwen
Eyes on stone.
monsi nasin tonsi tempo
Paths through time.i
akesi alasa (apeja) epiku (jasima) kalama kapesi kapeken kiljetesantakalu kipisi (kokosila) kulupu lanpan
(Shameful) hunt, chaotic sound, thieving group (speaking wrong languages).
misikeke monsuta namako palisi pakala pasila pimeja
Medicine for fear, broken simplicity.
sijelo sinpin sitelen soweli utala
Body forward, Night of drawn animals.

The tape of the first four songs has different pitch material played as simple MIDI instruments, overlapping the recorded speech and ambience sounds. The last, fifth song uses five different audio tapes suggested by different AIs, with the soprano improvising on top.

kalama musi lili was premiered 6.6.2026 as part of the Arabia kaikuu festival at the Technical museum, Helsinki. Iida Hirvola sang, and Onni Korhonen was the sound designer. Hirvola has a high, lyric soprano voice. A documentation of the the performance will be provided in the conference.

IV. RESULTS

Before the composition process, I had already decided not to simply compose the text, but rather to let the specific sounds and rhythms of Toki pona inspire and structure the songs. The first significant result of the project concerns the text generation in Toki Pona. When Perplexity was asked to list all Toki Pona words in an order of graduated phonetic similarity, each word resembled the previous one moderately, neither too much nor too little. The output revealed the strengths of AI performance in underrepresented linguistic domains. The resulting sequence had a certain uncanny plausibility at the local level: adjacent words frequently shared phonemes or syllable structures, perceptually audible.

The translation of English text into Toki pona and back usually overspecified the semantic content, making the results less interesting. Toki Pona's word 'moku', for instance, denotes any act or substance related to eating, drinking, or consuming, and any translation that renders it as a specific foodstuff or a particular verb of ingestion imposes a narrowing that the original resists. Perplexity's translations tended toward precisely such specificities, producing English texts that were grammatically fluent but tonally at odds with the ambient generality of the Toki Pona originals. After asking the Perplexity to do more free translations, the results became interestingly poetic.

The second phase included AI-assisted harmonic generation in 31-EDO and produced the most revealing results regarding AI aesthetic bias. I prompted Perplexity and Claude [24][25] to generate seven interesting chords, each containing three to five pitches, that were moderately consonant, along with five-chord progressions built from them. The initial results were excessively consonant: the systems produced chords that functioned essentially as slightly recoloured tonal triads, with almost inaudible microtonal variations. When I requested more dissonant or adventurous results, they did not emerge; instead, they suddenly began to look like hallucinations, as if they were absolutely out of scope for the provided tuning. The language models chose 12-EDO tonal harmonic logic with superficial microtonal labelling.

The notation problem proved equally illuminating. Since 31-EDO notation is not standardised, the AI tools were required to find their own means of communication. Both Perplexity and Claude adopted hybrid approaches combining conventional note names with cent deviations and verbal descriptors such as 'sub-minor-inflected' or 'neutral-harmonic seventh chord'. While these descriptions were often imprecise, they were sufficiently informative to guide me to write the chords in the Dorico notation program. From the AI outputs, I selected a progression characterised by a sub-minor quality and a neutral seventh sonority, deploying them in a fanfare-like gesture that underscored the ceremonial strangeness of the tuning system.

The third and fourth stages, which include composing the score and creating the tape, are not discussed here.

In the fifth phase, I used the audio files I created as references for AIs designed to generate music based on the user's prompt. The results were varying: Loudly, upon receiving the source audio, they used only and constructed over it a somewhat disturbing techno beat. Beatoven produced a heavily textured noise version, dense and disorienting, distant from the source material's clarity. Tempolor and Mubert produced outputs closer in atmosphere to the source, but with a generic, smooth production aesthetic that tended to neutralise the microtonal specificity of the original. Soundverse's output fell between these poles, retaining some melodic contour while adding harmonic material inconsistent with 31-EDO practice. Also, here, the outputs were either violently pushed towards an existing techno aesthetic or were almost completely random.

I layered these five versions with attention to structural articulation, the points at which one version cedes to another being selected for maximum perceptual contrast. The soprano's live improvisation over this layered material provides the human element that mediates among the divergent machine-generated aesthetics. At the beginning of this process, I was genuinely and maybe naively looking for a co-creation process that I had witnessed in co-composing the opera NOS or co-creating a study plan [26][27]. A better description might be composing with something that offers many new ideas but falls within a very limited aesthetic spectrum.

V. Discussion

The most consistent finding across all phases of the compositional process is what might be termed the conservatism gradient of AI creative tools. Finding more experimental results was sometimes possible, but it often yielded nonsensical results, such as words that did not exist in the language or chords that did not exist in that tuning system. At most stages, AI systems produced outputs that prioritised intelligibility, pleasantness, and generic coherence. This tendency is not accidental: it follows from the basic statistical logic of large language models and music generation systems, which are trained to produce outputs that are highly typical given their training distributions. An output that is maximally probable within a training distribution is, almost by definition, an output that resembles many prior examples. The radical constraints of Toki Pona and 31-EDO made this tendency unusually legible. In a conventional compositional context, AI-generated harmonies might pass as merely competent; in 31-EDO, the gap between the tool's output and the actual expressive possibilities of the tuning system was evident. By choosing domains in which AI competence is limited, the composer created conditions in which AI's own limitations became compositionally significant.

The evaluation of individual tools reveals significant differences in their failure modes. Perplexity and Claude, as general-purpose language models, attempted to reason their way through the 31-EDO notation problem, producing solutions that were logically coherent but aesthetically less developed. The music-continuation platforms operated from different architectures and produced more unpredictable results: Loudly's techno beat was clearly a product of genre-classification logic applied to a short audio segment without regard for the broader context; Beatoven.ai's noise output may reflect either a different segment selection or a different generative model that weighted textural density over tonal clarity. Tempolor and Mubert's generic smoothness reflects the production-music aesthetic that dominates their training data, optimised for background music that draws little attention.

The compositional strategy of layering these divergent outputs and selecting structurally productive juxtapositions is consistent with a broader practice in Schaefferian tradition of using collage and superimposition as compositional methods [28]. What is distinctive about the present work is that the collage elements were not pre-selected by the composer but generated by AI systems as a function of their individual limitations and biases. The resulting layer texture is not arbitrary: it reflects the different ways in which five distinct AI systems misread or misappropriated the source material, and the composer's judgment about which misreadings were musically interesting.

This suggests a model of human-AI collaboration that differs from the dominant paradigm of AI as assistant or co-author. In the present work, AI functions less as a collaborator with aesthetic intentions than as a source of material whose interest lies precisely in its systematic divergence from human compositional intent. The composer retains full responsibility for aesthetic evaluation and structural decision-making; the AI provides a variety of materials.

There are methodological limitations to the present study that should be acknowledged. The compositional process is described from the perspective of a single composer and has not been subject to external peer evaluation of its musical outcomes. The AI tools used were accessed in their public versions, without access to system prompts or model weights, and their outputs may not be representative of their full capacities. The Toki Pona and 31-EDO communities may have developed resources and practices relevant to this work, of which the composer was not aware.

VI. Conclusion and Future Work

This paper documents and analyses the compositional process behind *kalama musi lili*, a work that employs radical constraints on AI-assisted composition. The work contributes to the emerging discourse on human-AI creative collaboration by proposing an alternative to the dominant model of AI as an intelligent assistant. In the present framework, AI is a provider of biased, partial, and aesthetically legible material whose interest depends on the composer's capacity to recognise and exploit its characteristic distortions.

Several directions for future research suggest themselves. First, a systematic study of AI harmonic generation in alternative tuning systems would provide a broader basis for understanding the relationship between AI competence and tuning-system complexity. Second, the Toki Pona text-setting problem merits dedicated investigation: the language's semantic

structure raises questions about the relationship between lexical constraint and musical form, rhythm and meaning that extend well beyond the present work. Third, the layering methodology employed in the fifth movement might be formalised as a compositional technique and applied repeatedly, creating chains of machine misreading that the composer subsequently edits. Fourth, future research might address these limitations through comparative studies involving multiple composers, controlled prompt experiments, and community consultation.

kalama musi lili was not composed as a demonstration of AI's creative potential. This song cycle was composed to please my aesthetic style and need to create. The somewhat disruptive use of AI models was a spontaneous experiment, and it could have remained so had I not started writing this paper. It made the process somewhat more structured and allowed me to reflect on my own creative process, possibly also affecting the future compositions. The AI tools used here are not co-composers. They are, rather, powerful and somewhat broken instruments that have revealed their potential under unusual pressure.

REFERENCES

[1] S. Singh and S. Jadhav, "Music composition with ai", *World Journal of Advanced Research and Reviews*, vol. 25, pp. 1031–1037, Mar. 2025. DOI: 10.30574/wjarr.2025.25.3.0723

[2] P. Baxi et al., "Hybrid ai-human music composition for pedagogy", *ShodhKosh: Journal of Visual and Performing Arts*, vol. 6, Dec. 2025. DOI: 10.29121/shodhkosh.v6.i2s.2025.6695

[3] T. Aykut, D. Vera, C. Burgmair, and C. Kuhn, "Evaluating the role of composition and recording style in ai music", Sep. 2025. DOI: 10.1109/ResgenXAI64788.2025.11344059

[4] J. Wen et al., "Tuning immersion and performance with adaptive generative music in vr", *IEEE Transactions on Visualization and Computer Graphics*, vol. PP, pp. 1–11, Apr. 2026. DOI: 10.1109/TVCG.2026.3679907

[5] S. A. Salloum, "The architecture of generative ai and its role in the creative industry", in *Generative AI in Creative Industries*, Springer, 2025, pp. 13–29.

[6] V. Esposito, "Toki pona", *World Literature Today*, vol. 99, no. 1, pp. 12–13, 2025.

[7] M. Blahuš, "Toki pona: Eine minimalistische plansprache/toki pona: A minimalistic constructed language", *Spracherfindung und ihre Ziele. Beiträge der*, vol. 20, pp. 51–56, 2011.

[8] P. Welding, "Oops, i did it again! the humour of incongruity, risk-taking and creativity in art practice and everyday life", *Societies*, vol. 14, no. 3, p. 35, 2024.

[9] L. A. Hiller and L. M. Isaacson, *Experimental Music; Composition with an electronic computer*. Greenwood Publishing Group Inc., 1979.

[10] D. Cope and M. J. Mayer, *Experiments in musical intelligence*. AR editions Madison, WI, 1996, vol. 12.

[11] H. Borgdorff, "The production of knowledge in artistic research", in *The Routledge companion to research in the arts*, Routledge, 2010, pp. 44–63.

[12] P. Fournel, *Dear Reader*. Pushkin Press, 2014, p. 165.

[13] M. Hyytiäinen, "Composing anti-acousmatically: An autoethnography", *Trio*, vol. 14, no. 2, 2025. DOI: 10.37453/tj.159999

[14] D. Huang and H.-A. Joo, "A computational approach to analyzing language change and variation in the constructed language toki pona", *arXiv preprint arXiv:2508.10246*, 2025.

[15] P. Rapoport, "About 31-tone equal temperament", Accessed: May 15, 2026. [Online]. Available: https://www.huygens-fokker.org/docs/rap31.html

[16] "Zheanna erose's youtube channel", Accessed: May 15, 2026. [Online]. Available: https://www.youtube.com/@ZheannaErose

[17] E. Blackwood, "Modes and chord progressions in equal tunings", *Perspectives of New Music*, pp. 166–200, 1991.

[18] M. Sabat and W. Von Schweinitz, "The extended helmholtz-ellis ji pitch notation", *Plainsound Music Edition*, 2004.

[19] "Tempolor web page", Accessed: May 15, 2026. [Online]. Available: https://claude.ai/

[20] "Mubert: Human and ai music generator", Accessed: May 15, 2026. [Online]. Available: https://mubert.com

[21] "Soundverse web page", Accessed: May 15, 2026. [Online]. Available: www.soundverse.ai

[22] "Loudly web page", Accessed: May 15, 2026. [Online]. Available: www.loudly.com

[23] "Beatoven web page", Accessed: May 15, 2026. [Online]. Available: https://www.beatoven.ai/

[24] "Perplexity web page", Accessed: May 15, 2026. [Online]. Available: https://www.perplexity.ai

[25] "Claude web page", Accessed: May 15, 2026. [Online]. Available: https://claude.ai/

[26] M. Hyytiäinen, *Voice Map Method: enhancing composer-singer communication*. Taideyliopiston Sibelius-Akatemia, DocMustohtorikoulu, 2022.

[27] L. Fornhammar and M. Hyytiäinen, "Towards a collaborative higher music education curriculum: A singer-composer duoethnography", *Nordic Research in Music Education*, vol. 6, pp. 46–69, 2025.

[28] J. Dack, "Collage, montage and the composer pierre henry: The real, the concrete, the abstract in sound art and music", *Journal of Music, Technology & Education*, vol. 6, no. 3, pp. 275–284, 2013.