


Anna Bielska  <https://orcid.org/0000-0003-0614-354X>  
Akademia Humanistyczno-Ekonomiczna w Łodzi  
e-mail: [anna.bielska@ahc.email](mailto:anna.bielska@ahc.email)

# Who is a better translator – AI or human? A viewer-based study on AVT

## Abstract

This article investigates the differences between human and AI-generated translations of film dialogue, using *Shazam! Fury of the Gods* (2023) and its Polish dubbed version as a case study. By selecting linguistically and stylistically challenging fragments, such as idioms, humor, slang, and sarcasm, the study compares the two translations. The research is supported by a survey involving 80 Polish respondents, who were asked to guess which translations were done by a human and which by AI, and to indicate their preferences. Results show that most respondents could correctly identify the human version, though preferences were more varied. The analysis reveals that AI struggles with informal tone, idioms, and emotional depth, while human translations, despite occasional risks, better reflect the speaker's tone and cultural context. The study contributes to ongoing discussions in translation studies on the limits of machine translation in creative audiovisual texts.

**Keywords:** audiovisual translation, AI translation, dubbing, ChatGPT

## Introduction

In recent years, the rapid advancement of artificial intelligence (AI) technologies has led to significant changes in the field of translation, particularly in audiovisual media. With the growing availability of machine translation tools capable of generating fluent target-language content, an important question arises: can AI effectively replicate the creativity and cultural sensitivity of human translators, especially in a genre as linguistically dynamic as film dialogue?

This article shows the differences between AI-generated and human-translated film dialogue, on the basis of the 2023 film *Shazam! Fury of the Gods*, directed by David F. Sandberg and written by Henry Gayden and Chris Morgan, as a case study. The material

analyzed includes the original English dialogues and their Polish dubbed versions. Specifically, the study aims to show the contrast in translating dialogues by a professional translator, Jan Jakub Wechsile, under the direction of dubbing director Marek Robaczewski, with dialogues translated with the use of artificial intelligence.

The selected fragments were chosen based on their linguistic complexity and stylistic variety – particularly those containing idioms, humor, youth slang, and sarcasm – elements that often prove challenging for machine translation systems. Moreover, the AI-generated translations were created using ChatGPT with a prompt specifying the need to render the text humorously and naturally in Polish, which simulates an advanced user trying to guide the system toward more creative translation.

The study involves a comparative methodology, supported by a survey conducted among native Polish speakers (Appendix 1). 80 respondents, aged 18–42, were asked to read selected fragments of Polish translations, some produced by a human, others by AI, without being told which was which. They then indicated which version they preferred and which they believed to have been created by a human translator. Although language proficiency was not formally tested, participants were asked to confirm a minimum self-assessed English level of B1, which helped ensure they could adequately understand the original text while comparing the translated versions. The results of the survey reveal a general trend: while most respondents were able to correctly identify the human translation in the majority of cases, their preferences were more varied.

By focusing on a popular superhero film rich in humor, idiomatic expressions, and colloquial language, this article illustrates the strengths and limitations of current AI translation technologies. The aim is not only to identify the linguistic markers that distinguish human from AI translation but also to reflect on the implications for translation studies, dubbing practices, and the future of audiovisual localization.

## Human vs. AI translation

The distinction between human and machine translation has been an area of interest within translation studies. Traditionally, human translation has been valued for its sensitivity to context, culture and creativity. Human translators make decisions based on interpretation that reflect not only the semantic content of the source text (ST) but also its stylistic and pragmatic dimensions. These decisions can be made due to a complex understanding of both languages and cultures involved, and often draw on extensive professional experience.

Unlike human translation, AI is based on algorithms trained on massive collections of bilingual texts. Modern neural machine translation (NMT) systems such as those used in Google Translate, DeepL, or GPT-based tools – have demonstrated impressive fluency and syntactic accuracy. They handle EGP texts well and often sound natural in many situations. However, they still struggle with polysemy, irony, cultural references, idioms, and register shifts – elements that are central to the success of film dialogue, particularly in dubbed versions.

AI's main weakness is that it relies on patterns based on statistics, not human understanding or judgment. AI cannot infer meaning from extralinguistic factors such as tone of voice, visual cues, or character development. For example, a line delivered sarcastically may be rendered too literally, flattening the intended effect. Human translators are better at decoding such pragmatic signals and recreating them appropriately in the target language (TL).

In dubbing, the challenges are further amplified. The translated text must not only convey meaning but also fit the lip movements, rhythm, and emotion of the original performance. This requires a high degree of linguistic flexibility and creativity, which current AI tools cannot fully replicate. Moreover, the localization of humor, slang, or youth language often demands a deep understanding of social and cultural connotations that go beyond the lexical level.

Finally, from a theoretical perspective, the contrast between human and AI translation touches on broader debates in translation studies – particularly the tension between literal and dynamic equivalence (Nida, Taber, 1969), the translator's visibility (Venuti, 1995), and the ethics of representation in cross-cultural mediation. AI translation challenges the way we see translation by introducing a non-human factor into a fundamentally human process. It encourages scholars to reconsider what counts as authorship and real language use as technology continues to develop.

## Case study: human vs. AI translation in *Shazam! Fury of the Gods*

That first line is spoken by Billy Batson, the teenager who becomes Shazam, in a conversation with a therapist. In that moment, Billy is reflecting on how he, his foster siblings, and even the villain Dr. Sivana were all transformed, overpowered by 7 demons, and how it affected his family. This line highlights Billy's frustration with his family's drifting apart after they all gained powers. His speech is informal, disorganized, and intentionally over-the-top:

SL (00:06:50 – 00:06:57)	TL (Human translation)	ChatGPT
A doctor conjured seven demons and held my whole family hostage, so we had to rip out one of his eyeballs, and then everybody got superpowers, but now everyone's kind of, like, doing they own thing.	Jeden gościu wezwał siedem demonów i uprowadził mi rodzinę, więc wyrwaliśmy mu oko i wszyscy się zrobili supermocni, ale teraz każdy kręci swoje lody.	Jakiś lekarz przywołał siedem demonów i wziął całą moją rodzinę jako zakładników, więc musieliśmy wyrwać mu jedno oko, a potem wszyscy dostali supermoce, ale teraz każdy trochę robi swoje.

Source: own elaboration.

The chaotic tone, paired with colloquial and ungrammatical elements such as “they own thing”, contributes to the humorous and stylized character of the utterance. It reflects not only Billy’s personality and age but also the film’s playful narrative tone. The line combines exaggerated content, youthful syntax, and spoken errors, making it a challenging segment for any translator, especially one driven by AI.

At first glance, both translations preserve the sequence of events described in the original. However, significant differences emerge at the levels of tone, lexical choices, and idiomatic language.

The human version employs casual, teen language that matches the source character’s age and speech style. Phrases such as “jeden gościu” and “zrobili się supermocni” show the informality and special rhythm of the original. In particular, the translator chooses the idiomatic expression “kręci swoje lody” to mirror the phrase “doing they own thing.” While this idiom may not be universally familiar, it carries the same informal connotation and adds local flavor, making the character’s voice more colorful.

In contrast, the AI translation, while grammatically correct and coherent, opts for safer, more literal language. The choice of “lekarz” over the more casual “gościu” or “facet”, and “dostali supermoce” rather than “zrobili się supermocni”, exemplifies a consistent tendency toward lexical neutrality. The final phrase, “każdy trochę robi swoje”, is a direct but stylistically flat equivalent of the English idiom. It fails to convey the playful, expressive style of the original speech and reflects a more standard register.

This discrepancy is visible in the survey data. While 84% of respondents correctly identified the human translation, only 34% indicated a preference for it. This suggests that although readers recognized the human version as more idiomatic and expressive, they may have found some of its elements, such as “kręci swoje lody” either unfamiliar or too unconventional. This outcome highlights a key challenge in translation: the more creative and culture-specific a translation is, the greater the risk that it will confuse or distance some audience members, especially those who aren’t used to certain expressions or styles.

From a translational perspective, this example highlights several essential limitations of AI-generated texts. First, AI systems struggle with idiomaticity: they tend to simplify colloquial expressions or translate them literally, which makes the text less expressive. Second, AI does not fully understand the context in which language is used – especially how speech style reflects the speaker’s identity. A human translator can adjust the tone and level of formality to match a teenage character’s informal way of speaking, but AI often uses more neutral and general language, missing those personal and stylistic details. Third, AI translations often have a flat rhythm and lack natural flow. In contrast, the human version better reflects the fast-paced, building rhythm of the original, which helps show both the absurd tone of the story and the character’s emotions.

In dubbing, where timing, natural speech, and matching the character’s voice are essential, these weaknesses become more serious. AI may provide fast and fluent translations, but it lacks the human ability to interpret style and to adapt language in ways that make sense culturally and pragmatically.

The next example also comes from the therapy session scene, where Billy Batson, as a superhero, speaks with a therapist. This line shows Billy’s emotional struggle.

SL (00:06:11 – 00:06:15)	TL	ChatGPT
Can you help fix my thoughts so that they just, like, you know, stop happening?	Możesz zrobić coś mi w mózg, żeby przestał myśleć takie myśli?	Mógłbyś naprawić moje myśli, żeby po prostu, wiesz, przestały się pojawiać?

Source: own elaboration.

This sentence sounds deliberately clumsy, mirroring how people, especially teenagers, often speak when they are emotionally overwhelmed or unsure of how to express themselves. The use of filler phrases like “like” and “you know” highlights the character’s vulnerability and makes the dialogue sound more realistic and spontaneous. Although both versions communicate the general meaning, they differ significantly in tone, structure, and emotional effect.

The human translation adopts a highly informal and intentionally awkward phrasing: “zrobić coś mi w mózg” is not a standard expression in Polish and sounds childlike, which reflects the original speaker’s immaturity. This version attempts to mirror the character’s confusion by reproducing the imperfection of his speech. However, this unconventional phrasing might sound odd to some Polish-speaking viewers, especially because of its non-standard syntax.

The AI version, in contrast, is smoother and more grammatically correct. It uses a polite form “Mógłbyś”, a more abstract expression “naprawić moje myśli”, and a filler phrase “wiesz” that corresponds well to the English “you know.” Therefore, the AI version may sound more natural to many speakers, but it also feels more flattened. It does not fully reflect the insecurity or awkwardness of the original English line.

This difference reflects a broader distinction between grammatical correctness and emotional aspect. The AI version prioritizes standard language, producing a translation that is linguistically smooth but emotionally flatter. The human version, by contrast, embraces linguistic risk in order to imitate the psychological tone of the speaker more closely, even if the result is syntactically strange.

Survey responses confirm this ambiguity: 53% of participants correctly identified the human translation, and only 18% preferred it. These numbers suggest that while some respondents noticed the stylistic effort behind the human version, most were not convinced by its form. The sentence “zrobić coś mi w mózg” may have seemed confusing, ungrammatical, and odd, even if it was designed to match the emotional and conversational tone of the original.

This fragment illustrates one of the key challenges in adapting a speech that is deliberately ungrammatical or emotionally marked. Machine translation tends to correct such language, creating a smoother version that may be easier to follow but loses expressive details. Human translators, on the other hand, must balance the speaker’s voice, tone, and register with what the target viewers expect. In this case, the human version succeeded in mirroring the speaker’s emotional state, but at the cost of comfort for the viewer. This example shows that a natural translation can still lack emotional depth. While AI is better at producing clear and neutral sentences, human translators are

better at presenting the messy, emotional, and imperfect aspects of real speech. However, this does not always lead to greater appreciation from the audience and sometimes, readers or viewers prefer what feels more familiar and grammatically correct, even if it is less expressive.

This next line comes from a light, humorous moment when Billy Batson sets a cheerful, slightly ironic tone.

SL (00:09:22 – 00:09:25)	TL	ChatGPT
Man, it sure is a beautiful day in Philadelphia!	Mordki, a może wyskoczylibyśmy trochę na miasto?	Człowieku, ależ dzisiaj piękny dzień w Filadelfii!

Source: own elaboration.

Billy's tone is informal and expressive. The use of "Man" at the beginning is a typical conversational filler in American English, adding emphasis and relaxed enthusiasm. The reference to Philadelphia sets the scene in a specific location. While on the surface the sentence seems straightforward, translating it in a natural way into Polish, especially in the context of dubbing, presents specific challenges related to tone, localization, and cultural expectations.

The AI translation is literal and close to the SL. It reflects the core message and its grammatical structure. The phrase "Człowieku" corresponds to "Man", and the rest is a direct rendering of the comment related to weather, preserving the name of the city. However, this version sounds slightly unnatural in Polish and artificial in a dubbing context. While "Człowieku" is technically accurate, its tone feels less colloquial and more awkward than intended. Similarly, the expression "ależ dzisiaj piękny dzień" sounds overly formal, even literary, which weakens the original line's casual and cheerful tone.

In contrast, the human translation takes a bold creative turn. Instead of translating the sentence literally, it adapts it to suit the character's energy, the social context, and Polish informal speech. "Mordki" is a playful and highly informal term of address, roughly equivalent to "guys" or "dudes". The second part, "a może wyskoczylibyśmy trochę na miasto?", introduces a different action, suggesting an outing, rather than directly commenting on the weather. However, it fits the scene's tone and conveys the same emotional content: energy, spontaneity, and team spirit.

What we see here is a case of domestication vs. literal translation. The human translator sacrifices literal accuracy in favor of cultural and pragmatic equivalence, using expressions that sound natural to a Polish viewer, especially within a dubbed film targeted at younger audiences. The AI version, while accurate in content, lacks stylistic sensitivity and sounds unnatural.

Survey results clearly reflect this contrast: 97% of respondents correctly identified the human translation, and 47% preferred it. The high recognition rate suggests that viewers can intuitively detect literal translation, even if it appears grammatically fine. Nearly half preferred the human version, which indicates an appreciation for more creative, idiomatic language, though the preference was not overwhelming. This may show different levels of tolerance for overly informal expressions like "Mordki".

From a translation perspective, this fragment highlights the unique advantage of human creativity in localizing speech acts. A direct, word-for-word translation may deliver the surface meaning, but often fails to match the style, intention, and cultural references of the original. AI translations still lack the cultural and communicative awareness necessary to adapt language playfully or to adjust register in a way that suits both character and audience. In dubbing, where speech must sound natural, this difference is especially important. A human translator can imagine how a line will sound in the actor's voice, and adapt it to match the rhythm and tone of the character's world. AI, even if fluent, cannot reproduce this kind of interpretation.

The next line is delivered with contempt and sarcasm, mocking Billy's status as a superhero.

SL (00:27:52)	TL	ChatGPT
Billy Batson! You are a fool of a champion!	Billy Batson! Palant i ćwok a nie obrońca!	Billy Batson! Jesteś głupim czempionem!

Source: own elaboration.

The insult “fool of a champion” is formal, archaic, and theatrical. It is designed to match the tone of a mythological villain and mimics the style of classical or fantasy antagonists, giving the speaker an exaggerated pompous voice. At the same time, the phrase mocks Billy's supposed heroism by combining the grand word “champion” with the insulting term “fool.”

The AI version is straightforward and literal. It mirrors the original structure and vocabulary: “czempion” for “champion”, and “głupi” for “fool”. While this version is grammatically correct and semantically accurate, it lacks the stylistic and emotional touch of the original. The phrase “głupi czempion” is a flat combination and does not sound idiomatic or expressive in Polish. Moreover, the borrowing “czempion” is slightly artificial in this line.

The human translation, on the other hand, replaces the original wording with expressive Polish insults – “palant i ćwok” – that strongly convey disdain. These are informal, colloquial terms, commonly used in everyday speech to mock someone's intelligence or competence. The phrase “a nie obrońca” works well as a Polish equivalent of “fool of a champion”, preserving the contrast between supposed heroism and actual ridicule. This solution does not attempt to copy the English structure but instead highlights the intended tone and emotions in a way that sounds natural and powerful in Polish.

Survey responses clearly support the effectiveness of this approach: 90% of respondents correctly identified the human translation, and 90% also preferred it. This is the highest preference result across all fragments, indicating that the human version was not only more recognizable but also more convincing and satisfying for viewers. This example demonstrates how important stylistic adaptation is in film dialogue translation, especially when conveying sarcasm or insults. The human translator understands the dramatic tone of the scene and the villain's exaggerated speech style, and chooses to express it through strong, idiomatic Polish expressions that resonate with the target audience.

The AI translation, while technically correct, simply does not match the emotional tone or linguistic register expected in such a moment.

This fragment also reveals a broader pattern: AI often fails to fully translate expressive or dramatic language, focusing on lexical safety and structural imitation. It avoids risk and, as a result, produces flat or emotionally neutral text. Unlike AI, human translators can engage with the cultural and emotional aspects of the TL, employing strategies to maintain the line's original effect. In dubbing, where timing, voice acting, and emotional delivery are crucial, this difference is especially visible. A phrase like "palant i ćwok, a nie obrońca" is more likely to elicit the desired audience reaction, such as laughter, tension, or sympathy, than a literal construction like "głupi czempion".

The next fragment consists of a sarcastic and vulgar dialogue between several characters. The lines are delivered by Brett and Burke, minor antagonists who attempt to ridicule Freddy – a member of the superhero team. The characters speak with sarcasm and bravado, in an attempt to dominate the school environment.

SL (00:23:22 – 00:23:47)	TL	ChatGPT
<p>– I'm Brett. This is Burke. Listen, if you need anything, at all, we're your guys. We got it.</p> <p>– Brett and Burke Breyers. They got it all. I mean, what is it, narcissistic personality disorder, chlamydia, they got it. Plenty of it to go around, too.</p> <p>– This is Thundercrack's Make-A-Wish kid. Fanboy here thinks he's a big boy now.</p> <p>– Yeah, he doesn't know where he belongs.</p> <p>– Sure I do. I don't belong in that crash can. 'Cause that's... I'll smell like human garbage, do you realize that? And that's your brand, guys!</p>	<p>– Jestem Brett, a to Burke. Jakbyś czegoś potrzebowała, uderzaj do nas. Co tylko chcesz.</p> <p>– Brett i Burke Breyerowie. Wszystko co chcesz. Szereg zaburzeń emocjonalnych, choroby weneryczne, co-kolwiek. I to w hurtowych ilościach.</p> <p>– A to jest nasz fanboy specjalnej troski. Kolega Supera. Myśli, że jest kimś.</p> <p>– Nie wie gdzie jego miejsce.</p> <p>– Właśnie, że wiem. Na pewno nie w tym koszu, bo wtedy będę walił jak śmiecie. A od walenia jesteście wy, nie?</p>	<p>– Jestem Brett. To jest Burke. Słuchaj, jeśli potrzebujesz czegokolwiek, jesteśmy do usług. Załatwimy to.</p> <p>– Brett i Burke Breyers. Oni mają wszystko. Naprawdę wszystko. Co to było? Zaburzenie osobowości narcystycznej, chlamydia, oni to mają. I to w nadmiarze, starczy dla wszystkich.</p> <p>– To jest życzeniowy dzieciak Thundercracka. Fanboy tutaj myśli, że jest już dużym chłopcem.</p> <p>– Tak, nie wie, gdzie jest jego miejsce.</p> <p>– Oczywiście, że wiem. Nie należę do tego śmietnika. Bo wtedy... będę pachniał jak ludzki śmieć, rozumiesz?</p> <p>A to jest wasza marka, chłopaki!</p>

Source: own elaboration.



The English dialogue includes several distinctive elements: overlapping speech, casual tone, references to mental illness and STDs, and ironic insults. The speech is stylized, informal, and purposefully offensive but in a comic, not truly aggressive, way. For example: “Brett and Burke Breyers. They got it all. I mean, what is it, narcissistic personality disorder, chlamydia – they got it. Plenty of it to go around, too”. Or later: “I don’t belong in that crash can. ‘Cause that’s... I’ll smell like human garbage, do you realize that? And that’s your brand, guys!”. These lines are filled with sarcastic exaggeration and humor. This makes them particularly difficult to translate, especially for AI, which struggles with detecting tone, intention, and audience expectations.

Seemingly, the AI version is fluent and mostly accurate in meaning. It correctly identifies cultural references (Make-A-Wish, narcissistic personality disorder) and preserves sentence structure. However, it does that in a neutral tone and with limited adaptation. The expressions “jesteśmy do usług” and “naprawdę wszystko” sound flat and slightly outdated for a youthful conversation. The insult “he smelled like human garbage” is a literal translation, but it sounds weaker and less humorous.

In contrast, the human translation is bolder, more idiomatic, and better adapted to the social and stylistic register of the original scene. Expressions such as “uderzaj do nas”, “w hurtowych ilościach”, or “walić jak śmiecie” carry a stronger informal and urban tone that suits the characters’ social role. Moreover, the phrase “fanboy specjalnej troski” adds local cultural flavor while preserving the mocking tone. The closing line, “A od walenia jesteście wy, nie?”, is not a literal translation, but it captures the same sarcastic tone and arguably enhances it with humor tailored to Polish conversational style.

According to the survey, 76% of respondents correctly identified the human translation, and 68% preferred it. These are high scores, showing that viewers not only noticed the differences in tone and expression but also appreciated the more idiomatic and humorous approach. This supports the hypothesis that in longer, socially charged dialogues, human translators can more effectively manage shifts in tone, register, and group interactions than AI systems. The relatively high success of the AI version in this case may be attributed to its improved coherence over longer texts and its relatively accurate rendering of content. However, AI still struggles with emotional tone, mocking style, and the rhythm characteristic of comic insult.

This example illustrates the linguistic and cultural complexity of teenage speech and sarcasm. It highlights the need for translation strategies that go beyond lexical level and into areas such as intonation, informal conversations, and stylistic compensation. While AI can generate acceptable translations at the sentence level, it struggles with maintaining informal tone, especially in scenes involving confrontation, peer pressure, or humor. In dubbing, these aspects matter even more, as they directly affect how characters are perceived. The human version delivers a fuller, more localized translation, preserving the aggression of the scene, humor, and chaotic energy. In this case, as in many others, human creativity, pragmatics, and cultural intuition prove essential to successful audio-visual translation.

The last fragment analyzed in this article is a dialogue between Freddy and Billy – two central characters in the film. Freddy, frustrated and hurt, expresses a moment

of vulnerability: a girl is finally talking to him, and he wants the chance to talk to her without Billy's interference.

SL (00:31:28 – 00:31:36)	TL	ChatGPT
I have a girl talking to me, Billy. A girl. For the first time in my life! And I probably don't have a shot in hell with this girl, but I'll never find out if you keep trying to control everything I wanna do!	Odezwała się do mnie dziewczyna, Billy. Prawdziwa. Pierwszy raz ever. Pewnie nie mam co do niej stratować, ale nigdy się nie dowiem jak będziesz kontrolował każdy mój krok!	Dziewczyna ze mną rozmawia, Billy. Dziewczyna. Po raz pierwszy w życiu! I pewnie nie mam u niej żadnych szans, ale nigdy się nie dowiem, jeśli będziesz ciągle próbował kontrolować wszystko, co chcę zrobić!

Source: own elaboration.

This is a fast-paced monologue. The repetition of “a girl” and the emphatic “for the first time in my life” show Freddy's excitement and nervousness. The sentence also conveys a feeling of powerlessness – he wants control over his own life, but Billy, trying to protect him, unintentionally dominates his decisions. Freddy feels part joy, part frustration, part desperation. The translator's job is to reflect that complexity in a natural way in Polish.

At a first glance, both translations are correct in terms of meaning, but stylistically they differ. The AI version is smoother, syntactically clear, and preserves sentence structure. However, it uses more neutral language: “po raz pierwszy w życiu”, “nie mam u niej żadnych szans”, “ciągle próbował kontrolować wszystko”. These phrases are natural, but not particularly expressive. They only resemble Polish spontaneous speech. As a result, the emotional tone of Freddy's monologue is flattened.

The human version sounds more lively and fits better with how people actually speak. It includes “prawdziwa” – a short, emphatic exclamation that breaks the sentence rhythm effectively, and “pierwszy raz ever”, which includes a borrowing and reflects the expressive way of communicating common in youth language. The phrase “nie mam co do niej stratować” is a creative, informal expression that perfectly captures Freddy's emotions. Although slightly risky stylistically, this rendering feels alive and captures the characteristics of teenage speech.

According to the survey, 87% of respondents correctly identified the human translation, and 53% preferred it. These numbers suggest that viewers recognized the stylistic features and speech authenticity of the human version, but the preference was not overwhelming. This may be due to the use of expressions like “ever” or “stratować”, which, while natural in colloquial speech, may sound too informal or unusual for some viewers. This example shows the struggle between expressiveness and accessibility. While AI preserves the content and coherence of the original, it fails to match the speaker's tone and speech rhythm. The human version succeeds in imitating emotional flow and the stylistic flavor of teenage speech, but its use of informal language or borrowing may not appeal equally to all audiences.

This scene illustrates one of the core challenges in dubbing emotionally rich dialogue: how to capture the voice of a character, not just the content of what they say. The AI ver-

sion is safe, readable, and structurally safe, but too emotionally neutral for a moment filled with personal tension. The human translator takes risks by using slang and although not all viewers favored this approach, it better reflects the intent and style of the original scene.

## Conclusion

The analysis presented in this article confirms that, while artificial intelligence has made considerable progress in producing fluent translations, it still fails to capture many of the subtle, culture-bound, and stylistic aspects of film dialogue. In particular, AI-generated translations tend to remain literal, emotionally flat, and lexically safe. They may preserve meaning on a surface level, but often miss the tone, intention, and rhythm required in dubbed audiovisual texts.

In contrast, human translations, although not always preferred by all viewers, demonstrate greater creativity, cultural sensitivity, and responsiveness to context. Human translators are capable of adjusting language to reflect characters' voices, adapting idioms, slang, and humor to the target culture, and making bold choices that reflect the pragmatic function of a given line. Their work is often riskier, less grammatically perfect, but more expressive which is especially valuable in dubbing.

One of the main findings of this study is that most respondents were able to distinguish between human and AI translation, especially when idioms, humor or sarcasm was involved. However, knowing which was which did not always mean they liked the human version more. In some cases, participants favored the smoother, more neutral AI version – possibly because it sounded more correct or familiar. In other cases, particularly those involving jokes or insults, the human versions were both correctly identified and preferred, indicating that expressiveness and naturalness were more valued in those contexts.

These patterns suggest that viewers do not assess translations solely based on accuracy or fluency, but also on how naturally the lines sound in their language, whether the tone fits the character, and how effective the line is in context (e.g. whether it makes them laugh or feel the real emotions). The more style or emotion a line has, the easier it is to see the difference between AI and human translation. While there is no single language feature that always separates human from AI translation, certain signs appeared repeatedly in the examples.

Human translations were more likely to:

- use idiomatic or slang expressions,
- rephrase or restructure sentences freely,
- reflect the speaker's identity through lexical and syntactic choices,
- prioritize tone and pragmatic function over literal meaning.

AI translations, on the other hand, tended to:

- follow the structure of the source text more closely,
- use standard or formal vocabulary,
- avoid unusual or creative expressions,
- prioritize grammatical correctness and clarity.

Although these tendencies are not consistent in every case, they provide useful clues for telling human and AI translations apart. They also show the fact that, despite technological advancements, machine translation, particularly AVT, still cannot match the human ability to interpret, localize, and emotionally shape language. As AI continues to evolve, it may become better at imitating human expression. But for now, especially in tasks like dubbing that require rhythm, character voice, and humor, the role of the human translator remains important, if not indispensable.

## References

- Cambridge Dictionary (n.d.), <https://dictionary.cambridge.org/> [accessed: 27.05.2023].
- Díaz Cintas J. (ed.) (2009), *New Trends in Audiovisual Translation*, Bristol.
- Hauenschild C., Heizmann S. (eds.) (2011), *Machine Translation and Translation Theory*, Berlin.
- Nida E.A., Taber C.R. (1969), *The Theory and Practice of Translation*, Leiden.
- Norvig P., Russell S. (2021), *Artificial Intelligence: A Modern Approach, Global Edition*, London.
- Oxford Learners Dictionaries (n.d.), <https://www.oxfordlearnersdictionaries.com/> [accessed: 28.05.2023].
- Shazam! Gniew bogów* (2023), Filmweb, <https://www.filmweb.pl/film/Shazam+Gniew+bog%C3%B3w-2023-845693> [accessed: 26.05.2023].
- Shazam! Fury of the Gods* (2023), IMDb, [https://www.imdb.com/title/tt10151854/?ref\\_=fn\\_all\\_ttl\\_2](https://www.imdb.com/title/tt10151854/?ref_=fn_all_ttl_2) [accessed: 26.05.2023].
- Venuti L. (1995), *The Translator's Invisibility: A History of Translation*, London and New York.

## Streszczenie

### Kto jest lepszym tłumaczem – SI czy człowiek? Badanie opinii widzów na temat przekładu audiowizualnego

Artykuł przedstawia analizę różnic między tłumaczeniami wykonanymi przez człowieka a wygenerowanymi przez sztuczną inteligencję (SI) na przykładzie filmu *Shazam! Gniew bogów* (2023) i jego polskiej wersji dubbingowej. Do badania wybrano fragmenty szczególnie trudne dla systemów tłumaczeniowych – zawierające idiomy, humor, młodzieżowy slang i sarkazm. Porównano tłumaczenia profesjonalnego tłumacza z wersjami przygotowanymi przez SI (ChatGPT), kierując się promptem mającym na celu uzyskanie humorystycznego i naturalnego efektu. Badanie zostało wsparte ankietą, w której 80 uczestników oceniało tłumaczenia, wskazując ich prawdopodobne źródło (człowiek czy SI) oraz osobiste preferencje. Wyniki pokazały, że większość badanych poprawnie rozpoznała tłumaczenia wykonane przez człowieka, choć ich wybory preferencyjne były bardziej zróżnicowane. Analiza ujawnia, że SI ma trudności z odwzorowaniem stylu potocznego, idiomatycznego i emocjonalnego, podczas gdy tłumaczenia ludzkie lepiej oddają charakter postaci i kontekst kulturowy. Artykuł wpisuje się w aktualne dyskusje nad granicami zastosowania SI w tłumaczeniach audiowizualnych.

**Słowa kluczowe:** tłumaczenie audiowizualne, tłumaczenie maszynowe, dubbing, ChatGPT