**1 - Making Voices The Gendering of Pitch Correction and The Auto-Tune Effect in Contemporary Pop Music**

https://www.researchgate.net/publication/335667997\_Making\_Voices\_The\_Gendering\_of\_Pitch\_Correction\_and\_The\_Auto-Tune\_Effect\_in\_Contemporary\_Pop\_Music

<https://doi.org/10.1525/jpms.2019.312008>

The paper "Making Voices: The Gendering of Pitch Correction and The Auto-Tune Effect in Contemporary Pop Music" explores the intersection of gender, race, and technology in the use of pitch correction and the Auto-Tune effect in pop music. The study delves into how these technologies influence perceptions of authenticity, emotional delivery, and vocal labor, particularly focusing on gendered and racialized expectations.

**Research Methods:** The paper draws on ethnographic fieldwork conducted between 2014 and 2018 in Los Angeles and New York, engaging with music industry professionals. The author examines how pitch correction technologies like Auto-Tune are applied in studio settings and explores the implications for gendered labor in the creation of pop music voices.

**Key Findings:**

* **Gendered Perceptions:** Female voices, especially in pop music, are subject to pitch correction to maintain a standard of "perfection," whereas male voices using Auto-Tune are often seen as more artistic or emotionally expressive.
* **Technology and Emotion:** The technology of pitch correction affects the perceived emotional authenticity of the voice, with male voices using Auto-Tune considered more emotionally legitimate, while female voices are often framed as less authentic or too artificial.
* **Labor Dynamics:** The paper highlights how pitch correction technology can de-skill female singers while increasing the power of male engineers and producers, reinforcing gender inequalities in the music industry.

**Limitations:**

* The study primarily focuses on the pop music genre and may not fully represent the experiences in other genres or the broader music industry.
* While it provides in-depth analysis of gendered dynamics in vocal production, the paper may underplay the complexities of race and class in the context of music production and the use of Auto-Tune.

This paper contributes to the understanding of how modern music production tools reinforce gender and racial stereotypes and the ways in which these technologies shape public perceptions of vocal authenticity.

**2 - Auto-Tune as instrument: trap music's embrace of a repurposed technology**

https://www.cambridge.org/core/journals/popular-music/article/autotune-as-instrument-trap-musics-embrace-of-a-repurposed-technology/C5B160BE7E8AB29576E92CC7CD9D60B7

<https://doi.org/10.1017/S0261143024000369>

The paper "Auto-Tune as Instrument: Trap Music’s Embrace of a Repurposed Technology" examines how Auto-Tune, originally a pitch-correction tool, has been repurposed in trap music, particularly focusing on the unique use of the technology in this genre. The author argues that Auto-Tune, especially with its "zero-onset" setting (where pitch correction occurs almost instantaneously), functions more as a creative instrument than just a tool for pitch correction, transforming the vocal performance and enhancing emotional expression.

**Research Methods:** The article explores Auto-Tune's role in trap music through an analysis of its technical usage, its evolution, and its impact on vocal timbre and emotional expression. The author draws on examples from various trap artists, like Young Thug and T-Pain, to illustrate how the technology is used creatively in both studio recordings and live performances. The paper also contextualizes Auto-Tune's use within the broader history of repurposed technologies in hip-hop, such as turntables and samplers.

**Key Findings:**

1. **Instrumental Role of Auto-Tune:** In trap, Auto-Tune is not merely for pitch correction but serves as an expressive tool that modifies vocal timbre, allowing for a new form of vocal performance.
2. **Technological Repurposing:** Trap artists, much like earlier hip-hop producers, repurpose technology to create unique soundscapes, with Auto-Tune becoming central to the genre’s aesthetic.
3. **Emotional Expression:** The technology facilitates emotive vocal expressions, often blurring the line between human and machine, which challenges traditional notions of vocal authenticity.

**Limitations:**

1. The study primarily focuses on Auto-Tune’s role in trap music, which may limit its applicability to other genres or contexts within popular music.
2. It may also oversimplify the varying ways in which different artists use Auto-Tune, as the tool’s impact can differ significantly depending on the artist's intention and the production process.

Overall, the paper emphasizes how Auto-Tune in trap music has evolved beyond its initial purpose, becoming a fundamental part of the genre's sonic identity and enabling new forms of creative expression.

**3 - A System for Tuning Instruments Using Recorded Music Instead of Theory-Based Frequency Presets**

https://direct.mit.edu/comj/article-abstract/36/3/43/94398/A-System-for-Tuning-Instruments-Using-Recorded

<https://doi.org/10.1162/COMJ_a_00128>

The paper "A System for Tuning Instruments Using Recorded Music Instead of Theory-Based Frequency Presets" introduces a novel approach to tuning musical instruments, specifically designed for traditional and non-western music systems where theory-based presets may not be applicable. The method enables musicians to tune instruments based on a recording of a master musician, rather than relying on standard tuning systems. This system provides both audio and visual feedback to the user during the tuning process.

**Research Methods:** The paper proposes a system consisting of two main components: the analysis part, which extracts stable frames from a provided reference recording (like that of a master musician), and the application part, which uses these frames to help the musician tune their instrument. The system avoids traditional frequency estimation techniques, instead comparing the frequencies of two signals directly using an auto-difference function. The method also includes a unique approach to creating loopable audio segments to provide continuous playback for tuning.

**Key Findings:**

1. **User-Centric Tuning:** The system allows for user-provided reference recordings, making it adaptable to various musical traditions and avoiding the need for predefined tuning frequencies.
2. **Stable Frame Segmentation:** It automatically extracts stable frames from the reference recording, ensuring consistency in pitch during tuning.
3. **Visual and Audio Feedback:** The system provides visual feedback using the auto-difference function to show pitch discrepancies and audio feedback through looped reference frames.
4. **Cultural Sensitivity:** The method supports tuning in musical traditions with microtonal intervals, such as Turkish makam music, where theoretical tuning systems are often not suitable.

**Limitations:**

1. **Complexity for Novice Users:** The manual pruning of stable frames may be challenging for beginner musicians, and the system requires user input, which could be a barrier for non-expert users.
2. **Limited Testing:** While the prototype has been tested with traditional music, it lacks real-time feedback capabilities, which are critical for practical use.
3. **Prototype Limitations:** The system is still in development, and the feedback suggests it is more of a research tool than a practical, widely usable device at this stage.

The system shows potential for improving the tuning process in non-western music traditions and could serve as both a tool for musicians and a research instrument for exploring alternative tuning systems.

---------

4 - Auto-tuning mother nature: Waves in music and water

https://ieeexplore.ieee.org/abstract/document/6404781?casa\_token=KJxJPp\_dMB0AAAAA:J\_IivnaWw\_zeXYJWkkPVm5e5QssU2pyuDC-F\_o5AXUgXB8SpzEf3vBtN5vHByh2\_jHCpOBiO

10.1109/OCEANS.2012.6404781

The paper "Auto-Tuning Mother Nature: Waves in Music and Water" by Hunter C. Brown explores the application of the Fourier transform in two seemingly unrelated fields: music and oceanography.

**Research Content:**

The paper draws a parallel between the use of the Fourier transform in music production, particularly in the Auto-Tune technique, and its application in oceanography to study wave patterns. The author discusses how Auto-Tune, a technology popularized in the late 1990s, uses Fourier transforms to correct pitch in music. Simultaneously, the paper highlights how oceanographers have employed the same mathematical functions for decades to analyze the characteristics of ocean waves.

**Method:**

The paper provides a brief overview of the Fourier transform and its discrete version, the Discrete Fourier Transform (DFT), explaining how both are applied to analyze time-series data in sound and ocean waves. It also discusses the use of window functions to reduce spectral leakage in the Fourier transform process and applies this theory to both music and ocean waves, demonstrating the broad applicability of this mathematical tool across different fields.

**Limitations:**

While the paper provides an interesting cross-disciplinary perspective, it lacks an in-depth exploration of the technical challenges or limitations involved in applying Fourier transforms to both music and oceanography. Additionally, the practical implications or innovations in using these mathematical techniques together are not fully explored. The paper also doesn't address potential real-world complexities like noise and signal distortions in more varied or uncontrolled environments.

5 - Deep Autotuner: A Pitch Correcting Network for Singing Performances

https://ieeexplore.ieee.org/abstract/document/9054308?casa\_token=QYBjuoKZXHcAAAAA:zPQsFktdi1iyWWoBNR84ST-1kWvQ3xLD5YyfJe47w-ofZQh5zjxvp1gKRpt45v6mh6sPDc97

10.1109/ICASSP40776.2020.9054308

The paper "Deep Autotuner: A Pitch Correcting Network for Singing Performances" proposes a data-driven approach for automatic pitch correction of solo singing performances. The key feature of the model is its ability to predict note-wise pitch shifts based on the relationship between the spectrograms of the singing and accompaniment, without relying on a musical score. Unlike traditional pitch correction systems that map vocal notes to predefined pitches, this model treats pitch as a continuous value, allowing for more natural pitch variation, such as vibrato and pitch bending, while correcting unintended pitch deviations.

**Research Method:**

The authors train a deep neural network consisting of convolutional layers followed by gated recurrent units (GRUs) to handle sequential data. The model learns to correct pitch by analyzing the harmonic alignment between the vocals and the backing track. A unique aspect of the training involves synthetically detuning high-quality singing performances and using them to train the model to predict the necessary pitch shifts to restore the original pitch. The system is trained on a dataset of 4,702 karaoke performances with accurate intonation, treating slight deviations as corrections while preserving intentional pitch variations.

**Limitations:**

The model assumes that each vocal performance targets a specific pitch for every note, which may not hold true in more complex or expressive musical styles. Additionally, while the model is effective in correcting pitch shifts within a semitone, it struggles with larger deviations. The subjective listening tests indicate that the model works best when the performance is already reasonably in tune, and its effectiveness diminishes as the pitch deviation increases. Further refinement of the model is needed to handle more diverse singing styles and music genres robustly

6 - The oleaginous voice: Auto-Tune, linear predictive coding, and the security-petroleum complex

10.1109/WASPAA58266.2023.10248127

https://www.tandfonline.com/doi/full/10.1080/07341512.2024.2402580#abstract

https://doi.org/10.1080/07341512.2024.2402580

The paper "The Oleaginous Voice: Auto-Tune, Linear Predictive Coding, and the Security-Petroleum Complex" by Owen Marshall traces the historical and technological intersections between the oil industry, vocal processing, and national security. The paper argues that both geophysical exploration and digital vocal processing technologies, particularly Auto-Tune, emerged from similar signal processing techniques that were originally developed for oil exploration and later adapted for voice technologies.

**Research Content:**

The paper examines how signal processing techniques such as deconvolution filtering, developed for oil exploration, were adapted to speech processing technologies like Linear Predictive Coding (LPC) for secure voice communication. It explores the development of technologies like secure voice transmission systems, the Speak & Spell educational toy, automatic speech recognition, and Auto-Tune, and highlights the role of military and intelligence interests in fostering technological spillovers between the oil and voice industries.

**Method:**

Marshall utilizes a historical and sociotechnical approach, tracing the development of key technologies from the oil industry’s research into digital signal processing to the creation of vocal processing technologies like Auto-Tune. He emphasizes the political and economic relationships, particularly national security interests, that shaped the transfer of knowledge and innovation between the petroleum and voice sectors. The author discusses the role of organizations such as the NSA and the Department of Defense in facilitating these technological advancements, as well as the influence of military-industrial interests in shaping both industries.

**Limitations:**

The paper focuses on the broader historical connections but does not delve deeply into the technical details of the signal processing methods themselves. While it successfully outlines the socio-economic and political forces behind these technologies, it leaves gaps in explaining the specific technical innovations within the field of vocal processing. Additionally, the paper could expand on the contemporary implications of these technologies and how they have evolved in modern contexts beyond their military and oil industry origins.

------------------

7- Research on intelligent interactive music information based on visualization technology

https://www.degruyter.com/document/doi/10.1515/jisys-2022-0016/html

<https://doi.org/10.1515/jisys-2022-0016>

This paper focuses on intelligent interactive music information visualization by pairing music with images using deep learning techniques. The study integrates convolutional neural networks (CNN) and long short-term memory (LSTM) models to extract features from both music and images for effective music-image pairing. To enhance emotional accuracy, an emotion classification loss function was introduced, improving the alignment of images with the emotions conveyed by the music. Experimental results showed that the improved deep learning-based algorithm achieved higher matching accuracy compared to traditional methods and a non-improved version.

**Methodology**:

* **CNN** is used to extract features from images.
* **LSTM** is employed to capture emotional features from music signals.
* The proposed algorithm improves the loss function by incorporating sentiment classification loss to refine the music-image pairing process.

**Limitations**:

* The sentiment labeling in the training set was based on manual voting, which could introduce errors.
* The algorithm's performance depends heavily on the weight of the sentiment classification loss, with the optimal weight found to be 0.2, but further tuning might be needed for different datasets.
* The paper also acknowledges the limitations of subjective emotional interpretation in music-image matching.

8- Music dynamics visualization for music practice and education

https://link.springer.com/article/10.1007/s11042-025-20637-0

<https://doi.org/10.1007/s11042-025-20637-0>

This paper presents the development and evaluation of a Musical Dynamics Visualization Method (MDVM), aimed at enhancing music practice and education. The method provides real-time visualization of musical dynamics, allowing musicians to compare their performance against a pre-recorded track. The MDVM interface overlays the student's dynamic visualization on top of the teacher’s, enabling a detailed and intuitive comparison. The system significantly improved the accuracy of dynamic expression, reducing dynamic errors by approximately 256%, while enhancing emotional engagement and the overall learning experience.

**Methodology**:

* The MDVM visualizes sound amplitude in real-time and overlays the student’s performance with a teacher’s reference.
* The system uses a wave visualization of dynamic changes, allowing real-time comparison and analysis.
* It supports both individual practice and educational environments, enhancing understanding and expression of musical dynamics.

**Limitations**:

* The focus is primarily on dynamics, potentially overlooking other key musical elements like pitch and rhythm, which could also enhance the learning process.
* The precision of dynamic visualization is limited by the system’s general approach to amplitude changes, lacking the ability to replicate dynamic variations with exact precision.
* Future work should integrate other aspects of musical expression and refine the dynamic visualization for higher accuracy and more creative learning opportunities.

9 - Understanding the structure of musical compositions: Is visualization an effective approach?

https://journals.sagepub.com/doi/abs/10.1177/1473871616655468?casa\_token=I77PLiDtiIsAAAAA:0vbtJ2YnVpqbR53qoRndtI\_8CdIPX87rJeSj8ByLsUEGSf0KWaaOTXpxrrWyZEYvNRwMFA9fG\_\_8&casa\_token=XGgy0mBk85cAAAAA:xmoLYZZpGAkjmVTQOTCaUqr\_wo-KbDT781EXoHHByr1VXGU\_1hmx5YcZQwYCl4Z84K-IVcHdh-69

<https://doi.org/10.1177/1473871616655468>

This paper explores the effectiveness of visualization techniques in understanding and composing classical music, specifically in four-part chorale compositions. The study aims to assist individuals with limited music theory knowledge by using visual representations to highlight melodic errors during the compositional process, making it easier to follow classical music rules.

**Methodology**:

* A tool called **VisualMelody** was developed to visualize the melodic structure of music compositions.
* The tool helps users avoid errors such as parallel fifths, octaves, unisons, and voice crossings by providing real-time feedback on the correctness of their compositions.
* An evaluation was conducted with 40 participants, divided into two groups: one using the tool with visualization features and the other using it without. The study measured completion time and error rates.

**Findings**:

* Participants using the visualization tool were 60% faster in completing tasks and made 85% fewer errors than those using the tool without visualization.
* Feedback from participants indicated that the visualization significantly improved the compositional process and the overall learning experience.

**Limitations**:

* The tool focused only on a limited set of classical music rules, mainly related to melodic structure, and further expansion of these rules could enhance the tool’s usability.
* The study was conducted with a relatively small sample size, and further research involving a broader group could validate the tool’s effectiveness across different user demographics. Additionally, aesthetic choices such as colors and shapes could be refined to improve user experience.