

Veränderungen in männlichen und weiblichen Stimmen in den Billboard Top 100 Charts seit 1958

Fabian S. Klinke

2024-08-15

*Forschungsexposé — Computational Music Studies: Music Information Retrieval in der
Musikrezeptionsforschung (TU Berlin, Dr. Steffen Lepa). Stand: 2026-01-12.*

Gibt es einen Trend hin zu mehr Androgynität im Gesang?

Abstract

Die vorliegende Studie untersucht, ob sich das zunehmende Verschwimmen traditioneller Geschlechterrollen in der Gesellschaft auch in der Musik widerspiegelt, insbesondere in den Stimmen weiblicher und männlicher Sänger:innen. Dazu werden die wöchentlichen BB100 von 1958 bis heute analysiert, wobei der Schwerpunkt auf der Extraktion und Analyse von Pitch und Timbre der Gesangsstimmen aus frei zugänglichen 30s Previews liegt. Mittels State-of-the-art Techniken wie Demucs für die Audioseparation und CREPE für die *Pitch*-Analyse sowie der Nutzung von MFCC zur *Timbre*-Analyse werden die Songs untersucht. Eine anschließende Regressionsanalyse sowie K-Means-Clustering ermöglichen es, Trends in der Annäherung der Stimmlagen und Timbre-Cluster der Geschlechter

über den Testzeitraum zu identifizieren. Die Ergebnisse dieser Studie könnten aufzeigen, inwieweit sich die stimmlichen Merkmale von Sänger:innen in den letzten Jahrzehnten verändert haben und wie diese Veränderungen mit gesellschaftlichen Entwicklungen zusammenhängen.

Einleitung

Musik verändert sich ständig. Sie spiegelt unsere Gesellschaft und ihre Werte wider, beeinflusst diese jedoch auch stark. Über die vergangenen Jahrzehnte hat sich die Musikwelt stark gewandelt: Neue Technologien haben die Produktion und den Klang von Musik stark verändert, neue Instrumente und Techniken haben völlig neue Welten eröffnet. Doch nicht nur die Musik selbst hat sich verändert, sondern auch wie sie gehört wird (von Vinyl bis hin zu Streaming-Diensten), aber vor allem auch wer sie macht.

So hat sich beispielsweise der Anteil von Frauen in der Billboard Hot 100 Year-End Chart von 2012 bis 2022, beginnend mit 22% auf 30%, erhöht (Smith et al., 2023). Frauen haben in den letzten 70 Jahren einen bemerkenswerten Anstieg ihrer Präsenz in den Charts verzeichnet. In den 1960er Jahren machten weibliche Künstlerinnen nur etwa 22% der Nummer-1-Hits aus, während dieser Anteil in den 2010er Jahren auf über 46% stieg (Anderson, 2024).

Diese Veränderungen weisen auf einen allgemeinen Wandel in der Gesellschaft im Hinblick auf Gleichstellung und Akzeptanz von Frauen hin und sind somit auch ein erstes Indiz hin zur allgemeinen Veränderung von Geschlechterrollen. Bereits (Twenge, 1997) hat in einer Metaanalyse von BSRI und PAQ Messungen zwischen 1973 und 1994 festgestellt, dass traditionelle Geschlechterstereotypen allmählich an Bedeutung verlieren, insbesondere bei Frauen, die zunehmend Eigenschaften übernahmen, die zuvor als tra-

ditionell männlich galten, während bei Männern zur Zeit der Studie noch kein Effekt gemessen werden konnte. In ihrem Buch “The Unfinished Revolution” vertieft Kathleen Gerson dieses Verständnis und beschreibt, wie sich diese Entwicklung bis in die heutige Zeit fortsetzt. Insbesondere erzählt sie vom Wandel in den Perspektiven der jüngeren Generation in Bezug auf Familie, Arbeit und der immer stärkeren Aufweichung traditioneller Geschlechtergrenzen (Gerson, 2010). Auch die steigende Akzeptanz von LGBTQ+-Communities, besonders von non-binären und trans-gender Identitäten, könnte einen maßgeblichen Einfluss auf das Aufweichen und langsame Verschwinden von Geschlechtergrenzen, sowie deren Auslebungen haben (Adamczyk & Liao, 2019).

Dieser stetige Wandel in unserer Gesellschaft wirft die Frage auf, ob auch das zunehmende Verschwimmen von Geschlechterrollen und Stereotypen und deren Wahrnehmung einen Einfluss auf die Musik hat, vorwiegend, ob es einen Trend hin zu mehr Androgynität im Gesang gibt. Durch eine Analyse der Billboard 100 können unbewusste Veränderungen in der Art und Weise, wie erfolgreiche Interpret:innen ihre Stimme einsetzen, genauer untersucht werden. Hier ermöglichen vor allem Fortschritte im Bereich des Machine Learning und der Rechenleistung eine tiefgehende computergestützte Untersuchung von *Pitch* und *Timbre*. Die vorliegende Arbeit versucht, genau diesen Fragen auf den Grund zu gehen, indem die Veränderungen des durchschnittlichen Pitches und der Timbre-Feature-Vektoren beobachtet werden.

Stand der Forschung

Obwohl es zahlreiche Studien gibt, die einzelne Aspekte der Fragestellung beleuchten und Hinweise auf mögliche Effekte liefern, fehlt bisher eine umfassende Untersuchung, die sich gezielt mit der langfristigen Entwicklung der weiblichen und männlichen stimmlichen

Merkmale in der Popmusik beschäftigt. Insbesondere eine Trendanalyse, die sowohl Pitch als auch Timbre über einen längeren Zeitraum hinweg geschlechtsspezifisch betrachtet, wurde bisher nicht durchgeführt.

Doch zuerst soll der Begriff der Androgynität geklärt werden. (Bem, 1974) definiert Androgynität als das Halten von sowohl männlichen als auch weiblichen Eigenschaften. Viel mehr werden nach dieser Arbeit Maskulinität und Femininität nicht als gegensätzliche Enden eines Spektrums betrachtet, sondern als unabhängige Dimensionen. Dies drückt sich also primär durch eine erhöhte Flexibilität für den Einzelnen aus, da es erlaubt je nach Situation und Umständen mehr feminine oder vermehrt maskuline Eigenschaften anzunehmen, wodurch die Grenzen der traditionellen Geschlechterrollen verschwimmen und aufbrechen. Androgynität ist somit ein Zustand zwischen den Geschlechtern, eine Fähigkeit, sich zu wandeln und anzupassen. Bei der Beobachtung dieses Phänomens über eine größere Gruppe von Menschen würde sich diese Flexibilität effektiv in einer Annäherung der Geschlechter in ihren Ursprungsmerkmalen ausdrücken.

Im Folgenden verschaffen wir uns nun einen Überblick über bisherige Untersuchungen, die Hinweise darauf geben könnten, ob ein solcher Effekt vorliegen könnte, und uns bei der Auswahl von Methoden unterstützen, um diesen möglichen Effekt zu untersuchen.

Veränderungen in Musik und Sprache

(Mauch et al., 2015) haben untersucht wie sich die Musik in den BB100 im Zeitraum von 1960 bis 2010 verändert hat. Sie analysierten hierzu 30s lange Previews von ca. 17.000 Songs und untersuchten diese auf verschiedene musikalische Merkmale, vor allem auf harmonische Merkmale und Timbre Eigenschaften. Sie fanden vor allem heraus, dass die Musik in diesen 50 Jahren an Diversität nicht abgenommen hat, sondern von drei

großen “musikalischen Revolutionen” geprägt ist. Für unsere Überlegungen besonders relevant jedoch stellten sie im Timbre der Musik eine Veränderung der wahrgenommenen Geschlechter fest. So sind männliche Merkmale relativ gleich geblieben mit einem kurzzeitigen Rückgang in den 80ern, während weibliche Merkmale stetig zugenommen haben. (Interiano et al., 2018) entdecken einen ähnlichen Effekt im Zeitraum von 1985 bis 2013, in dem sie zeigen, dass die wahrgenommene Männlichkeit von Musik in UK Charts abgenommen hat.

(Ali et al., 2020) untersuchte die Veränderungen der Vokale der BB100 von 1949 bis 2019 und fand unter anderem heraus, dass es Geschlechterübergreifend keine signifikante Veränderung im durchschnittlichen Pitch gab. Die könnte mehrere Ursachen haben. Da einerseits immer mehr Frauen in den Charts vertreten sind, müsste, angenommen Frauen singen im Schnitt höher als Männer, der durchschnittliche Pitch der Charts merklich steigen. Eine Veränderung im Pitch von männlichen und weiblichen Sänger:innen könnte jedoch diesen Effekt ausgleichen und somit zu einem konstanten durchschnittlichen Pitch führen. Alternativ könnte es auch sein, dass eine allgemeine Unterscheidung nach Stimmlagen von Männern und Frauen in den Charts nicht möglich ist, da die Unterschiede in den Stimmlagen nicht groß genug sind, um statistisch signifikant zu sein. Dies erscheint jedoch als eher unwahrscheinlich, da zumindest in der Sprache der Unterschied in der Stimmlage zwischen Männern und Frauen deutlich ist (Patel & Prasad, 2013).

Es stellt sich also die Frage, ob diese Veränderungen in der Geschlechterwahrnehmung und der Stimmen durch einzelne Künstler:innen zu beobachten sind, oder ob diese eher, aufgrund der steigenden Anteile an weiblichen Künstlerinnen, in der Musik als Ganzes zu finden sind. Um dem auf den Grund zu gehen, betrachten wir zunächst die Veränderungen von Stimmen im Allgemeinen, die Ursachen und mögliche Fortführungen schon

bestehender Trends.

(Pemberton et al., 1998) verglich Aufnahmen von Stimmen von Frauen aus den 1940er Jahren mit Aufnahmen aus den 1990er Jahren und konnte zeigen, dass die Stimmen der Frauen insgesamt tiefer geworden sind. Sie weisen darauf hin, dass dies vor allem auf eine Veränderung der Geschlechterrollen und -stereotypen zurückzuführen sein könnte.

Für Männer gibt es kaum vergleichbare Studien, jedoch lassen sich auf Basis von Untersuchungen der Sprache und dem Zusammenhang mit der Selbstwahrnehmung von Geschlecht einige Schlüsse ziehen. (Puts et al., 2006) untersuchte den Einfluss von “männlichem Dominanzwettbewerb” auf die Stimmlage und fand heraus, dass Männer in Situationen, in denen sie sich dominant fühlen, ihre Stimme unbewusst vertiefen. Der BSRI beispielweise zeigt, dass Dominanz eine traditionell eher als männliche Eigenschaft betrachtet wird (Bem, 1981). Eine Veränderung in der Gesamtgesellschaftlichen Wahrnehmung von Geschlechterrollen könnte also auch zu einer Veränderung in der Stimmlage von Männern führen.

Zusammenhang zwischen Geschlecht und Timbre

Nicht nur der *Pitch* jedoch kann ein Indikator für *Geschlecht* sein, sondern auch das *Timbre* der Stimme. Doch was ist Timbre überhaupt?

Timbre ist ein komplexes multidimensionales Konzept, das die Klangfarbe eines Tons beschreibt. Es ist das, was es uns ermöglicht zu erkennen, ob “zwei Töne, die auf ähnliche Weise präsentiert werden und die gleiche Lautstärke und Tonhöhe haben, unterschiedlich sind” (Moore, 2012). Hier gibt es verschiedene Methoden das *Timbre* zu quantifizieren, unter anderem die MFCCs.

Es stellt sich die Frage, ob es grundsätzlich überhaupt möglich ist, *Geschlecht* anhand von *Timbre* zu erkennen. (Nashipudimath et al., 2021) trainierten erfolgreich Klassifikationsmodell, welches mit mindestens 98% Genauigkeit zwischen männlichen und weiblichen Sprachaufnahmen unterscheiden konnte. Auch (Harb, 2003) konnte anhand ähnlicher Methoden Geschlecht anhand von Timbre-Features erkennen.

Die Geschlechtererkennung anhand Timbre funktioniert nicht nur in der Sprache, sondern auch in der Musik. (Jitendra & Radhika, 2021) trainierten auf Basis von MFCCs, Pitch und Delta-MFCCs ein Klassifikationsmodell, dass mit einer Genauigkeit von bis zu 94% zwischen männlichen und weiblichen Sänger:innen unterscheiden konnte.

Timbre ist also ein wichtiges Merkmal, das uns dabei hilft Klänge einzuordnen und zu untersuchen und kann uns dabei helfen, die Veränderungen von Stimmen genauer zu untersuchen.

Methoden in der Untersuchung von Pitch und Timbre

Audio Separation beschreibt einen Machine-Learning Ansatz, mittels dessen aus einer vollständig gemischten Audio datei verschiedene Elemente, sowie auch beispielsweise die Stimme, extrahiert werden können. Hier gibt es verschiedene Modelle und Ansätze, die sich in ihrer Genauigkeit und Geschwindigkeit unterscheiden. Demucs ist ein solches Modell, welches 2021 von einem Team von Forschern bei Facebook AI entwickelt wurde und sich im Vergleich zu anderen verbreiteten Modellen, wie Wave-U-Net, durch eine hohe Genauigkeit auszeichnet ((Défossez, 2021; Rouard et al., 2023)).

Pitch Extraction ist ein Prozess in dem die Tonhöhen (*Pitch*) aus Audiodateien extrahiert werden. Hierbei werden die Grundfrequenzen der Töne ermittelt und in einer Zeitreihe gespeichert. (Kim et al., 2018) entwickelte mit CREPE einen neuartigen

Ansatz, der direkt aus den rohen Audiodaten, im Gegensatz zu einem Spektrogramm, die Tonhöhen mittels eines Convolutional Neural Network extrahiert und dabei eine hohe Genauigkeit erzielt.

Timbre Feature Extraction will die Timbre-Features aus den Audiodaten extrahieren. (Terasawa et al., 2005) hat verschiedene Methoden zur Extraktion von *Timbre*-Features untersucht und dabei festgestellt, dass die MFCCs am ehesten der menschlichen Wahrnehmung von Timbre entsprechen. Da die MFCCs jedoch eine hohe Dimensionalität haben, wird oft eine PCA durchgeführt, um die Dimensionalität zu reduzieren und die Daten besser visualisieren und schneller verarbeiten zu können (Nashipudimath et al., 2021). Noch dazu wird durch die PCA der Vektor auf seine wichtigsten Merkmale reduziert, wodurch die Daten besser interpretierbar werden und weniger anfällig für Noise sind.

K-Means Clustering wird verwendet um Datensätze auf Strukturen und Gruppen zu untersuchen. Es handelt sich um einen unsupervised Machine-Learning Algorithmus, der Daten in Gruppen einteilt, wobei die Anzahl der Gruppen vorher festgelegt werden muss (Steinley, 2006). Diese Cluster können dann analysiert werden, um Muster und Strukturen in den Daten zu erkennen. Ein *Centroid* ist der Mittelpunkt eines Clusters und wird oft verwendet, um die Cluster zu charakterisieren und zu vergleichen.

Die vorliegende Arbeit versucht diese Methoden zu kombinieren um die Veränderungen in den Stimmen von männlichen und weiblichen Sänger:innen in den BB100 im Hinblick auf die Frage, ob es einen Trend hin zu mehr Androgynität im Gesang gibt, zu untersuchen.

Geplante Methoden

Um diese Fragestellung zu beantworten, müssen zunächst Daten erhoben werden und anschließend ausgewertet werden, wobei auch eine Analyse des Pitches und des Timbres der Songs durchgeführt wird.

Hierzu stellen wir uns folgende Hypothesen:

- Es besteht ein signifikanter Unterschied im Pitch-Umfang zwischen männlichen und weiblichen Gesangsstimmen in den BB100.
- Der durchschnittliche Pitch im Gesang hat sich seit 1958 zwischen männlichen und weiblichen Künstler:innen einander angenähert.
- Die MFCC-basierten *Timbre* Cluster der Songs können mit hoher Genauigkeit zwischen männlichen und weiblichen Gesangsstimmen unterscheiden.
- Die *Centroiden* der MFCC-basierten *Timbre* Cluster, die männliche und weibliche Gesangsstimmen repräsentieren, haben sich seit 1958 einander angenähert.

Erhebung des Datensatzes

Zunächst werden die wöchentlichen BB100 von 1958 bis heute aus dem frei zugänglichen GitHub-Repository von (Hollingshead, 2023) heruntergeladen und in eine Datenbank eingepflegt (<https://github.com/mhollingshead/billboard-hot-100>). Dieser Datensatz enthält wöchentliche Informationen von allen Songs, den Künstler:innen und dem Zeitraum und die Position ihrer Chart-Platzierung. Anschließend werden die gesammelten Einträge um relevante Metadaten erweitert.

Zunächst wird jeder Song in der MusicBrainz-Datenbank gesucht, um von dort ausgehend eine standardisierte und vollständige Liste der Künstler:innen zu aufzubauen. Die

Einträge in den ursprünglichen Daten sind, insbesondere bei Featured Artists, häufig unvollständig oder inkonsistent. Ebenso wird für jeden Song ein *Release-Datum* und das *Genre* ermittelt und gespeichert, wobei für das *Genre* die Kategorisierung in übergreifende Kategorien noch nicht feststeht, da sich diese erst aus dem Auslesen des vollständigen Datensatzes ergibt.

Sobald die nun vollständige Liste aller Künstler:innen vorliegt, wird das *Geschlecht* der Künstler:innen ermittelt. Hierbei wird ein ähnlicher Ansatz wie in (Aguiar et al., 2018) angewendet. Falls das *Geschlecht* auf MusicBrainz nicht verfügbar ist, wird der Vorname des Künstlers in entsprechenden Datenbanken verglichen und bei einer Wahrscheinlichkeit von 90% übernommen¹. Bei Gruppen und Bands wird versucht eine Liste aller Mitglieder zu erstellen, die Geschlechter der Mitglieder festzustellen und die Gruppe schließlich als männlich, weiblich oder gemischt zu klassifizieren. In allen anderen Fällen, in denen eine automatisierte Zuweisung nicht möglich war, wird das *Geschlecht* manuell zugewiesen. Hierzu werden die Pronomen der Künstler:innen auf Spotify, Social Media, Wikipedia und anderen Quellen analysiert. Es wird somit eher das wahrgenommene Geschlecht als das selbst identifizierte Geschlecht erfasst. Der letztere Schritt der manuellen Zuweisung wird womöglich, wenn trotzdem genug Daten vorhanden sind auch weggelassen.

Auf Basis des *Geschlechts* der jeweiligen Interpret:innen wird als nächstes pro Song wieder eine ähnliche Methodik wie in (Aguiar et al., 2018) angewendet. Ein Song wird entweder **first male/female**, **all male/female** oder **mixed** kategorisiert.

Dieser Datensatz umfasst nun insgesamt 31.432 Songs, denen jeweils ein *Geschlecht*, in *Genre* und ein Zeitpunkt zugewiesen wurde. Diese werden ähnlich wie auch in (Mauch et al., 2015) als 30s Previews von Spotify heruntergeladen.

¹Analyse mittels <https://gender-api.com/>

Da wir in dieser Studie ausschließlich an Veränderungen des Gesangs interessiert sind, erfolgt zunächst eine Audioseparation mithilfe von demucs, wie beschrieben in (Défossez, 2021) und (Rouard et al., 2023). Anschließend wird der Pitch-Umfang des extrahierten Gesangsparts durch eine Pitch-Analyse mit crepe (Kim et al., 2018) ermittelt. Auf Basis dieser Pitch-Daten werden der durchschnittliche Pitch sowie die Standardabweichung für jeden Song berechnet. Um die Veränderung des Timbres der Stimmen zu analysieren, werden die zeit-gefensterten MFCCs der Songs berechnet, gemittelt und anschließend durch eine PCA in ihrer Dimensionalität für spätere bessere Visualisierung reduziert.

Auswertung der Daten

Bevor mit der detaillierten Datenanalyse begonnen wird, sollen zunächst deskriptive Statistiken für den gesamten Datensatz von 31.432 Songs erstellt werden. Hierbei wird der durchschnittliche *Pitch* und die Standardabweichung für männliche und weibliche Stimmen in den verschiedenen *Genres* und Jahren aufgezeigt, sowie Verteilung von *Geschlecht* pro *Genre* und über den Testzeitraum hinweg.

Zum Testen unserer ersten Hypothese H1.1 wird eine ANOVA durchgeführt, um zu untersuchen, ob signifikante Unterschiede in den Stimmlagen zwischen den *Geschlecht*-Gruppen pro *Genre* und über alle Genres hinweg existieren. Dieser Test soll klären, ob es grundsätzlich systematische Unterschiede in den Stimmlagen zwischen unseren Gender-Gruppen gibt.

Im Anschluss wird eine Regressionsanalyse pro *Genre* und *Geschlecht* durchgeführt, um zu erkennen, ob es einen Trend in der Veränderung der Stimmlagen gibt. Hiermit können wir zur Überprüfung von H1.2 quantifizieren, ob sich möglicherweise die Stimmlagen im Pitch der beiden *Geschlecht*-Gruppen aneinander annähern, voneinander entfernen oder

konstant bleiben.

Im Anschluss werden mittels K-Means-Clustering die durch PCA reduzierten MFCC-Daten der Songs in 2 Cluster (männlich und weiblich) eingeteilt. Hierbei wird zuerst die Genauigkeit des Clustering-Modells überprüft, indem die Cluster mit dem tatsächlichen *Geschlecht* der jeweiligen Interpret:innen verglichen werden. Dies gibt uns Auskunft darüber, ob unser Modell überhaupt in der Lage ist, männliche und weibliche Stimmen zu unterscheiden (H2.1) oder ob die Unterschiede in den Timbre-Features nicht ausreichen, um eine Unterscheidung zu ermöglichen, und weitere Features oder Preprocessing-Schritte notwendig sind.

Zum Schluss wird zur Überprüfung von H2.2 eine Regressionsanalyse auf der Distanz zwischen den Centroiden der Cluster über unseren Testzeitraum durchgeführt. Hierbei wird untersucht, ob sich die Centroiden der Cluster, die männliche und weibliche Stimmen repräsentieren, im Laufe der Zeit einander annähern, voneinander entfernen oder konstant bleiben. Hierzu müssen jedoch zuerst die jährlichen Centroiden der Cluster berechnet werden, was durch eine einfache Mittelung der MFCC-Daten der Songs in den jeweiligen Clustern zum jeweiligen Jahr geschieht.

Abkuerzung	Bedeutung
BB100	Billboard Top 100 Charts
MIR	Music Information Retrieval
MFCC	Mel-Frequency Cepstral Coefficients
PCA	Principal Component Analysis
BSRI	Bem Sex Role Inventory

Abkuerzung	Bedeutung
PAQ	Personal Attributes Questionnaire

References

- Adamczyk, A., & Liao, Y.-C. (2019). Examining public opinion about LGBTQ-related issues in the united states and across multiple nations. *Annual Review of Sociology*, 45(1), 401–423. <https://doi.org/10.1146/annurev-soc-073018-022332>
- Aguiar, L., Waldfogel, J., & Waldfogel, S. (2018). *Playlisting favorites: Is spotify gender-biased?* JRC Digital Economy Working Paper. <https://doi.org/10.3386/w29017>
- Ali, A., Azzaretto, J., & Moltz, K. (2020). The impact of vocal features on song success. *Journal of Student Research*, 9. <https://doi.org/10.47611/jsrhs.v9i2.1217>
- Anderson, T. (2024). *Leading ladies: A look at women's no. 1 success in the hot 100's history*. <https://doi.org/10.4324/9780203104026-11>
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2), 155. <https://doi.org/10.1016/b978-0-12-590241-0.50015-0>
- Bem, S. L. (1981). Bem sex role inventory. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/t00748-000>
- Défossez, A. (2021). Hybrid spectrogram and waveform source separation. *Proceedings of the ISMIR 2021 Workshop on Music Source Separation*.
- Gerson, K. (2010). *The unfinished revolution: How a new generation is reshaping family, work, and gender in america* (Vol. 16). Oxford University Press New York. <https://doi.org/10.14431/aw.2012.12.28.4.117>
- Harb, H. (2003). *Gender identification using a general audio classifier*. <https://doi.org/10.1109/ICME.2003.1221721>
- Hollingshead, M. (2023). *Historic billboard hot 100 data*. <https://doi.org/10.32614/cran.package.billboard>

- Interiano, M., Kazemi, K., Wang, L., Yang, J., Yu, Z., & Komarova, N. (2018). Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society Open Science*, 5, 171274. <https://doi.org/10.1098/rsos.171274>
- Jitendra, M. S., & Radhika, Y. (2021). Singer gender classification using feature-based and spectrograms with deep convolutional neural network. *International Journal of Advanced Computer Science and Applications*, 12(2). <https://doi.org/10.14569/IJA-CSA.2021.0120218>
- Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). CREPE: A convolutional representation for pitch estimation. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 161–165. <https://doi.org/10.1109/ICASSP.2018.8461329>
- Mauch, M., MacCallum, R. M., Levy, M., & Leroi, A. M. (2015). The evolution of popular music: USA 1960–2010. *Royal Society Open Science*, 2(5), 150081. <https://doi.org/10.1098/rsos.150081>
- Moore, B. C. (2012). *An introduction to the psychology of hearing*. Brill.
- Nashipudimath, M. M., Pillai, P., Subramanian, A., Nair, V., & Khalife, S. (2021). Voice feature extraction for gender and emotion recognition. *ITM Web of Conferences*, 40, 03008. <https://doi.org/10.1051/itmconf/20214003008>
- Patel, K., & Prasad, R. (2013). Speech recognition and verification using MFCC & VQ. *Int. J. Emerg. Sci. Eng.(IJESE)*, 1(7), 137–140.
- Pemberton, C., McCormack, P., & Russell, A. (1998). Have women’s voices lowered across time? A cross sectional study of australian women’s voices. *Journal of Voice : Official Journal of the Voice Foundation*, 12 2, 208–213. [https://doi.org/10.1016/S0892-1997\(98\)80040-4](https://doi.org/10.1016/S0892-1997(98)80040-4)

- Puts, D., Gaulin, S., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27, 283–296. <https://doi.org/10.1016/J.EVOLHUMBEHAV.2005.11.003>
- Rouard, S., Massa, F., & Défossez, A. (2023). Hybrid transformers for music source separation. *ICASSP 23*. <https://doi.org/10.1109/ICASSP49357.2023.10096956>
- Smith, P., Pieper, K., Hernandez, C., & Wheeler, S. (2023). Inclusion in the recording studio? Gender & race/ethnicity of artists, songwriters & producers across 1,100 popular songs from 2012 to 2022. *USC Annenberg Inclusion Initiative*.
- Steinley, Douglas. (2006). K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, 59(1), 1–34. <https://doi.org/https://doi.org/10.1348/000711005X48266>
- Terasawa, H., Slaney, M., & Berger, J. (2005). The thirteen colors of timbre. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 323–326. <https://doi.org/10.1109/ASPAA.2005.1540234>
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, 36(5-6), 305–325. <https://doi.org/10.1007/BF02766650>