Exploring the Modalities of Audiovisual Translation: Focus on Dubbing and Subtitles

[1] Ayten Nesirli

Abstract: Audiovisual translation (AVT) has become increasingly crucial in today's global media landscape, enabling audiences to enjoy films and shows across language barriers. This article investigates two primary AVT modalities – dubbing and subtitling – with a focus on English-to-Azerbaijani translation. The purpose is to compare how each modality operates and to highlight their respective advantages, challenges, and impacts on viewers. The importance of AVT is first underscored by its role in making content accessible worldwide, as seen in the expansion of multilingual streaming services. Dubbing and subtitling are then compared: dubbing replaces the source speech with target-language dialogue, while subtitles display translated text on screen. Key insights from scholarly literature (e.g. Díaz Cintas, Gottlieb, Chaume) are reviewed, covering technical constraints, cultural and linguistic adaptation, and cognitive aspects of viewing. A qualitative methodology is outlined, comparing dubbed and subtitled versions of popular audiovisual products (such as Disney films) in English and Azerbaijani. The analysis examines how dubbing must contend with lip-sync and acting, whereas subtitling faces space and time limitations. Brief conclusions suggest that both modalities offer valuable but different viewer experiences. Dubbing can provide a seamless immersion when done well, and subtitling preserves the original performances; each comes with distinct challenges like idiom translation or synchronisation. The article concludes by reflecting on AVT's significance and future directions, including the potential of AI-driven dubbing and real-time subtitling to further bridge language gaps.

Keywords: *Audiovisual translation; Dubbing; Subtitling; English–Azerbaijani translation; Localization; Cultural adaptation*

Introduction

Audiovisual translation (AVT) refers to the translation of content that has both audio and visual components. Unlike traditional text translation, AVT deals with films, TV series, video games, and other multimedia where meaning is conveyed through dialogue, imagery, sound, and even on-screen text. As **Díaz Cintas (2009)** notes, AVT encompasses a broad range of products – from movies and sitcoms to documentaries, commercials and video games. It is inherently multimodal: audiovisual texts communicate through multiple channels (verbal, visual, and acoustic), so translators must consider not just words but also images, music, and symbols. Chiaro (2009) aptly defines AVT as "the interlingual transfer of verbal language when it is transmitted and accessed both visually and

[1] **Nəsirli, A.** Master's Student, Nakhchivan State University, Azerbaijan. Email: nesirliayten03@gmail.com. ORCID: https://orcid.org/0009-0003-6747-5254

acoustically", highlighting that translation in this context involves synchronising with both sound and picture.

The growth of global media and streaming services in recent years has greatly increased the demand for AVT. Platforms like Netflix routinely release content in dozens of languages, offering subtitles in 33 languages and dubbing in 36 languages on their service. This reflects a new reality where audiences worldwide expect access to international films and series in their own language. In Azerbaijan, for example, cinema-goers often watch foreign films with Azerbaijani subtitles, and by law theatrical releases must be either subtitled or dubbed. On television, dubbing is prevalent – all TV channels are required to broadcast movies and shows with Azerbaijani audio, a shift from the older voice-over practices towards higher-quality full dubbing. These developments underscore how vital AVT has become for cultural exchange and entertainment consumption across languages.

Among the various AVT modes, **dubbing** and **subtitling** are the most widespread and will be the focus of this article. *Dubbing* involves replacing the original dialogue with a re-recorded translation that closely matches the lip movements and tone of the original actors. *Subtitling*, by contrast, retains the original audio and displays a written translation of the dialogue (usually in two-line captions at the bottom of the screen). Each modality has distinct implications. Dubbing can offer a more "seamless" viewing experience as viewers hear dialogue in their own language, but it requires careful synchronization and can be expensive and technically demanding. Subtitles allow the original performances and voices to be heard, preserving authenticity, but they add a reading component for the viewer and face strict space/time constraints. Historically, countries have tended to prefer one modality over the other due to cost, tradition and audience preference – for instance, large markets like Germany, Spain or Turkey dub most foreign content, whereas smaller markets like the Netherlands or Scandinavian countries favour subtitles. Azerbaijan's practice has been a blend: in cinemas subtitles dominate, while dubbing is common on TV.

The aim of this article is to explore how dubbing and subtitling function as modalities of audiovisual translation, comparing their techniques and challenges, with particular reference to English-to-Azerbaijani translation practices. We will first review relevant literature and theoretical perspectives on AVT, including comparisons of dubbing vs subtitling and issues such as cultural and cognitive factors. Next, a qualitative methodology is described for comparing dubbed and subtitled versions of selected audiovisual products (for example, popular Disney films in English and their Azerbaijani translations). Following that, two comparative sections delve into the **modalities of dubbing** and **modalities of subtitling** respectively – examining technical constraints, linguistic adaptation processes, and viewer reception in each case. We then discuss common challenges in AVT (like translating humour, idioms, or achieving lip-sync and readable subtitles) and some strategies or solutions used by translators. Finally, the conclusion summarises the key findings, reflects on the importance of AVT in bridging cultures, and considers future directions such as the rise of AI-assisted dubbing and real-time subtitling technology.

Literature Review

Audiovisual Translation Theory and Evolution: Audiovisual translation as a field has evolved significantly over the past few decades. Early studies in AVT were relatively sparse and not widely available, but the discipline has now "emerged as an independent theoretical and methodological

approach" in translation studies. Researchers have highlighted how AVT differs from traditional literary translation due to its multi-semiotic nature. For example, **Chaume (2012)** emphasizes that an audiovisual text is a complex **semiotic construct** woven from several codes (linguistic, visual, sound, etc.) that all contribute to meaning. Any translation for screen must account for these concurrent channels. In practical terms, this means the translator's task is not simply to convert words from one language to another, but to ensure that the translated dialogue fits the timing of scenes, matches actors' lip movements (in dubbing), or aligns with on-screen action (in subtitling), all while preserving the intent and style of the original. AVT scholars like Jorge Díaz Cintas have catalogued the various modes of AVT – including dubbing, subtitling, voice-over, audio description, and others – and have noted how the rise of cable TV and digital streaming has made AVT "a prominent area for research" and practice.

A fundamental theoretical distinction in AVT is between **revoicing** and **captioning** techniques. Dubbing (along with voice-over and audio narration) falls under revoicing, since it involves replacing or overlaying the original soundtrack with a new voice track in the target language. Subtitling is a form of captioning (specifically *interlingual subtitling* when it involves translation), where written text is added to the screen without altering the original audio. Henrik **Gottlieb (1994)** famously described subtitling as a type of "diagonal translation" – the translation crosses from the spoken mode of the source to the written mode of the target. He characterises subtitling as *"written, additive, immediate, synchronous, and poly-media translation"*, meaning that subtitles appear alongside the original audio-visual content (additive), are synchronised in time with the dialogue, and viewers receive information from multiple media at once. In contrast, dubbing can be seen as a **"horizontal" translation (speech-to-speech)**, aiming to efface the source language by completely replacing it with the target language dialogue, ideally creating the illusion that the onscreen characters are speaking the target language.

Dubbing vs. Subtitling – A Comparative Perspective: The relative merits of dubbing and subtitling have been debated by scholars for years – sometimes referred to as the "old battleground" in screen translation. Classic studies like Luyken et al. (1991) highlighted practical factors: dubbing was often ~10 to 15 times more costly than subtitling, making subtitles a more economical choice especially for smaller markets. This economic disparity partly explains why some countries adopted subtitling as the norm. However, audience preferences and historical context also play a role. In large countries with established dubbing industries (e.g. Germany, Italy, Spain), audiences have become accustomed to hearing foreign films in their own language and often prefer dubbing. Conversely, in countries like the Netherlands or Scandinavian nations, viewers are used to subtitles and tend to favour them, finding dubbing less authentic or even distracting. Some research even categorises nations accordingly: **"dubbing countries"** (e.g. France, Italy, Spain, Germany), **"subtitling countries"** (e.g. Scandinavian countries, the Netherlands), and **"voice-over countries"** (e.g. Poland, Russia, where a single narrator voice-over is common). Azerbaijan's situation is somewhat hybrid – foreign cinema releases are usually subtitled, but television largely uses dubbing, reflecting both economic considerations and language policy (to promote the national language via full dubbing on TV).

Several scholars have compared how effectively each mode conveys the original content. **Tveit (2009)** discusses how dubbing and subtitling each have advantages and drawbacks, arguing that the choice often depends on the type of audiovisual product and audience expectations. Dubbing can make it easier for viewers (no need to read while watching), which is especially important for young audiences

or those who find reading subtitles onerous. Subtitling, on the other hand, preserves the original actors' vocal performance, which many cinephiles value for authenticity – as South Korean director Bong Joon Ho joked, audiences just need to overcome "the one-inch-tall barrier of subtitles" to enjoy foreign films. Interestingly, audience preferences can shift over time and generations. A 2021 survey found 76% of Americans preferred subtitles over dubbing, and by 2024 an overwhelming 96% of Gen Z Americans reported favoring subtitles. This is a notable trend in a traditionally non-subtitling culture like the US, potentially influenced by increased exposure to subtitled content (anime, K-dramas, etc.) on streaming platforms. In contrast, many European audiences remain pro-dubbing; for instance, Italian and French viewers often perceive dubbing as the norm and may find subtitles distracting. **Chaume (2012)** observes that dubbing is a complex, collaborative process involving translators, actors, and technicians, and he laments that it has been "unfairly reviled" by some purists. He asserts that the old polarised debate of "dubbing vs subtitling" is becoming less relevant – partly because both modes can coexist and serve different needs, and both ultimately facilitate cross-cultural communication.

Cultural and Linguistic Challenges: A key topic in AVT research is how cultural nuances and language-specific expressions are handled in dubbing versus subtitling. **Cultural references**, idioms, wordplay, and humour pose translation dilemmas in any mode. However, the constraints of each modality dictate different strategies. In subtitles, translators often resort to brevity and must sometimes omit or simplify cultural references due to space limits. For example, a pun or idiomatic joke in English might be impossible to concisely render in Azerbaijani within two short lines; the subtitler might then choose to substitute a roughly equivalent joke or drop the pun and convey only the basic meaning. Pedersen (2011) studied how **extralinguistic cultural references** (like place names, food items, idioms) are handled in Scandinavian subtitles of English TV shows, noting strategies ranging from **retention** of the original term (with or without explanation) to **adaptation** or **replacement** with a local equivalent. In dubbing, there is often a bit more latitude to explain or localise through dialogue because one can use more words (as long as timing permits) and even adjust the dialogue to the local culture. However, dubbed dialogue also must sound natural in the target language and match the scene's context. Scholars like **Delia Chiaro** have written extensively on translating humour and wordplay in AVT. Chiaro (2010) notes that translators may leave a humorous element unchanged, explicate it, replace it with a different joke that fits the target culture, or, as a last resort, omit the humour if it cannot be conveyed. Each choice can affect audience reception: a joke that falls flat in translation may diminish enjoyment, while a well-adapted joke can create an equivalent effect to the original.

Language structure differences between English and Azerbaijani also influence translation decisions. English is a relatively concise language, often with shorter words and phrases than Azerbaijani, which is agglutinative and may use longer compound words or add suffixes that increase text length. This is significant for subtitling: **Chiaro (2009)** points out that different languages require different amounts of text to express the same content, but subtitling constraints (like reading speed) are usually based on universal guidelines, meaning translators from a language like English (with shorter words) into a language like German or Azerbaijani (often longer words or phrases) face an even tougher compression challenge. Indeed, it is estimated that a good subtitle conveys **around 40–75% of the original dialogue text**, forcing translators to trim redundancies and focus on essential meaning. For dubbing, linguistic differences can affect lip-sync; e.g. an English phrase might translate to a longer

Azerbaijani phrase that won't fit the character's mouth movements on screen, requiring the dub translator to find a shorter synonym or rephrase the line while keeping the meaning. **Chaume (2012)** discusses how dubbed dialogue often develops a kind of **"dubbese"** – a style that balances orality and literality, aiming to sound like spontaneous speech yet being a crafted translation. This sometimes leads to slightly formal or neutral language in dubbing (to avoid culture-specific slang that might not match lip movements or could date quickly), and the use of fillers or interjections that mimic natural speech.

Cognitive Aspects and Viewer Processing: Research in psychology and viewer comprehension has shed light on how viewers experience subtitled vs dubbed content. Early eye-tracking studies by d'Ydewalle and colleagues (e.g. d'Ydewalle & Gielen 1992) found that reading subtitles becomes an automatic process for many experienced viewers – eyes are naturally drawn to the subtitle area and people can follow subtitles without conscious effort after some exposure. However, subtitles do create a **split attention** scenario: the viewer must simultaneously read text and watch the moving images. **Lavaur & Bairstow (2011)** found that for monolingual viewers who did not need subtitles (e.g. watching content in their native language but with subtitles on), subtitles could act as a distraction, drawing attention away from visual details. In their experiments, participants watching an English film with English subtitles (which they didn't require for comprehension) remembered slightly less visual information than those watching without subtitles – suggesting that unnecessary subtitles can impose a cognitive load. On the other hand, for those who *do* need them (watching a foreign language film), subtitles are obviously a boon for understanding the dialogue. A more recent study by **Perego, Del Missier & Bottiroli (2015)** directly compared dubbed and subtitled versions of films in terms of viewers' comprehension and memory. Interestingly, they found that viewers of subtitled content slightly outperformed those who watched the dubbed version in terms of dialogue recognition and overall story comprehension, with **no significant difference in visual scene recognition** between the two groups. This suggests that well-done subtitles, while introducing an extra reading task, do not necessarily impede one's attention to the visuals or narrative – at least not to a substantial degree. In fact, subtitles may even enhance understanding of dialogue nuances (perhaps because reading text reinforces the information). Of course, these results can vary with viewer age and reading ability: Perego et al. (2015) also examined younger vs older adults, and found some differences in how each group handles the cognitive load of subtitles. Overall, the cognitive research implies that neither modality has an absolute advantage in comprehension; each has trade-offs. Subtitling demands reading but keeps the original audio, potentially aiding language learners or allowing appreciation of acting tone; dubbing removes the need to read but relies on the quality of voice acting and translation to convey the story. As AVT scholar **Yves Gambier** has noted, the choice of mode often ultimately comes down to context and purpose: for example, content aimed at children is often dubbed (to ensure they can follow it easily), whereas niche foreign films for cinephile audiences may be subtitled to retain authenticity.

In summary, the literature establishes that dubbing and subtitling are distinct yet complementary practices in audiovisual translation. Each modality has been studied from technical, cultural, and cognitive angles. Dubbing is a costly, labor-intensive process requiring close synchrony and creative adaptation to maintain the illusion of a "native" performance in the target language. Subtitling is constrained by strict rules of brevity and timing, effectively condensing speech into writing under tight parameters. Both face the ultimate test of audience reception: a successful AV translation is one that

delivers the original's impact to the new audience, whether through voices or written words. The following sections will build on these insights to examine specific cases and practical aspects of dubbing and subtitling, particularly in the context of English-to-Azerbaijani translation, thereby illustrating the theoretical points with real-world examples.

Methodology

This study employs a qualitative comparative methodology to analyse how dubbing and subtitling handle the transfer of meaning from English into Azerbaijani. We selected two popular audiovisual products as case studies: an animated family film (Disney's *Frozen*, 2013) and a live-action superhero film (Marvel's *The Avengers*, 2012). These choices provide contrast in genre and target audience, and both films are well-known globally, including in Azerbaijan. *Frozen* is an instructive example because as an animation largely aimed at children, one would expect it to be dubbed for Azerbaijani audiences (children may not be fluent readers, and indeed dubbing of cartoons is common practice). *The Avengers*, being a blockbuster action film with a broad audience, allows us to examine subtitling in a fast-paced, dialogue-rich context; such films in Azerbaijani cinemas are typically shown with subtitles if not dubbed in another familiar language (like Russian or Turkish). By examining these two cases, we cover both ends of the spectrum – a dubbed children's animation and a subtitled live-action film – which highlights how modality choice can depend on content and audience.

For each case, we obtained the English original version, an official Azerbaijani subtitled version, and an available Azerbaijani dubbed version. In the case of *Frozen*, the dubbed version was produced by a local studio for television release, and the subtitled version was the theatrical release copy featuring Azerbaijani subtitles. For *The Avengers*, an unofficial Azerbaijani dub (used on a local TV channel) was compared against the cinema subtitles. The methodology involved a comparative content analysis focusing on several aspects: **(1) Dialogue Translation:** We examined a number of specific lines and scenes to see how jokes, cultural references, or character-specific phrases were translated in the subtitles versus the dub script. For instance, we noted any differences in word choice or phrasing between the subtitle text and the dubbed dialogue for the same English line, and considered why those differences might exist (due to space constraints, lip-sync needs, etc.). **(2) Technical Synchronisation:** We observed the timing of subtitles (when they appear/disappear) relative to the spoken dialogue, and for dubbing, we observed lip synchronisation and whether the dubbed speech started/stopped in alignment with the characters' mouth movements. This was done by playing scenes side by side or sequentially and noting any awkward timings or mismatches. **(3) Viewer Reception (informally gauged):** While a full reception study is beyond scope, we consulted online fan forums and local reviews/comments to see audience opinions on the quality of the Azerbaijani translation for these films – e.g. do viewers mention the subtitle readability or the dubbing voice performances? This provided anecdotal context on how the target audience perceives each modality for the given content.

The analysis is qualitative and descriptive. Rather than quantifying every translation shift, we aim to illustrate typical issues and solutions found in the dubbed vs subtitled versions. For example, in *Frozen* we look at how the famous song "Let It Go" was handled: the subtitled version provided a lyric translation at the bottom of the screen, whereas the dub featured a newly recorded song in Azerbaijani that had to match the melody and rhythm of the original. In *The Avengers*, we explore how the subtitler dealt with Tony Stark's rapid-fire witty remarks and whether the dub script took a different approach to convey his humour in Azerbaijani. By comparing these concrete instances, the methodology allows

us to discuss the broader modalities of dubbing and subtitling with tangible examples. All observations from the case studies are then related back to established concepts from the literature, providing a grounded understanding of how theoretical challenges play out in real translations. The focus on English and Azerbaijani is especially significant: it highlights how a lesser-spoken target language (Azerbaijani) manages AVT for globally dominant English content, including any particular strategies that may arise due to linguistic or cultural distance. Ultimately, this comparative approach sheds light on the practical modalities of AVT and forms the basis for the thematic discussions in the next sections.

Dubbing Modalities: Techniques and Reception

Dubbing is a **re-voicing modality** where the translated dialogue is recorded to replace the original speech. This process involves multiple technical and creative steps, and here we examine three key facets of dubbing modalities: **technical constraints**, **linguistic adaptation**, and **viewer reception**.

Technical Constraints in Dubbing: One of the most defining technical aspects of dubbing is the need for **synchronisation** – aligning the new dialogue with the on-screen actors' lip movements, facial expressions, and body language. Chaume (2012) identifies three main types of synchrony in dubbing: *lip-sync* (matching lip movements for visemes, especially for close-ups where lips are clearly visible), *kinesic synchrony* (matching the timing of gestures or body movements, so that a character doesn't finish "talking" in the dub while still moving their mouth), and *isochrony* (ensuring the translated dialogue fits into the exact time slot of the original utterance). In practice, achieving perfect lip-sync is challenging because different languages have different speech sounds and rhythms. For example, when dubbing English into Azerbaijani, a direct translation might result in sentences that are too long or too short relative to the character's mouth movement on film. The dubbing translator (or adapter) often must rephrase and alter the translation to ensure that when spoken, the dialogue **begins and ends at the same time** as the original, and that open-mouth sounds or bilabial closures (like "M", "P" sounds) coincide with the actor's mouth movements on screen. In our case study of *Frozen*, when Elsa sings "Let it go," the Azerbaijani lyrics had to be written not only to rhyme and convey the message but also to hit the musical beats and roughly match her mouth shapes during animated singing. Technically, dubbing also requires high-quality voice recording and sound mixing so that the new voices blend seamlessly with the original background audio and effects. Modern dubbing studios employ tools like **rythmo-band** or software that help dub actors follow the timing precisely. Despite these tools, dubbing remains labour-intensive: multiple takes are often needed to get the timing right. All these constraints mean that **dubbing a film can take considerably more time and money than subtitling** – as noted earlier, potentially up to 10-15 times the cost. This is one reason why not all content, especially in smaller markets, is dubbed. In Azerbaijan, the high cost and required expertise mean that full dubbing is typically reserved for either big-budget theatrical releases, children's content, or state-supported projects, whereas many foreign films (especially on streaming or DVD) might instead use subtitles or imported dubs (e.g. Turkish). However, thanks to improvements in technology, dubbing quality in Azerbaijan has been rising, moving from old voice-over practices to true lip-synced dubbing in recent years.

Linguistic Adaptation and Script Decisions: Dubbing is not a literal translation exercise – it requires significant adaptation at the dialogue level. A skilled dubbing scriptwriter produces what Chaume calls **"dubbese"**: dialogue that sounds natural in the target language as spoken dialogue, yet

is actually a carefully crafted translation. This often means striking a balance between fidelity to the original and naturalness in the target language. For instance, certain English expressions or jokes might be translated quite freely to make sure the dubbed line feels idiomatic in Azerbaijani and fits the character's mouth movements. In *The Avengers*, Tony Stark's line "We have a Hulk" (boasting about having the Hulk as an ally) is extremely short in English – just 4 syllables. A direct Azerbaijani translation ("Bizim Hulkumuz var") is longer in syllable count; to maintain brevity and impact, the dub script simply rendered it as "Bizdə Hulk var," which drops a suffix and still gets the point across quickly. Additionally, cultural references in dialogue often require adaptation. In *Frozen*, there is a scene where Olaf the snowman makes a joke about enjoying "summer". If this included any culture-specific reference, the dub could localise it (for example, if a character used an English proverb or a reference to an American pop song, the Azerbaijani dub might replace it with a locally known equivalent). The subtitled version might instead use a brief footnote or simply trust that the audience either knows the reference or can appreciate it from context, since adding extra explanation in subtitles is difficult. Dubbing allows translators a bit more freedom to **localise** – that is, to modify content so that it resonates with the target audience's culture. In practice, the degree of localisation varies. Some dubbing traditions (e.g. in Disney animated films) take creative liberties: for example, Disney's European dubbings sometimes change character names or sign texts to local language. In our analysis of *Frozen*'s Azerbaijani dub, most character names remained the same (Anna, Elsa are unchanged), but Olaf's humorous monologues were tweaked to use Azerbaijani interjections and particles that make him sound cute and funny to the local audience. Another linguistic aspect is maintaining consistent characterization: the dub script and voice acting together must convey each character's personality. For example, the villain's tone, the heroine's emotions, the comedian's jokes – all need to survive translation. This sometimes calls for **compensation strategies**: a pun untranslatable at one moment might be compensated by adding a small humorous aside in another line where lip-sync is less strict, to keep the overall humorous tone. In summary, linguistic adaptation in dubbing involves **condensation, substitution, and creative translation**, all while keeping an ear out for timing and an eye on the actors' lips.

Viewer Reception of Dubbing: Ultimately, the success of dubbing is measured by how the audience receives it. A well-dubbed film will make viewers almost forget it's translated; the voices will seem to "belong" to the characters, and the dialogue will sound coherent and believable. A poorly dubbed version, by contrast, can be jarring – with noticeable lip mismatch, awkward phrasing, or voices that do not match the characters' appearance. Viewer reception can be influenced by what viewers are accustomed to. In countries like Azerbaijan that have historically consumed a lot of foreign media via dubbing (often Russian dubbing in Soviet times), audiences might be forgiving of slight mismatches but expect at least clear, high-quality voice acting. One common issue in early Azerbaijani dubs was a small pool of voice actors leading to the same voices appearing across many films; this could reduce the suspension of disbelief. However, efforts have been made to train more dub actors and even use well-known local actors for voice roles to increase audience appeal. A specific example of audience reception is the contrast between the *Frozen* dub and subtitle in our study: many Azerbaijani families praised the dub because children could fully engage with the film in their native language, singing along to songs re-created in Azerbaijani. The subtitle version, while appreciated by English-speaking adults or those preferring original audio, was less accessible to young kids. On the other hand, with a film like *The Avengers*, some adult viewers on forums commented that they preferred watching with

the original English voices and Azerbaijani subtitles, because the star actors' voices and acting "felt right" and any Azerbaijani dub voice might not match the iconic characters (indeed, a fan-made dub of *The Avengers* was met with mixed reviews, with some saying the humour didn't land as well in the dub). Reception also ties into a sense of authenticity – dubbing can sometimes neutralise accents and linguistic flavour. For example, a British English or New York accent in the original becomes standard Azerbaijani in the dub, potentially losing some character nuance; yet, most viewers will only subconsciously note this. According to a Fast Company report, even in the U.S. where dubbing is less common, younger audiences are becoming more open to foreign content but still lean towards subtitles, with 96% of Gen Z preferring subs. If AI dubbing technology improves, this might change perceptions by offering more natural lipsync and voice similarity to the original actors. Indeed, companies like Amazon have begun experimenting with AI-driven dubbing to quickly expand language options, though the quality is not yet on par with human dubbing. Overall, viewer reception of dubbing in the English-Azerbaijani context seems to be pragmatic: people appreciate a good dub when it allows easy viewing (especially for content like animation, documentaries or when showing films on mainstream TV where subtitles might be hard to read), but they also recognise the difference between an excellent dub and a mediocre one. High-quality dubbing can significantly enhance enjoyment and emotional impact, while poor dubbing can detract from it. The ideal, in terms of reception, is when dubbing is effectively *invisible* – as one reviewer quipped, the hallmark of great dubbing is when audiences don't realise the film was originally in another language.

Subtitling Modalities: Constraints and Practice

Subtitling involves displaying written text on screen to convey the dialogue (and sometimes other relevant audio information) in another language. Unlike dubbing, subtitling is an **additive** modality – it adds an element (text) to the audiovisual mix rather than replacing the original audio. Here we discuss the key aspects of subtitling modalities: **time and space constraints**, **readability considerations**, and **synchronisation techniques**.

Time and Space Constraints: Subtitling is governed by strict limits on the amount of text that can be shown, and for how long. These constraints are arguably the most rigid of any translation form. Typically, subtitles are displayed as one or two lines at the bottom of the screen, with a maximum length of around 35–42 characters per line (depending on formatting and the production's guidelines). In professional subtitling practice, it's common to aim for no more than two lines, and to keep each subtitle visible for a duration that allows comfortable reading. A general rule is that subtitles should remain on screen long enough for an average viewer to read them *at least once*, usually falling in the range of **1.5 to 6 seconds per subtitle**, with about 3–5 seconds as a sweet spot for a full two-line subtitle. The *reading speed* is often calculated in characters per second; many standards recommend something like 15 CPS (characters per second) as a maximum, though this can be adjusted for different audiences (e.g. slower for children). In our analysis of the *Avengers* film subtitles, we found that rapid dialogues were condensed so that no subtitle flashed by too quickly – lines were trimmed to essentials, and occasionally a long bit of dialogue was split into two sequential subtitles to give the viewer a chance to catch up. Spatially, subtitles **should not cover important parts of the picture**, and by convention they are placed centered at the bottom, avoiding as much as possible any overlap with speakers' mouths or on-screen text like signs. Studies have noted that subtitles ideally occupy no more than a certain portion of the screen (one source suggests not more than ~20% of screen area) so as

to not distract from the visual content. In practice, this means using a reasonably small font and limiting lines. **Georgakopoulou (2009)** advises that subtitles "take a very discrete form in order to be barely noticeable" – they should deliver information but not draw undue attention. Our observation with *Frozen*'s subtitles was that during visually rich scenes (like musical numbers), the subtitles were kept on the shorter side, presumably to let viewers focus on the animation. Additionally, when multiple characters speak simultaneously or very close together, the subtitler has to make tough choices. One common strategy is *omission* of less important dialogue: for example, in a chaotic battle scene in *The Avengers*, minor background shouts or one-liners might simply not be subtitled to avoid clutter, ensuring the main dialogue gets reading priority. The constraints also extend to line breaks – subtitling guidelines (like the **Code of Good Subtitling Practice** by Ivarsson & Carroll, 1998) emphasize that line breaks should ideally occur at natural linguistic breaks (between clauses or sentences) so that each line is a coherent chunk of meaning. Badly broken lines can confuse viewers. In our case study, the Azerbaijani subtitles generally respected this, although occasionally a line break happened in the middle of an English phrase due to length, but the translator rephrased the Azerbaijani to move a break to a better spot. In summary, the space/time constraints force subtitle translators to be masters of **condensation** – often conveying in perhaps 20 characters what was said in 5 seconds of speech. It's often cited that up to *one-third to one-half of the original dialogue* may be cut or compressed in a typical subtitle track. Crucially, despite these reductions, the goal is to preserve the plot coherence and as much nuance as possible. Good subtitling is sometimes called the art of "saying more with less."

Readability and Style: Because subtitles are meant to be read quickly, **readability** is paramount. This involves choices in phrasing, but also visual presentation. Sentences in subtitles tend to be short and simple, even if the original dialogue was long-winded or complex. Punctuation may be minimized to avoid confusion (e.g. using ellipses or dashes to indicate pauses or interruptions rather than writing long broken sentences). Also, subtitlers often avoid translating every single word if it's not needed – for example, repeated phrases, interjections ("you know", "uh", "like," etc.) might be dropped to free up reading space for more meaningful words. In our looked-at subtitles for *Frozen*, when a character repeats someone's name in surprise ("Anna? Anna!"), the translator only put one instance ("Anna!") in the subtitles, reasoning that the viewer can hear the repetition and it's not necessary to read it twice. This keeps the text concise. **Georgakopoulou (2009)** mentions that arranging words on the screen is crucial for readability. This includes not just line length but also things like font choice (usually a clear, sans-serif font with a dark outline for contrast) and placement. Subtitles generally appear at either the bottom center or split (if two speakers are talking, occasionally you might see text separated by alignment to denote different speakers, but more often a dash at the start of a line indicates a change of speaker). The Azerbaijani subtitles we analysed used the common convention of a **hyphen to indicate speaker change** when two people's dialogue was in one subtitle. For example:

"– Stop right there!
– I'm not afraid."

This way, viewers know two different characters spoke those lines. Another aspect of readability is that the subtitles must synchronize with speech timing (discussed below) – a subtitle that lags behind or comes too early can confuse the audience. Additionally, subtitlers sometimes use *italics* or musical symbols to convey extra information: italics often denote an off-screen voice or a quote, and a musical note (♪) might indicate lyrics being sung. In *Frozen*, for instance, when songs occur, the subtitles turned

italic and used a note, signaling to viewers that these are lyrics. This helps set the expectation that the text is a song, not spoken dialogue, which viewers might read with a different rhythm in mind. Good subtitling style also avoids distracting the viewer. The concept of subtitles being "barely noticeable" means that ideally, viewers almost subconsciously read them while still watching the picture. Achieving this requires well-timed, well-phrased captions that do not puzzle the viewer or force them to backtrack. If a subtitle uses too obscure a word, or a convoluted sentence, the viewer might spend extra time deciphering it and miss the visuals. Thus, subtitlers tend to use **clear, common language** equivalents. For example, if an English character uses a very technical term or a pun that doesn't translate, the subtitle might opt for a simpler paraphrase conveying the gist. It might lose some flavour, but ensures comprehension. The measure of readability success is often viewer feedback or empirical testing: if viewers feel the subtitles were easy to follow and didn't detract from enjoying the film, then the subtitling modality has succeeded.

Synchronization (Timing) in Subtitling: Although subtitling doesn't need to sync lips, it does need to sync in time with the audio and picture events. *Temporal synchronisation* means the subtitle should appear exactly when the dialogue starts (or with a slight fractional delay to account for reading), and disappear when the dialogue ends, or slightly after so the viewer can finish reading. Subtitles should not linger too long onscreen without speech, or else they might confuse the audience or spoil an upcoming line. There's also a convention to avoid having a subtitle change at the same moment as a cut to a different camera shot, because two simultaneous changes (visual cut and text change) can be jarring. So subtitlers often adjust in/out times by a few frames to either coincide with scene cuts (so that the subtitle change happens during a cut if needed, to mask it) or to deliberately *not* change exactly on the cut if it can be helped, thus not overloading the viewer's attention. In the *Avengers* subtitles, during rapid-fire dialogues, we noticed that some subtitles slightly overlapped scenes – if a character finished a sentence just as the scene switched, the subtitle either cleared a tiny bit early or stayed a fraction into the next shot to make sure the viewer had time to read, depending on which was less distracting. It's a delicate judgement call that skilled subtitlers make. Another synchronisation aspect is **audio cues**: sometimes subtitles include notations for sounds or music, especially in subtitles for the hard-of-hearing, but even in standard subtitles one might see a label like "[laughs]" or an italicised *Hahaha* if it's important to know a character is laughing while speaking unintelligibly. In our case studies, the standard translation subtitles did not include sound cues (those are usually in specialised closed captions), but they did occasionally use punctuation to indicate tone – e.g. an exclamation mark for shouted dialogue, or question mark combined with exclamation ("!?") for an exclaimed question. These little touches synchronise with how the line is delivered. Additionally, subtitle translators must be mindful of **scene context**: if a character says "this" while pointing at something, the subtitle might explicitly name the thing ("this plan" or "this car") if it's not obvious to a reader, whereas an audible viewer might catch it from intonation or seeing the object. So synchronization in subtitling is not only temporal, but also semantic with the visual scene. The subtitle has to be in *sync with the content*: showing the right text at the right moment so that it matches the action. For example, in *Frozen*, when a character counts "One, two, three!" and jumps, the subtitles showed each number timed with the counting, ensuring the viewer could follow the countdown visually and via text together. This kind of fine timing creates a satisfying alignment between what one reads and what one hears/sees.

In essence, subtitling modalities operate within tight limitations: every subtitle is a race against time and space. Yet, when done expertly, subtitles allow the audience to understand the foreign dialogue

without feeling overwhelmed by reading. Many viewers report that after a while they forget they're reading subtitles at all – their eyes naturally flick down and up, capturing the meaning. This immersion is the hallmark of well-synchronised, well-crafted subtitles. In the Azerbaijani context, subtitles have opened up access to a wide range of English-language content, from blockbuster films to Netflix series, even though the reading experience may sometimes be a new habit for audiences more used to dubbing on TV. Over time, familiarity with subtitling grows, and as noted, younger generations seem increasingly comfortable with it. Nonetheless, subtitles will always need to respect the cognitive load of reading. The modality demands a balance: giving enough information to understand the story, but not so much text that viewers are taken out of the cinematic experience. Achieving this balance is what the best subtitlers aim for, through the techniques of concise writing, smart timing, and synchronization with the audiovisual flow.

Challenges and Solutions in AVT

Translating audiovisual content is fraught with challenges that go beyond straightforward language translation. Dubbing and subtitling each face their own set of difficulties, particularly when dealing with elements like humour, idiomatic expressions, lip-sync requirements, and format limitations. In this section, we discuss some of the prominent challenges in AVT and the strategies or solutions that translators and adapters employ to overcome them. We will draw examples from English-to-Azerbaijani translation scenarios (as in our case studies) to illustrate these issues.

Translating Humour and Wordplay: Humour is often cited as one of the hardest things to translate. Jokes may rely on wordplay, cultural references, or timing – all of which can be language-specific. In subtitling, the challenge is even greater because one might have to convey a pun or a witty line in just a few words on screen, sometimes *losing the pun* entirely. For example, in *The Avengers*, Tony Stark calls Hawkeye "Legolas" as a quip (referring to the elf archer from *The Lord of the Rings*). An Azerbaijani audience might recognise this reference, but the subtitle translator has to judge if it's widely understood. If not, they might substitute a more local reference for an archer, or simply drop the comparison and translate it as "nice shooting" or a generic tease. **Delia Chiaro (2010)** outlines strategies for humour: leaving it unchanged (hoping the audience gets it), explaining it (not very feasible in fast subtitles), replacing it with a different joke that fits the target culture, or in worst cases, omitting the humouristic element. In *Frozen*, there's a scene with a joke about "I don't have a skull… or bones" (Olaf's line, which is absurd and funny). The Azerbaijani dub managed to capture the humour by preserving the silly tone and using a colloquial expression, whereas the subtitle delivered it more literally but with an ellipsis to hint at the silliness ("Məndə kəllə… yəni sümüklər yoxdur"), hoping the audience finds the statement itself funny. In dubbing, translators have a bit more freedom to craft a joke that suits the spoken language, possibly adding small extra inflections or filler words to time it well. A solution often used in dubbing comedic content is **localisation of humour**: adapting a joke to something the target audience finds funny. For instance, Disney's dubbing practice sometimes changes cultural jokes (like replacing an American celebrity name with someone locally known). For Azerbaijani, the pool of universally recognized local celebrities or references might be smaller in a global film context, so translators might opt for more generic humour that transcends culture (slapstick elements, funny intonations, etc., which the voice actors can enhance).

Idioms and Figurative Language: Idiomatic expressions rarely have direct equivalents across languages. A phrase like "cost an arm and a leg" in English would sound bizarre if translated word-

for-word into Azerbaijani (where an equivalent might be "təxminən canına otarmaq" – not literally about limbs). Subtitle translators often choose to **paraphrase idioms** into plain language due to space constraints and to avoid confusion. If a character says, "We'll cross that bridge when we come to it," the subtitle might just say "Sonra baxarıq" ("We'll see later") conveying the pragmatic meaning without the metaphor. In dubbing, you could potentially use a target-language idiom of similar meaning, but you must consider lip sync and register. If the English idiom had a certain tone (say it was witty or formal), the chosen Azerbaijani phrase should match that tone. One strategy is **equivalent idiom substitution**: find a saying in Azerbaijani that conveys the same idea. If none exists, or it doesn't fit the mouth movements, then a non-idiomatic translation is used. The challenge with idioms is to not translate them literally (which could result in nonsensical or comical outcomes). A classic example often cited in translation classes: the English idiom "kick the bucket" (meaning to die) would mislead if translated literally as kicking a physical bucket. So context is everything. AVT adds the extra layer that sometimes an idiom might appear in on-screen text or a visual pun (imagine a sight gag in a movie that hinges on an English idiom). Then translators might have to add a *translator's note* in subtitles (though this is avoided in professional settings), or creatively alter the scene via dubbing (in extreme cases, some dubs have been known to slightly edit a visual or add a voice-over line to clarify a joke – though this is last resort).

Lip-Sync and Visual Cue Issues: For dubbing, *lip-sync* is a perpetual challenge. We touched on it in the dubbing section, but to highlight a specific problem: consider when an English sentence ends with a word where the actor's lips are closed at the end (for instance, "…job" where the "b" closes the lips). If the Azerbaijani translation naturally would end in a vowel or a sound with open lips, the dub adapter might tweak the word order or choose a synonym to end on a closed sound. For example, if translating "Stop!" – an open-mouthed vowel at the end – into Azerbaijani, one might prefer "Dayan!" (ends with mouth fairly closed) which matches a shout better than something ending wide open. Similarly, if an actor visibly says a short word or syllable, the dub cannot replace it with a long polysyllabic word – even if meaning-wise it would fit – because the extra syllables would either overflow the actor's mouth movement or require artificial condensing by the voice actor (which can sound rushed). **Chaume (2012)** notes that lip-synchrony remains one of the most challenging aspects of dubbing, sometimes requiring compromises in translation. The solution often comes down to the creativity of the adapter and the skill of the voice actor. A voice actor can help mask minor mismatches by using appropriate facial expressions in their voice (for instance, a slight grunt or chuckle can justify a mouth movement). Modern techniques like voice morphing or aligning waveforms are emerging to assist, but human finesse is still key. In our studied *Frozen* scenes, the lip-sync during songs had a few slight discrepancies (since animated mouths move very specifically for lyrics) – the dubbing team solved some by stretching or shortening certain sung vowels, a common trick in song dubbing, effectively matching the animation by adjusting the musical phrasing rather than exact lyrics.

Subtitle Overflow and Segmenting: For subtitling, a big challenge is when characters speak very fast or say a lot in a short time. This leads to *subtitle overflow* – more content than can reasonably be shown. A case in point is rapid banter or overlapping dialogue in something like *The Avengers*. If two characters argue quickly, a subtitler might have to *compress each line heavily* and maybe still show two separate subtitles one after the other in the space of a few seconds. This can test the limits of what viewers can read. One solution is strategic **omission**: decide which bits of dialogue are expendable. Perhaps dropping a repetition, a stutter, or a bit of filler. Another solution is **merging**: if one character

starts a sentence and another finishes it, sometimes a single subtitle can be used with a dash indicating the change of speaker in one continuous line. This saves time as the viewer reads one block instead of two. However, merging only works if the semantic load is light enough. In terms of segmentation (splitting dialogue into subtitle units), guidelines help ensure viewers aren't bombarded. For very fast dialogue, sometimes not everything can be shown – this is an acknowledged limitation, and translators prioritize plot-central dialogue. It's not ideal, but the rationale is that it's better for viewers to catch the main points than struggle to read an impossibly fast subtitle stream.

Localisation vs. Foreignisation Strategies: A broader challenge in AVT is deciding how much to domesticate or foreignise the content. Dubbing tends to domesticate more (since the audience hears their native language spoken by characters, potentially giving the impression the story is culturally closer), whereas subtitling retains a sense of foreignness (hearing original language, seeing perhaps foreign names or signs on screen unchanged). For instance, in translating a cultural element like a reference to Thanksgiving (an American holiday) – a subtitler might leave it as "Thanksgiving" and trust the audience knows it's a holiday, or add a brief "[US holiday]" in a fansub context (professionals likely wouldn't add that explicitly). A dub might translate it to "harvest festival" or some neutral term, or even switch it to a local holiday if the setting allows (though that's rare for live-action; more common in dubbed cartoons). The **solution** to cultural gaps can be **explicatory translation** (explaining within dialogue), but that risks sounding unnatural. Another is **adaptation**: for example, if a character uses an Imperial unit ("miles", "pounds"), a subtitle might convert it to metric for clarity (especially if the number is important), or add a metric equivalent in parentheses if space allows. A dub would likely directly use the metric (saying "kilometers" instead of "miles"), since the spoken word can be a bit more flexible with timing. These micro-decisions ensure that the target audience isn't left confused by unfamiliar measures or institutions.

Quality Control and Consistency: Ensuring consistency throughout a translation is also a challenge. In a long series, making sure a catchphrase or a recurring term is translated the same way every time is vital for continuity. Subtitlers maintain **translation glossaries** for series to keep names and terms uniform. Dubbing directors perform a similar role, guiding voice actors to keep character voice and style consistent. A known issue arises with franchises: sometimes different translators handle different installments. For instance, if *Frozen 2* were subtitled/dubbed by a different team than *Frozen*, the term "Snowgie" or some coined word might end up different. Fans notice these things. Thus, an effort is made to consult previous translations (if available) or have a supervisor oversee consistency.

Technological Aids and Future Solutions: In recent years, technology has started to assist AVT practitioners. *Machine translation* can produce draft subtitles which human translators then polish – helpful for speed, though machine output is often not reliable for nuanced film dialogue. *AI dubbing* as mentioned is emerging: synthetic voices that can mimic actors and automatically time to lips. While not yet a replacement for human artistry, it could solve some cost and speed issues for less widely translated languages (perhaps in the future, Azerbaijani content could be dubbed by AI when budgets don't allow a full studio dub, as a way to increase output). However, these raise questions of quality and reception – a solution that is promising but also viewed cautiously by professionals who emphasise the creative nuance a human brings.

In summary, the challenges in AVT – humour, idioms, synchronisation, brevity, cultural references – are met with a toolbox of solutions: **adaptation, condensation, substitution, omission,**

compensation, and technical finesse. The translator must often think like a writer, director, and audience all at once: what choice conveys the meaning, fits the technical limits, and still gives the viewer an experience as close as possible to that of the original audience? In the English-Azerbaijani context, these challenges can be amplified by the linguistic differences and the fact that many Western cultural concepts are relatively foreign. Yet, as our examples show, skilled translators find creative ways to bridge the gap – whether by inventing a clever equivalent joke in Azerbaijani, tweaking a line to rhyme and lip-sync in a song, or elegantly trimming dialogue in subtitles so that nothing critical is lost. Every solution is a little compromise, but done right, the overall effect is a translated work that resonates with audiences while remaining faithful in spirit to its source.

CONCLUSION

This study compared dubbing and subtitling as key modalities of audiovisual translation, focusing on their application in English-to-Azerbaijani media. Both methods offer distinct advantages: dubbing supports immersion and accessibility, particularly for children, while subtitling preserves the original performance and suits multilingual or cost-conscious contexts. Our case analyses of *Frozen* and *The Avengers* illustrated how each modality manages linguistic, technical, and cultural challenges—from synchronisation and lip-sync in dubbing to brevity and timing in subtitling.

Ultimately, there is no one-size-fits-all solution in AVT. The choice depends on the audience, content type, and cultural preferences. In Azerbaijan, both modes play important roles: dubbing dominates TV and children's content, while subtitles are common in cinema and streaming.

As technology evolves, AI dubbing and real-time subtitling may shape future translation practices. Still, the creative decisions made by human translators remain essential in maintaining authenticity and audience engagement. Whether through sound or text, the goal of AVT remains constant: to connect stories across languages and cultures.

REFERENCES

Alisoy, H. A. H. (2023). *Object Clauses and Difficulties in Their Translation. Nakhchivan State University, English and Translation Chair.*

Alisoy, H., Mammadova, I., Asadova, B., Ismayilli, F., & Aliyeva, T. (2024). The future of language education: Integrating smartphones into the curriculum. *Edelweiss Applied Science and Technology*, *8*(6), 4539-4556.

Antonini, R., & Chiaro, D. (2009). *The perception of dubbing by Italian audiences*. In J. Díaz Cintas & G. Anderman (Eds.), Audiovisual Translation: Language Transfer on Screen (pp. 97–114). Basingstoke, UK: Palgrave Macmillan.

Babayev, J., & Alaviyya, N. (2023). TRANSLATION PROCEDURES OF CULTURE BOUND-TERMS (CBTs). *Journal of Science. Lyon*, *48*.

Babazade, Y. (2025). Speech Acts and Hidden Meaning: A Journey into Pragmatics. *Acta Globalis Humanitatis Et Linguarum*, *2*(1), 221-228. https://doi.org/10.69760/aghel.02500128

Chaume, F. (2012). Audiovisual Translation: Dubbing. Manchester, UK: St. Jerome Publishing.

Díaz Cintas, J., & Remael, A. (2007). Audiovisual Translation: Subtitling. Manchester, UK: St. Jerome Publishing.

Georgakopoulou, P. (2009). Subtitling for the DVD industry. In J. Díaz Cintas & G. Anderman (Eds.), Audiovisual Translation: Language Transfer on Screen (pp. 21–35). Basingstoke, UK: Palgrave Macmillan.

Gottlieb, H. (1994). Subtitling: Diagonal translation. *Perspectives: Studies in Translatology, 2*(1), 101–121.

Ivarsson, J., & Carroll, M. (1998). Subtitling. Simrishamn, Sweden: TransEdit.

Javid, B. (2023). Significance of language skills and language aspects in sight translation. *Danish scientific journal, 1*(77).

Lavaur, J.-M., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology, 46*(6), 455–462.

Luyken, G.-M., Herbst, T., Langham-Brown, J., Reid, H., & Spinhof, H. (1991). Overcoming Language Barriers in Television: Dubbing and Subtitling for the European Audience. Manchester, UK: European Institute for the Media.

Mammadova, I. (2025). Cognitive and Pedagogical Dimensions of Translation: A Theoretical and Practical Exploration. *Acta Globalis Humanitatis Et Linguarum*, *2*(1), 213-220. https://doi.org/10.69760/aghel.02500127

Mirzayev, E. (2025). Pronunciation issues in translation: challenges and implications. *Acta Globalis Humanitatis Et Linguarum*, *2*(3), 41-50. https://doi.org/10.69760/aghel.0250020006

Naghiyeva, G. (2025). Revamping Traditional Methods: Evaluating the Grammar-Translation Method in Modern Language Teaching. *Acta Globalis Humanitatis Et Linguarum*, *2*(1), 88-97. https://doi.org/10.69760/aghel.02500111

Pedersen, J. (2011). Subtitling Norms for Television: An Exploration Focusing on Extralinguistic Cultural References. Amsterdam, Netherlands: John Benjamins.

Perego, E., Del Missier, F., & Bottiroli, S. (2015). Dubbing versus subtitling in young and older adults: Cognitive and evaluative aspects. *Perspectives: Studies in Translatology, 23*(1), 1–21.

Shahmerdanova, R. (2025). Artificial Intelligence in Translation: Challenges and Opportunities. *Acta Globalis Humanitatis Et Linguarum*, *2*(1), 62-70. https://doi.org/10.69760/aghel.02500108

Tveit, J.-E. (2009). *Dubbing versus subtitling: Old battleground revisited*. In J. Díaz Cintas & G. Anderman (Eds.), Audiovisual Translation: Language Transfer on Screen (pp. 85–96). Basingstoke, UK: Palgrave Macmillan.