DOI: 10.61782/fa.2025.0058

FORUM ACUSTICUM EURONOISE 2025

ACOUSTIC AND COMPARATIVE ANALYSIS IN THE PROCESS OF TRANS-VOICE HORMONAL TREATMENT

Alonso Nogales, Jone[1*]

[1]University of Cádiz, Andalusia, Spain

ABSTRACT

This study focuses on the analysis of the changes that the voice undergoes in the process of taking sex change hormones, with the aim of comparing these changes to create a transitional vocal passport. For this purpose, three samples were taken from the informants: before the hormones were taken, during the hormone treatment and once the treatment was finished; within the informants we found changes of sex from female to male and changes from male to female with the intention of making a comparative analysis. This work is necessary because there is almost no evidence or studies of trans voices, which is a problem because they are absolutely indispensable for many fields, such as forensic linguistics and clinical linguistics. These sciences are growing in these last decades, so it is of prime necessity to continue advancing in them and to cover all those lines of research that are unknown and are so important, as is the case of this study for authorship attribution, for psychoacoustics...

Keywords: *voice, hormones, transgender, linguistics.*

1. INTRODUCTION

Voice is an identifying feature, no two voices are exactly alike anywhere in the world (Nolan and Oh, 1996) [1] , hence it is often compared to a fingerprint [2] . So it makes sense that it is often referred to as a 'vocal passport', which is the creation of linguistic profiles. The production of these passports is crucial for forensic linguistics, as it helps to find the person suspected of a crime. Less recognised is the contribution of this task to clinical linguistics, but it can also be of great help in assessing speech or language disorders, which specialists such as phonoaudiologists or speech therapists will then use to apply a treatment, in addition, linguists can subsequently pass test batteries related to the pathology being treated to see how the evolution of the disease is 'attacking' the language.

Speaking of pathologies, the World Health Organisation in [2] claimed that transsexuality was a mental illness and encapsulated it in section 17, 'conditions related to sexual health'. It was given the name gender incongruence and explained that "is characterised by a marked and persistent incongruence between an individual's experienced gender and the assigned sex. Gender variant behaviour and preferences alone are not a basis for assigning the diagnoses in this group". Fortunately, we can see that the WHO rectifies in [3] and does not include transsexuality as a disease. However, in section b 180, "Experiencias relacionadas con uno mismo y con el tiempo", the following is mentioned:

*__Corresponding author__: jone.alonsonogales@alum.uca.es.

[1] Cited by Elena Garayzábal and Mercedes Reigosa in [1].

[2] According to Lawrence G. Kersta (1962), "initial successes indicated that a voiceprint identification system analogous to existing fingerprint identification is a possibility. Subjective matching of voiceprints has resulted in success scores of better than 97%".

> Funciones mentales específicas relacionadas con la conciencia de la propia identidad, del propio cuerpo, de la posición de uno mismo en su mundo y en el tiempo. Incluye: funciones de la experiencia de uno mismo, de la imagen corporal y del tiempo. [3]

Trans people are those who do not identify with their biological sex and this can lead to a mental health problem due to different factors, such as social pressure, discrimination, isolation, lack of security etc. However, gender discordance itself is not an illness. In other words, this condition can lead to mental health problems, but not necessarily; sometimes it occurs due to external factors, sometimes it has to do with non-identification with oneself and sometimes it is usually the sum of the two. The organisation Transaludes in [4] devotes a section to the relationship between trans people and mental health, let's look at the tables of the results obtained:

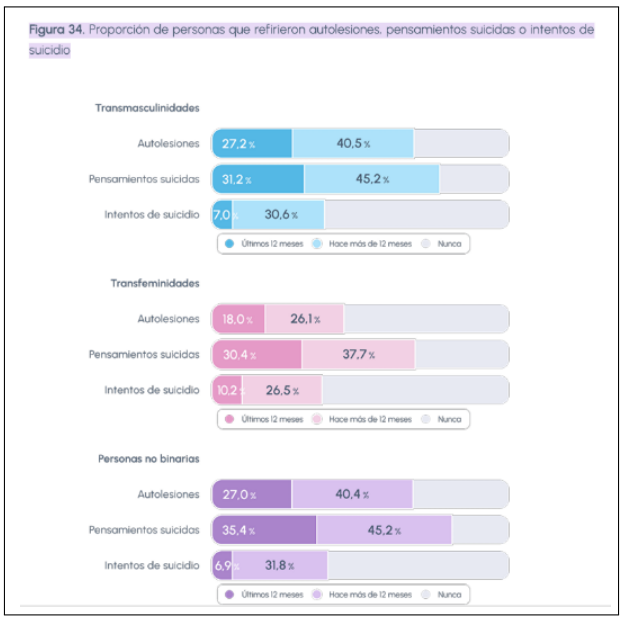

Figure 1. Proportion of people who reported self-harm, suicidal thoughts or suicide attempts

In [4] they explain that Fig. 1 and Fig. 2 give an idea of the serious consequences of exclusion, discrimination

[3] This could lead one to think that it refers to the process of transition from male to female and female to male, but from a much more respectful approach.

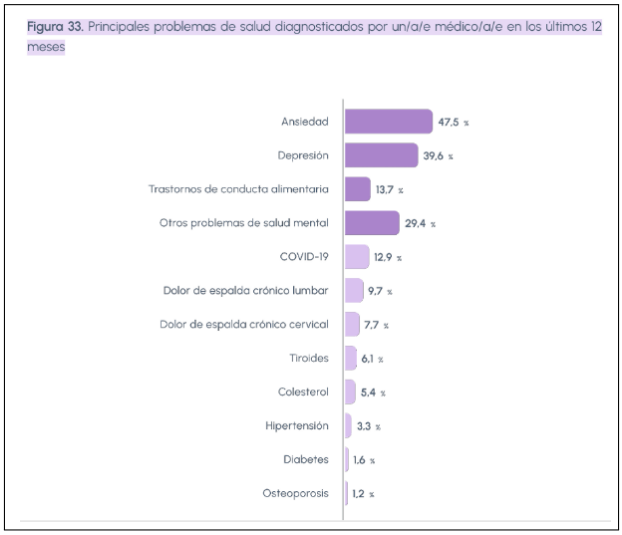

Figure 2. Main health problems diagnosed by a doctor in the last 12 months

and violence experienced by trans and/or non-binary people. In the first figure we can see the proportion of people of each gender identity who report having experienced self-harm, suicidal thoughts and attempts. Furthermore, we can see that the two main mental health problems they experience are anxiety and depression, anxiety at the top with 47.5% and depression with 39.6%.

For this reason, it is important that we approach the topic from different disciplines. It has been studied for some years from different fields of knowledge such as medicine in general, and more specifically from endocrinology, psychology, and to a lesser extent from otorhinolaryngology. However, there is little research on trans people from the humanities and the main findings have been logopaedic. That is why I consider that linguistics has to address this "new" [4] reality, with the advantage of being inter-, transdisciplinary. When I came to this conclusion, I started to review the background and I was obliged to include research from very different areas, so that the present study would give a current overview and to be able to support the results obtained.

With this, it is worth noting that the human brain, biologically speaking, is exactly the same in women and

[4] We have been living with trans people for centuries, but today it is a socially accepted reality or should be accepted because there are laws and regulations that protect the rights of transgender people.

men [5] [5], so it is reasonable to think that the change of sex does not imply a change in the cognitive level. The physical differences are the most noticeable, as they are the ones we can see first through sight, which is one of the senses most used by human beings. However, when it comes to perception, specifically that related to language, hearing is also of great value. Acoustic and auditory perception are fundamental in our daily lives: "Listeners are able to identify talkers by their voices with a great deal of accuracy and very little trouble. Even when voices are unfamiliar, listeners have very consistent impressions of a talker's gender, height, weight, etc." [6]. In the same article, they rightly point out that:

> The most important finding is that voice gender is not stored in abstract male and female voice representations. Instead, voice gender appears to be stored in the form of auditory-based perceptual representations. These representations, in all probability, contain specific auditory information about acoustic voice parameters relevant to gender.The results of experiment 4 suggest that these representations are not based on one isolated parameter like F0 or formant frequencies. Instead, the representations are probably an auditory composite of the various acoustic factors relevant to voice gender like F0, formant frequencies,breathiness, etc. Although there is a close relationship between phonetic coding and voice coding processes during perception, the representations of phonemes and voices appear to be qualitatively different in that phonetic representations may not be as detailed [6].

So the challenge is to get the voice to modify so much that it corresponds to the gender they identify with and, as we can see, this is not an easy task because many parameters have to be taken into account. In this article, we will analyse the fundamental frequency, since it is a pilot test of a work that will be extended later on. There are many articles that concern other disciplines such as glottoplasty operations, type III thyroplasty, type IV, injections of filling material, etc. Therefore, we will prioritise voices that have only been treated with hormone replacement therapy [7]: testosterone for masculinisation and oestrogens for feminisation. There is much more background on vocal feminisation as it is a more complex process than masculinisation, as testosterone reaches the voice better than oestrogen, in other words, oestrogen does not change the pitch of the voice significantly, whereas testosterone does produce a thickening of the vocal cords. In voice masculinisation processes, hormone treatment is usually sufficient, while trans women usually go to speech therapists and ENT specialists to remove tension from the vocal cords and continue to raise the fundamental frequency until they feel comfortable with their voice.

2. HYPOTHESIS AND OBJECTIVES

Considering the facts previously stated, the following hypothesis and objectives have been established:

HYPHOTESIS:

- H1: Trans voices undergo changes at different stages.
- H2: The fundamental frequency of trans voices, after HRT, approach the mean fundamental frequency of cis voices.

OBJECTIVES:

- O1. Examine trans-voice changes at different stages.
- O2. Check that the fundamental frequency of trans voices are close to the mean F0 of cisgender voices.

3. METHODOLOGY

This research focuses mainly on the analysis of trans voices at three stages: before starting hormones, during treatment [8] and after the treatment is "finished" [9]; the fundamental frequency of the voice at the different stages will

[5] Neuroscientist Gina Rippon analysed the brains of babies who were not yet exposed, due to age, to social roles and showed that they were exactly the same in both sexes in [5].

[6] In addition, it must be taken into account that the vocal cords of male voices are biologically thicker from adolescence onwards, boys and girls have the same vocal cord structure, which is why it is often difficult to recognise the sex of an infant.

[7] And that they have received vocal therapy if they have received it, i.e. that they have not undergone any surgery affecting the voice at the time the audios were recorded.

[8] Comprising approximately an intermediate stage in the process. Samples have been taken from the tenth month for this stage.

[9] The inverted commas are used here because hormone replacement therapy is sometimes a lifelong treatment, as the change can be reversed and thus the effectiveness of the hor-

be analysed and compared using the free-to-use *Praat* tool and, afterwards, a comparison will be made between the process of voice masculinisation and vocal feminisation. As this is a sensitive topic and, as we have seen, can be traumatic for many people, the search for informants became a challenge. So I decided to take the samples from videos posted on the public network, in this case, Youtube, which I then converted into MP3 audio using the Y2Mate converter and trimmed with Audacity. Each sample is thirty seconds long.

These audios correspond to trans people who have narrated their process in hormone therapy (HRT) and have shared it publicly online; it was a decision taken based on the need for the sample to be sufficiently representative for a pilot test, which will later be expanded, as it is hoped that this research will give visibility to the collective and improve the quality of life of these people and their environment. Both vocal feminisation and masculinisation will be studied, but specific treatment will not be considered [10] , because some informants in the network do not specify the treatment they are following and I do not have access to communication with these informants. Samples were chosen without too much background noise to ensure maximum reliability of the results. The sample is composed of 4 trans women and 4 trans men, with the intention of achieving equity and balance. All of them are English speakers, it is not known if they are all native speakers, but they are fluent in English, varieties of the language will not be taken into account. No specific age range is specified, but the informants must be young people with a stabilised voice, since during adolescence there are many vocal changes, and as we reach a certain age the voice varies.

3.1 Terminology

The term 'trans' will be repeated or, failing that, alternated with the term 'transgender' [11] instead of 'transsexual', as not all people who do not identify with their biological sex have the desire to change their physical appearance [12] . As this is a work that deals with sensitive data and narrates a process that can lead to mental health issues, I intend to be as cautious as possible, including transgender people, transsexual people, and non-binary people. However, only people whose voice has been modified by taking hormones will be studied, as the aim of this work is none other than to analyse and observe the change in voice in three different time periods.

4. RESULTS

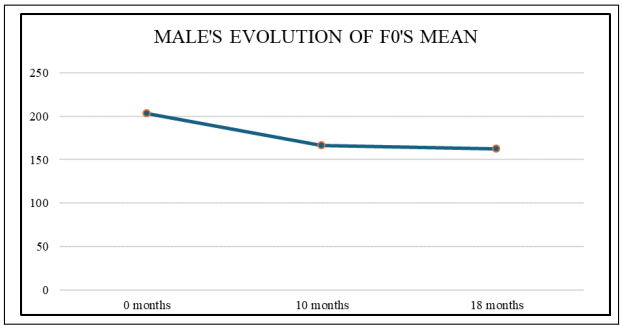

Figure 3. Male's evolution of F0's mean.

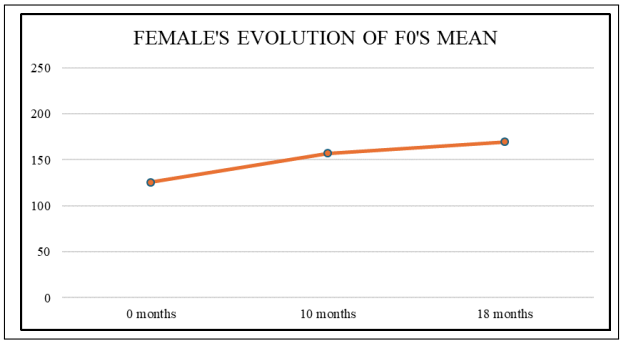

Figure 4. Female's evolution of F0's mean.

It can be seen in Fig. 3 that trans males in the transition from 0 months to 10 months lower the mean fun-

mones taken earlier is lost. By "finish" I mean the moment when the voice stabilises, we have taken samples from month eighteen.

[10] There are many types of treatment: gel, injections, sprays...

[11] As [7]: "we use the phrase transgender and gender diverse (TGD) to be as broad and comprehensive as possible in describing members of the many varied communities that exist globally of people with gender identities or expressions that differ from the gender socially attributed to the sex assigned to them at birth. This includes people who have culturally specific and/or language-specific experiences, identities or expressions, which may or may not be based on or encompassed by Western conceptualizations of gender or the language used to describe it".

[12] "No todas las personas transexuales desearían cambiarse de sexo ni estarían incómodas con sus genitales. existe una categoría de personas dentro de la transexualidad conocidas como transgénero, que se sentirían identificadas con el género contrario al de su sexo genital pero no desearían modificarlo" (Mejía, 2006. Cited by Janet Noseda in [8]).

damental frequency significantly, more specifically, 36.85 Hz on average. Afterwards, it continues to drop, but the change is practically imperceptible: 4.3 Hz. Therefore, we can infer that from the intermediate to the 'final' stage, from 10 months to 18 months, there is a stagnation or stability. In the case of trans women, see Fig. 4, something similar occurs: the most noticeable changes occur in the first 10 months. From the pre-oestrogen stage to the intermediate stage (10 months) they increased on average 31.25 Hz. Then, from 10 months to 18 months the change is more noticeable than in men, but it is not a big change either: 12.77 Hz, while in men there was only a 4.3 Hz difference.

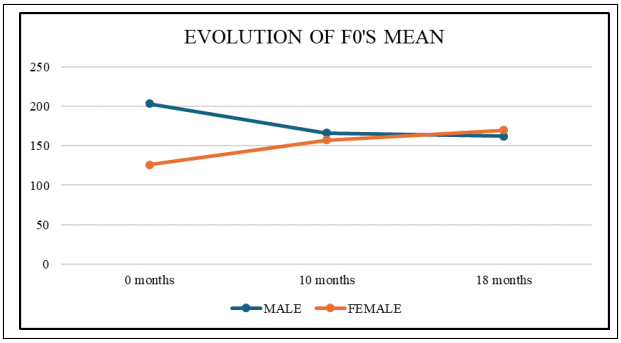

Figure 5. Evolution of F0's mean.

In Fig. 5, it is interesting to see that in the intermediate phase both genders tend towards a similar fundamental frequency of around 150 Hz. This is interesting because the average F0 for males is at 120 Hertz and 210 Hertz for cisgender females, especially in the case of Western languages [9]. Therefore, at 10 months trans women and trans men reach an intermediate range between the average values of cisgender men and women.

As mentioned above, trans women aim to raise the hertz of their voice to achieve a higher pitched voice, while trans men, in vocal masculinisation, aim to lower the hertz to make the voice sound deeper and more similar to what we perceptually consider 'masculine voices'.From this information, the following questions were calculated in Tab. 1: the progress of trans people (men and women) from 0 months on hormones to 18 months into the process, the distance between the average fundamental frequency of cisgender voices and trans voices before starting hormone therapy and, finally, the distance between the average values of F0 in cisgender people and the average value obtained from trans people in the 18th month of HRT.

Looking at Tab. 1, we can see that the average F0 of trans men from month 0 to month 18, i.e. from the first stage (pre-testosterone) to the last (18 months on testosterone), has dropped by 20.24%. And that at the beginning (0 months), F0 is 40.99% above the cis-male average, but that after 18 months, it approaches the typical cis-male range, with the difference reduced to 26.01%. On the other hand, trans women experience an increase in fundamental frequency of 35.00% between the early and late stages. In the pre-oestrogen phase, F0 is 66.98% lower than the average for cisgender women. However, after 18 months of hormone therapy, the gap is reduced to 23.69%, approaching the cisgender female range, but not quite reaching it.

4.1 Clinic linguistics

These results should be provided during treatment, not only tests to measure testosterone and oestrogen levels, but also giving patients the right to receive detailed information about the changes in their voice. In addition, endocrinologists in the first contact before starting any treatment should use some of these examples to explain that it is not a linear process, that it takes time, but that there really are noticeable changes in the voice; with data and real cases it shows that it really is possible.

The figure of the linguist together with that of the speech therapist is essential, as long as there is no overlap: the speech therapist is in charge of vocal therapy and the linguist is in charge of narrating the changes in auditory perception in consultation with the *Praat* tool to provide real and accurate data. In addition, it is essential to consider other aspects, such as the tension of the vocal cords, the appearance of nodules and other conditions that may appear due to the effort to modulate the voice. A vocal passport of before, during and after should be made to demonstrate that the changes are real, even though some patients continue to suffer from vocal dysphoria after treatment.

4.2 Forensic linguistics

The realisation of such vocal passports in their three stages could make a significant contribution to forensic linguistics, which faces a major challenge in identifying speakers who have modified their voice. However, as we have seen, voice is a personal thing and there are parameters that remain virtually unchanged [13]. However, wheeled voice recognition would not be a reliable test in the case of a

[13] It will be taken into account in the extension of this work.

Table 1. Longitudinal Comparison of Vocal F0 in Cis and Trans Speakers: Progression at 0, 10, and 18 Months and Relative Distance to Cisnormative Values

Subjects	0 months	10 months	18 months	F0 mean cisgender voices	Progress 0-18M	Distance 0M-cisgender	Distance 18M-cisgender
Male	203,3385 Hz	166,486 Hz	162,188 Hz	120 Hz	20,24%	40,99%	26,01%
Female	125,764 Hz	157,0145 Hz	169,782 Hz	210 Hz	35,00%	66,98%	23,69%

suspect whose voice has been modified by HRT, because an untrained ear could be overlooking many details that are fundamental to recognise that the voice it is hearing belongs to the same person who committed the crime in the past.

As [1] point out, "se ha de recordar que en este proceso selectivo es necesario tener en cuenta las variables básicas relativas al autor que la víctima aportó durante su entrevista (sexo, edad, variedad dialectal, etc.), pero por encima de todo el grado de similitud con la voz del sospechoso". The "sex" section in this case is a challenge, not because the suspect does not identify with his or her biological sex, but because it no longer corresponds to what the victim saw or perceived as such a long time ago and because the offender might therefore have changed his or her physical appearance and other features such as voice. This is not to say that trans people have an advantage over non-trans people when it comes to offending, but that trained ears, i.e. linguists, are needed to deal with the situation. This is why the figure of the forensic linguist has been in vindication for decades, but has received so little recognition.

5. CONCLUSIONS

We can see from the figures that there is a significant change in fundamental frequency and that, if we take into account the three-stage process, we can see that the voice becomes more and more similar to that of the sex with which they identify. After these first results, it would be useful to add one or two more stages after the eighteen-month stage, to see if there is still an evolution or if, on the contrary, there is a stagnation from the 10-month stage until the new stage is added. The most important inference, in my opinion, is to obtain real data showing the change in voice and thus to be able to determine that vocal feminisation and masculinisation through HRT is a process that contributes to the well-being of trans people and consequently to the people around them. There is a need for linguists in hospitals to accompany trans people in this life-changing process and for linguists in the judicial field who are able to determine the almost unalterable basic features of people regardless of any treatment [1].

6. REFERENCES

[1] E. G. Heinze and M. R. Riveiros, "Reconocimiento de voz en rueda: una vuelta de tuerca," *Del Español: Revista de Lengua*, vol. 1, pp. 101–133, 2023.

[2] World Health Organization, *Gender incongruence, unspecified (ICD-11 code: HA6Z)*. World Health Organization, Geneva, 11th revision ed., 2025. Accessed on April 4, 2025.

[3] Organización Mundial de la Salud, *Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud*. Ginebra: Organización Mundial de la Salud, 2001.

[4] M. J. Belza, P. González-Recio, S. Moreno-García, M. Donat, A. Iglesias, J. M. Guerras, E. Granda, D. Palma, and R. Barba, *Transaludes: salud en personas trans y/o no binarias en España*. Madrid: Instituto de Salud Carlos III, 2024.

[5] G. Rippon, *The gendered brain : the new neuroscience that shatters the myth of the female brain / Gina Rippon*. London: The Bodley Head, 2019.

[6] J. W. Mullennix, K. A. Johnson, M. Topcu-Durgun, and L. M. Farnsworth, "The perceptual representation of voice gender," *The Journal of the Acoustical Society of America*, vol. 98, pp. 3080–3095, 12 1995.

[7] E. Coleman, A. E. Radix, W. P. Bouman, G. R. Brown, A. L. C. de Vries, M. B. Deutsch, R. Ettner, L. Fraser, M. Goodman, J. Green, A. B. Hancock, T. W. Johnson, D. H. Karasic, G. A. Knudson, S. F. Leibowitz, H. F. L. Meyer-Bahlburg, S. J. Monstrey, J. Motmans, L. Nahata, T. O. Nieder, S. L. Reisner, C. Richards, L. S. Schechter, V. Tangpricha, A. Tishelman, M. A. A. V. Trotsenburg, S. Winter, K. Ducheny, N. J. Adams, T. Adrián, L. R. Allen, D. Azul, H. Bagga, K. Başar, D. S. Bathory, J. J. Belinky, D. Berg, J. U. Berli, R. O. Bluebond-Langner, M.-B. Bouman, M. L. Bowers, P. J. Brassard, J. Byrne,

L. Capitán, C. J. Cargill, J. M. Carswell, S. C. Chang, G. Chelvakumar, T. Corneil, K. B. Dalke, G. D. Cuypere, E. de Vries, M. D. Heijer, A. H. Devor, C. Dhejne, A. D'Marco, E. K. Edmiston, L. Edwards-Leeper, R. Ehrbar, D. Ehrensaft, J. Eisfeld, E. Elaut, L. Erickson-Schroth, J. L. Feldman, A. D. Fisher, M. M. Garcia, L. Gijs, S. E. Green, B. P. Hall, T. L. D. Hardy, M. S. Irwig, L. A. Jacobs, A. C. Janssen, K. Johnson, D. T. Klink, B. P. C. Kreukels, L. E. Kuper, E. J. Kvach, M. A. Malouf, R. Massey, T. Mazur, C. McLachlan, S. D. Morrison, S. W. Mosser, P. M. Neira, U. Nygren, J. M. Oates, J. Obedin-Maliver, G. Pagkalos, J. Patton, N. Phanuphak, K. Rachlin, T. Reed, G. N. Rider, J. Ristori, S. Robbins-Cherry, S. A. Roberts, K. A. Rodriguez-Wallberg, S. M. Rosenthal, K. Sabir, J. D. Safer, A. I. Scheim, L. J. Seal, T. J. Sehoole, K. Spencer, C. S. Amand, T. D. Steensma, J. F. Strang, G. Taylor, K. Tilleman, G. G. T'Sjoen, L. N. Vala, N. M. V. Mello, J. F. Veale, J. Vencill, B. Vincent, L. M. Wesp, M. A. West, and J. A. and, "Standards of care for the health of transgender and gender diverse people, version 8," *International Journal of Transgender Health*, vol. 23, no. sup1, pp. S1–S259, 2022.

[8] J. N. Gutiérrez, "Muchas formas de transexualidad: diferencias de ser mujer transexual y de ser mujer transgénero," *Revista de Psicología*, vol. 21, no. 2, pp. 7–30, 2012.

[9] H. Traunmüller and A. Eriksson, "The size of f0 excursions in speech production and perception," *Working papers/Lund University, Department of Linguistics and Phonetics*, vol. 43, pp. 136–139, 1994.