

Lip Synchrony of Rounded and Protruded Vowels and Diphthongs in the Lithuanian-Dubbed Animated Film *Cloudy with a Chance of Meatballs 2*

Indrė Koverienė

Vilnius University, Kaunas Faculty
Institute of Languages, Literature and Translation Studies
Muitinės g. 8, LT-44280 Kaunas, Lietuva
Email: indre.koveriene@knf.vu.lt
ORCID iD: <https://orcid.org/0000-0002-3571-2564>
Research interests: audiovisual translation, dubbing, visual phonetics, English and Lithuanian phonology

Kristina Čėdaitė

Vilnius University, Kaunas Faculty
Institute of Languages, Literature and Translation Studies
Muitinės g. 8, LT-44280 Kaunas, Lietuva
E-mail: k.ceidaite@gmail.com
ORCID iD: <https://orcid.org/0000-0001-9152-1642>
Research interests: film dubbing, lip synchrony, phonetic differences between languages

Abstract. In this article, the problems of dubbing, especially related to lip synchrony as one of the most challenging aspects of audiovisual translation, are scrutinised. Contrarily to the traditional focus on bilabials and open vowels, the object of this research is lip synchrony of both rounded and protruded vowels and diphthongs since lip rounding is a visibly marked feature, which cannot be neglected especially in close-ups. The study aims at determining the inaccuracies in lip synchrony of the mentioned phonemic group in the dubbed animated feature film *Cloudy with a Chance of Meatballs 2* from English to Lithuanian. Qualitative and quantitative analysis is carried out by employing a comparative method. The research methodology is based on the theoretical insights and assumptions provided by Frederic Chaume (2004, 2006, 2012), Richard Barsam & Dave Monahan (2010), and Indrė Koverienė (2015). The research findings demonstrate the main issues of lip synchrony a translator might face while adapting a piece of audiovisual material for the target language audience. Also, it provides insights into the quality of the overall translation of the chosen film.

Keywords: dubbing; lip synchrony; viseme; vowels; diphthongs.

Introduction

With the development of information technologies, a considerable number of changes regarding the mainstream audiovisual translation modes have occurred recently. As the

Submitted 22 April 2020 / Accepted 30 July 2020
Iteikta 2020 04 22 / Priimta 2020 07 30

Copyright © 2020 Indrė Koverienė, Kristina Čėdaitė. Published by Vilnius University Press. This is an Open Access article distributed under the terms of the [Creative Commons Attribution License CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits unrestricted use, distribution, and reproduction in any medium provided the original author and source are credited.

findings of the research conducted on behalf of MESA Europe Content Localisation Council (2017) suggest, the rising popularity of video-on-demand platforms such as Netflix or Amazon Prime has contributed significantly to the increasing number of requests for the translation of audiovisual products all over the world. As a result, the recent trend has witnessed the unexpected viewers' preferences for dubbing referred to as the "revolution of dubbing" (as cited in Ranzato et al., 2019, pp. 2–3). This approach is successfully adopted by Netflix to satisfy the growing consumer needs for dubbing by streaming not only the "authentic" subtitled versions but also the highly required dubbed options by default even to the English-speaking countries like the United Kingdom and the United States (Roxborough, 2019).

Considerable growth in the number of viewers who opt for the dubbed audiovisual products can be related to the heightened entertainment value experienced by watching the production, which embraces the mother-tongue of the target audience the most (Koverienė et al., 2018, p. 85). The abovementioned value plays a part, however, provided that the quality of dubbing is satisfactory. Nevertheless, the quality of dubbed foreign audiovisual production is sometimes beyond the limits of the viewer's tolerance, and there is no surprise that such companies as Netflix tend to redub entire series or talk shows to avoid "dubby"¹ lines (Goldsmith, 2019). In seeking the answer, whether it is challenging to avoid poor dubbing, it is worth focusing on the features of dubbing as an audiovisual translation mode.

Dubbing is an eminently complex phenomenon where a change of the text delivered by the translator is inevitable. It allows and even encourages the translators to get out of their comfort zone, not to be attached to the original text and instead concentrate on the spectators and how they are going to perceive the finished translation (Chaume, 2006, p. 7). Nevertheless, the main difficulty for the translators is posed by the absence of satisfactory synchronisation, and mainly the flaws and discrepancies in lip synchrony result in constant criticism when confronted with dubbing.

The problem of lip synchrony is widely addressed by a significant number of audiovisual translation scholars (Vöge, 1977; Delabastita, 1989; Whitman-Linsey, 1992; Zabalbeascoa, 1997; Barbe, 1996; Chaume, 2004, 2012). Even though nowadays, lip synchrony is not "the absolute dogma", as stated by Whitman-Linsen (1992, p. 21), it remains one of the most challenging aspects of dubbing. However, contrarily to subtitling, there are relatively few in-depth and systemic inquiries as well as recent advances concerning lip synchrony and methodology of teaching proper lip synchrony, except for Istvan Fodor's seminal study *Film Dubbing. Phonetic, Semiotic Esthetic and Psychological Aspects* (1976).

Thus, to contribute to the growing demand of research in this field, the study focuses on lip synchrony and the application of the viseme theory to provide the findings of the analysis of lip synchrony in the Lithuanian dubbed film. The study aims at determining the inaccuracies in lip synchrony of a single group of sounds – the rounded and protruded vowels and diphthongs – as the distinct visual aspects of their pronunciation are easily noticed and recognised by viewers.

¹ The term "dubby" has a negative connotation as "not speech-like – jarring diction or awkward wording – or is conspicuously out of time with how the actors' mouths are moving onscreen." (<https://www.nytimes.com/2019/07/19/arts/television/netflix-money-heist.html>).

Since animated films are most commonly dubbed in Lithuania, the chosen material for this research is an animated feature film *Cloudy with a Chance of Meatballs 2* (*Debesuota, numatoma mėsos kukulių kruša 2*) (2013) dubbed from English to Lithuania. Qualitative and quantitative analysis is carried out by employing a comparative method. Methodological basis is grounded on the theoretical insights and assumptions provided by Frederic Chaume (2004, 2006, 2012), Richard Barsam & Dave Monahan (2010) and Indrė Koverienė (2015).

1. Dubbing as an audiovisual translation mode

Dubbing, as an audiovisual translation mode, encompasses a change of the record of the source text (ST) speech by the record of the target text (TT) speech (Chaume, 2012, p. 11). The translator has to recreate analogical speech structure, fit the duration of the utterances and select sounds that correspond to the lip movements (Luyken et al., 1991, p. 73). Therefore, it is essential to gather the knowledge of what is needed for proper dubbing and how to increase its quality continuously.

Audiovisual translation scholars suggest different typologies of synchronies. As seen in the table below, lip synchrony falls under the title of phonetic synchrony as an umbrella term in Fodor's typology along with Chaume's isochrony and kinetic synchrony and Linsen's even more detailed syllable articulation, length of utterance and kinetic synchrony. Fodor was not only the first to name the types of synchronies, but he also developed visual phonetics as a separate area of study which analyses "the mouth articulatory movements of the screen actor and the phonemes that the translator should fit his or her mouth" (Chaume, 2004, p. 38). These studies are important as they seek to find a way to make the translation fit the original image and preserve the feeling of reality for the audience.

Table 1. Typologies of synchronies

Fodor (1976)	Phonetic				Character					Content
Whitman-Linsen (1992)	Visual/Optical				Audio/Acoustic					(-Content)
	Lip	Syllable articulation	Length of utterance	Kinetic	Idiosyncratic vocal type	Paralinguistic elements	Prosody	Cultural variations	Accents and dialects	
Chaume (2004; 2012)	Lip	Isochrony		Kinetic	(-Character)					(-Content)

Source: Koverienė, 2015, p. 12.

This study employs Chaume's typology of synchronies, including kinetic synchrony, isochrony and lip synchrony. Kinetic synchrony implies the maintenance of the logical relation between the words and gestures of the characters that are known by the audience

to avoid contradictions (Chaume, 2012, pp. 70–71). Isochrony, the lack of which is the foremost noticeable indicator of a poorly dubbed product, addresses the matching of the time between the ST and TT speech phrases and pauses (Chaume, 2012, p. 73). However, the primary attention is focused on lip synchrony as the most challenging feature of the dubbing modality that requires specific knowledge and extreme attention to details.

1.1 Lip synchrony

Lip synchrony involves adapting the TT to the articulatory movements of the on-screen character or actor of the audiovisual product. While trying to achieve lip synchrony and creating the final translated script, the dialogue writer needs to take the actor's lip movements into account. Mainly, if bilabial consonants, labio-dental consonants and open vowels occur in the ST, as they are the most visible on the lips of the speaker (Chaume, 2004, p. 10). However, there is no need to use only the same phonemes or very similar sounding ones (Chaume, 2012, p. 74). Special classifications can be applied to determine what types of phonemes can replace other phonemes (the viseme theory).

Possible techniques for lip synchrony include repetition (using words of similar origin), change of word order (so that similarly pronounced sounds match the lip movements), substitution (employing a synonymic / antonymic words or changing the meaning if it does not distort the overall meaning of the film), reduction or amplification and omission or addition (eliminating or adding text) (Chaume, 2012, p. 75). The recommended strategy for the lip synchrony is to choose utterances with visually similarly pronounced sounds in SL and TL (Chaume, 2012, p. 74).

Lip synchrony is a priority over a precise semantic meaning when the mouths of the characters can be clearly seen. The scale (or proximity to the camera) determines how much of the character can be visible (also, how much of the mouth movements can be visible). In total, seven basic types of shots in terms of the subject's closeness to the camera (or the supposed distance to the camera if the film is animated) can be distinguished: extreme long shot, long shot, medium-long shot, medium shot, medium close-up, close-up, extreme close-up (Barsam & Monahan, 2010, pp. 232–235). The frequent absence of extreme close-ups, as well as the low number of close-ups in live-action or animated feature films and TV series, can be explained by the producers' attempts at avoiding to even comply with different types of synchronies at the very initial stages of film production. Recently, however, the number of extreme close-up and close-up shots has increased considerably, as stated by Romero Fresco (2019), and peaks at nearly 75% for such TV series as *EastEnders* (BBC, 1985–1985) (as cited in Sánchez-Mompeán, 30, 2019).

Nevertheless, when it comes to animated feature films, there are nearly no extreme close-ups and only a few close-ups that correspond regular classification of shots, since animation genres are mostly action-oriented. However, differently from the standard practice to adapt lip synchrony mainly in close-up and extreme close up shots, the analysis of this study focuses on medium close-up and medium shots when the subject is seen from head to the middle part of the chest or even about half of the subject's body is visible, as

well. Such choice is justified by the fact that the heads and the mouths of the characters in the animated film *Cloudy with a Chance of Meatballs 2* are enlarged, making the lip movements explicitly visible to the audience.

It is also important to notice where the head of the character is turned. The needed accuracy for synchronisation varies if the character is facing the screen directly (indicated as front), looking to the side (profile) or have their head slightly turned (3/4).

1.2 The viseme theory

The term viseme was first used to differentiate between the visual aspects of consonant pronunciation (Fisher, 1968). Later, it was noticed that there were some similarities and patterns between the pronunciations of different phonemes with respect to visual perception; therefore, visemes were created. The classification of phonemes into visemes is done by relying on linguistics (Saenko, 2004, p. 48). However, there are many nuances and diversity in the pronunciation of phonemes; consequently, various classifications exist.

Therefore, as this study focuses on dubbing from English to Lithuanian, a suitable system of visemes, in this case, is provided by Indrė Koverienė in her dissertation *Dubbing as an Audiovisual Translation Mode: English and Lithuanian Phonemic Inventories in the Context of Visual Phonetics* (2015). All phonemes of both languages are distributed to a particular group according to similarities in visible aspects of their pronunciation. The six visemes are presented in Table 2.

The first viseme, Considerable separation of jaws, includes sounds that are open front, open back and central, and are produced with jaws parted to some extent (Koverienė, 2015, pp. 92–93).

The second viseme, Neutrally open mouth, consists of phonemes that are pronounced with a neutrally open mouth (Koverienė, 2015, p. 95).




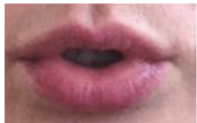
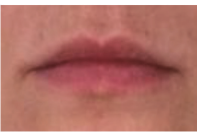
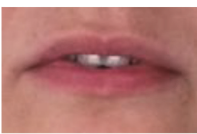
The third viseme, Spread lips, contains close, close-mid (or open-mid), front and central monophthongs and diphthongs, during their pronunciation lips are pulled to sides – spread (Koverienė, 2015, pp. 103–104).

The fourth viseme, Rounded and protruded lips, includes sounds that are produced with rounded lips pushed front: close, close-mid or open-mid, back or central monophthong or diphthong vowels (Koverienė, 2015, pp. 114–115) as well as a bilabial sonorant and velar sonorant [w] (Koverienė, 2015, pp. 133–134).

However, the sound [w] is not analysed in the empirical part of the paper, for the object of this study is exceptionally the vowel and diphthong phonemes. It is also important to note that AmE has a diphthong [oʊ] used instead of BrE diphthong [əʊ]. Moreover, [oʊ] will be considered belonging to the fourth viseme of rounded and protruded lips, for it has the corresponding features.

The fifth viseme, Closed lips, involves sounds whose pronunciation require lips held together – at least in the first part of the utterance (Koverienė, 2015, p. 129).

Table 2. Visemes of English and Lithuanian phonemes

Viseme	Visual Features	The Lithuanian phonemes	The English phonemes
1.	 Considerable separation of jaws	[ā:], [âː], [âːu], [æ:], [æː]	[ɑ:], [aɪ], [aʊ], [æ]
2.	 Neutrally open mouth	[e], [ɛ̃], [ɛ̃ː], [ɛ̃], [ɛ̃ː], [d], [dʰ], [dʒ], [dʒʰ], [dʒ̃], [dʒ̃ʰ], [ʒ], [ʒʰ], [ɳ], [ɳʰ], [s], [sʰ], [ʃ], [ʃʰ], [r], [rʰ], [ʁ], [ʁʰ], [z], [zʰ], [ʒ], [ʒʰ], [x], ç, [g], [j], [y], [j̃], [j̃], [k], [c], (allophones [ŋ] and [ŋʰ])	[ʌ], [ə], [θ], [ð], [tʃ], [dʒ], [l], [r], [s], [ʃ], [tʃ], [z], [ɹ], [ʒ], [g], [h], [j], [k], [ŋ]
3.	 Spread lips	[i], [i:], [eɪː], [eɪː], [ie], [ie], [ɛ], [e:]	[i], [ei], [iə], [e], [eə], [ɜ:]
4.	 Rounded and protruded lips	[u], [u:], [ôɔ], [uɔ], [ôɪ], [uɪ], [eɪː], [eɪː], [ɔ], [o:], [ôɪ], [ôɪ]	[u], [u:], [ʊə], [əʊ], [ʊ], [ɔ:], [ɔɪ], [w], [ʊʊ]
5.	 Closed lips	[p], [pʰ], [b], [bʰ], [m], [mʰ]	[p], [b], [m]
6.	 The lower lip touching the upper teeth	[f], [fʰ], [v], [vʰ]	[f], [v]

Source: Koverienė, 2015, p. 143.

Finally, the sixth viseme, The lower lip touching the upper teeth, consists of sounds that are achieved when the lower lip and the upper teeth touch (Koverienė, 2015, pp. 135–136).

Due to a visibly marked feature of lip rounding, rounded and protruded vowels are crucial in the view of dubbing an audiovisual product. Therefore, rendering of these sounds into the TL will be analysed in the empirical part of the study.

2. Lip synchrony of rounded and protruded vowels and diphthongs in film *Cloudy with A Chance of Meatballs 2*

In the empirical part of this study, the lip synchrony of rounded and protruded vowels and diphthongs is analysed in the film *Cloudy with a Chance of Meatballs 2* (2013). Based on theoretical implications specific tables for the analysis were constructed:

- the International Phonetic Alphabet (IPA) is applied to the transcription of the Lithuanian speech which serves as “tertium comparationis” and allows the comparison of the phonetic inventories of two different languages;
- each example of the phoneme from the fourth viseme contains two pictures (for shot type, and mouth position indication);
- the phoneme is indicated, the utterance in which it appears is shown, and the transcription of the word with the phoneme is provided;
- the Lithuanian translation of the utterance and the transcription of the word with the corresponding phoneme are included, and the viseme of that particular phoneme is indicated.

The material for this study comprises 476 cases of stressed English vowels and diphthongs from the fourth viseme of rounded and protruded sounds, which together with other phonemes constitute 1 hour 35-minute long film. There were identified precisely 165 cases of [u:], 151 – [oʊ], 92 – [ɔ:], 47 – [ʊ], 15 – [ɔɪ] and 6 – [ɒ] (the findings are shown in Figure 1). The calculated percentage of the overall results is as follows: 35% of [u:], 32% of [oʊ], 19% of [ɔ:], 10% of [ʊ], 3% of [ɔɪ] and 1% of [ɒ].

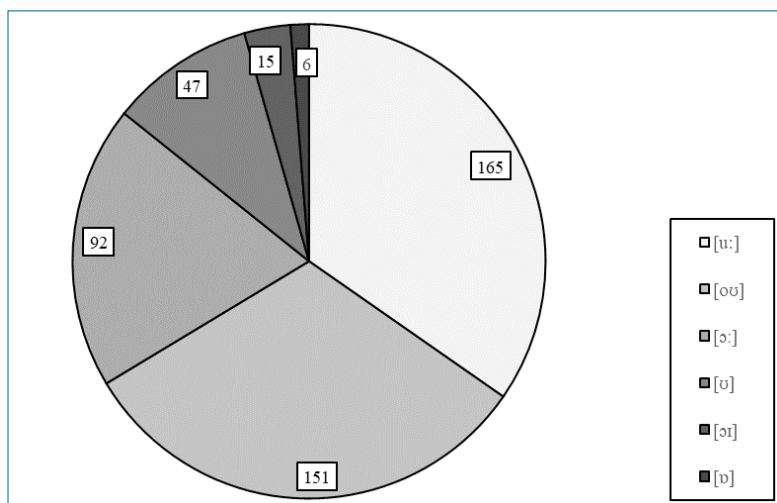


Fig. 1. The fourth viseme phonemes found in the original film²

² All remaining tables and figures in this article have been produced by the authors.

Further on, zero instances of diphthongs [oə] and [əʊ] were detected. The absence of [əʊ] can be explained by the fact that all words having such a diphthong in BrE pronunciation are replaced with a diphthong [oʊ] in AmE pronunciation. Similarly, there are only 6 cases with [ɒ] sound, for it also mostly emerges in BrE pronunciation. The reason why there are nonetheless instances with [ɒ] is that two characters have distinct accents.

The detailed findings concerning TT equivalents for ST fourth viseme phonemes are summarised in Table 3. The first section indicates the English phoneme; the second reveals the exact number of instances with the phoneme, then the calculations of matching/mismatching cases of the stress are shown (yes – the stress corresponds to ST stress, no – the stress does not correspond). Further on the number of visemes used for the translation of each phoneme is displayed and the number of particular shot types (e – extreme close-up, c – close-up, m – medium close-up, M – medium shot) and different face positions (f – front, $\frac{3}{4}$ – 3/4, p – profile) is summed up.

Table 3: Data about TT equivalents for the fourth viseme phonemes from ST

Phoneme	Quantity	Stress matching (yes/no)	Viseme 1	Viseme 2	Viseme 3	Viseme 4	Shot type (e/c/m/M)	Face position (f/ $\frac{3}{4}$ /p)
[u:]	165	107/58	11	21	50	83	0/29/49/87	89/47/29
[oo]	151	110/41	17	14	61	59	1/26/35/89	72/56/23
[ɔ:]	92	59/33	14	13	20	45	0/10/24/58	42/36/14
[ʊ]	47	30/17	3	8	12	24	0/12/11/24	23/14/10
[ɔɪ]	15	13/2	2	1	7	5	0/2/5/8	7/6/2
[ɒ]	6	4/2	1	0	2	3	0/1/2/3	5/1/0
All in all:	476	323/153	48	57	152	219	1/80/126/269	238/160/78

In the majority of cases, the stress in the TT matched the stress in the ST: 323 (68%) matching cases and 153 (32%) mismatches. While such results are not ideal, they do not necessarily indicate a low-quality dubbing. The corresponding accentuation was found in the cases of phonemes [ɔɪ] (87% of matched cases) and [oo] (73% of matched cases). In comparison, the most inadequate cases were detected with phonemes [ʊ] (64% of matched cases) and [ɔ:] (64% of matched cases). Although in some cases there was the noticeable illogical intensity of lip movements, the overall result of the Lithuanian dubbing was entirely credible.

It is worth mentioning that 56% of the analysed instances appeared in medium shots (269 cases). Therefore, this justifies the decision to include this type of shot into the analysis. In terms of face position when characters were depicted speaking, they usually – 50% of the time – faced the screen directly (front), sometimes with a comparatively turned head (34%) and least frequently by profile (16%). Hence, in the majority of cases, the mouths of the characters were visible quite clearly.

The most crucial aspect in the table is corresponding visemes. There are no sections for the fifth and sixth visemes, for they were not used as equivalents for the fourth viseme phonemes. Out of all equivalents for rounded and protruded vowels the fourth viseme phonemes were most frequently applied in the cases of [ʊ] (in 24 out of 47 cases, thus 51%) and [u:] (in 83 out of 165 cases – 50%). In contrast, the rarest matches appeared with diphthong vowels [ou] (in 59 out of 151 cases – 39%) and [ɔɪ] (in 5 out of 15 cases – 33%). Contrary to all analysed phonemes, in the case of diphthong [ou] the third viseme was employed more often (in 40% of the cases) than the fourth.

According to Table 3, a diagram (Figure 2) was created to give a more precise presentation of the percentage of each employed viseme. The first and second visemes were used only in 10% and 12% of all the cases. The most commonly used viseme for translation of ST rounded and protruded phonemes was the fourth one (46%); however, there was also a considerable number of the usage of the third viseme (32%).

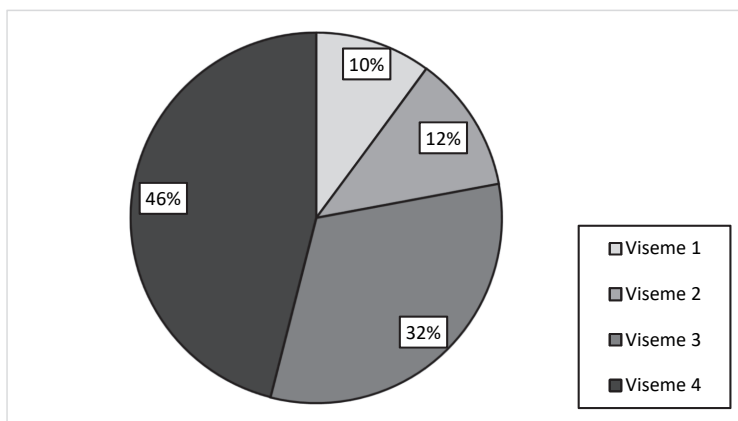


Fig. 2. TT equivalent visemes for ST fourth viseme

2.1 Mistakes in word usage and pronunciation




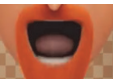
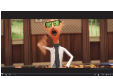
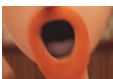
In the film *Cloudy with a Chance of Meatballs 2*, the exact number of syllables was of secondary importance as there were many cases when instead of the intended lip movement for one syllable in ST there were inserted a few syllables in the TT. Therefore, in the cases of apparent additions of multiple syllables, the first inserted sound was considered the equivalent to the ST sound to avoid confusion and distortion of the analysis results. In the tables containing examples, shot types are indicated with symbols (analogically to Table 3). The section “Time code” points out the moment when the phoneme from the fourth viseme is pronounced.

In the first example (Table 4) from the film, the stress is corresponding in all three cases, and the phonemes [ou] and [ɔ:] are replaced by phonemes from the same viseme (except in the third case). However, the overabundance of syllables contradicts the visual

display. The character has a big mouth and a very clearly conveyed articulation; thus, the audience would surely notice that the voice in translated version begins to speak earlier and stops later than the character's lips start/stop moving.

Although some changes are allowed in translation, in this case the semantic meaning is the opposite of the original one. Thus, the requirement of faithfulness is not met. The expression *mokslo niekada nebūna per mažai* indicates that *you can never have too little of science*. The suggestion would be: *Atsiminkite – mokslui nerūpi amžius. Kurti gali visi*. The letters in bold represent the corresponding phonemes that substitute the fourth viseme phonemes from ST, the number of syllables is the same as in original, and the requirement of faithfulness is met.

Table 4. Examples of inaccurate semantic meaning in the translation

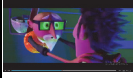

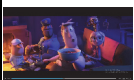

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face position
1.			oo	00:01:38	But, remember, there's no such thing as small science.	noo	Bet atsiminkite – mokslo niekada nebūna per mažai.	'mo:kɫo:	4	M	f
2.			o:	00:01:38	But, remember, there's no such thing as small science.	smo:l	Bet atsiminkite – mokslo niekada nebūna per mažai.	p'e:'bu:ɫe	4	M	f
3.			oo	00:01:40	Only small scientists!	'oonli	Yra tik maži mokslininkai!	i:'rɛ	2	M	f

There are other mistakes left by the translator or voicing actor (Table 5). According to the rules of Lithuanian accentuation, in the first example word *mušeika* should have a stress on the last vowel (VDU Kompiuterinės lingvistikos centras). Still, the voicing actor of the TT placed it on the second syllable. This mistake could have been avoided had the actor pronounced the whole word emphasising each syllable (that way, the correct viseme would have fit in the place of [ʊ]), for the pace of the speech in the scene was relatively slow.

In the second example, the usage of the word *bliūdukas* is incorrect for it is unmotivated (not justified by a specific situation or language style of the character) barbarism (VLKK Konsultacijų bankas). Instead, the word order should be changed and reduction used: *Ir tau palikau*. This way the number of syllables remains the same and the diphthong in bold corresponds with the same viseme in the ST.

Although such mistakes decrease the credibility of TT, these discussed issues are the only cases with such type of improperly used/pronounced words; therefore, they do not have a significant impact on the overall effect of the film.



Table 5. Examples of mistakes in pronunciation and word choice

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face
1.			o	00:44:56	A bully can never be your friend.	'boli	Mušėika niekada nebus tavo draugas.	mo'ʃeɪ'ke	3	M	¾
2.			oo	00:46:25	I saved you a bowl .	bool	Palikau ir tau bliūduka .	bli'u:'doka	4	M	f

2.2 The advantage of visual obstructions

There are several cases in this animated film, where the lips of the characters are barely visible, or the viewer attention is distracted by other objects. In these cases, the translator has more freedom to manipulate the text. This can be illustrated by an example in **Table 6**.

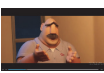


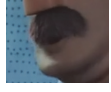
Table 6. Examples of TT during fast-paced speech and action

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face position
1.			u:	00:22:23	I'm going to Swallow Falls to destroy the FLDSMDFR, which is creating deadly food monsters, which are trying to learn to swim so they can attack Lady Liberty!	fu:d	Aš grįžtu į Kramturijuką sunaikinti FLDSMDRo, kuris gamina baisius monstrus, kurie mokosi plaukti, kad užpultų Laisvės statulą.	bai'ʃioŋ	4	M	f

The example depicts an extremely excited character moving vigorously and speaking very quickly. Nonetheless, the translator managed to match isochrony, stress and a phoneme from the fourth viseme. There are some effectively employed techniques for lip-synchrony: substitution – using a synonym *grįžtu* (coming back) for verb “going”; repetition – repeating the same coined term “FLDSMDFR” (only modifying it by TL rules); reduction – instead of “deadly” (TL *mirtini*) a word *baisūs* is used, while “trying to learn to swim” is reduced to *mokosi plaukti* (SL “learning to swim”); and omission – words “food”, “trying” and “can” are eliminated. Thus, the translation is done skillfully.



Table 7 below provides examples of the character speaking with a partially visible mouth. In both cases we have chosen for illustration purposes the stress does not coincide with the original stress, and in the second case, the TT phoneme belongs to the third viseme. Inaccuracies can be justified due to obstructing moustache through which the mouth can be barely seen, thus it is acceptable to put any vowel phoneme regardless of its viseme. Moreover, the shot type is a medium close-up, the lighting in both cases is relatively dim, and character's head is turned at 3/4 – all these elements make it difficult to notice the lip movements.

Table 7. Examples of translation in the case of barely visible lips and mouth

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face position
1.			o	00:18:00	Is it winter on your torso and summer on your arms?	jor	Čia ką – tavo liemeniui žiema, o rankoms vasara?	'rɛŋko:mɔ	4	M	¾
2.			u:	00:39:58	Had to be looters .	'lu:toɾz	Tikriausiai plėšikai.	ʧi'crɛ:ˈoʂeɪ	3	M	¾

In **Table 8**, on the other hand, the TT phoneme belongs to the first viseme. However, the shot is middle close-up, the whole body of the character is covered in colourful confetti, and he is moving his hand. Due to a distracting visual setting, the audience is unlikely to notice any inaccuracy regarding synchrony. Therefore, the choice of not matching the appropriate TL phoneme can be partly justified. Moreover, in this case, the lips are neither exceptionally rounded nor protruded. Thus, monophthong [ɑ:] is quite a convincing equivalent.

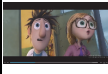
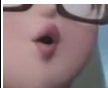

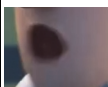

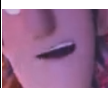
Table 8. Examples of translation conditioned by visual distractions

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face
1.			oo	01:20:18	Now hand over the FLDSMDFR.	'ooʋər	Gražink man FLDSMDRą.	mā:ŋ	1	M	f

2.3 Special cases

The examples in **Table 9** contain words that due to cultural differences, have additional meanings in the ST. Saying “This is totally... bananas!” means not only that the characters are seeing bananas but also that they think the situation is unbelievable. Similarly, “We’re... toast” not only indicates that one character interrupts the other to inform about a toast of enormous size, but it also serves as a joke that the characters are in trouble. Both of these words in the TL have only a literal meaning. The translator chose to render only the literal meaning. Apparently, this is the best possible solution that does not obstruct isochrony and lip synchrony in this close-up scene, although part of the playfulness is lost.

Table 9. Examples of translation of culturally-related words

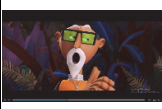
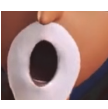
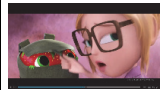
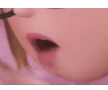
No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face
1.			oo	00:26:52	We're... toast!	toast	Tai... duona!	ˈd̪ɔ̃n̪ə	4	c	f
2.			oo	00:26:55	This is totally ... bananas!	'tootoli	Juk tai tikras... bananas!	ˈt̪ɪkr̪ɛʂ	3	c	p
3.			u:	00:43:31	They're my Wedgie- proof underwear.	pru:f	Jie nesulenda į šiknotarpi.	ʃɪkˈnɔːt̪ɛrˈpɪː	3	M	f

While the TL [ʊ̃] fits the SL [oo] in the first example, the ST [oo] in the second example is replaced with monophthong [ɪ] from the third viseme. Considering that the character's mouth is clearly visible the mismatch of viseme should be especially avoided. Thus, a more suitable translation would be *Ar matau... bananus!*, in this case, the number of syllables can be manipulated as the dubbing actor can start to speak earlier and the word “bananas” does not have to follow lip synchrony, for it is pronounced by a special electronic device.

In the third example, the word “wedgie” refers to either underwear being stuck between one's buttocks or to a prank when someone purposely pulls other person's underpants up (*Collins Dictionary*). This is a culturally-related term, and it exists in source culture while it is absent in target culture. The translator chooses the former meaning, creates and uses an informal word šiknotarpis (“in-between-buttocks”), thus preserving naturalness and faithfulness. Moreover, the continuity between the visual and oral channels is maintained. The only improvement which seems necessary in this case is accentuation and viseme. The utterance in the ST can be improved by eliminating the unnecessary pronoun *jie* – this way the matching stress and viseme are reached (*šiknotarpi*).

The examples in **Table 10** can be used to compare substitutions from the first and second visemes. In the first case, monophthong [ɔ:] is replaced by monophthong [ɑ:] from the first viseme. Although the differences are not extreme ([ɔ:] – back rounded moderately open vowel; [ɑ:] – central neutral open vowel), a better translation would be achieved by omitting redundant pronoun *tą* and using a shorter form of verb *nešiojasi*: *Jaunasis Lokvudai, kodėl tavo draugė savo kuprinėje nešasi pasiutusį braškių?* This way, the phoneme would be a proper substitution from the fourth viseme and the lip synchrony together with isochrony would be improved.

Table 10. Examples of substitutions from the first and the second visemes

No.	Picture 1	Picture 2	Phoneme	Time code	EN utterance	EN transcript	LT utterance	LT transcript	Viseme	Shot type	Face
1.			ɔ:	00:38:03	Young Lockwood, why was your lady friend carrying this rabid little strawberry in her backpack?	'strɔ:beri	Jaunasis Lokvudai, kodėl tavo draugė savo kuprinėje nešiojasi tą pasiutusį braškių?	ɫɑ:	1	m	f
2.			oo	00:51:39	Okay, Barry, time to go home now.	goo	Braškiau, dabar jau metas namo.	'miæ:ɫɐʂ	2	e	¾

The only case of the fourth viseme phoneme being pronounced while showing the character in an extreme close-up is shown in the second example (however, the shot is somewhat between an extreme close-up and a close-up). Thus, lip synchrony and isochrony are essential in this case. Even though such a choice is not ideal – [oo] substitution by unstressed [ɐ] from the second viseme – the neutral vowel does not distort the coherence between visual and auditory channels, for the character speaks and moves quickly. However, a better translation would include the fourth viseme phoneme: *Gerai, braškiau, skuosk namo.*

Conclusions

In order to produce a high-quality dubbing translator has to follow the requirements of credibility, coherence and faithfulness, and manage to match kinetic synchrony, isochrony and lip-sync.

Articulatory features of phonemes allow to classify them into visemes. Rounded and protruded vowels and diphthongs belong to the fourth viseme: LT (in IPA symbols) – [ɔ], [u:], [ʊɔ], [ʊɔ], [ʊɪ], [ʊɪ], [eʊ], [eʊ], [ɔ], [o:], [ɔɪ], [ɔɪ], EN – [ʊ], [u:], [ʊə], [əʊ], [ɔ], [ɔ:], [ɔɪ], [oʊ] (the latter is included by the authors of this paper, considering the fact that in American English (in which film characters speak) [əʊ] is usually replaced by [oʊ]).

There are 476 cases of rounded and protruded English vowels and diphthongs in the film *Cloudy with a Chance of Meatballs 2*. The total number of each phoneme was

calculated: [u:] – 165 (35%), [oʊ] – 151 (32%), [ɔ:] – 92 (19%), [ʊ] – 47 (10%), [ɔɪ] – 15 (3%), [ɐ] – 6 (1%).

The results of the corresponding accentuation: matches in 323 cases (68%), mismatches in 153 cases (32%). The stress matching was most frequently achieved in diphthongs [ɔɪ] (87%) and [oʊ] (73%), and least frequently in monophthongs [ʊ] (64%) and [ɔ:] (64%).

The corresponding fourth viseme was most common in the cases of monophthongs [ʊ] (51%) and [u:] (50%), and least common in the cases of diphthongs [oʊ] (39%) and [ɔɪ] (33%). The overall employment of visemes in translation: the first viseme – 48 cases (10%), the second viseme – 57 cases (12%), the third viseme – 152 (32%), the fourth viseme – 219 cases (46%).

TT that most accurately corresponded to accentuation and lip movements was translated when employing techniques of substitution, change of word order, reduction, omission and repetition. Moreover, in the cases of distracting action / setting, the translator could manipulate the text more freely (although in the film, such opportunity was rarely used).

There is a problem with the number of words and redundancy in the Lithuanian version of the film. The importance of isochrony and lip synchrony (and possible techniques of omitting deictics and choosing shorter synonyms/antonyms) should be emphasised more when preparing new translators.

It was noticed that if there is no possibility to match rounded and protruded vowels and diphthongs from ST with the fourth viseme phoneme in TT, phonemes from the first and second visemes were a better fit than from the third viseme.

The effort to match lip synchrony in the film is evident. However, the moderately frequent usage of the third viseme phonemes as equivalents for the fourth viseme phonemes proves that specialists that provided the translation for dubbing lacked the knowledge of visual phonetics and the viseme theory.

Although one film is not enough to warrant generalisations about dubbing practice in Lithuania, with more attention and further analysis, it can become a significant source of useful data for dubbing development.

Sources

Cloudy with a Chance of Meatballs 2. 2013. [film] Directed by C. Cameron, K. Pearn. Culver City: Sony Pictures Animation.

Debesuota, numatoma mėšos kukulių kruša 2. 2013. [LT translation DVD]. Directed by C. Cameron, K. Pearn. Kaunas: Acme Film.

References

Barbe, K., 1996. Dubbing in the Translation Classroom. In: *Perspectives: Studies in Translatology*, 4 (2), pp. 255–274. <https://doi.org/10.1080/0907676X.1996.9961291>

Barsam, R., Mohanan, D., 2010. *Looking at Movies: An Introduction to Film*. 3rd ed. New York: W. W. Norton & Company.

Chaume, F., 2004. Synchronization in Dubbing: A Translational Approach. In: *Topics in Audiovisual Translation*. Ed. P. Orero. Amsterdam / Philadelphia: John Benjamins, pp. 35–52. <https://doi.org/10.1075/btl.56.07cha>

Chaume, F., 2006. Dubbing. In: *Encyclopedia of Language and Linguistics*. 2nd ed. Ed. K. Brown. Oxford: Elsevier, pp. 6–9. <https://doi.org/10.1016/B0-08-044854-2/00471-5>

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

Collins Dictionary. [online] Available at: <<https://www.collinsdictionary.com/dictionary/english/wedgie> [Accessed 20 March 2020]>

Delabastita, D., 1989. Translation and Mass-communication: Film and T.V. Translation as Evidence of Cultural Dynamics. In: *Babel*, 35 (4), pp. 193–218. <https://doi.org/10.1075/babel.35.4.02del>

Fisher, C. G., 1968. Confusions Among Visually Perceived Consonants. *Journal of Speech and Hearing Research*, 11 (4), pp. 796–804. <https://doi.org/10.1044/jshr.1104.796>

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

Goldsmith, J., 2019. Wants to Make Its Dubbed Foreign Shows Less Dubby. The New York Times, 19 July. [online] Available at: <<https://www.nytimes.com/2019/07/19/arts/television/netflix-money-heist.html>> [Accessed 2 April 2020]

Koverienė, I., 2015. *Dubbing as an Audiovisual Translation Mode: English and Lithuanian Phonemic Inventories in the Context of visual Phonetics*. PhD thesis. Vilnius: Vilnius University.

Koverienė, I., Satkauskaitė, D., 2018. Lithuanian Viewers Attitude towards Dubbed Animated Films. In: *Readings in Humanities*. Eds. O. Andreica, A. Olteanu. Vol 3. Cham: Springer International Publishing AG, pp. 67–84. https://doi.org/10.1007/978-3-319-66914-4_5

Luyken, G., Herbst, T., Langham-Brown, J., Reid, H., Spinhof, H., 1991. *Overcoming Linguistic Barriers in Television. Dubbing and Subtitling for the European Audience*. Manchester: EIM.

Ranzato, I., Zanotti, S., 2019. The Dubbing Revolution. In: *Reassessing Dubbing: Historical Approaches and Current Trends*. Eds. I. Ranzato, S. Zanotti. John Benjamins Publishing, pp. 1–14. <https://doi.org/10.1075/btl.148.00ran>

Roxborough, S., 2019. Netflix's global reach sparks dubbing revolution: "The public demands it". *The Hollywood Reporter*, 13 August. [online] Available at: <<https://www.hollywoodreporter.com/news/netflix-s-global-reach-sparks-dubbing-revolution-public-demands-it-122976>> [Accessed 2 April 2020]

Saenko, E., 2004. *Articulatory Features for Robust Visual Speech Recognition*. Master's Thesis. Massachusetts Institute of Technology. [online] Available at: <<https://pdfs.semanticscholar.org/2ca1/b5706152f5a956a35109c3c2f287b3d3f40e.pdf>> [Accessed 20 March 2020]. <https://doi.org/10.1145/1027933.1027960>

VLKK Konsultacijų bankas. [online] Available at: <http://www.vlkk.lt/konsultacijos/414-bliudas-dubuo?fbclid=IwAR0XeqhiNuZCP_Ezm9UGRqUXsjOmxffKtLSsofhV9SjYIM3WNBBzOMrvmmo> [Accessed 20 March 2020]

VDU Kompiuterinės lingvistikos centras. [online] Available at: <http://donelaitis.vdu.lt/main.php?id=4&nr=9_1> [Accessed 20 March 2020]

Vöge, H., 1977. The Translation of Films: Sub-Titling Versus Dubbing. In: *Babel*, 23 (3), pp. 120–125. <https://doi.org/10.1075/babel.23.3.05vog>

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

Zabalbeascoa, P., 1997. Dubbing and the Nonverbal Dimension of Translation. In: *Nonverbal Communication in Translation: New Perspectives and Challenges in Literature, Interpretation and the Media*. Ed. F. Poyatos. Amsterdam / Philadelphia: John Benjamins, pp. 327–342. <https://doi.org/10.1075/btl.17.26zab>