A comparative acoustic analysis of non-verbal vocalisations in film dubbing

Artur Gudmanian [a], Iryna Struk [a, *], Halyna Yencheva [a], Alla Sitko [b]

[a] State University of Information and Communication Technologies, Ukraine
[b] State University "Kyiv Aviation Institute", Ukraine

Received October 28, 2025; Revised November 7, 2025; Accepted November 10, 2025

Abstract. This study explores the translation of non-verbal vocalisations in dubbing and examines their role in maintaining the authenticity and emotional depth of audiovisual performance. It addresses a notable gap in translation and psycholinguistic research by focusing on vocal behaviours such as swallowing, sneezing, stuttering, wheezing, laboured breathing, and groaning, which have often been neglected within studies of paralanguage and kinesics. The corpus was compiled from four seasons of *Friends* (1994–2004) and their Ukrainian-dubbed versions, obtained from the Kyivstar TV platform. It included 1,050 instances of non-verbal vocalisations along with their corresponding dubbing solutions. A mixed-methods approach was adopted. Qualitative analysis focused on contextual use and acoustic characteristics, while quantitative analysis employed a corpus-based design. Acoustic parameters, including pitch, duration, stress, and amplitude, were measured using Praat, and Audacity was used to isolate background audio from dubbed tracks. The analysis also drew on sound analysis, functional equivalence, error analysis, and contrastive analysis frameworks. The findings demonstrate that vegetative vocalisations, such as coughing, sneezing, and swallowing, were frequently modified or omitted in the Ukrainian dub, resulting in the loss of emotional and comedic nuance. Aural vocalisations, including wheezing and laboured breathing, were similarly underrepresented, reducing synchronisation with visual cues and weakening emotional resonance. The study concludes that preserving the acoustic and emotional integrity of non-verbal vocalisations is essential for achieving greater authenticity and expressiveness in audiovisual translation.

***Keywords*:** acoustic phonetics, voice acting, vegetative and aural sounds, Praat software, English, Ukrainian, sitcom *Friends*.

Гудманян Артур, Струк Ірина, Єнчева Галина, Сітко Алла. Порівняльний акустичний аналіз невербальних вокалізацій у дублюванні фільмів.

У статті розглядається переклад невербальних вокалізацій у дубляжі та аналізується їхній вплив на автентичність і емоційну виразність аудіовізуального тексту.

* Corresponding author. Iryna Struk, 0000-0002-8128-0205, i.struk@duikt.edu.ua
© *Gudmanian, Artur; Struk, Iryna; Yencheva, Halyna; Sitko, Alla, 2025.* This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International Licence (http://creativecommons.org/ licenses/by/4.0).
East European Journal of Psycholinguistics, 12(2), 122–140, https://doi.org/10.29038/gud

Автори дослідження звертають увагу на малодосліджені вокальні прояви, зокрема ковтання, чхання, заїкання, хрипіння, утруднене дихання та стогін, які є важливими для створення комунікативного ефекту і розкриття характеру персонажа. Такий підхід дозволяє заповнити значну прогалину в сучасних перекладознавчих і психолінгвістичних дослідженнях. Матеріалом для дослідження став корпус, укладений на основі чотирьох сезонів серіалу *Friends* та їхніх українських дубльованих версій, отриманих з платформи Kyivstar TV. Корпус містить 1050 прикладів невербальних вокалізацій разом із відповідними перекладацькими стратегіями. Дослідження здійснено з використанням інтегрованого підходу, який поєднує якісні й кількісні методи аналізу. У межах якісного підходу автори дослідили контекстуальне використання й акустичні характеристики вокалізацій, тоді як кількісний аналіз грунтувався на корпусному матеріалі. Акустичні параметри, зокрема висота тону, тривалість, наголос і амплітуда, вимірювались за допомогою програми Praat. Для виокремлення фонових звуків з дубльованих треків було використано програму Audacity. Інтерпретація результатів здійснювалась на основі методів звукового аналізу, функціональної еквівалентності, аналізу помилок та контрастивного аналізу. Порівняння оригінальних і дубльованих фрагментів показало, що невербальні вокалізації, такі як кашель, чхання та ковтання, часто зазнавали змін або були вилучені, що призводило до втрати комічних і емоційних відтінків оригіналу. У меншій кількості були представлені ауральні вокалізації, зокрема хрипіння та утруднене дихання, що негативно позначалось на синхронізації з візуальним рядом і знижувало емоційне наповнення одиниць. Результати дослідження підтверджують необхідність добірного відтворення акустичних і емоційних характеристик невербальних вокалізацій, оскільки саме вони забезпечують коректність, глибину та природність дубльованого аудіовізуального тексту.

Ключові слова: акустична фонетика; дубляж; вегетативні та ауральні звуки; програмне забезпечення Praat; англійська мова; українська мова; серіал “Friends”.

Introduction

The study of sneezing, coughing, swallowing, and other involuntary and semi-voluntary physical actions in audio-visual translation has long been overlooked in academic research. However, this trend appears to be changing, as rare studies and publications on this topic have appeared over the last few decades. Initially, the concept emerged through research on paralanguage (Chaume, 1997; Zabalbeascoa, 1997; Poyatos, 2009; Alós, 2023; Padilla, 2023; Cruz, 2023), where scholars examined the role of non-verbal elements in communication. Over time, the academic focus shifted more towards the direct integration of non-verbal information in dubbing. This evolution is evident in subsequent studies (Matamala, 2011; Trouvain & Truong, 2012; Oktem, Farrús, & Wanner, 2017; Sánchez-Mompean, 2020; Zárate, 2021; Longari, 2023), that analyse how careful attention to non-verbal information contributes to the expressive quality of dubbed speech anong with a naturalistic viewing paradigm (Bellini et al., 2024). In particular, Sánchez-Mompeán (2020) provides a detailed analysis of prosodic features, highlighting how subtle vocal cues can be even more

effective than facial expressions and gestures in conveying meaning. This body of work emphasises the importance of prosody as a 'silent' yet powerful element of dubbing. This aspect is becoming increasingly important. However, the classification of physical actions still remains somewhat unclear. Poyatos (2009) refers to them as "ingressive or egressive, voluntary or involuntary, conscious or unconscious, independent, single or compound nonverbal segmental utterances" highlighting their complexity and variability. Other scholars use different terms to describe similar phenomena. For example, Longari (2023) refers to them as 'paralinguistic qualifiers', Matamala (2011) compares them to 'incidental speech', and Chaume (1997) uses the term 'paralanguage alternant', particularly when discussing cues that lack a standard written form. In the context of dubbing films, Trouvain and Truong (2012) accurately categorise behaviours related to physical actions as non-verbal vocalisations. This term refers to aspects of communication that accompany speech through non-verbal signals transmitted via the acoustic channel. Despite these terminological differences, there is a growing effort to analyse such vocal behaviours, especially those that are involuntary or semi-voluntary, not just as physiological reactions such as slips of the tongue or irrepressible coughing (Longari, 2023), but as meaningful elements within character speech. These physical actions often serve a prosodic function and can, in some cases, be represented in written form or fully dubbed, depending on their acoustic and contextual characteristics. Their inclusion in the dubbing process is crucial because they often convey a character's emotional or psychological state. This contributes to the depth, realism, and emotional authenticity of the performance. As a result, viewers are more likely to relate to the character, which creates believable audio-visual experience.

Method

Research Model

The research process includes the idea of assessing the naturalness of a dubbed version by analysing the acoustic features of non-verbal vocalisations. This is based on the concept that most people may experience the same states equally, such as swallowing, sneezing, stuttering, wheezing, and more. Drawing on their similar characteristics, the practical section of the article takes a two-sided approach. Firstly, it describes the contextual characteristics of non-verbal vocalisations, focusing particularly on the factors in the original passage that could have a natural effect on the dubbed version. Considering these factors, we closely examine the acoustic features of both the dubbed and original

versions. These features determine the subtle nuances that sound directors and translators can pay close attention to during the redubbing process.

Corpus-Based Approach to Term Selection and Categorisation

The challenge of providing practical and theoretically sound definitions of non-verbal vocalisations stems from the limited understanding of their acoustics and the diverse communicative functions they may serve. Firstly, they generally lack clear articulation and primarily consist of audible air frictions (Poyatos, 2009). Secondly, they can be represented in writing through textual manifestations of audible elements (textual paralanguage) (Alós, 2023). Finally, while non-verbal vocalisations are primarily involuntary, they can sometimes be intentionally triggered or modified, especially for comedic or stylistic purposes. Taking all of their functions into consideration, the most suitable classification seems to be the one proposed by Trouvain and Truong (2012) especially highlighting some of them as: (1) vegetative sounds (e.g. swallowing, coughing, sneezing and stuttering); (2) interjections; (3) filler sounds (hesitation sounds) and (4) melodic utterances (humming, singing and whistling).

Before applying this classification to our research corpus, we considered three main behavioural factors related to physical actions and emotions: (1) audible reflexive acts (such as real or mimicked coughs, sighs, and laughter), (2) textual representations (e.g., [exhales loudly], [swallows hard]). The third factor includes verbal expressions that accompany physical actions (for example, "Abso...[swallows hard]...lutely"). They are partially integrated into spoken delivery for realistic effect. Collectively, these manifestations fall under the broader category of functions designed to evoke emotions and elicit physiological reflexes while also preserving the pragmatic intention of the original line. That is why our analysis of non-verbal vocalisations focuses on involuntary and semi-voluntary behaviours related to physical actions. We also pay careful attention to emotions when they play a vital role in describing the characters' physical states. Based on the results of our research and the frequency of various types of non-verbal vocalisations in *Friends*, we will prioritise vegetative sounds and expand our research to include aural sounds. These sounds can greatly contribute to conveying non-verbal messages in dubbing. Aural sounds encompass any sound perceived by the ear, including laboured breathing, wheezing, or other sounds resulting from, for example, allergic reactions or a quick walk. Unlike vegetative sounds and other types of non-verbal vocalisations, these sounds can convey physical states or emotions without the use of words, yet still carry deep meaning and the potential for textual translation.

So, we adopt non-verbal vocalisations as the central analytical term, following Trouvain and Truong (2012). They are defined as distinct, non-lexical acoustic events that accompany or interrupt speech and convey physiological or emotional states. They represent a narrower category within paralanguage, which broadly encompasses prosody, intonation, and other vocal features (Poyatos, 2009). This classification enables us to treat non-verbal vocalisations as acoustic physical actions, which means that they are neither purely linguistic nor purely visual, but which play a significant role in dubbing by shaping character realism and emotional authenticity.

Collection Procedure and Data Analysis

Sample movie data was collected from four seasons of the American television sitcom *Friends* (1994-2004) and its Ukrainian-dubbed version. The data was sourced from the Kyivstar TV platform. Additionally, we used English subtitles transcribed by Eric Aasen and Mindy Mattingly Phillips. Since there are no official dual-language (English-Ukrainian) transcripts of *Friends* available, we used Transkriptor to generate subtitles and gather additional data, such as conversation sentiment and speaking time.

The Ukrainian version of *Friends* was dubbed using voice-over translation instead of full voice acting. Therefore, we used the audio editing software Audacity to isolate the background audio from the Ukrainian track in order to capture the subtle details of non-verbal vocalisations in the dubbed version. The files were first uploaded in WAV format, and the Voice Extractor tool was applied to remove background sounds, leaving only the vocal track. This process produced a clean audio stem suitable for precise phonetic analysis. Audacity enabled the extraction of low-intensity non-verbal cues, which are critical for studying differences between the original English track and the Ukrainian dub. Without this step, many subtle acoustic events would have been obscured, reducing the accuracy of spectrographic and quantitative analyses in Praat. To analyse non-verbal vocalisations in detail, all audio segments were imported into Praat, where they were examined using spectrogram, pitch, intensity, and formant displays. A TextGrid was created to label each non-verbal vocalisations (such as breath, sigh, cough) and perform precise measurement of timing, duration, and acoustic features. Spectrogram settings were adjusted to highlight subtle acoustic details. For example, breathing sounds typically have a higher pitch compared to sneezing, while swallowing generally produces a lower tone than stuttering speech. Paying close attention to these subtle acoustic differences can enhance the authenticity and clarity of dubbing, ensuring that non-verbal cues are conveyed effectively.

Following these steps, we collected samples of non-verbal vocalisations from various episodes and analysed their acoustical features. To this end, we used qualitative methods (observations), quantitative methods (corpus-based approach) and specific dubbing and translation technique analysis (such as analysing sound, functional equivalence, error analysis, contrastive analysis). The study hypothesised that non-verbal vocalisations were often omitted in the dub version because voice-over dubbing prioritised meaning over the specific phonetic details.

Results

A total of approximately 1,050 non-verbal vocalisations were identified across four seasons of *Friends* and their Ukrainian dubbed versions. Vegetative sounds dominated the corpus (40%), followed by aural sounds (26%), filler/interjections (19%), and melodic utterances (15%) (see Table 1). This confirms that reflexive physiological sounds constitute the most frequent and therefore most critical category in dubbing practice.

Table 1
Distribution of Non-Verbal Vocalisations Across the Corpus

Category	**Tokens (n)**	**% of total (1,050)**	**Examples**
Vegetative sounds	420	40	Coughing, sneezing, swallowing, stuttering
Aural sounds	275	26	Wheezing, laboured breathing, groaning
Filler/interjections	195	19	“Uh”, “erm”, hesitation noises
Melodic utterances	160	15	Singing and whistling
Total	**1050**	**100**	

Vegetative Sounds

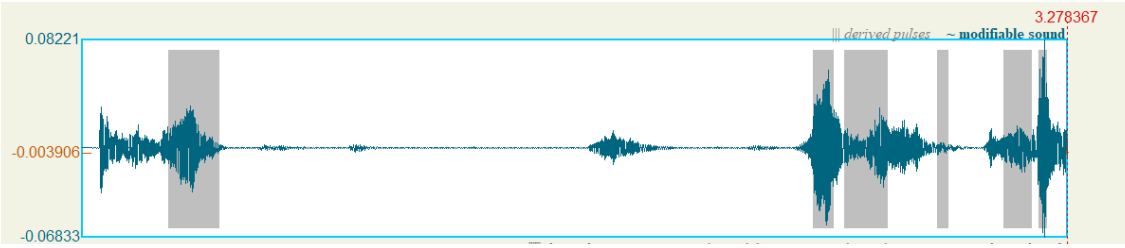
Vegetative non-verbal vocalisations were often altered or reduced in the Ukrainian dub. For instance, Monica’s forced swallowing during “Absolutely” was noticeably shortened (3.28s – 2.49s), with lower amplitude and energy (see Table 2). This diminished the perception of strain that reinforced the comedic effect of eating “Mockolate”:

(1) 06:11 Monica: **Abso...[swallows hard]...lutely.** See, I love creating new recipes. I love Thanksgiving. And, well, now, I love Mockolate.

06:11 Моніка**: Абсо...лютно.** Я люблю створювати нові рецепти, люблю День Подяки і тепер люблю Маколат.

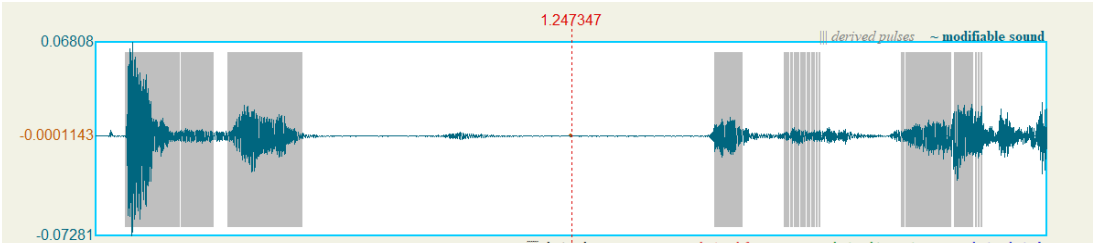
This is more clearly demonstrated in the Pratt program, where the first spectrum corresponds to the original text (see Fig. 1), and the second to the Ukrainian adapted version (see Fig. 2).

Figure 1
Spectrogram for the word "Absolutely" as Spoken by Monica in the Original Version



Monica was responsible for promoting a very sticky chocolate with a chewy texture, which was difficult to eat and swallow. In fact, this effect was partially achieved by equalizing the number of syllables, but the spectrums clearly show differences in length and stress of these words, even considering their similar stress patterns with a primary stress on the third syllable (comp. *absolutely* /ˌæb.səˈ**lu**ːt.li/; *абсолютно* [абсоул`ýтно]).

Figure 2
Spectrogram for the Word "Absolutely" as Spoken by Monica in the Dubbed version



The English token lasted 3.28 seconds, compared to 2.49 seconds in the Ukrainian dub (Table 2). The dubbed version also showed lower amplitude (.073 Pa vs. .082 Pa) and total energy (40 %). These differences reduce the

perceived strain and rhythmic emphasis of the original, which conveyed Monica's difficulty swallowing Mockolate (Figure 1).

Table 2
Acoustic Comparison of "Absolutely"

Parameter	English	Ukrainian	Difference
Duration (s)	3.28	2.49	- .79
Max amplitude (Pa)	.082	.073	- .009
Total energy (Pa^2s)	8.37e-05	4.99e-05	- 40%

A similar reduction occurred in nasalized speech caused by Monica's cold ("I'm fine-d"). The Ukrainian version not only shortened delivery (4.88s – 1.88s) but also reduced amplitude, weakening both comedic rhythm and lip synchronisation (see Table 3).

(2) 02:31 Monica: (wiping her nose) *I have not been sick in over three years!* (Sneezes.)
02:36 Chandler: *I'm gonna grab you some tissue.*
02:37 Monica: *I don't need a tissue!* ***I'm fine-d!***
02:40 Ross*: When you put a 'D' at the end of 'Fine' you're not fine.*
02:47 Monica: ***I'm fine-d. I'm fine-d!*** *Y'know, it's a really hard word to say.*

02:32 Моніка: *Я понад три роки не хворіла!*
02:36 Чендлер: *Я принесу серветки.*
02:37 Моніка: *Обійдусь без серветок! Я в нормі.*
02:41 Рос*: Коли ти кажеш не «в нормі», а «**в дорбі**» це уже ненормально.*
02:47 Моніка: ***Я в днормі, я в днормі****! Слово дуже важке!*

Sneezing speech might change the way vowels and consonants are pronounced. It affects lip movements or timings, making the dubbed speech harder to sync with the original lip movements. For instance, in American English, the letter "d" typically represents the voiced alveolar stop consonant sound /d/. This sound is produced by placing the tip of the tongue on the ridge behind the upper front teeth and briefly stopping the airflow before releasing it, while also vibrating the vocal cords. From the other hand, the Ukrainian letter "б" in the phrase «в дорбі» represents a bilabial voiced plosive sound produced by bringing both lips together and then releasing them with a burst of air while the vocal cords vibrate. Consequently, it does not interfere with lip-syncing. The phrase "в днормі" with a voiced dental plosive /d/ seems to be a better variant for creating the effect of sneezing speech. Nonetheless, the

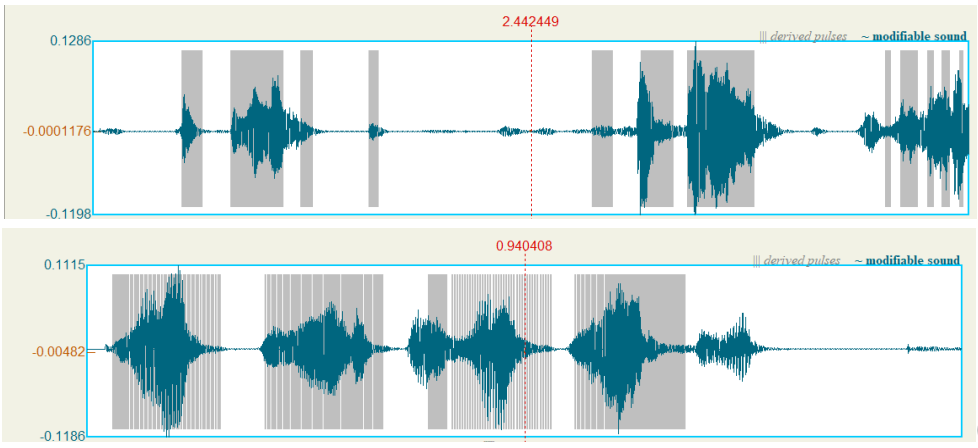
dubbed phrase was delivered at a significantly faster rate, with no synchronization to the character's lip movements.

Table 3
Acoustic Comparison of "I'm fine-d"

Parameter	English	Ukrainian	Difference
Duration (s)	4.88	1.88	- 3.00
Max amplitude (Pa)	.129	.111	- .017
Δ Sound level (dB)	_	≈ 1.2 dB	-

One more factor we should consider in the context of dubbing sneezing speech is the loudness and intensity of the sound wave. Since a hyponasal voice is typically caused by a blockage in the nasal passages, which prevents air from escaping properly, the character would have difficulty pronouncing consonants that require high air pressure (Watson, 2018). Conversely, they might try to intensify pressure consonants or "mouth sounds" in an effort to be clearer and more understood. Monica in this case has a slightly higher-pitched voice with maximum amplitude of .13 Pascal, which proves her struggle in pronouncing certain phonemes. In the Ukrainian dubbed version, the maximum amplitude is .12 Pascal. The difference indicates a subtle variation in sound intensity of nearly .02 Pascal. The intensity of the sound wave is represented on spectrograms by the blue colour.

Figure 3
Spectrogram for the Phrase "I'm fine-d. I'm fine-d!"
(a) Intensity of the Sound Wave in English; (b) Intensity of the Sound Wave in Ukrainian



Stuttering, a recurrent feature of Ross's speech, was often omitted. In Season 6, when Ross announced his lectures at NYU, the dub replaced his exaggerated stutter with a flat delivery. This resulted in the loss of character-specific rhythm and emotional nuance.

(3) 1:30 Ross: (on phone): No-no-no, that's great! I'll be there Monday. And thank you again! (Listens) Okay. (Hangs up) (To All) Umm, that was the head of the Paleontology department at **N uh, Y, uh U**!
1:45 Monica: Wow! **Uh** what, did he **uh**, say

1:30 Рос: Ні, ні, влаштовує у понеділок буду! Ще раз Вам дякую, добре! Це був начальник кафедри палантеології Нью-Йоркського університету.
1:44 Моніка: І що він сказав?

In the Ukrainian translation, the dubbing actor simply overlooked the character's non-verbal vocalisations and went with a robotic tone instead. However, in similar situations, it is important to use alternative methods to avoid losing the phonographic peculiarities that were deliberately chosen for specific reasons in the original work (Costabile 2024, p. 41). In this context, achieving can be done by stretching and emphasising certain words to reflect the characters' feelings. For example, the pitch could be raised on the letters "Н(ью-ю)-Й(о-оркськo-ого)" and "Щ(о-о) В(і-ін)", placing equal emphasis on Monica's mockery and Ross's non-verbal vocalisations. To accurately perform prolongation according to the principle of isomorphism, the phrase 'Нью-Йоркського університету' (New York University) can be rephrased as 'університету Нью-Йорку' (University in NYC). Neutralizing the speech pattern can sometimes affect the phonological unit, but even small changes to pitch, tempo, length and rhythm can improve the overall meaning of the line. Therefore, the primary goal is to grasp the character's way of speaking, diction, emotions and even mood (Gudmanian et al., 2022).

Aural Sounds

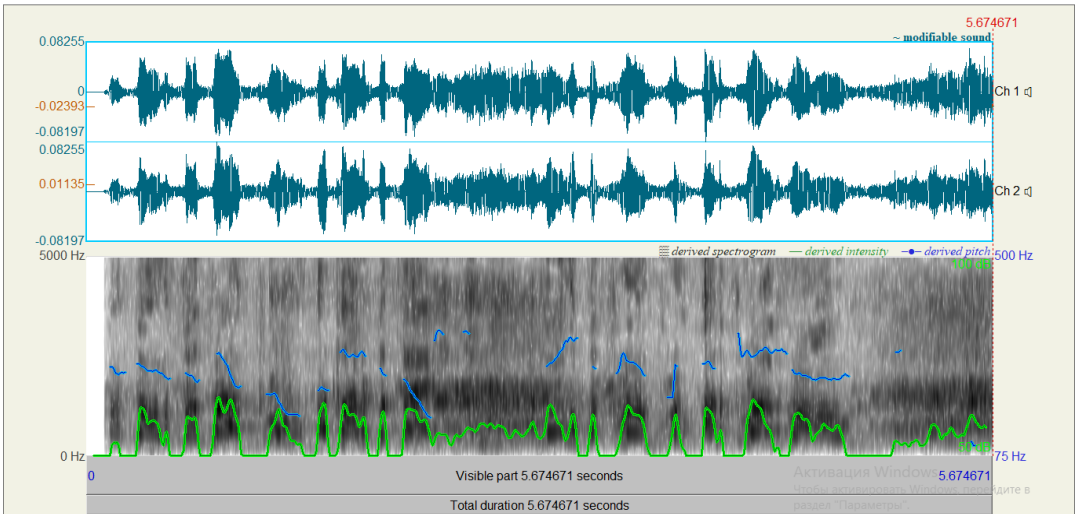
Aural non-verbal vocalisations such as wheezing and laboured breathing posed further challenges. In a scene where Ross suffers an allergic reaction, his slurred and obstructed speech ("thro up a thro thro") was left untranslated, creating a semantic and acoustic gap. The English delivery lasted 5.68s with rapid tempo and multiple pauses, whereas the Ukrainian version omitted these features entirely (see Fig. 4):

(4) 03:45 Ross: Ok, well, if you do take him out for his walk, you might wanna bring his hat, and there's extra milk in the fridge, and there's extra diapers in the bag.

03:52 Joey: Hat, milk, got it.
3:59 Ross: ??? (speech garbled)
Thro up a thro thro--a thro thro!
04:02 Joey: Consider it done.
04:03 Chandler: You understood that?
04:05 Joey: Yeah, my uncle Sal has a really big tongue.

03:45 Рос: Якщо захочете піти з ним на прогулянку, то не забудьте надіти йому шапочку, в холодильнику є ще молочко, а памперси в сумці.
03:52 Джої: Шапочка, молочко, зрозумів…
3:59 Рос: -
04:02 Джої: Вважай, що все зроблено.
04:03 Чендлер: Ти це зрозумів?
04:
05 Джої: Так, у мого дядька теж був великий язик.

Figure 4
The Praat-Generated Spectrum of the Phrase "thro up a thro thro--a thro thro"



The naturalness of a dubbed phrase can also be measured by the speech pitch of the character, as indicated in blue on the screenshot. The adapter could also imitate the speech changes caused by allergies. All individuals, regardless of language, may experience the same diseases equally. For example, the effect of the character's slurred speech could be achieved by using short bursts with pauses, pronouncing vowels instead of consonants (or vice versa), placing emphasis on the wrong parts of words, speaking rapidly, or using a nasal, raspy or strained voice. For a more precise phonetic analysis, one could focus on sounds that are acoustically similar and relevant in the Ukrainian language: *thro up* [трохи] *a thro thro--a thro thro* [три рази].

In another example (Season 8), Ross's wheezing and heavy exhalation during class were absent from the dub. Acoustic analysis revealed 7,609 frames

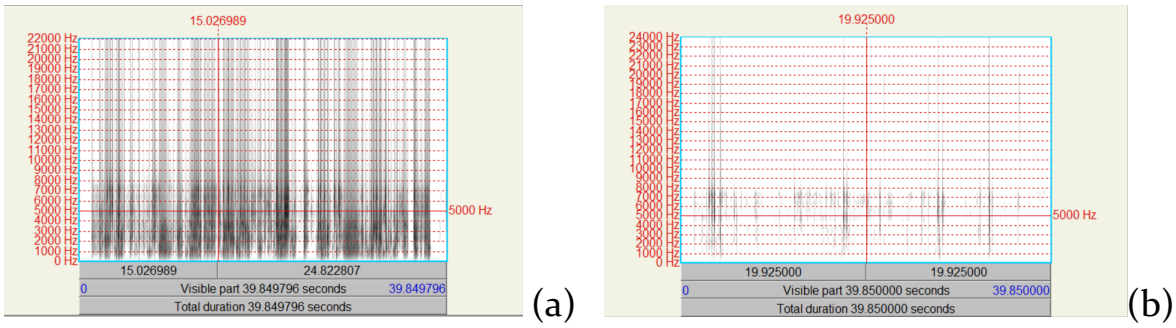
in the English version versus 7,964 in the Ukrainian. However, instead of expressive non-verbal vocalisations, the added frames contained extended neutral speech, stripping the scene of its intended breathlessness:

(5) [Scene: N.Y.U, Ross's new class, he is entering out of breath.]
17:28 Ross: Yeah!!!! Yes, I made it! I'm on time! (Grabs a bottle of water from a student, takes a drink, and splashes some on his face like the marathon runners do.) Okay, why don't we all **uh,** (**Exhales loudly**) open our books to page 23. Where… (**Exhales again**) Where you will see…a **uh**…a bunch of **uh**…red spots. Okay, (Closes his book.) **umm**, why don't, why don't you all start to read, while I – (Passes out and collapses.)

17:23 Рос: Так! Я встиг (short pause) вчасно! Добре! Розгорніть, будь ласка, свої підручники на 23 сторінці, де ви побачите (short pause) червоні цятки! Ви почніть читати, а я поки що…

In the original English version, these moments are carefully timed and acoustically rich, with subtle variations in pitch, breathiness, and duration.

Figure 5
Pitch-Dependent Spectrogram for Aural Sounds in (a) English and (b) Ukrainian

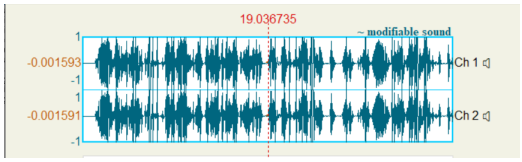


The spectrogram of these sequences shows a detailed and dynamic acoustic profile, with a clear increase in frame count due to the fine temporal structure of the wheezes and loud exhales[1]. However, when comparing this to the Ukrainian dubbed version, it becomes apparent that many of these non-verbal, expressive cues were completely omitted. Instead, the dubbed performance relies on neutral, monotonous speech, which strips the scene of its original

[1] The spectrograms also show the precise start point of a wheezing sound in the original recording versus the dubbed version (selection start). The timestamp of a selected segment in the original line starts earlier at 15.026989 seconds, while in the dubbed version it starts at 19.924897959183674 seconds. This difference is due to the original sound line remaining unchanged in the dubbed version. Therefore, we will not take into account the selection start and selection end shown in red on both diagrams. Instead, we only include this information to clarify the numbers on the spectrograms.

rhythmic variation and expressive sound detail. This creates a mismatch with the vocals and significantly reduces the narrative impact. Furthermore, stereo channels present in English (see Fig. 6a) were collapsed to mono in Ukrainian (see Fig. 6b), reducing immersion and detail.

Figure 6
Spectrograms for Audio Streams (Number of Channels) in (a) English and (b) Ukrainian



(a)

(b)

They demonstrate the difference in the number of audio channels in CH1 and CH2, as abbreviated on the spectrogram (a). The English version has two channels, left and right, which provide a more immersive listening experience. In contrast, the Ukrainian version only has a single channel, so sounds like wheezes, sniffles and exhales are not distinguished. Consequently, the audio feels artificial and less immersive.

In general, the vegetative and aural sounds in the English and Ukrainian versions reveals significant differences in the dubbing strategies that effect deeply original speech rhythm and emotional nuance of character's speech. The omission of wheezing, slurred speech, and labored breathing also weakens the connection between the characters' physical states and their speech patterns. In scenes like Ross's allergic reaction and heavy exertion (e.g., the class scene), the original non-verbal vocalizations are replaced with more neutral speech, removing the physicality of the performance. This substitution diminishes the sensory and emotional engagement that the original sounds provided. In general, the translators should be mindful of the texture of their texts, recognizing that the message is communicated to the audience through a blend of visual and auditory elements. For instance, Zabalbeascoa (1997) suggests that dubbing translation strategies can be either purely verbal (adding, substituting or deleting words in the target language, paraphrasing, defining, or increasing verbal redundancy), purely non-verbal (adding, changing or deleting nonverbal signs in the form of pictures, colours, gestures, mimicry, sounds), or combination of both to preserve meaning across various language levels. Nevertheless, in the analysis of non-verbal vocalisations, both vegetative and aural sounds tend to be either omitted or reduced (see Table 4). This results in a reduction of physical struggle and emotional intensity, and in some cases, a loss of humor.

Table 4
The Translators' Methods for Non-Verbal Vocalisations

Sound Type	**Sound Feature**	**Methods**	**%**	**Content Impact**
Vegetative sounds	Stuttering	Omission	25	Loss of character-specific rhythm and emotional nuance; replaced with neutral tone.
	Swallowing	Reduction	20	Loss of comedic strain and rhythm, reducing the impact of the scene.
	Sneezing	Substitution	10	Substituted more neutral phrasing.
	Coughing	Replacement with neutral speech	15	Weakened comedic rhythm and lip-syncing issues due to faster speech rate.
Aural sounds	Slurred speech	Paraphrasing	10	Rephrased hesitation into a more neutral, less hesitant phrasing.
	Wheezing	Replacement with neutral speech	5	Reducing the intended physical struggle and humor.
	Heavy exhalation	Replacement with neutral speech	15	Loss of intended breathlessness and physical struggle. Creates an acoustic gap and weakens the emotional intensity of the scene.

The primary issue in both categories is a tendency to neutralize the original speech features, which affects not only phonetic authenticity but also the characterization and comedic timing intended by the original performance. To maintain the naturalness of the dub, translators could consider imitating the speech changes caused by conditions like allergies or stuttering using techniques such as short bursts, pauses, or nasal, raspy, or strained voices. The following example illustrates a translation decision that employs an effective and appropriate technique.

(6) 15:59 Monica: (entering) Chandler, I think I'm sick.
16:03 Chandler: Really? **Struck down in the pribe of libe!**

15:59 Моніка: Чендлере, по-моєму, я захворіла
16:03 Чендлер: Серйозно! **Злягла у досквіті сил.**

In the line "struck down in the pribe of libe" Chandler imitates Monica's sneezing, and the Ukrainian dub "злягла у досквіті сил" attempts to replicate the nasalized sounds, although the syllables could be slightly prolonged to better match the original rhythm. In this specific example, translatour preserves the character's rhythm, emotions, and humor by using similar sounds in Ukrainian to match the original speech features. Although these translation choices are numerically limited, they are crucial for redubbing, as they enhance phonetic authenticity, preserve emotional nuance, and sustain the original comedic effect. This demonstrates that small, deliberate techniques are valuable tools for maintaining character identity and performance integrity in dubbing.

Discussion

This study examined how non-verbal vocalisations are treated in the Ukrainian dubbing of *Friends*, focusing on vegetative and aural sounds. Acoustic analysis demonstrated that dubbing often reduced or omitted these cues, with measurable effects on duration, amplitude, and energy. The results focused on understanding how similar physical actions maintain consistent acoustic characteristics across English and Ukrainian languages. The findings offered insights into the relationship between context and the acoustic properties of non-verbal vocalisations. This partially sheds light on how these elements are adapted in the dubbing process. This is done in order to decode the meaning and emotional impact of the non-verbal vocalisations in the target language. The initial finding pertains to vegetative non-verbal vocalisations, including swallowing and nasalised speech, which were consistently attenuated through shortening and weakening. Such modifications diminished the perceptible strain, rhythm, and comedic timing characteristic of the original performance. In a similar vein, Ross's stutter, an essential marker of characterisation, was frequently neutralised, yielding a less dynamic delivery. These observations corroborate earlier assertions that dubbing practices tend to privilege semantic transfer over phonetic fidelity (Fresco, 2009; Miggiani, 2019). Aural non-verbal vocalisations, such as wheezing and laboured breathing, were also frequently omitted or substituted with monotone speech. These alterations disrupted synchrony with visual cues and reduced the overall sense of immersion. Spectrographic analysis further revealed a loss of stereo richness in the Ukrainian version, resulting in a flattened soundscape. Such omissions align with Chaume's (1997) observation that non-verbal elements are often relegated to a secondary status despite their communicative significance. As a result, our hypothesis that people experience the same states in the same way does not

fully apply to the counterparts we analysed. The main reason is the linguistic nature of English, which tends to be more concise than other languages. English often requires amplification to match the length of the line. This results in a discrepancy between the length of vocal cues and the length of the spoken word. Moreover, English is a stress-timed language characterised by reduced vowels, extensive use of weak forms, and strong rhythmic compression. In contrast, Ukrainian is syllable-timed, maintains fuller vowel quality, and relies more heavily on melodic intonation contours. These differences not only affect the duration of utterances but also impact the treatment of pauses, bursts of air, and breath-based non-verbal sounds. Secondly, differences in speech tempo between English and Ukrainian contributed to mismatches in rhythm and lip synchronisation. English typically has shorter vowel durations and quicker articulatory transitions, which naturally accommodate gasps, exhales, and stutters in speech. In contrast, Ukrainian demands more precise articulation of consonants, including palatalised forms, leaving less room for short vegetative sounds without sacrificing clarity. This phonetic difference likely played a role in translatours choosing to reduce certain non-verbal cues. Thirdly, the main reason for the naturalisation of non-verbal vocalisations is that voice-over dubbing prioritises meaning over specific phonetic details. Nevertheless, we have provided a statistical overview of each non-verbal cue analysed in terms of its length, breadth, tempo and sound intensity. In separate cases, we found that we could reproduce the non-verbal vocalisations in the original message by copying acoustic characteristics of various physical states (for example, substituting vowels for consonants, matching the number of syllables, emphasising the wrong parts of words). These findings indicate that higher fidelity can be achieved when translators pay attention to acoustically analogous phonetic elements in Ukrainian. This can be done by adjusting vowel openness, modifying consonant resonance, or using controlled nasal, raspy, or strained vocal qualities. These strategies help accurately reproduce physical states, such as allergies, nasal obstruction or laboured breathing. This would definitely enhance synchronisation and ensure that each non-verbal vocalisations carries the same depth and authenticity, thereby preserving the acoustic and emotional intent of the original performance.

Conclusions

Overall, the findings from the data analysis suggest that translators should carefully consider the phonographic features of the script and provide precise translations of both physical speech characteristics and emotional expressions

to ensure that the intended meaning is conveyed accurately. Most non-verbal vocalisations tend to be more directly tied to their sound or emotional expression. That is why it is important to note that translation of non-verbal vocalisations can be a complex and nuanced task that requires careful attention to tone, rhythm, intonation, and stress patterns in both the source and target languages. Additionally, the emotional context and certain involuntary bodily actions, such as swallowing, sneezing, stuttering, wheezing, and others must also be carefully examined. These elements add depth and realism to a character, making them more believable and relatable to the reader. Without accurate sound matching translation outputs may contain significant inaccuracies. This study highlights how vegetative or aural sounds depicted in the TV show *Friends* posed specific challenges in the process of dubbing. It also finds that emotional content had to be carefully translated to produce a similar emotional response in the target audience as in the original. Accordingly, it is hoped that the current scientific article will contribute to the existing body of work on this issue by shedding light on other types of non-verbal vocalisation that have been overlooked in translation studies.

Acknowledgements

The study is carried out within the framework of the project *Methods and Tools for Optimising the Parameters and Resources of Electronic Communication Networks and Information and Communication Systems Using Artificial Intelligence (AI)* at the Educational and Scientific Institute of Information Technologies, State University of Information and Communication Technologies, Ukraine.

Disclosure Statement

The authors reported no potential conflicts of interest.

References

Alós, J. (2023). Paralanguage in the translation of children's graphic novels into Arabic: Jeff Kinney's Diary of a wimpy kid. *Children's Literature in Education, 56*(2), 256–272. https://doi.org/10.1007/s10583-023-09558-4

Bellini, C., Del Maschio, N., Gentile, M., Del Mauro, G., Franceschini, R. & Abutalebi, J. (2024). Original language versus dubbed movies: Effects on our brain and emotions. *Brain and Language, 253*, 105-424. https://doi.org/10.1016/j.bandl.2024.105424

Brannon, W., Virkar, Y., & Thompson, B. (2023). Dubbing in practice: A large scale study of human localization with insights for automatic dubbing. *Transactions of the Association for Computational Linguistics, 11,* 419–435. https://doi.org/10.1162/tacl_a_00551

Chaume, F. (1997). Translating non-verbal information in dubbing. Benjamins Translation Library, 315–326. https://doi.org/10.1075/btl.17.25cha

Chaume, F. (2020). The language of dubbing: A matter of compromise. Audiovisual Translation: Dubbing, 81–99. https://doi.org/10.4324/9781003161660-5

Consonants and vowel sounds. (n.d.). Scribd. Retrieved December 5, 2025, from https://www.scribd.com/document/646908639/Consonants-and-Vowel-Sounds

Costabile, L. (2024). English-language cyberdubbing in the age of artificial intelligence [Unpublished master's thesis]. Università degli Studi di Pavia. https://unitesi.unipv.it/retrieve/2e56bf1d-b07c-4307-b664-e6ec45e49dd5/tesi%20definitiva%20formato%20pdfA.pdf

Đurić, D. M. (2020). Free software tools for computational linguistics: An overview. Zenodo. https://doi.org/10.5281/zenodo.4073971

Fresco, P. R. (2009). Naturalness in the Spanish dubbing language: A case of not-so-close *Friends*. *Meta, 54*(1), 49–72. https://doi.org/10.7202/029793ar

Gudmanian, A., Struk, I., & Myshko, A. (2023). Lip-synch-translation of onomatopoeic words in the American sitcom "F.R.I.E.N.D.S." In K. Lisetskyi (Ed.), Third International Online Conference "Corpora and Discourse." National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." https://corpora.kamts1.kpi.ua/cad-2023/paper/view/29305/17153

Hassan, A. Y., Abdullah, S. S., Al-Rawe, M. F. A., & Hasan, M. A. (2024). The use of relevance and euphemisms in translating taboo expressions in the subtitles of the TV show "Friends." *Kurdish Studies, 12*(2), 337–345. https://doi.org/10.58262/ks.v12i2.028

Matamala, A. (2011). Dealing with Paratextual Elements in Dubbing: A Pioneering Perspective from Catalonia. *Meta,56*(4), 915-927. https://doi.org/10.7202/1011260ar

Miggiani, G. S. (2019). Dialogue writing for dubbing: An insider's perspective. Palgrave Macmillan.

Miggiani, G. S. (2021). English-language dubbing: Challenges and quality standards of an emerging localisation trend. *Journal of Specialised Translation, 36,* 2–25. https://doi.org/10.26034/cm.jostrans.2021.054

Öktem, A., Farrús, M., & Wanner, L. (2017). Automatic extraction of parallel speech corpora from dubbed movies. In S. Sharoff (Ed.), Tenth Workshop on Building and Using Comparable Corpora (pp. 31–35). Association for Computational Linguistics. https://doi.org/10.18653/v1/w17-2506

Orozco-Jutoran, M. (2001). Translation and dubbing: Words, voices and images. *Studies in Translatology, 9*(1), 45–46.

Padilla Cruz, M. (2022). Paralanguage and ad hoc concepts. *Pragmatics: Quarterly Publication of the International Pragmatics Association, 33*(3), 343–367. https://doi.org/10.1075/prag.21059.pad

Poliakova, O. V., & Struk, I. V. (2017). Rozmovna pryroda animatsiinykh filmiv [The spoken nature of animated films]. In A. Malek (Ed.), International Research and Practical Conference Modern Philology: Relevant Issues and Prospects of Research. Maria Curie-Skłodowska University. https://er.nau.edu.ua/items/b2403c5d-d843-42af-a8c3-a67ffd480265

Sánchez-Mompeán, S. (2020). The Prosody of Dubbed Speech. Beyond the Character's Words. Palgrave Macmillan.

Sisto, A. (2014). Film Sound in Italy: Listening to the Screen. Palgrave Macmillan.

Supardi, M., Putri, D. A. (2018). Audio-Visual Translation Techniques: Subtitling and Dubbing of Movie Soundtrack in Frozen: Let it Go. *Buletin Al-Turas, 24*(2), 399–414. https://doi.org/10.15408/bat.v24i2.8621

Trouvain, J., & Truong, K. P. (2012). Comparing non-verbal vocalisations in conversational speech corpora. In L. Devillers, B. Schuller, A. Batliner, P. Rosso, E. Douglas-Cowie, R. Cowie, & C. Pelachaud (Eds.), Fourth International Workshop on Corpora for Research on Emotion Sentiment and Social Signals (ES3 2012) (pp. 36–39). European Language Resources Association. http://www.lrec-conf.org/proceedings/lrec2012/workshops/18.Proceedings%20ES3%202012.pdf

Zabalbeascoa, P. (1997). Dubbing and the nonverbal dimension of translation. In F. Poyatos (Ed.), Nonverbal communication and translation (pp. 327–342). John Benjamins Publishing Company. https://doi.org/10.1075/btl.17.26zab

Zárate, S. (2021). Captioning and subtitling for d/Deaf and hard of hearing audiences. UCL Press. https://doi.org/10.14324/111.9781787357105

Sources

Audacity Team. (n.d.). Audacity [Computer software]. Retrieved December 5, 2025, from https://www.audacityteam.org/

Corbari, E. (n.d.). Friends' scripts. Retrieved December 5, 2025, from https://ederson-corbari.github.io/friends/

Friends [Television series]. (n.d.). Kyivstar TV. Retrieved December 5, 2025, from https://tv.kyivstar.ua/ua

Transkriptor. (n.d.). Video to text converter. Retrieved December 5, 2025, from https://transkriptor.com/video-to-text/

Watson, S. (2018, September 18). What it means to have a nasally voice. Healthline. https://www.healthline.com/health/nasal-voice