Indrė Koverienė
Vytautas Magnus University, Lithuania

LIP SYNCHRONY OF BILABIAL CONSONANTS IN THE LITHUANIAN DUBBED LIVE-ACTION FILM *A DOG'S WAY HOME*

Annotation. This article explores the intricate nature of lip synchrony in dubbing as an audiovisual translation mode, highlighting its significance beyond aesthetic considerations. Often overlooked during the translation process for dubbing, the mismatch between auditory and visual signals caused by unsynchronized lip movements can have a negative impact on speech perception. Moreover, with the constant rise of audiovisual content, achieving satisfactory lip synchrony remains a crucial challenge that demands attention. Although scholars such as Fodor (1976) and Chaume (2012) have recognized the importance of bilabial consonants in maintaining lip synchrony, there is a notable research gap specifically focusing on this aspect within the Lithuanian dubbed industry. To address this gap, this study investigates the lip synchrony of bilabial consonants in the Lithuanian dubbed version of the live-action film *A Dog's Way Home* (2019). By employing a comparative research approach that integrates qualitative and quantitative analyses, the study draws on theoretical perspectives presented by McGurk and MacDonald (1976), Fodor (1976), Chaume (2004, 2012), and Koverienė (2015). The analysis of the cinematographic shots reveals that only a small percentage (19%) of the bilabial phonemes in the source language utterances were visually prominent instances, and consequently chosen for detailed examination of lip synchrony. This finding suggests that strict adherence to lip synchrony may not be crucial in numerous instances, allowing translators to have greater flexibility in their approach. The target language utterances demonstrate a relatively high percentage (approximately 76%) of synchronous cases for bilabials, particularly in proper names and international words. However, the research also reveals additional 94 instances of bilabials in the target language that were not found in the source language utterances, leading to dischronemes. Therefore, audiovisual translators are urged to consider not only how to maintain the lip synchrony of bilabial consonants in the target language, but also how to avoid the emergence of new bilabials.

Keywords: audiovisual translation; bilabial consonants; dischroneme; dubbing; lip synchrony; visemes.

Introduction

The maintenance of lip synchrony (LS) holds significant importance across diverse areas, including speech recognition technology, computer facial animation, and language teaching. However, considering the rising influence of audiovisual (AV) culture on a global scale, the need to achieve LS in

the translation of television series, live-action or animated feature films becomes increasingly apparent.

The issue of poor LS becomes particularly noticeable in dubbing, where the original speech is replaced with a new track to create the illusion that the AV content was initially produced in the viewer's native language. Despite the substantial role of AV culture in the lives of many individuals, the impact of unsynchronized lip movements and the translated version is often overlooked. However, this discrepancy not only distorts the perception of speech sounds but also impedes speech comprehension and gives rise to cognitive dissonance among viewers (Massaro, Cohen & Smeele, 1996; Romanski & Hwang, 2012). Therefore, to avoid negative consequences and achieve a natural viewing experience in dubbing, it is significant to ensure satisfactory LS, especially in close-up and extreme close-up shots that demand particular attention to detail (Chaume, 2012, 68).

Although LS has been a subject of research for many scholars in AV translation, such as Fodor (1976), Whitman-Linsen (1992) or Chaume (2004, 2012), the study of LS in the Lithuanian context remains limited. Despite having a 19-year history of dubbing in independent Lithuania since the release of the first Lithuanian dubbed animated feature film, Shrek 2, in 2004, comprehensive research on this topic is lacking. Analysing LS involves a complex examination of the complete sound inventories in both the source language (SL) and the target language (TL), which poses a challenge. To address this, the primary focus of this study is to investigate the LS of bilabial consonants, as they are the most visually prominent phonemes. The choice of this research object is also motivated by the absence of systematic studies in the field, particularly in relation to consonant phonemes.

The existing research on LS in Lithuanian-dubbed content has mainly focused on two studies. The first one is Koverienė's dissertation titled *Dubbing as an Audiovisual Translation Mode: English and Lithuanian Phonemic Inventories in the Context of Visual Phonetics* (2015). However, this study examines the phonemic inventories of the Lithuanian and English languages, and how they are grouped into visemes, rather than specifically analysing lip synchrony of an AV product. The second study, *Lip Synchrony of Rounded and Protruded Vowels and Diphthongs in the Lithuanian-Dubbed Animated Film*

Cloudy with a Chance of Meatballs 2 (Koverienė & Čeidaitė, 2020), focuses only on LS related to protruded vowels and diphthongs. These studies provide valuable insights but do not cover LS of visually prominent consonants. Therefore, this research aims to fill in this gap by focusing specifically on the LS of bilabial consonants, which play a crucial role in maintaining visual coherence in dubbed content.

The empirical analysis of the present research focuses on the English and Lithuanian dubbed versions of the live-action film *A Dog's Way Home* (2019, directed by Charles Martin Smith). The methodological framework employed in this study encompasses both qualitative and quantitative research approaches, drawing on the theoretical insights provided by McGurk and MacDonald (1976), Fodor (1976), Chaume (2004, 2012), and Koverienė (2015).

Rational Behind: Auditory-Visual Phenomenon

Lip synchrony and lip reading are interconnected as they both contribute to effective communication and understanding of spoken language. The early studies of speech perception consider the role of a visual component or lip-reading as an alternative or complementary mode to hearing (Pelson & Prather, 1974; Dodd, 1977). However, even though visible information of sound articulation often referred to as lip-reading or speechreading is the most useful for the hard of hearing, interlocutors in a noisy environment as well as identification of a foreign accent or complicated speech (Rosenblum, Yakel & Green, 2000, 1), it affects perception even when communication conditions are excellent (Irwin, Avery, Brancazio, Turcios, Ryherd & Landi, 2018, 40).

This "previously unrecognised influence of vision upon speech" was first proved in one of the most critical studies of speech perception called the *McGurk effect* (McGurk & MacDonald, 1976, 746). In this study, McGurk and MacDonald underlined a close relationship between visual and audio signals by demonstrating that lip-reading information is automatically integrated into speech perception, should the context of the conversation be ideal. The McGurk effect has been demonstrated in many languages for people of all ages. However, this illusion persists even if the experiment participants

are informed about the results in advance (McGurk & MacDonald, 1976). Overall, AV speech perception is the result of integrating lip-reading and auditory information.

Extensive research in the scientific literature only confirms the importance of audio and visual information in speech perception, which is conditioned by the movements of speech production organs (Dodd & Campbell, 1987; Massaro & Simpson, 1987). To explain this illusion, Summerfield (1987) claimed that visual and acoustic information complement each other and are called VPAM (Visual: Place; Auditory: Manner). Visual information improves the perception of consonants by 57%, vowels by 30%, monosyllabic words by 39%, and short expressions by 17% (O'Neill, 1954). Also, lip-reading is more effective when not only the lips but also the entire face of the interlocutor together with important facial expressions are visible. Teeth, tongue, and oral cavity can also become an essential clue in the process of reading from the lips (Williams, Rutledge, Katsaggelos & Garstecki, 1998).

Comparing syllable structural parts (onset, nucleus, coda), the onset of a syllable is generally easier to identify than the coda, particularly in short monosyllabic words. This is because the onset provides more information about the identity of the syllable and is more perceptually salient, with more acoustic energy and a stronger neural response. However, in longer utterances, other factors such as context and lexical stress may also play a role in determining the salience and identification of onset and coda (Greenberg, Carvey, Hitchcock, & Chang, 1997; Cutler, 2012). Finally, shorter stimuli are better perceived because they require less mental effort to process and remember (Miller, 1956; Hughes, Vachon & Jones, 2005).

The scientific literature supports the critical role of LS in AV translation. Beyond its role in enhancing naturalness and entertainment value, LS plays a crucial impact in ensuring accurate speech perception. By maintaining visual coherence between lip movements and spoken language, LS minimizes confusion and facilitates effective communication. The integration of auditory and visual information significantly contributes to accurate speech perception. Therefore, the importance of LS in AVT extends beyond aesthetic considerations, emphasizing its fundamental role in facilitating clear and meaningful communication.

Lip Synchrony in Dubbing

Dubbing is an interlinguistic AVT mode that entails substituting the original language soundtrack with a translated soundtrack (Matkivska, 2014, 39–40). It guarantees the translator greater flexibility in rendering the original text. However, it also imposes limitations to ensure synchronization between the audio and visual elements. This necessitates substantial modifications to the translated text (Martinez, 2004, p. 5).

Synchresis or a close relation between auditory and visual phenomena occurring simultaneously, does not always presuppose perfect synchrony (Chion, 1994, p. 5–6). It leads to the emergence of various levels and typologies of synchronisation (Fodor, 1976; Whitman-Linsen, 1992; Chaume, 2004; 2012). In his book *Film Dubbing: Phonetic, Semiotic, Esthetic and Psychological Aspects* (1976), Fodor introduced three key types of synchrony: phonetic synchrony, which focuses on matching the lip movements of the characters on screen; character synchrony, which involves aligning the dubbing actor's voice in terms of timbre, tempo, and power with the visual characteristics of the speaking character; and content synchrony, which aims to transfer the main idea of the source text (ST) into the target text (TT) while maintaining its original meaning (Fodor, 1976, p. 10). The achievement of *satisfactory synchrony* is crucial to ensure the attractiveness and appeal of the dubbed AV product to the target audience.

Instances of mismatch between the mouth movements of a character on screen and the corresponding sound result in cases of *dischrony* (Harris, 1996, p. 75). The term *synchroneme* is employed to denote a single occurrence of synchrony, while *dischroneme* refers to a case of dischrony (Fodor, 1976, p. 10). Fodor describes the sporadic nature of these non-coincidences as *discontinuous dischrony*, whereas *continuous dischronies* are present throughout an entire film or program and are thus more easily perceptible to the audience (Fodor, 1976, p. 80).

The level of sensitivity to dischrony varies among individuals and is influenced by their memory type. Viewers with a visual-motor memory type are more likely to notice instances of lip dischrony compared to those with a dominant auditory memory type. Individuals whose motor memory primarily

relates to their limbs tend to be less sensitive (Fodor, 1976, p. 51). However, individuals whose memory is associated with speech tend to find dischrony more irritating. These individuals pay close attention to their way of speaking (Fodor, 1976, p. 50–51).

This research is based on Chaume's typology of synchronies, encompassing isochrony, kinesic synchrony, and lip synchrony. *Isochrony* denotes the synchronization of translation duration with characters' on-screen utterances (Whitman-Linsen, 1992, p. 28). *Kinesic synchrony* pertains to aligning the translation with the actors' body movements. In contrast, *lip synchrony* involves aligning the translation with the articulatory movements of on-screen characters, particularly noticeable in close-up and extreme close-up shots (Chaume, 2012, p. 68–69). Chaume stresses the importance of open vowels and bilabial consonants for dubbing, however, the scholar proposes (2004, p. 74) to substitute SL bilabial consonants (/p/, /b/, /m/) with not only TT bilabial but also labiodental consonants (/f/, /v/).

Chaume suggests a few techniques which could be employed to preserve LS in dubbing including repetition of identical or similar words in both the SL and TL utterances; changing the word order to align with similar or identical phonemes; substituting target words with synonyms, antonyms, hypernyms, hyponyms, or other stylistic resources while maintaining the original meaning; reducing or amplifying words, phrases, or sentences as needed; employing omission or addition of elements, even though they may be categorised as translation errors in written translations, is permissible in the context of dubbing (2012, p. 74–75).These techniques are applied with the primary aim of achieving phonetic equivalence and are commonly practised in professional dubbing.

Visemes and their Application in Dubbing

The process of speech production involves several organs, including the lips, teeth, tongue, jaw, velum, larynx, and lungs, but only a few of these are visible to the audience, such as the lips, teeth, jaw, and tongue. When considering consonants in dubbing, it is important to focus on phonemes whose internal structure is informative in terms of visual expression. Labial and
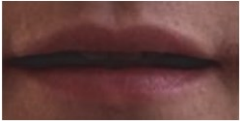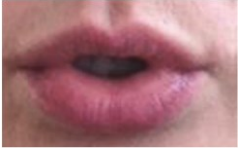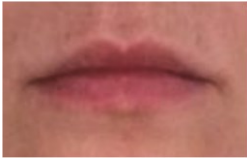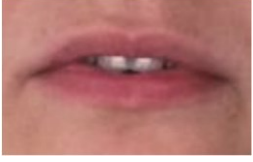
labiodental consonants are particularly prominent to visual cues based on their place of articulation. Bilabial consonants in English and Lithuanian (/p/, /b/, /m/, /pʲ/, /bʲ/, /mʲ/) are produced by pressing the lips together, while labiodental consonants (/f/, /v/, /fʲ/, /vʲ/) involve the lower lip touching the upper front teeth (Roach, 2009, p. 40; Pakerys, 2003, p. 73). The lips and teeth are particularly important since they are the most visually prominent organs during the articulation of the mentioned sounds.

An English consonant phoneme /w/ shares the same active organ of speech as bilabial and labiodental consonants, which is lips, thus it can also be regarded as a visually prominent sound. However, /w/ is at the same time a velar sonorant or labiovelar articulated with the rounded lips, and the raised back of the tongue. The visual expression of the sound /w/ corresponds to the position of the protruded lips, and even though articulating the consonant /w/ the lips open forward to a lesser degree, the visual expression of this sound is closer to the close back vowel phoneme /uː/ (Roach, 2009, p. 50). In Lithuanian, the consonant phoneme /w/ does not exist, but occurs as the allophone of a consonant phoneme /v/, for example, (was) [ˈbʊvʷɔ:] (Aprijaskytė-Valdšteinienė, 1960, 184). Due to its different articulation a labiovelar /w/ cannot be used interchangeably with bilabial or labiodental sounds in dubbing; therefore, this sound has not been included in research.

Phonemes encompass both vowels and consonants, serving as fundamental units in language that contribute to the meaning of words. When examining sounds, it is important to shift our focus from individual letters to phonemes (Roach, 2009, p. 2). Based on the assessment of sound production, it is possible to group phonemes into several categories to know which sounds can be used interchangeably in dubbing. *Visemes* consist of phonemes that share a similar lip appearance (Bear & Harvey, 2017, 1). While each phoneme belongs to a specific viseme category, multiple phonemes can be represented by a single viseme. The ability to effectively employ visemes can greatly enhance the dubbing process, enabling AV translators to identify and interchange similar sounds more efficiently.

Table 1

Visemes of English and Lithuanian phonemes (Koverienė, 2015, 17–18)

Viseme	Visual Features	Lithuanian phonemes	English phonemes
1.	Considerable separation of jaws	[äː], [âˑɪ̯], [âˑʊ̯], [æː], [æ̂ˑɪ̯]	[ɑː], [aɪ], [aʊ], [æ]
2.	Neutrally opened mouth	[ʋ], [t̑s̑], [t̑s̑ʲ], [t̑ʃ̑], [t̑ʃ̑ʲ], [d̪], [d̪ʲ], [d̑z̑], [d̑z̑ʲ], [d̑ʒ̑], [d̑ʒ̑ʲ], [ˈɫ], [lʲ], [n̪], [n̪ʲ], [s̪], [s̪ʲ], [ʃ], [ʃʲ], [r̄], [rʲ], [t̪], [t̪ʲ], [z̪], [ʒ], [ʒʲ], [x], [c̜], [g], [ɟ], [ɣ], [j], [ɉ], [k], [c̑], (allophones [ŋ] and [ŋʲ])	[ʌ], [ə], [θ], [ð], [tʃ], [dʒ], [l], [n], [s], [ʃ], [t], [z], [ɹ], [ʒ], [g], [h], [j], [k], [ŋ]
3.	Spread lips	[ɪ], [iː], [eɪ̯ˑ], [ʋɪ̯ˑ], [îe], [ɪĕ], [ɛ], [eː]	[ɪ], [eɪ], [ɪə], [e], [eə], [ɜː]
4.	Rounded and protruded lips	[ʊ], [uː], [ôɔ], [ʊɔ], [ôɪ̯], [ʊɪ̯], [eŭˑ], [ʋŭˑ], [ɔ], [oː], [ɔ̂ɪ̯], [ɔ̂ʊ̯]	[ʊ], [uː], [ʊə], [əʊ], [ɒ], [ɔː], [ɔɪ], [w], [oʊ]
5.	Closed lips	[p], [pʲ], [b], [bʲ], [m], [mʲ]	[p], [b], [m]
6.	The lower lip touching the upper teeth	[f], [fʲ], [v], [vʲ]	[f], [v]

As shown in Table 1, Lithuanian and English phonemes can be classified into six main categories based on their articulation, including the level of jaw separation, lip shape, and mouth features. The number of visemes applied in dubbing is comparatively limited in contrast to other fields like computer facial animation (Bozkurt, Erdem, Erzin, Erdem & Ozkan, 2007, p. 2). Nevertheless, this limited number is essential to ensure their application by AV translators in dubbing.

This research focuses on the analysis of the fifth viseme containing Lithuanian /p/, /pʲ/, /b/, /bʲ/, /m/, /mʲ/ and English /p/, /b/, /m/ bilabial consonants which are articulated with the closed lips. It is important to note that labiodental Lithuanian /f/, /fʲ/, /v/, /vʲ/ and English /f/, /v/ consonants fall into the sixth viseme characterised by the lower lip touching the upper teeth. The English labiovelar /w/ can be replaced by a number of Lithuanian phonemes as all of them are pronounced with rounded and protruded lips: /ʊ/, /uː/, /ʊ̃/, /ô̰ɔ/, /ʊ̯ːˑ/, /ô̰ɪ̯/, /eǔ̯ˑ/, /ɐǔ̯ˑ/, /ɔ/, /oː/, /ô̰ɪ̯/, /ô̰ʊ̯/ (Koverienė, 2015, p. 17–19).

It is crucial to examine the occurrence of sounds that are easily distinguishable not only at the onset and coda of words but also when adjacent to vowels. Bilabial and labiodental consonants are distinctive due to their manner of articulation, making them easier to differentiate from adjacent vowels, whereas consonants articulated further back in the oral cavity are less perceptible (Fodor, 1976, p. 34).

To achieve satisfactory synchronisation, it is possible to replace a phoneme of one viseme with a different one as long as it belongs to the same viseme. In this research, translation will be considered satisfactory if a phoneme is replaced with one belonging to the same viseme.

Analysis of Bilabial Sounds in the Live-Action Feature Film
A Dog's Way Home

This empirical study focuses on analysing bilabial consonants assigned to the fifth viseme (Koverienė, 2015) in the live-action film *A Dog's Way Home* (2019). It emphasizes the significance of conducting a close analysis to identify specific episodes where translators must prioritize LS over semantic synchrony

and demonstrate high accuracy. To achieve this, close-up shots that capture
the subject from the shoulders to the head were specifically chosen to limit
the number of phonemes included in further LS research. Additionally,
the identification of on-screen and off-screen episodes was crucial in detecting
instances where LS should not be preserved. Consequently, this focused
approach has resulted in a reduced frequency and distribution of bilabial
consonants in the SL utterances.

Figure 1

Total Number of Bilabial Phonemes in the SL Utterances Vs Total Number of
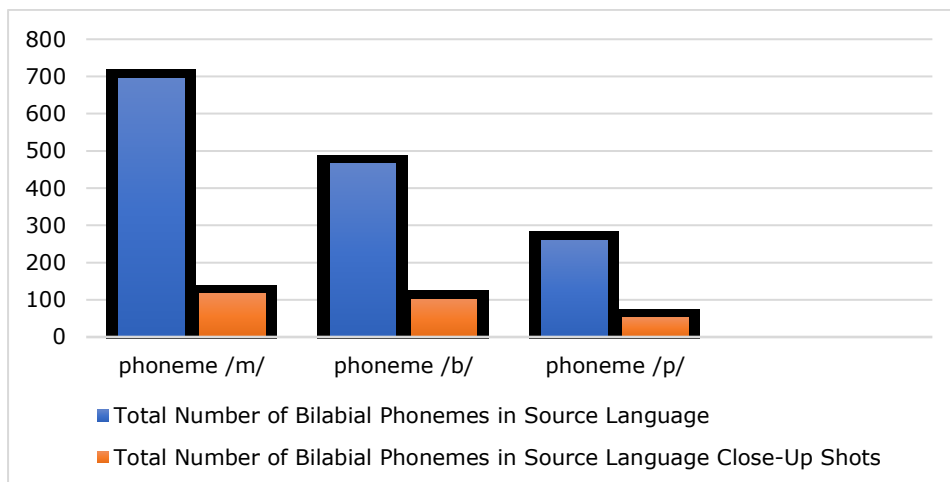*Bilabial Phonemes in the Close-Up Shots of the TL Utterances*

Figure 1 presents a comprehensive analysis of bilabial phonemes (/m/,
/b/, /p/) in the SL utterances. The analysis involved a total of 1,419 identified
phonemes[1], out of which 271 were selected for further examination of LS,
based on careful consideration of on/off screen episodes and shot sizes.

Further examination of the data reveals that the phoneme /m/ appeared
695 times; however, only 117 instances (17% of occurrences) were visually
prominent and considered for analysis. This phoneme was notably present in

---

[1] This research emphasizes the analysis of phonemes rather than graphemes. This is
because the English orthography, or written representation, may not always perfectly
align with the corresponding phonetic pronunciation as it is exemplified by instances such
as /fəʊn/ for *phone* and /ˌrɛkəˈmɛnd/ for *recommend*.

key phrases related to the leitmotif of *home* such as go *home* as well as in the context of possessive determiners like *my*. Similarly, the bilabial phoneme /b/ appeared 465 times, with 102 visually significant instances. The higher frequency of the /b/ phoneme (accounting for approximately 22% of occurrences) can be attributed to the inclusion of specific lexical items. For example, the protagonist's name, *Bella* was mentioned 100 times in close-up conversations, contributing to the frequent occurrence of the /b/ sound. Additionally, the conjunction *but* was consistently retained in the translated utterances, further contributing to the prominence of the /b/ phoneme. On the other hand, the least frequent phoneme was /p/, which occurred 259 times. Out of these occurrences, only 52 (approximately 20%) were included in the LS analysis. The presence of the /p/ phoneme was notable in lexical items such as *pit bull*, *property*, *puppy* and others.
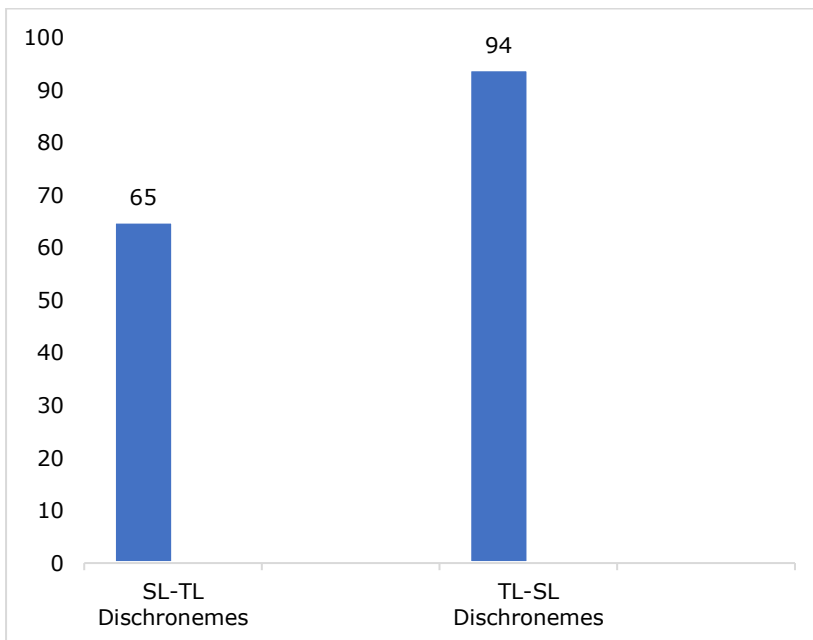
When considering all the bilabials chosen for further analysis, the findings indicate that LS was visually prominent in a mere 19% of the bilabial cases. This relatively low percentage serves as a valuable facilitator for translators during the process of dubbing, as it suggests that maintaining strict LS may not be a crucial factor in many instances.

To evaluate LS, a specific procedure was implemented. The first criterion for maintaining LS involved replacing English consonants with Lithuanian consonants from the fifth viseme category. For instance, the phoneme /b/ in the English word *body* /ˈbɒdi/ was substituted with the phoneme /m/ in the Lithuanian word *žmogus* /ʒmoːˈgʊs̪/ (*human*). The second criterion focused on LS being preserved even when the TL utterances contained a single bilabial instead of two consecutive bilabials found in the SL utterances. For example, in the SL word *combined* /kəmˈbaɪnd/ from the utterance *Chuck, I get more complaints about you than all the other ACOs combined*, or the separate words *am pulling* /əm ˈpʊlɪŋ/ from the utterance *I am pulling you out of the field*, the consecutive phonemes /m/ and /b/ or /m/ and /p/ serve as the coda and onset of two syllables, with the lips remaining close during their production. Hence, it is sufficient to render them as a single phoneme /b/ in the TL word *nebegali* /n̪ʲɛbʲɛˈɡäːlʲɪ/ (*can't*) in the TL utterance *Čakai, dėl tavęs aš gaunu daugiausiai skundų ir ilgiau taip tęstis jau nebegali* (*Chuck, you are the reason I get the most complaints and it can't go on like this any longer*),

or a single phoneme /m/ in the word *nuimu* /ˈn̪ʊ-ɪmʊ/ (*taking off*) of the TL utterance *Aš nuimu tave nuo tarnybos* (*I'm taking you off duty*).

Figure 2

SL-TL and TL-SL Dischronemes

Based on the analysis of the TL utterances, it can be concluded that LS was maintained in approximately 76% of cases, with 65 dischronemes and 206 synchronemes identified out of the 271 visually prominent bilabial consonants. These findings suggest that the translators demonstrated adherence to the requirements of LS during the dubbing process. However, a significant observation arises from Figure 2, which alters the initial interpretation. It is evident that the TL utterances include 94 newly introduced TL bilabials that are absent in the SL utterances. This introduces two types of dischronemes: TT-ST and SL-TL dischronemes. The presence of these dischronemes results in a mismatch between the auditory and visual signals, potentially leading to negative implications for speech perception and the overall quality of the dubbing.

Specific Occurrences of Synchronemes and Dischronemes

Further analysis provides evidence regarding the translation of specific instances of synchronemes and dischronemes.

Table 2

Examples of Synchronemes of the Bilabials /p/, /b/ and /m/

No, facial view	Screenshot	Time code	SL utterance	TL utterance / Back translation	SL transcription	TL transcription	SL phoneme	TL phoneme	SL – TL Visemes
1. front		20.12	As you may know, **pit bulls** are illegal in the City of Denver.	Manau žinote, kad **pitbuliai** yra uždrausti Denveryje./ I think you know that pit bulls are banned in Denver.	ˈpɪtˌbʊlz	ˈpɪtbʊliɛĭ	p	p	5-5
2. ¾		19.45	They do when our **board** member calls.	Jie gali, nes **prašo** tarybos narė./ They can because the councilwoman asks.	bɔd	ˈprɑʃoː	b	p	5-5
3. front		10.57	My son isn't gonna back off, *Mr. Beckenbauer.*	Sūnus nesitrauks, **pone Bekenbaueri.**/ The son will not withdraw, Mr. Beckenbauer.	ˈmɪstə ˈbɛkenˈbauer	ˈpɔɲɛ ˈbɛcɛɲbâˈŭɛrɪ	m b	p b	5-5

Table 2 includes examples that meet the three criteria of satisfactory synchronisation. The first example does not raise any difficulties for translators due to phoneme correspondence in both languages. In the latter case the lip movements of the on-screen individual, an animal control officer, are matched with the articulation of two bilabial phonemes /p/ and /b/ by the voice actor, with

the SL word /pɪtˌbʊlz/ *pit bulls* being replaced by the phonemes /pʲ/ and /b/ in the TL version *pitbuliai* /ˈpʲɪdbʊlʲeɪ̯/ *(pit buls)*. The visual articulation of the phoneme remains unchanged despite the palatalization of the SL phoneme /pʲ/. As well as in this case, a repetition technique is applied in dubbing when the SL and TL words are similar, as noted by Chaume (2012, p. 74). There are a lot of examples where LS was maintained in this film due to the repetition of proper names such as *Beckenbauer*, *Blanche*, *Basra*, *Molly*, *Millie*, *Mack* or the name of the dog *Bela*, which was repeated 100 times throughout the film. Also, LS of bilabial phonemes was preserved in a number of international words such as *information*, *biometric*, *racism*, *business*, *captain* and others. This type of translation, involving the repetition of SL words, is often viewed as appropriate and effective in maintaining LS, meaning and connotations of the original text.

Regarding isochrony, the SL utterance *As you may know, pit bulls are illegal in the City of Denver* consists of 17 syllables, while the TL version *Manau žinote, kad pitbuliai yra uždrausti Denveryje* (*I think you know that pit bulls are banned in Denver*) contains 18 syllables. To decrease the number of syllables in the TL and, most importantly, to avoid dischroneme with the bilabial /m/ at the onset of the first TL syllable, an alternative translation is proposed: *Turbūt žinote, pitbuliai yra uždrausti Denveryje* (*As you may know, pit bulls are banned in Denver*). While this suggested version is not the optimal translation, it does have the advantage of not starting the first syllable with the bilabial /m/, thus avoiding potential dischroneme and helping maintain better synchrony between the SL and TL utterances. Kinesic synchrony is maintained as the animal controller is depicted engaged in a conversation at the door and pointing towards the listeners while addressing Bela's owners.

In the second example provided in Table 2, it is evident that the bilabial phoneme /b/ in the SL word *board* /bɔːd/ is successfully replaced by the phoneme of the fifth viseme /p/ in the TL word *prašo* /ˈprɑːʃoː/ (*asks*). Both of these phonemes occur at the beginning of the sixth syllable in the entire utterance and function as the onset of the syllables. Therefore, the visual representation of these phonemes is equally significant in comparison to the bilabials in the coda of the syllable.

The last utterance presented in Table 2 demonstrates examples of

synchronemes and dischronemes, illustrating cases where the SL utterance contains five bilabials, but only three of them are accurately preserved in the TL. The adaptation of the TL word *pone* /ˈpɔɲ̟ɛ/ (*mister*) and the proper name *Bekenbaueri* /bʲɛcɛn̪ˈbâ·ʊɛrʲɪ/ (*Beckenbauer*) produces synchronemes with the bilabials /p/ and /b/. However, the initial bilabial in the SL utterance, represented by the consonant phoneme /m/ in *my* /maɪ/, becomes a dischroneme in the TL due to the presence of the consonant phoneme /s/ in the word *sūnus* /s̪uːˈn̪ʊs̪/ (*son*), assigned to the second viseme. Due to coarticulation the phoneme /uː/ strongly influences the visual expression of /s/, causing the word to start with rounded and protruded rather than closed lips.
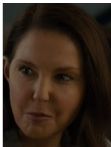
To achieve phonetic equivalence and maintain synchrony, a literal translation of the SL word *my* into the TL word *mano* /ˈmɐn̪oː/ (*my*) would be sufficient. However, this modification would result in an increased number of syllables, specifically 14 syllables in the TL compared to 13 syllables in the SL. Nevertheless, this change aligns with the principles of isochrony as the visual expression of the final short syllable *ri* /rʲɪ/ in the TL remains unchanged and maintains the same lip position as the corresponding syllable *er* /ɛr/ in the SL.

Another improvement that could be made to maintain synchrony of the bilabial /b/ in the word *back* /bæk/ is to replace it with /p/ by adding a prefix *pa* /pɐ/ to the TL word *nesitrauks* /n̪ʲɛs̪ʲɪˈt̪rɐ·ʊks̪/ (*will not withdraw*), resulting in the utterance *Mano sūnus nepasitrauks, pone Bekenbaueri* (*My son will not withdraw, Mr. Beckenbauer*). This modification would preserve the original meaning and enhance the natural flow of the conversation. However, it should be noted that the suggested alternative would result in a higher number of syllables in the TL, specifically 15 syllables. Therefore, the decision to implement this change would depend on the film director's judgment regarding the voice actor's ability to articulate the sounds at a faster pace.

If the discroneme discussed above remains unchanged, it can be considered a minor mistake, as this is a long utterance and there are even five bilabials present, with one of them being close to the analysed phoneme /b/. Finally, it is challenging for viewers to detect dischronemes especially when there are several bilabials next to each other present and one is a case of a synchroneme.

Table 3

Instances of Dischronemes of the Bilabials /p/ and /b/

No, facial view	Screenshot	Time code	SL utterance	TL utterance/ Back translation	SL transcription	TL transcription	SL phoneme	TL phoneme	ST – TT Visemes
1. front		16.51	I'm gonna **bring** her again tomorrow.	Aš **atsivesiu** ją vėl rytoj./ I will bring her back tomorrow.	brɪŋ	ɐˈt̪sʲɪvʲɛˌsʲʊ	b	vʲ	5-6
2. ¾		19.22	**Property's** certified. No cats.	**Valdos** patikrintos. Jokių kačių./ The holdings are checked. No cats.	ˈprɒpətiz	ˈvaɫd̪os	p	v	5-6
3. ¾		1.26.11	**Wait.**	**P**ala./ Wait.	weɪt	ˈpaɫɐ	w	p	4-5

The first and second occurrences in Table 3 depict the pronunciation of the bilabial consonants /b/ and /p/ with closed lips, which are incongruent with the phonemes /vʲ/ and /v/ in the TL utterances, characterized by the lower lip touching the upper teeth and assigned to the sixth viseme. In close-up shots, translators aim to match an open or closed mouth with corresponding open vowels or bilabials in the translation; however, as Chaume suggests even in the more challenging case of bilabials, the consonants in the TL do not have to be identical to the SL consonants; for example, the bilabial /p/ can be replaced not only with the bilabials of the fifth viseme but also with labiodentals /f/ or /v/ belonging to the sixth viseme (Chaume 2012, p. 74). Having this in mind, the above-given examples can be treated as insignificant cases of dischrony since the difference in articulation between the two is rather minimal as compared to the visual difference in articulating open vowels and bilabials.

However, it is important to consider that the perception of dischronemes by viewers can vary depending on their position within the utterance and the overall length of the utterance. In the first example provided, involving the dischroneme /b/ – /vʲ/, it may not be immediately noticeable due to the bilabial being positioned in the middle of a nine-syllable TL utterance. The placement of the bilabial in this context makes it a less prominent and easily identifiable stimulus, making it challenging to perceive the onset of the syllable /vʲɛ/.

Nevertheless, to maintain satisfactory LS and meet the needs of the most attentive viewers, it is beneficial to substitute SL bilabial consonant /b/ in the word *bring* /brɪŋ/ from the utterance *I'm gonna bring her again tomorrow* with the TL bilabial /mʲ/ in the word *pasiimti* /pɐˈsʲiːmʲtʲi̯/ from the TL utterance *Gal man pasiimti ją vėl rytoj?* (*Can i pick her up again tomorrow?).*

By making this substitution, the first dischroneme of the SL bilabial phoneme /m/ in *I'm* /aɪm/ is effectively avoided, as it is modified to the identical phoneme /m/ in the TL word *man* /ˈmaːn̪/ (*to me*). This alteration ensures better synchrony and enhances the overall quality of the dialogue for the audience.

The second example provided involves the bilabial consonant /p/ pronounced at the beginning of the SL utterance, making it easily visible and distinguishable from the substitute /v/ in the first TL syllable /vɛl̪/. The initial syllable's acoustic prominence and neural response are significant to viewers (Greenberg et al., 1997; Cutler, 2012). To ensure LS, the suggested translation focuses on preserving the syllable onset and retains the bilabial /p/ in the word *pastatas* /ˈpäːs̪t̪ɐt̪ɐs̪/ (*building*) in the utterance *Pastatas patikrintas. Jokių kačių* (*The building is inspected. No cats*). However, both the original and suggested translations exceed the required eight syllables of the SL utterance. To address this, Chaume (2012, p. 91) suggests modifying "suffixes, prefixes, and shortening processes" when minor changes are required to achieve isochrony. Therefore, a version with eight syllables, such as *Pastatas tuščias. Kačių nėr* (*The building is empty. There are no cats*) can be proposed to meet the requirements of isochrony. This example illustrates that reducing vowels helps to minimise the number of syllables, particularly in the final words of an utterance.

The last example in Table 3 illustrates a case of bilabial translation,

which can be classified as a TL-SL dischroneme. In the SL utterances, there were no bilabial phonemes present, but they appeared in the TL utterances. Specifically, the translation of the word *wait* /weɪt/ was rendered as *palauk* /pɐˈl̪ɑʊ̯ˑk/ or its reduced form *pala* /pɐˈl̪ɐ/, reflecting the characteristic language style of Lithuanian dubbing known as *dubbese*. Throughout the film, these reduced forms of *palauk* (*wait*) were used 15 times in brief utterances consisting of one to three words, sometimes repeated twice.

However, this translation approach proves to be unsuccessful due to the lack of synchronization between the /w/ and /p/ phonemes. The consonant /w/ belongs to the fourth viseme, characterized by rounded and protruded sounds, which do not correspond with the closed lips position required for the /p/ sound.

This case is particularly noteworthy because the consonant phoneme /p/ occurs at the syllable onset, and the segment of the utterance is very short. To achieve synchronization, one suggestion is to employ alternatives such as *stok* /ˈs̪t̪oˑk/ or *sustok* /s̪ʊs̪ˈt̪oˑk/, both meaning *stop*, where the only dominating visual expression is in the position of the rounded and protruded lips, corresponding to the articulation of the vowel phoneme /o/ and a velar sonorant or labiovelar /w/. The articulation of other phonemes belonging to the second viseme remains invisible to the viewers.

Conclusions

The analysis of bilabial consonants in the live-action film *A Dog's Way Home* reveals several key findings. The research focused on the significance of LS in translation and identified specific episodes where translators must prioritize this type of synchrony. The close-up shots of the film were specifically chosen to limit the number of phonemes included in the analysis.

The analysis of bilabial phonemes (/m/, /b/, /p/) in the SL utterances showed that out of 1,419 identified phonemes, 271 visually prominent instances were selected for further examination of LS. The phoneme /m/ appeared most frequently (695 times), followed by /b/ (465 times) and /p/ (259 times). However, only a small percentage of these occurrences (approximately 19%) were visually prominent and considered for LS analysis.

This indicates that strict LS may not be crucial in many instances, providing flexibility for translators during the dubbing process.

The analysis of the TL utterances revealed that LS was maintained in approximately 76% of cases, with 65 dischronemes and 206 synchronemes identified out of the 271 visually prominent bilabial consonants. However, 94 newly introduced TL bilabials were absent in the SL utterances, resulting in dischronemes. This mismatch between auditory and visual signals may have negative implications for speech perception and the overall quality of the dubbing. Therefore, audiovisual translators are urged to consider not only how to maintain the lip synchrony of bilabial consonants in the target language, but also to avoid the emergence of new bilabial dischronemes.

Specific instances of synchronemes and dischronemes were examined. Repetition of proper names and international words contributed mostly to maintaining LS and meaning in translation.

In conclusion, while LS holds importance, strict adherence may not always be necessary. When considering synchronemes with bilabial onsets, preserving the LS of bilabials proves to be crucial for translators. Furthermore, it is not essential to maintain two bilabials in the TL utterances if the first bilabial serves as the coda and the second as the onset of the syllable, even if they belong to different words. Finally, dischrony involving bilabials becomes difficult to notice in situations such as long utterances with multiple bilabials, or dischronemes occurring in the middle of a lengthy expression.

References

Aprijaskytė-Valdšteitienė, R. (1960). Anglų ir lietuvių kalbų priebalsių lyginimas. *Kalbotyra*, II, 167–185.

Bear, H. L., & Harvey R. (2017). Phoneme-to-viseme mappings: the good, the bad, and the ugly*. Speech Communication*, *95*, 40–67. https://doi.org/10.1016/j.specom.2017.07.001

Bozkurt, E., Erdem, C. E., Erzin E., Erdem, T., & Ozkan, M. (2007). Comparison of phoneme and viseme based acoustic units for speech driven realistic lip animation. 2007 *IEEE 15th Signal Processing and Communications Applications,* SIU, 1-4, https://doi.org/10.1109/SIU.2007.4298572.

Chaume, F. (2004). Synchronisation in dubbing. A translational approach. In P. Orero (Ed.), *Topics in Audiovisual Translation* (pp. 38–50). John Benjamins Publishing Company.

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

Chion, M. (1994). *Audio-Vision: Sound on Screen.* Columbia University Press.

Cutler, A. (2012). *Native Listening: Language Experience and the Recognition of Spoken Words.* MIT Press. https://doi.org/10.7551/mitpress/9012.001.0001

Dodd, B. (1977). The role of vision in the perception of speech. *Perception*, *6*, 31–40. https://doi.org/10.1068/p060031

Dodd, B., & Campbell, R. (1987). *Hearing by Eye: The Psychology of Lip-Reading. London:* Lawrence Erlbaum Associates.

Fodor, I. (1976). *Film Dubbing. Phonetic, Semiotic, Esthetic and Psychological Aspects.* Hamburg: Helmut Buske.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (1997). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics*, *25*(2), 237–270.

Harris, G. T. (1996). *On Translating French Literature and Film.* Amsterdam: Rodopi.

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 913–927. https://doi.org/10.1037/0278-7393.33.6.1050

Irwin, J., Avery, T., Brancazio, L., Turcios, J., Ryherd K., & Landi, N. (2018). Electrophysiological indices of audiovisual speech perception: beyond the McGurk effect and speech in noise. *Multisensory Research*, *31*, 39–56. https://doi.org/10.1163/22134808-00002580

Koverienė, I. (2015). *Dubliavimas kaip audiovizualinio vertimo moda: anglų ir lietuvių kalbų garsynai vizualinės fonetikos kontekste.* Vilnius: Vilniaus universitetas.

Koverienė, I., & Čeidaitė, K. (2020). Lip synchrony of rounded and protruded vowels and diphthongs in the Lithuanian-dubbed animated film 'Cloudy

with a Chance of Meatballs 2', *Respectus Philologicus*, *38*(43), 214–229. https://doi.org/10.15388/RESPECTUS.2020.38.43.69

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. https://doi.org/10.1038/264746a0

Martinez, X. (2004). Film dubbing. Its process and translation. In P. Orero (Ed.), *Topics in Audiovisual Translation*. John Benjamins.

Massaro, D. W., & Simpson, J. A. (1987). *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry* (1st ed.). Psychology Press. https://doi.org/10.4324/9781315808253

Massaro, D. W., Cohen, M. M. & Smeele, P. M. (1996). Perception of asynchronous and conflicting visual and auditory speech. *The Journal of the Acoustical Society of America, 100*, 1777–86. https://doi.org/10.1121/1.417342

Matkivska, N. (2014). Audiovisual translation: conception, types, characters' speech and translation strategies applied. *Studies About Languages*, *25*, 38–44. https://doi.org/10.5755/j01.sal.0.25.8516

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

O'Neill, J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech and Hearing Disorders*, *19*, 429–439. https://doi.org/10.1044/jshd.1904.429

Pakerys, A. (2003). *Lietuvių bendrinės kalbos fonetika*. 3-iasis leid. Enciklopedija.

Pelson, R. O., & Prather, W. F. (1974). Effects of visual message-related cues, age and hearing impairment on speechreading performance. *Journal of Speech and Hearing Research*, *17*, 518–525. https://doi.org/10.1044/jshr.1703.518

Roach, P. (2009). *English Phonetics and Phonology*. *A Practical Course*. (4th ed). Cambridge University Press.

Romanski, L. M., & Hwang, J. (2012). Timing of audiovisual inputs to the prefrontal cortex and multisensory integration. *Neuroscience*, *214*, 36–48. https://doi.org/10.1016/j.neuroscience.2012.03.025

Rosenblum, L. D., Yakel, D. A., & Green, K. P. (2000). Face and mouth

inversion effects on visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 806–819. https://doi.org/10.1037/0096-1523.26.2.806

Smith, Ch. M. (Director). (2019). *A Dog's Way Home.* Sony Pictures Entertainment.

Smith, Ch. M. (Director). (2019). *Belos kelionė namo.* Acme Film/Sony Pictures Entertainment.

Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audiovisual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by Eye: The Psychology of Lip-Reading* (pp. 3–51). Lawrence Erlbaum Associates, Inc.

Whitman-Linsen, C. (1992). *Through the Dubbing Glass*. Frankfurt: Peter Lang.

Williams, J. J., Rutledge, J. C., Katsaggelos, A. K., & Garstecki D.C. (1998). Frame rate and viseme analysis for multimedia applications to assist speechreading. *Journal of VLSI Signal Processing*, *20*, 7–23. https://doi.org/10.1023/A:1008062122135

Indrė Koverienė
Vytauto Didžiojo universitetas, Lietuva
indre.koveriene@vdu.lt

ABILŪPINIŲ PRIEBALSIŲ LŪPŲ SINCHRONIJA LIETUVIŠKAI DUBLIUOTO FILMO *BELOS KELIONĖ NAMO* VERTIME

Anotacija. Šiame straipsnyje nagrinėjama sudėtinga lūpų sinchronijos problema dubliuojant – vienoje iš audiovizualinio vertimo modų, pabrėžiant ne tik estetinę lūpų sinchronijos svarbą, bet dažnai ir nepakankamai įvertinamą nesinchronizuotų lūpų judesių daromą neigiamą įtaką kalbos suvokimui. Šio sinchronijos tipo aktualumas visame pasaulyje didėja augant audiovizualinio turinio kiekiui. Audiovizualinio vertimo teoretikai seniai pripažįsta abilūpinių priebalsių svarbą lūpų sinchronijai (Fodor, 1976; Chaume 2012), vis dėlto išsamiai šį aspektą analizuojančių tyrimų Lietuvoje neatlikta. Siekiant užpildyti šią spragą, tyrime nagrinėjama abilūpinių priebalsių lūpų sinchronija lietuviškai dubliuotoje vaidybinio filmo *Belos kelionė namo* (2019) versijoje. Taikant lyginamąjį tyrimo metodą, atliktos kokybinė ir kiekybinė analizės, kuriose remiamasi H. McGurk ir J. MacDonald (1976), I. Fodor (1976), F. Chaume (2004, 2012) ir I. Koverienės (2015) teorinėmis prielaidomis. Kinematografinių planų tyrimas atskleidė, kad tik nedidelė dalis (19 proc.) abilūpinių priebalsių buvo svarbūs tolesnei lūpų sinchronijos analizei vertimo kalboje (tai reiškia, kad vertėjams suteikiama didesnė laisvė dubliuojant) ir lūpų sinchronija nėra būtina daugumoje iš nagrinėtų abilūpinių priebalsių atvejų. Nustatyta, kad iš visų vizualiai svarbių abilūpinių priebalsių lūpų sinchronija vertimo kalboje buvo išlaikyta net 76 proc. nagrinėjamų garsų, sėkmingiausiai sinchronizuojant abilūpinius priebalsius tikriniuose daiktavardžiuose ir tarptautiniuose žodžiuose. Ypač svarbu išlaikyti abilūpinių priebalsių lūpų sinchroniją trumpuose pasakymuose ir skiemens pratare. Tačiau tyrimo rezultatai atskleidė ir tai, kad vertimo kalboje buvo panaudoti net 94 abilūpiniai garsai, kurių originalo kalboje nebuvo. Todėl audiovizualinio turinio vertėjai raginami atsižvelgti ne tik į tai, kaip vertimo kalboje išlaikyti originalo kalbos pasakymų abilūpinius priebalsius, bet ir atsakingiau vertinti vertimo kalbos pasirinkimus siekiant išvengti naujai atsirandančių abilūpinių priebalsių.

Pagrindinės sąvokos: audiovizualinis vertimas; abilūpiniai priebalsiai; dischronema; dubliavimas; lūpų sinchronija; vizema.