

A critical review of English pronunciation teaching on  
YouTube and in mobile applications

Bachelor's thesis

Alexi Siidoroff

Juho Sundberg

University of Jyväskylä  
Department of Language and Communication Studies  
English  
May 2018

## JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty <b>Humanistis-yhteiskuntatieteellinen tiedekunta</b>	Laitos – Department <b>Kieli- ja viestintätieteiden laitos</b>
Tekijät – Authors <b>Aleksi Siidoroff, Juho Sundberg</b>	
Työn nimi – Title <b>A critical review of English pronunciation teaching on YouTube and in mobile applications</b>	
Oppiaine – Subject <b>Englannin kieli</b>	Työn laji – Level <b>Kandidaatintutkielma</b>
Aika – Month and year <b>Toukokuu 2018</b>	Sivumäärä – Number of pages <b>43</b>
Tiivistelmä – Abstract <p>Vaikka natiivinomainen puhe ei ole ollut kielenoppimisessa erityisen painotettu osa-alue nykyään vallitsevilla opetusmetodeilla, on ääntäminen silti tärkeä osa englannin kielen oppimista. Se vähentää väärinymmärrysten määrää kommunikaatiotilanteissa. YouTube-videot ja mobiilisovellukset sisältävät paljon materiaalia, joiden käyttöä ei ole tutkittu kattavasti, varsinkaan YouTube-tapauksessa. Teknologian ja mobiilisovellusten käytöstä opetuksessa on saatavilla jo jonkin verran aiempaa tutkimusta. Tämä tutkimus tuo näkökulmaa siihen, millaista materiaalia on saatavilla englannin kielen ääntämisen opiskeluun YouTubeissa ja mobiilisovelluksina. Lisäksi tutkimuksessa pyrittiin selvittämään näiden materiaalien käyttökelpoisuutta opetuksessa ja oppimisen apuna. Tutkimukseen sisältyi 25 YouTube-videota ja kolme mobiilisovellusta, joista tehtiin kuvaukset ja analyysit valittuun kirjallisuuteen pohjautuen. Valitsimme viisi suprasegmentaalista elementtiä tarkasteltavaksi. Kyseiset elementit olivat sanan ja lauseen paino, rytmi, intonaatio ja ‘connected speech’. Tarkastelemastamme aineistosta selvisi, että tämä materiaali on hyvin vaihtelevaa ja ennalta-arvaamatonta. Tästä teimme johtopäätöksen, että sitä tulisi kohdella toissijaisena materiaalina opetuksessa ja opetuksessa. Tällöin ennalta-arvaamattomuutta voisi tasapainottaa luotettavilla ja virallisilla lähteillä, kuten opetusmateriaaleilla kouluissa. Selvitimme kuitenkin, että nämä materiaalit ovat käyttökelpoisia oppimisen tukena, ja pystyvät tuomaan oppijoille erilaista ja autenttista materiaalia. Tämä päätelmä tuki tutkimuksen hypoteeseja.</p>	
Asiasanat – Keywords <b>e-materials, YouTube, pronunciation teaching, education, mobile applications</b>	
Säilytyspaikka – Depository <b>JYX</b>	
Muita tietoja – Additional information	

## **Table of contents**

<b>1. Introduction</b>	3
<b>2. Background</b>	4
2.1 Overview of English pronunciation teaching	4
2.2 English pronunciation teaching in Finland	7
2.3 Use of electronic materials and applications in teaching	7
<b>3. Research questions</b>	11
<b>4. Data &amp; Methods</b>	11
4.1 Data collection	11
4.2 Description of data	12
4.3 Methods of analysis	12
<b>5. Review and analysis</b>	13
5.1 Videos	13
5.1.1 Word stress	13
5.1.2 Sentence stress	15
5.1.3 Rhythm	16
5.1.5 Intonation	18
5.1.6 Connected speech	21
5.2 Applications	23
5.2.1 KEPHAM English pronunciation	23
5.2.2 TFlat English Pronunciation	26
5.2.3 Thomson’s English Accent Coach	27
<b>6. Discussion and conclusion</b>	29
6.1. Videos	29
6.2. Applications	33
6.3. Conclusion	36
<b>REFERENCES</b>	39

## 1. Introduction

Pronunciation teaching has seen many differences of opinions and its importance or the overall role in language teaching and how it is best taught. Some of these approaches put more emphasis on pronunciation than others. For example, if we look at the direct method of language teaching, the learners' native language is not suggested to be used in teaching. Instead, all teaching happens in the target language (Hişmanoğlu, 2006: 104). However, the direct method is seen as unrealistic because ideally, in this method the teacher should be a native or near-native speaker (Hişmanoğlu, 2006: 105). Naturally, the vast majority of English teachers are nonnative in countries where English is not a majority language, so this ideal is not met. Currently, the communicative approach is the most prominent of the teaching methods (Celce-Murcia et al., 2010: 8). The emphasis on pronunciation teaching has shifted from teaching segmentals to suprasegmentals as well as focusing on the needs of an individual learner (Hişmanoğlu, 2006: 104).

Internet sources and mobile applications can help provide a wide range of different types of learning possibilities with ease of access and a chance for students to learn, not only guided by the teacher, but also independently. New possibilities offered by the growing industry of mobile applications and video sharing sites, such as YouTube require more attention. As Celce-Murcia et al. (2010: 360) emphasizes, internet provides the future of pronunciation technology; web sites offer materials for pronunciation that include sounds, videos, graphics and more. These sources serve teachers and students many resources that are not tied to classroom teaching, and also allow easier emphasis on individual learners. This includes access to authentic material that is provided by native English speakers, such as in the case of YouTube.

Teachers can include more diverse and interesting learning methods using different applications. However, a review of these materials is necessary as it is not a given that the materials are suitable for use in teaching and independent learning. The materials can be both disadvantageous and advantageous to the user or inappropriately made, hindering learning instead of supporting it. Thus, the overall aim of this study is to provide some insight to how useful these sources are for educational use and how well they agree with already existing theoretical background.

This study also includes an overview of English pronunciation teaching and the use of mobile applications in teaching. While a number of studies about the use of technology and applications in teaching languages (e.g. Derwing, 2015) have been conducted, the same cannot be said about the use of YouTube materials, even though it is a relevant issue due to digitalization and the popularity of social media and internet content in general.

## 2. Background

### 2.1 Overview of English pronunciation teaching

Even though human speech can be a highly sophisticated and complex tool, people tend to assign it a very simplified dichotomy consisting of consonants and vowels (Odisho, 2003: 53). However, this dichotomy alone is unable to convey the complexity of human speech. When observing language a bit more closely, we can make out different features. Segmental features, such as phonemes, are important to language, but they alone cannot explain the whole language. As Odisho (2003: 53) points out, this is why we need suprasegmental elements of language. These suprasegmentals consist of features, such as intonation, stress, rhythm. Odisho (2003: 53) also reminds that there are different linguistics schools that view these elements differently. We have chosen word and sentence stress, rhythm, intonation and connected speech as the suprasegmental elements we focus on our study. In contrast, as the applications had an inherent emphasis on phonetics themselves, it was the main focus. It should be noted, that the efficiency of phonetic training itself is not the result of the amount of exposure to sounds received in L2 language. Instead, the effectiveness could be seen to depend on how successfully phonetic training guide the learners' attention towards the phonetic cues natives attend to. The most promising and successful of paradigms that guide the learners' attention towards these phonetic cues is argued to be high variability phonetic training (HVPT) (Aliaga-García, 2007: 10.). The HVPT method uses high-variability stimuli in different phonetic contexts to create an environment where these phonetic cues are best realized, such as using multiple different speakers' voices in training instead of just one.

The study benefited greatly from the book *Teaching Pronunciation: A course book and reference guide* (Celce-Murcia et al., 2010), as it provided a good general reference for English pronunciation and its teaching. Pronunciation teaching has great theoretical background which should not be ignored when setting sights on new methods of pronunciation teaching. The forming of Phonetic Association and the creation of International Phonetic Alphabet (IPA) acted as inspiration to the

reform movement in the 1890. This resulted in pronunciation teaching suddenly seeing a variety of theories and approaches as it tried to establish the role and importance of pronunciation in the field of language teaching. Communicative Language Teaching (CLT), also known as the Communicative approach, settled as the most prevalent of the approaches. CLT is still used today in language teaching. According to it, communication in the language should be the central part of language instruction in the classroom. Celce-Murcia et al. (2010: 8) summarize the urgency of pronunciation teaching, noting that there is both, empirical and anecdotal evidence to support the importance of pronunciation.

There seems to be a threshold level of pronunciation for English speakers that are nonnative. Should the level of pronunciation fall below this level, they will have problems with oral communication regardless of their grammar competence or the size of their vocabulary. The goal, according to Celce-Murcia et al. (2010: 9), should thus not be to teach learners to sound like native speakers of English, but to help them surpass this threshold level so it would not obstruct their ability to communicate.

IPA is still widely in use today, as it can be used to represent the sounds made in any language in a visual form (Celce-Murcia et al., 2010: 3). The use of phonetic transcriptions in teaching a foreign language varies greatly. Sometimes these tasks are used in a systematic manner to help students with their pronunciation skills whereas sometimes they are not used at all. It depends on the teacher's competence in phonetics and the students' understanding of the topic (Heselwood, 2013: 253). Phonetic transcription can help students with the pronunciation of a single word and offers great help in teaching connected speech (Heselwood, 2013: 254). When it comes to using transcription tasks, it is advised that a balance between the level of detail in phonetic transcription and the students' proficiency level must be found before applying them to teaching.

There has been limited research on the use of IPA in pronunciation teaching. One clear take on the subject was provided by Lintunen (2005), where Finnish university students were observed while they studied phonemics and pronunciation. His findings state that there is a clear indication for phonemic transcription having a positive influence on learners who are accustomed to orthography. He suggests that the phonemic transcription could make learners more aware of this level of language, and therefore benefit from it in terms of pronunciation learning (Lintunen, 2005: 1).

Celce-Murcia et al. (2010: 9) offer a rather comprehensive list of teaching methods that were present even before the communicative method gained favor. They are presented in the table below.

Table 1. Collection of teaching methods

Method	Description of the method
Listening and imitating	A method where the students listen to a model provided by the teacher and then repeat after it. This model can be a recording as well and it is suitable for computer labs.
Phonetic training	A method that uses articulatory descriptions, diagrams and the IPA in teaching
Minimal pair drills	Realized using words that differ only by one sound from each other, emphasizing that a phoneme can be a very slightly distinctive sound
Contextualized minimal pairs	pairs are given a context, for example “The blacksmith hits/heats the horseshoe” paired with a cued answer “with the hammer/in the fire”.
Visual aids	Include tools that can help the teacher to visualize the teaching, such as sound-color charts, fidel charts or other props and visual methods.
Tongue twisters	Used in strategies for speech correction
Developmental approximation drills	Encourage ESL speakers to retrace the steps taken by L1 speakers when they first are learning the language, such as acquiring /w/ before /r/, “wed → red”
Vowel shifts	Includes differentiating the long and short vowels for example in the words “mime” and “mimic”, where the first one has a long i and the second a short i. It is given a sentence context, such as “Mimes often mimic passersby”
Stress shifts	Emphasizes the use of stress in words such as “PHOtograph” and “phoTOGraphy”. They are also given a context, such as “Judging by these photographs, you are good at photography”. Reading aloud or using recitation is another way to draw focus on stress, timing and intonation, especially on texts made to be spoken
Recording learner’s production	Can prove to be useful as it offers an opportunity to go back to the material and receive feedback from the teacher or peers.

However, since the new approach started to shift towards teaching language as communication, majority of the aforementioned methods were deemed unworthy of the communicative method. Teachers and material developers focused on coming up with better methods of teaching pronunciation in the communicative frame. The result was to start paying attention towards the suprasegmental elements of the language. The current situation seems to have found a balance between segmental and suprasegmental features. However, new research indicates that focus on prosody, and thus suprasegmentals can improve the comprehensibility and intelligibility of the learner more than focus on segmental features (Celce-Murcia et al., 2010: 33.).

We have discussed prosody and suprasegmental features without delving deeper into their meaning. Let us examine the elements that are relevant to this paper a bit closer. When talking about stress, the focus is directed towards a stressed syllable. It can be realized by the speaker using a greater

physiological effort when producing the stressed syllable. It is often louder than an unstressed one (Odisho, 2003: 55). Stress can occur in a word or within a sentence. Whereas stress occurs in a certain syllable within a word, in the sentence it occurs as prominence in a stressed word (Odisho, 2003: 55).

Linda Lane (2010) provides a practical approach to teaching pronunciation. She lists four main goals for pronunciation teaching: intelligibility, comprehensibility, accent and voice quality. Intelligibility is defined as a listener being able to recognize words, phrases and utterances. She argues that correct use of suprasegmentals seems to be the most impactful aspect for intelligibility. Furthermore, she mentions that native and non-native listeners do not always find the same elements equally important for intelligibility. Comprehensibility, in her book, is defined as being able to understand non-native speakers. Here, the important factors are both segmental and suprasegmental elements. For accent, Lane (2010: 3) explains that the teachers should pay attention to distracting, stigmatized and stereotyped pronunciations. Accent, in her definition, is a combination of segmental, suprasegmental and voice quality aspects. Voice quality is defined as pronunciation features that are generally accepted as parts of native speech, such as the level of pitch. This is important because, for example, some languages may be spoken in higher or lower pitch than others (Lane, 2010: 88).

## 2.2 English pronunciation teaching in Finland

To get more familiar with the topic, we are going to view the state of pronunciation teaching in Finland to tie our research more closely to the curriculum of English teaching in Finland.

In her doctoral dissertation consisting of four studies, Tergujeff (2013) found that Finnish teachers use a lot of the traditional techniques for pronunciation teaching. These techniques include imitation, reading aloud and phonetic training. According to Tergujeff (2013), there can be found eight prevalent types of materials regarding pronunciation teaching in text books, of which phonetic training is the most common one, followed by reading aloud and listen and repeat -practices. However, despite the suggested methods presented in pronunciation literature previously, the textbooks of EFL in Finland seem to be short of teaching materials on connected speech, rhythm and intonation (Tergujeff, 2013: 42). In addition, there seems to be some controversy in views about the usefulness of knowing phonetic symbols. In her study, Tergujeff found that they were regarded essential by some learners, but the other half found them more confusing than helpful (Tergujeff,



2013: 44). When assessing pronunciation, the most frequently used tasks include reading tasks, oral performances and exams.

Furthermore, Tergujeff (2013) found that the techniques used by teachers were very varying. Most common methods were imitation tasks and teacher corrections, as most of the teachers seemed to utilize them. It was also found that teachers point out issues in the learners' pronunciation and corrects them. The teaching seems to be heavily focusing on segmental features, especially those that might be difficult for Finnish learners, such as sibilants and affricates (Tergujeff, 2013: 46). However, the studies concluded that learners do not aspire to pronounce like a native speaker and instead the aim was on fluent, intelligible speech.

### 2.3 Use of electronic materials and applications in teaching

Over the years, internet has become one of the most useful platforms for a multitude of purposes. One of such is education. Only a few years ago most of the teaching materials were in written form, even when it came to communication learning. Only when the use of internet became more frequent as bandwidth costs got cheaper, audio and video content on the internet began increasing tremendously. As a result, audio and video blogging, for example has gained an entryway to aid education in multiple fields, including foreign language education (Chan, 2011: 25). One of the most important skills in today's schools is to learn how to use all the available media around us Jenkins (2006:19) brings up that adolescents need to learn how to integrate knowledge from multiple different sources. Jenkins lists these sources as music, video and online databases as well as other media. As Jenkins (2006: 19) mentions, the very definition of twenty-first century literacy is to recognize that set of skills where visual, aural and digital literacy interlace with each other.

The appeal of user-created content in YouTube and its popularity has been on the rise for a long time and it is important for teachers and students to realize its potential. For example, 450,000 people watch over 1900 educational TED Talks content and the TED Talks YouTube channel has over 8,9 million subscriptions as this paper was written (The Conversation, 15.01.2012). The same goes for mobile applications. Students lead an especially hectic life balancing between school and social life. Mobile applications are a popular on-the-go method to spend time in an engaging way. Mobile gaming is a rapidly growing industry; there are already applications that are not only for fun and play, but are also making education more attractive to learners. Instead of sitting in the classroom, in an ideal world the applications would work in tandem with classroom teaching and learning and should be realized as potential and powerful learning aids. According to Statista.com,

an U.S. research found that in 2016, gamers played 3.6 mobile games per month, which translated to 1.3 games daily. This figure is only expected to rise in the coming years. Regarding educational applications in the field of language learning, Duolingo, a language learning application has achieved the mark of 200 million active users according to DMR business statistics (Smith, 2018).

Derwing & Munro (2015) raise an interesting point when it comes to the level of technology. They point out that with the current technology, it is easier to gain access to activities that would have required considerable effort to produce in the past. As quoted by Derwing & Munro (2015: 123), Goodwin (2013, as cited in Derwing & Munro, 2015) reviewed the use of shadowing and mirroring in digital multimedia tools, such as YouTube. In shadowing tasks, the learners imitate what a model is doing slightly after or at the same time. Mirroring means that the learner is imitating speech and body movements. Derwing & Munro (2015) reminds that such activities have been viewed as backbones of pronunciation classes for years, but they also have gained critique for not being very interesting or being downright demotivating. Naturally, technology is not the answer to all of these problems, but it offers easy, diverse and natural content for such purposes (Derwing & Munro, 2015: 123). Furthermore, as Thomson (2012: 1232) states, empirical research that is merged into materials of L2 pronunciation is rare to come by despite all the advances in the field of understanding L2 speech learning. Current study indicates that a computer-mediated approach to L2 speech learning could bring results in L2 speech perception gains that are not allowed in other, more naturalistic learning environments, such as classrooms.

Teachers have many electronic platforms available for use in their teaching. These platforms can offer many advantages to pronunciation teaching on individual level. Derwing & Munro (2015: 124) argue that teachers can use materials, such as audio and video clips, links and website where students can practice materials that they have already discussed in class. Furthermore, teachers can point students to focus on areas that they have problems with individually, thus saving this time in class. This also gives students some freedom, as they can decide themselves when and where to do those assignments (Derwing & Munro, 2015: 124).

The use of video games in language learning and learning in general has been researched quite extensively. Multiple incentives, such as games reducing anxiety and improving motivation have been given as arguments for their use as teaching tools (Reinders 2012: 1). Reinders (2012) mentions that as early as in the 1990's it was predicted that digital games could become an integral part of teaching. In addition, he reports views that argue the digital games answer to the new expectations from "digital native" learners and new possibilities new types of learning. It is also mentioned that young people are likely to transfer their creative skills to use digital technologies in

out-of-school practices and that game environments provide a possibility to explore learner's skills and improved engagement more effectively. Furthermore, it is suggested in the article that digital games in teaching could provide solutions for such problems as boredom, dropouts and disengagement in schools (Reinders 2012: 50).

As mobile applications are often gamified at least to a degree, with features such as having scores and goals to achieve, some insight into using games in education is needed. Peterson (2013: 15) provides essential characteristics of *play* in games: voluntary nature, the sense of it occurring outside of everyday, absorption, satisfaction and enjoyment. While traditional educational practices may find difficulty in fulfilling such characteristics, applications that borrow from game-like features may achieve some of it. Furthermore, some games can be classified as play governed by a collection of rules (Piaget, 1961, as cited in Peterson, 2013: 33). This nature of rule governed environment can be directly applied to applications in this study, as this acting in this type of environment entices cognitive development (Peterson, 2013: 33). He further argues through multiple studies, that rule-based games can increase motivation of the learner, as the rules are closely connected with *goals*. Ma et al. (2011: 10) show that the popularity of gaming and internet have encouraged the emergence of serious games and applications for education. They support the view that games have a potential to motivate and engage learners through problem-solving and game-like activities and argue that games can also improve self-monitoring abilities (Ma et al., 2011: 12).

Kucirkova & Falloon (2017) provide direct insight for the use of applications in children's education. They illustrate how devices, such as smartphones, have become integrated part of people's lives, thus offering potential for learning regardless of age, culture and place. The mobile technology is an effective learning tool for active and independent learning and can improve learner's self-regulatory and metacognitive skills (Kucirkova & Falloon, 2017: 103). However, they also show that the amount of information provided to parents and educators about available applications is highly inconsistent (Kucirkova & Falloon, 2017: 46).

One of the most popular language learning applications to day is the educational language game Duolingo. In Munday's (2016) study, it was used as a part of a program in two Spanish university courses. One of them was a beginner level course and the other an intermediate level course. In the study Duolingo was assessed from multiple points of view and possible ways of incorporating Duolingo in foreign language teaching. It was found that the students found Duolingo easy to use,

helpful and enjoyable application when learning their target language. In the study there was also a clear preference in using Duolingo over traditional textbooks and formats. Duolingo was also preferred due to its game-like features (Munday, 2016: 96). There were many pros in using Duolingo in regards of motivation. However, using this type of an application also had some cons, such as the lack of accuracy in some of Duolingo's translations (Munday, 2016: 94).

E-learning in general has been increasing in importance and popularity in recent years. Liu et al. (2002: 75) argue that Internet based online learning has become the main alternative for distance education. In this case it means providing students learning materials through internet and using virtual environments in education. According to their article, they give the students a rich and engaging learning space that is not bound to a classroom.

### 3. Research questions

This study has a focus on YouTube videos and mobile applications where one can learn English pronunciation. We had two research questions: firstly, the study aspired to find what kind of English pronunciation teaching materials are available on YouTube and as mobile applications. Secondly, we wanted to consider how suitable the materials are for learning English pronunciation. This can be considered in both, classroom teaching and self-study options.

We expected to discover that there are plenty of different and interesting materials in YouTube that could aid in pronunciation learning, and that they could also be used as a pronunciation aid in classrooms and as a self-study option. However, we expected that some of the material analyzed would not be suitable for educational purposes as YouTube material is not monitored in the same way as official materials. For example, as YouTube content is user created the title of the video may not describe the actual content of the video or the content may include inappropriate elements from the point of view of pronunciation learning or educational environment in general. As for applications, we expected to find that the applications found for mobile devices could also be utilized in pronunciation teaching and learning. We expected that this material is more task than information oriented and could be used in conjunction with the YouTube videos or with classroom practices. Furthermore, we expected the applications would be limited in some areas.

## 4. Data & Methods

### 4.1 Data collection

The data was collected from the social media material provided by different content creators in YouTube, as well as reviewing educational applications available for mobile devices. The criteria for choosing the videos and apps was basically two folded; their apparent suitability for use in education and availability for users. Our preference regarding the selection of applications was those that do not require subscription and are free to use at least initially, as the applications may demand payment to proceed further in using them, or gain access to extra features. Depending on the data and the course of the research determined by the data we wanted to keep all of our options open until we obtained a clear understanding of our data. As for the analysis method, we used qualitative content analysis to break the study down in smaller pieces that are easier to work with. They were categorized into, for example, videos that focus on stress or intonation. The applications were compared and labeled into categories according to themes stemming from the findings.

The data for the study was collected from YouTube video materials and applications available for mobile devices. The YouTube videos were selected by searching the keywords determined by the categorized criteria. For example, searching with the keyword intonation brought up a certain number of videos, five of which were chosen as sources for the data. The same was repeated with all the chosen themes. This resulted in a total sample size of 25 videos for the data. The applications were analyzed by testing the features they provided, as well as their functionality and emphasis on the same themes that were used for the videos.

### 4.2 Description of data

The study included a total of 25 YouTube videos that were divided into five categories: word stress, sentence stress, rhythm, intonation and connected speech. On top of the videos, we chose 3 free applications available for mobile devices. The reason for choosing free applications was their availability for a large quantity of people and therefore their increased probability to be used by learners. It is less likely for English learners to pay for the applications than to choose an application that is free to use. However, it is worth to mention that some of the applications included the possibility for paid content via microtransactions within the app, which could mean

that the experience of using the app or its benefits may not be the same for all the users. The chosen applications were: KEPHAM English Pronunciation, TFlat English Pronunciation and Thomson's English Accent Coach. The first two of the application are available on android, while the Thomson's English Accent Coach is only available as an internet application and on iOS-devices.

### 4.3 Methods of analysis

The main method for analyzing the data was qualitative data analysis. Seidel (1998:3) displayed noticing, collecting and thinking as important parts for the process of qualitative content analysis. This model's idea is to show how the data collection method is a manifold process that does not always progress directly from stage to stage. Researcher might begin the analysis by taking important themes into account, as we have done with our suprasegmental themes and gather them into analysis. After this step, the researcher can contemplate the content that was collected in the research which might lead to new ways of perceiving the phenomena.

The categorized data was compared with the guidelines provided by previous studies and literary concerning English pronunciation teaching and the use of e-materials in educational practices. The videos were analyzed by first describing them and then reviewing them based on the background material. For the videos we had decided the categories beforehand, concentrating on suprasegmental elements of pronunciation teaching. However, as the data was collected it became clear that some of the original categories needed revising, and so some of them were changed or removed. Mobile applications were similarly compared with the guidelines of the studies. Same theoretical background in addition to studies concerning the use of e-materials in teaching were applied in the case of the mobile applications. In contrast to the videos, the features of the applications were tested instead of observed.

## 5. Review and analysis

### 5.1 Videos

#### 5.1.1 Word stress

Videos within the word stress theme mostly used a similar structure where the rules of word stress were explained, though the rules were rarely the same. For example, the word stress videos 1 and 3 had the basic rules of word stress usage. Video 1 had five rules for word stress, presented by a speaker facing the camera without any props. However, to visualize the lesson, examples and words are edited to appear in sync with the speech. Words with suffixes -ee and -oo usually stressed the last syllable. Video 3 had only two rules explained, but they took similar approaches with the other videos. In video 3, the rules given were that word's suffixes often determine the stress, such as -logy, -ical, -tic or -ate. The second rule, however, is similar to rules one and two in the word stress video 1; with noun words the stress is on the first syllable and verb words the stress is on the second syllable. Videos 3 and 4 explained word stress partly through IPA chart and demonstrated how the stress is marked on the words written in phonetic alphabets. All of the videos within the theme gave clear examples of words and where the stress is by displaying it in the video, whether it was with animation or just by bolding parts of the word. Within word stress theme one videos used whiteboard, two videos had a visible person on the screen explaining the topic and used post-editing to further clarify the explanations. Two videos in that category only had animated slides with examples while the topic was explained by a speaker. Many of the videos in word stress category utilized a structure of tips or simple rules of how to learn to recognize where the stress on the word is.

For example in video 1 there were five rules: nouns and adjectives with two syllables, the stress is on the first syllable, most verbs with two syllables, the stress on the last syllable, thirdly, words ending in “-TION”, “-SION”, “-CIAN”, are usually stressed on second last syllable (e-du-CA-tion), words ending in “-IC” are usually stressed on second last syllable (e-co-NO-mic) and lastly, words ending in “-EE” and “-OO” are usually stressed on the last syllable (em-ploy-EE). These rules are simplified versions of general stress rules found in our selected literature. Noun-verb doublets are one of the most appropriate ways of teaching English (Odisho, 2003: 99). The three latter examples are also highlighted in the selected literature. It is likely better to not confuse the learners with too

much information. Rules three and four are relevant, because with many words, the suffixes cause stress to shift to the penult. However, this happens with many more suffixes than the three mentioned (such as “-cial / -tial” or “-geous / gious”). It seems like the video could go more in depth, but it also might overwhelm the learner. In the last rule, the need for stress is caused because it is one of the stress-demanding suffixes (Celce-Murcia et al., 2010: 190-191), which are also more varied than the video gives them credit for. The second video mostly concentrates on noun-verb stress patterns with more examples than the first video.

In video 3, there are only two rules mentioned: the word endings determining stress, and noun-verb patterns. For emphasis, the examples use bubble method of showing where the stress of the word is. Video 3 only had two levels presented with the bubbles, unstressed and stressed, thus leaving out if there were lightly stressed or strongly stressed parts. An interesting method of teaching word stress was introduced; working backwards, which did not show up in other videos that were reviewed. The working back method also does not seem to be prevalent in any of our literature. In this method the learner reads the word backwards syllable by syllable, such as “experimental” (/ek ʃpersə'mentəl/) counted backwards syllable by syllable, starting from “/mentəl”.

However, one might question the authenticity of IPA methods used in the video, since, for example, the Cambridge dictionary offers different phonetic transcription for the word “experimental” (/ɪk,spɛr.ə'men.tʃəl/). The video encourages the students to learn the places of word stress by using dictionaries and seeing where the superscripts and subscripts are placed in phonetic transcriptions. Video 4 had the focus on three syllable words and their stress patterns. According to Odisho (2003: 102), it is easier for learners to familiarize themselves with stress patterns by using nonsensical structures that contain two or more syllables, such as <da> and <DA>. The video utilizes this very method, demonstrating the long and short syllables, or in other words, stressed and unstressed syllable. The video also introduces the learner to superscripts and encourages learners to seek them in dictionaries to practice the word stress. The main focus of the video, however, is the rhythmic patterns, such as “da-da-DA” with a counterpart word, “everyday”. The literature seems to emphasize communicative approaches more than nativelike speech, but in this video, for example it is encouraged to learn these stress patterns to sound more American. Video 5 taught word stress through finding syllables in each word but did not go in detail about any rules of how to find or recognize the stress. The method is almost purely auditory, clapping while pronouncing each syllable. These auditory techniques, such as clapping or tapping can be useful for stress perception teaching, too (Odisho, 2003: 95). Interestingly, none of the videos specifically taught the three main



levels of stress and only stayed with the dichotomy of stressed and unstressed syllables whereas they could have also explained the differences between strong and light stress.

### 5.1.2 Sentence stress

Most of the sentence stress videos were very similar in presenting the topic. Furthermore, the videos in the category were similar in their way of explaining what sentence stress is. Video 6 presents it as louder or longer portions on the sentence. In contrast, video 7 describes it otherwise similarly, but adds melody and pause to it. All of the videos used the same approach with their examples. The videos examined sentences by stressing different words within the sentences. It was then examined how the meaning of the sentence changed according to the stressed word. Video 7 is a good example of this method. It has a sentence “Jack will cycle to the restaurant tonight” which is then broken into 6 parts, and each part was discussed with a stress on a different word. Three videos used whiteboards as visual aids with a person explaining the topic in the video. Videos 7 and 10 had a slideshow in the background. However, video 10 had a more engaging approach in its slideshow with illustrative pictures. There were many differences with the presentators as well. For example, videos 8 and 9 had similar approaches. Both had a person with a whiteboard, but while the video 9 was very straightforward and serious, video 8 had a more personal and casual approach with enthusiasm and relatable examples.

The videos under sentence stress do not vary from each other much. Videos 6 to 9 are very similar in their presentation of how sentence stress works. Because the stress does not go in small units such as syllables, sentence stress is viewed in a broader aspect. The videos seem to drive the meaning of sentence stress through in a simpler manner than the literature. Video 6 presents a claim on a whiteboard, “I like your painting”, and then presents three different meanings that are achieved by placing the stress on different parts of the sentence. The meanings are, “it’s an okay painting”, “my friends don’t like it” and “I only like your painting, not the others”. The rest of the videos went through similar practices and explained how the meaning of the sentence changes when different words in the sentence are stressed. The examples are different, but the basic content is almost identical. Video 10 offers variety to this pattern, since it goes deeper than giving examples on how the meaning of the sentence change when different words are stressed. English is a stress-timed language, rather than syllable-timed, meaning that the syllables of a word are not the same length in comparison to each other, which might be important for students to understand. Overall, the video explains that often content words - verbs and nouns - are stressed in sentences. This also falls in place with our selected literature, but the question is if this is really necessary for the learners to

know. In the end, the video does fall in line with the rest of the videos in showing how the meaning of the sentence changes when a different word is being stressed. It seems that all but video 10 approach the issue through a very task-oriented view in trying to make the learner understand how important word stress is and how it can make sentences change their meaning almost completely. Video 10 takes a more theoretical approach, explaining how to recognize sentence stress in a more broad way.

### 5.1.3 Rhythm

The videos within the rhythm theme were more varied in terms of their views on teaching rhythm and thus we will give examples of each separately. Video 11 uses the iambic pentameter to illustrate rhythm and displays its usage through boxing according to the rhythm patterns while speaking. Video 12 demonstrates rhythm via clapping to the right rhythm and thereby shows how it affects the flow of a sentence. For example, how to find syllables in words such as “desk” or “printer”. The former is realized by one clap while the latter with two claps according to their syllable count. Video 13 gave a more academic approach revolving around one phrase: “Stick out your tongue”. This phrase was then examined in different ways to demonstrate how the use of rhythm affects it. First, the content words were identified to find out which words to stress. These words were identified as stick, out and tongue, further categorized as verb, adverb and noun. Second, the function word to de-stress was identified, which in this case was ‘your’. Third, the connect and link areas were established, which help with the flow, efficiency and musicality of the phrase. Video 14 had example words that were often stressed, such as verbs, nouns, adjectives, adverbs and question words in a sentence. Unstressed words were articles, prepositions, pronouns, auxiliaries and conjunctions. None of the above were discussed more thoroughly in other videos. Video 14 also touched on word compression to produce a correct rhythm. This was similar with the third rule in video 13. The final video in this category, video 15 tried to explain rhythm through the use of chunking and dechunking. The video demonstrates the difference in rhythm by using chunking in a sentence and reading it aloud, and then comparing it to a non-chunked sentence. Some practices are also offered in the description of the video to help viewers follow the practice part of the lecture. Two videos, 11 and 12, in this category had only the person speaking in front of the camera while explaining the rhythm. Video 14 had a person speaking and a whiteboard to demonstrate the topic. Video 15 also had a person in front of the camera, but to aid with teaching an embedded slideshow was used. Video 13 only had a slideshow with use of different colors and phonetic alphabet to emphasize the important parts.

The videos in this category were understandably very closely related to videos about stress. One of the major differences between teaching stress and teaching rhythm was when it came to rhythm, poetic forms and similar rhythmic examples of language were present. Video 11 described rhythm through the use of iambic pentameter, which is used in poetry and translates to a weak beat after a strong beat. The examples were drawn from music and shakespearean poetry. According to Celce-Murcia et al. (2010: 350), a very effective way to teach the stress-timed nature of English language is to use poetry to illustrate this. These traditional rhythmic feet, such as iambic pentameter which was present in videos 11 and 14, but also other rhythmic feet such as hexameters etc. are used to illustrate rhythm. According to video 11 this can be achieved by counting beats or thinking the flow of speech through music. Video 12 emphasizes the patterns by clapping. As mentioned before, these kinds of methods can bring emphasis on the lesson. However, this seems more like a video about word stress rather than rhythm. One should note that rhythm is very closely related to stress and they can often be confused with one another. These misunderstandings are relatively common. Rhythm results from alternation between content words and grammatical words, whereas in word stress the alternation is between the syllables of a specific word (Linda Lane 2010: 45).

Video 13 includes more suprasegmentals than just rhythm, but rhythm is still represented as part of the overall pronunciation teaching. Video 13 firstly guides the learner to identify the content words of the sentence which they will stress, and secondly, identify the function words which they will de-stress. For example, the word used in the video is “stick out your tongue”, where “stick, out, tongue” are stressed and “your” will get de-stressed, and thus appear in the sentence in faster pace. This does not seem to be the most effective way to demonstrate rhythm as a separate entity, as rhythm is best realized separately using poetry, as previously mentioned. However, it seems to work as a more broad guide to teach this particular aspect of English pronunciation. Step three in video 13 goes more into the area of connected speech, that is expanded further later on in the paper.

Video 14 offers a more traditional view of teaching rhythm. Even though the video itself does not stand out among others as it is simply a teacher in front of a white board. The lesson, however includes a short poem to illustrate the use of rhyme. The video describes rhythm as sentence stress and word or syllable stress, which highlights their close relation to each other. Although, this might confuse some learners as the difference and the relation of those terms is left unclear. However, this video seems to be the most closely related to the literature containing the teaching of rhythm, going through a poem and even mentioning the use of iambic pentameter.

Video 15 introduces the learner to the term chunking, which is repeated throughout the video. Chunking is a new term as it was not explored in the literature, but it seems that it is closely related to linking and other aspects of connected speech. Introducing new terms is questionable, as it might confuse learners that have previous knowledge about the terminology. However, if chunking is a term that better explains these aspects of pronunciation, learners might find it more useful or innovative than already existing literature that does not involve it. It can also be compared to Lane's (2010: 47) tip for teaching rhythm to students; helping students find content words by lengthening the end of the group before to make the distinction.

### 5.1.5 Intonation

Intonation videos were more varied in their content and even provided instructions for two different accents; British English and American English. Video 16 had the levels of intonation labeled falling, rising, rise-fall and fall-rise. These were then gone through by using example sentences including one or multiple of these. The examples had visual aids pointing out the pattern of intonation. Video 17 was a podcast with no visible person on the screen, but instead an animated slideshow. The video explained pitch and gave illustrative examples by comparing sentences. It also visually demonstrated rising and falling intonation with visual aids in the video. Video 18 focused on American English intonation, briefly explaining intonation and pitch. The video stressed the importance of music and emphasis in speech. It demonstrated the difference in intonation between a statement and closed question. It is an interaction between a teacher and a student that is acted out. This is the only one of its kind in the all of the categories of videos selected. Video 19 defined intonation by referring to stressed and unstressed syllables, which form a pitch pattern into a sentence or a phrase. It also explained that the melody, basically meaning the intonation, of speech is important as it conveys different meaning and attitudes. Furthermore, the importance of correct intonation to sound natural and American was emphasized. In contrast, video 20 showed British intonation and focused on how word stress and syllables affect intonation. The main point of the video was that weak forms that are not stressed in speech. It used visual aids and humor to support the teaching. Additionally, it employed some practices and repetition.

English is an intonation language, which means that it uses pitch to demonstrate the combination of different syntactic, semantic and attitudinal traits of an utterance (Odisho, 2003: 59). Intonation is used to identify and highlight important information to the listener (Lane, 2010: 85). According to Odisho (2003: 59), pitch patterns are essential in learning intonation, and thus, pronunciation. These are realized with the terms pitch height and pitch direction. The pitch levels are most commonly

labeled as high, mid, low, mid-high and mid-low. Pitch directions are high fall, low fall, high rise, low rise, rise-fall, fall-rise, high level and low level.

Video 16 demonstrates only four of these patterns, explaining the use of falling for understanding, rising for doubt, fall-rise for surprise and rise-fall for hesitation. The example sentences have varying intonation patterns, which are demonstrated with arrows pointing to the direction of intonation, marking which of the patterns it represents. To emphasize the lesson, the speaker uses hand gestures, while repeating the sentence. It is worth noting, that body gestures are important to emphasize the teaching and might help the learners to pay attention to the important parts through the physiological efforts or body gestures of the speaker (Odisho, 2003: 96). Overall, the lesson is easy to follow with simplified theoretical background. The learners can, for the most part, recognize the emotion contours (such as implying uncertainty by using the rising pattern) of intonation and use them to their advantage. Getting familiar with intonation and the varying pitch patterns is important from the communicative point of view as well. As Celce-Murcia et al. (2010: 248) point out, nonnative speakers of English can often be misinterpreted as rude or uninterested etc. because of the lack of these prosodic elements in their pronunciation.

Video 17 begins the lesson with explaining that intonation generally connects to pitch. However, the video claims that there are only three intonation patterns in English: rise, fall and partial fall. Rise and fall are explained as before, rising intonation meaning surprise, and falling intonation meaning statement. Falling intonation is also demonstrated to appear on wh -questions, such as “what”, “where”, or “who” as a side note, which might have been important to explain more clearly than in just as a side note. Partial fall was a term that did not appear in any of the other videos or in the chosen literature, but it resembles nonterminal fall, which is used to demonstrate uncertainty (Celce-Murcia et al., 2010: 239). It should be mentioned that this very simplified demonstration of intonation is questionable, as it does not give the completely picture of English intonation, despite the claim in the video. However, it could also be done on purpose, to keep the lesson easier to understand.

Video 18 explores the intonation through five sentence types that are presented in the video and have a predictable intonation pattern. These sentence types represent statement, closed question, open question and clarification question. The final sentence type is not mentioned by name at the same time with the others, which might confuse the viewer if the viewer is interested in getting to know these terms. At the very end of the video, the teacher of the video names it as an “either-or-question”, which we identified as closed-choice alternative question based on literature. In the video, two terms are introduced, “stepping” when the pitch of voice rises before falling intonation,

and “gliding”, which was used to demonstrate the musicality of the fall in the intonation. These terms are not prominent in the literature, and seem unnecessary, since intonation itself implies the falls. The method used in the video resembles the traditional method for teaching intonation, which teaches the most frequent intonation patterns via associations to the intuitive native-speaker patterns. This method is criticized as it does not always reflect the intonation in natural speech (Lane, 2010, 92).

Video 19 seems to be a simplified explanation video that does not have many visual aids to help the viewer. Intonation is explained to be the rise and fall in pitch, and closely related to melody. The speaker demonstrates the intonation by using a software called Praat, that shows spectrographs of exact patterns of intonation in recorded speech. This might be interesting for a learner to look at, especially if the learners could analyze the speech of authentic native English speakers, but as a teaching tool it might be too complicated.

Video 20 claims to address intonation, but instead addresses two issues very closely related to intonation: sentence stress and rhythm. This is shown by the main focus of the video being on what words to stress in a sentence. For example, the speaker explains that in the sentence “I want a coffee” the words “I” and “a” are not important to the meaning of the sentence and are thus reduced to schwa-sounds in natural speech. This type of phenomenon is what results in the rhythm of English (Lane, 2010: 45), as the grammatical and function words are unstressed. However, since the YouTube search was done specifically on intonation, the learner might get confused on the terms.

#### 5.1.6 Connected speech

Connected speech was explored through the use of catenation, intrusion, elision, assimilation and geminates, although only video 21 explicitly mentioned all five of these terms, giving examples in text and voice. Videos 22 and 24 also mentioned all but geminates. Video 23 mentioned intrusion, catenation and R-linking which was not in other videos. We also notified that the video 23 used song lyrics, quotes and actors in the examples. In that video R-linking was used as an example and it was not mentioned in any of the other videos. Video 25 did not mention the technical terms related to connected speech but explored them through examples of rhythm and flow as well as pitch. Most of the videos in this category, except video 25, used the IPA symbols to represent sounds that they were highlighting on screen. Only video 25 did not have a person in front of camera explaining the lesson. Instead, video 25 used text and animations, as well as visual pictures to demonstrate what was being said. Video 23 stood out as it demonstrated a more social media

approach, with the speaker her free-time, instead of being exclusively a teaching video. All of the videos had visual aids to demonstrate the important parts, using text, color and animation.

Video 21 has a very interesting title, as it insinuates that the learners will learn five secrets to fast and native pronunciation. The speaker implies that these secrets may not have been learned in the viewers' English classes. These secrets were revealed to be catenation, that is also known as linking, intrusion, elision, assimilation and geminates. For each of these terms the speaker gives a few examples and rules. These sound simplifications go well together with previous literature and research. To minimize the effort of articulation, adjacent sounds are often made similar to each other, which is called assimilation (Dalton & Seidlhofer, 1994: 27). The video also talks about elision, which means that it is produced by dropping a sound out. For example, in the video "next door" becomes "nexdoor" when spoken by using elision. More familiar term for elision could be deletion, or another alternative as presented by Celce-Murcia et al. (2010: 172), omission. Video 21 is very closely in line with basic connected speech rules, and also support the suggested methodology of teaching connected speech. For example, Celce-Murcia et al. (2010: 175) instruct that areas such as consonant-to-vowel linking, vowel-to-vowel linking, consonant assimilation and palatalization should be highlighted in teaching, as those occur in spoken English very frequently. All but palatalization occur in the rules given in the video - however, examples of assimilation also included palatalization, even though the term itself was not covered. In order, they would be, catenation, linking with insertion, geminates and assimilation (including palatalization). Thus, the video theoretically followed connected speech teaching suggestions well, even though the secrets were worded differently.

Video 22 begins from the very basics, explaining what consonants and vowels are. The point of the video is to teach how to link consonant sounds, consonants to vowels and two vowel sounds, as well as how to link words to speak more fluently in English. The video gives very specific examples of each. For example, the first part about consonant sounds, or geminates in previous video and literature, "look cool" linked together because of the shared /k:/. However, the video does not display the phoneme that they share and instead displays the regular letters connected with an underscore. If the learners are familiar with phonemes, displaying these phonemes might help them remember the rules better. Unlike other videos, this video also goes into exceptions to these rules, such as not linking /tʃ/ and /dʒ/ sounds, such as in the word "orange juice". These exceptions can provide learners a more in-depth view of the topic. Overall, the video mostly reviews catenation features of connected speech, which is why it can go more in depth without getting too tedious to follow.

Video 23 explored connected speech through R-linking. The video does not explain what R-linking actually is, but it falls under linking methods. However, it should be noted that R-linking is not needed in American English, since in American English the postvocalic /r/ is pronounced nonetheless (Dalton & Seidlhofer (1994: 123). This could be especially important for students that want to learn British English specifically, but for students that are focusing on American English, R-linking is most likely not worth practicing. We already established the importance of catenation, but the video also emphasizes intrusion, which is one of the linking methods. When using intrusion, two vowels are met with each other and the consonant glides, resulting /j/, /w/, or /r/ to be inserted (Dalton & Seidlhofer (1994: 123). However, Dalton & Seidlhofer (1994: 123) argue, that topics such topics as /j/ and /w/ -linking (intrusion in the video) are better left to more advanced students, and instead focus should be solely on catenation.

Video 24 also has a title including that it contains four secrets to speaking quickly and fluently by using connected speech. In this video, the secrets are catenation, intrusion, elision and assimilation. Catenation is explained similarly to other videos. Intrusion includes some tips that were not present in other videos, such as when pronouncing the sentences, learners are encouraged to pay attention to the form their mouth make: when their mouth is wide /j/ sound is often needed for intrusion between the words in sentence and when their mouths are round, /w/ sound is needed instead. For example, “we /j/ all play /j/ out” or “go /w/ out to /w/ open”. These kind of tips of using one’s own body to emphasize the lesson were already previously deemed important and they provide good mnemonics for learners, especially in harder parts of connected speech-like intrusion, as was pointed out earlier by Dalton & Seidlhofer. For elision, the exact same example was used as in video 21, “next door”, as was largely the case with assimilation. The video was also largely in line with the suggested methods of teaching connected speech as was pointed out earlier.

Video 25 only explains catenation. The example sentence, “Take a bit out of an apple” is pronounced and repeated. It is explained that the words are connected rather than taking a breath between each word. The video includes slides with animations to emphasize the content. The consonant sounds are shown to be pushed to the start of the next word to demonstrate it. It is also instructed to pay attention to the words that carry meaning and how they are stressed, as well as pitch which was higher to the most important words of the sentence. Overall, the video seemed to show a very simplified view of connected speech. The videos in this category seemed to be similar, using the similar terminology, although emphasizing different aspects of connected speech. All of them were straightforward in their teaching, giving examples of each term. The videos could have gone deeper, such as in the case of assimilation it could have been divided into further categories,



such as progressive or regressive assimilation with further examples. However, these details might only produce more confusion to the already difficult topic and especially for YouTube video content, it might be best to stay at basics.

## 5.2 Applications

### 5.2.1 KEPHAM English pronunciation

One of the applications chosen for the study was KEPHAM's English Pronunciation (KEPHAM). This application concentrates mainly on teaching and producing individual sounds and utilizes an IPA-chart for 44 unique sounds in total.

Four main topics in the application for practice were identified as; short (such as, /ɪ/) and long vowels (such as, /i:/), double vowels (such as, /ɪə/), voiceless (such as, /p/) and voiced consonant (such as, /z/) and other consonants (such as, /m/). The topics could be practiced with single word, phrase or sentence examples. The application also had practice options for learning to count syllables for pronunciation and how to correctly stress a word. The application's main sections were repetition tasks after a sound clue provided by the application. There are also detailed instructions for how to physically produce the different sounds. The division does give a easier to approach way of describing sounds than having all of them as a table together, and the descriptions offered on how to pronounce each sound are useful. However, the application could offer more accurate information on the description of phonemes. For instance, /k/ is a voiceless consonant as stated in the application, but learners might also benefit from knowing that it is more precisely a voiceless velar plosive. Knowing these elements gives the learner more information on recognizing other sounds that are voiceless, velar or plosives as well, and highlight the differences and similarities (McMahon, 2016: 23). As the application states, it is close to the /g/ sound, which is a voiced velar stop, giving more information on why the /g/ is similar to the learner. The application also utilizes the HVPT method. The learner can distinguish examples spoken with male and female voices, although the same recording is always played using the same voice. At least three different female voices can be made out, but for male voices there seems to be only one example. Furthermore, the examples can be heard by using American and British English which highlights their differences. However, the sentence sections are all read by the same voice.

The application offers listening and imitation processes (table 1) and gives direct feedback when the user utilizes the record function. However, since the application uses an algorithm to analyze if the pronunciation is correct or not, a learner discretion is advised. Furthermore, the application has links to YouTube videos, providing further instruction for the pronunciation. For actual pronunciation testing, the application has a record function, which allows one to record their own pronunciation of a word, to then compare it to the example provided by that application. When tapping the 'practice' button, the application gives six practice options for the sound of their choosing. This is most likely intended to do after listening the words and phrases. For first option, the user sees a word written in phonetic, which they have to type out using their regular keyboard. As stated before by Lintunen (2005), transcription could make learners more aware of their use of language, and therefore learn more effectively. The second task is to record their voice pronouncing a word not written in phonetics. The application gives maximum of three stars depending on how well the word is pronounced. Third task is to select a different sound between four words. A visual sound cue is heard if the task is correct or not, giving immediate feedback to the learner. The same is true for task 4, which asks the learner to choose words from six different options that include the sound they are practicing at that time. Fifth task is similar to first task, but in this case, the user is asked to type out the phonetic of the word by using the phonetic keyboard offered by the application. The last task is to type out phonetics of a word pronounced by the application, using the phonetic keyboard offered by the application.

The application also included nine rules for English stress and provided a number of different examples for practicing them. The rules in the application were more extensive than the ones found in videos regarding word stress. The stress was marked by red and green text in the phonetic transcriptions, marking the places for primary stress in red and unstressed syllables in green. The application also pronounced each of these words when the speaker icon was tapped. The rules are rather extensive, so they are gathered in the table below.

Table 2. Stress rules in KEPHAM English Pronunciation application

<b>Rule no.</b>	<b>Rule</b>	<b>Examples</b>
1.	Two syllable nouns and adjectives	Nouns, adjectives, adverbs
2.	Two-syllable verbs and prepositions	Verbs, prepositions
3.	Stress on the second from the end syllable	Suffixes: ion, ial, ien, iu, ual, ic, ish, ity, ify, logy, pathy, eous, graphy, lysis

4.	Stress on the third from the end syllable	Suffixes: graph, ise/ize, ator, ute/ude
5.	Stress on the last syllable	Suffixes: ee, eer, ese, ique, ain, ental, aire, ival, itis
6.	Three-syllable words	Ending with Y, a consonant + OR, two consonants + OR, a consonant + Age, two consonants + Age
7.	Four-syllable words	Ending with ATOR, ending with ARY
8.	Sentence Stress (P1)	Nouns, main verbs, adjectives, adverbs, Wh questions, negative auxiliary verb, determiners (this, those, these...)
9.	Sentence stress (P2)	Articles (a, an, the), prepositions (for, on, in, to, at, from, with, of...), pronouns (I, her, him, she...), conjunctions (but, as, and, or, than, so...), auxiliary verbs (has, can, have...)

These rules were partially in correspondence with both, videos and literature about word stress. It can be noticed that there is similarity with the general rules of teaching stress, such as two-syllable nouns and adjectives having their stress on the last syllable and two-syllable verbs on the first syllable. However, when it comes to suffixation, the rules begin to vary between the literature and applications. The rules within the application are not as accurate as in the literature. For example, when examining rule 3 of the category within the app, the stress can arguably be wherever before the presented suffixes, The suffix in the app “-ion” sees the stress shift to penult (second last syllable from the end), and “-ify” to the antepenult (third last syllable from the end) (Celce-Murcia et al., 2010: 191), which are all presented under the rule 3 tab even though the rule says it should only consist of words with the stress on second from the end syllable. However, rule 4 is supposed to present words that have the stress on antepenult. Thus, the application does not present this information in a consistent manner. This can get confusing to the learner very fast, especially if the learner does not open the tabs containing example words and just checks the titles of the drop list.

As the application includes examples and practices for learning the topics it presents, it provides an interactive rule and goal-oriented environment, that can motivate learners (Peterson, 2013). It offers a different and game-like environment for completing the same type of shadowing and imitation practices that as are used in normal teaching practices in classroom (Derwing & Munro, 2015). However, it is questionable whether the approval of the application to reward achieving the goals is

enough to keep the learner interested, despite the application introducing an alternative environment from the classroom. As the application also records the audio examples uttered by the user, it can develop the self-regulatory skills pointed out by Kucirkova & Falloon (2017) and Ma et al. (2011). Thus, the application offers the possibility for limited self-motivated learning independent from the teacher, as some feedback and self-monitoring methods are provided for the user. Here, it is also important to note, that using the application happens while using a device with possible access to internet. As Kucirkova & Falloon (2017: 108) show, this gives the learner access to the internet as an effective resource tool, providing easy to access and find - information and material. This point can mitigate the flaws and limitations of the application.

Critically reviewing the application's functionality, we decided that the application was quite versatile in terms of offering practice and instructions for the topics it introduced, but was difficult to use and navigate around, as some of the features seemed almost hidden behind tabs and links. However, despite being non-intuitive to use, it served well as on-demand instructional application, where the user could find instructions for a very specific topic in pronunciation learning. KEPHAM's application provides a deep and detailed description of its topics. It could serve better as an aid when a specific issue occurs, rather than the main tool for pronunciation learning or teaching, the former being the impression for what it was made to be. However, the application has some performance issues that make it hard to use at times. For example, the application has some delay related issues, which makes completing certain tasks overall use of the application difficult. It also seems to change pages and tasks unintentionally when using the 'practice' function. There is, in addition, an overwhelming amount of advertisements, but this is expected from a free application.

### 5.2.2 TFlat English Pronunciation

TFlat English Pronunciation teaches English through the use of phonetic alphabets. The user can choose which phonetic alphabet they want to concentrate on. The alphabets are divided into six categories, similar to KEPHAM application. These categories are short vowels, long vowels, double vowel sounds, voiced consonants, voiceless consonants and other consonants. When tapping one of the phonetic alphabets, the application leads to another page. This page displays a simple picture of where the sound is approximately formed inside the mouth. One can also tap the loudspeaker icon next to the picture to hear an example of the sound without it being attached to any words. Under the picture there are instructions on how to produce the sound and some notes one should take into account when producing them. Below the info section, there are words with which one can practice

the sound chosen. For instance, the words that are linked with /θ/ are “thief”, “thin”, “teeth” etc. up to 16 different words. Each word has the whole word in phonetics next to the text, such as “/θi:f/” for “thief”.

The practice happens by listening to the word by using the loudspeaker icon in the same box where the word is. The user hears how to pronounce the word. After hearing the word, the user can then tap the microphone icon, which allows the user to record their own sound. The app then rates the pronounced word between 1 and 5 stars depending on how the app deems its correctness. As noted before in the table 1, listening and imitation practices can be used to learn pronunciation. However, as was the case with KEPHAM, the application uses an algorithm to decide whether the pronunciation was good and may not always analyze the sample with perfect accuracy. It also seems that the whole word does not need to be pronounced correctly as long as the sound at hand is pronounced correctly.

The application also offers other features, such as reading, writing and listening, which all are aimed to teach the IPA and phonetic transcription. In the reading section, the user gets a word written in IPA and they have to type out the written version in a box below. The writing tasks work the same way as reading, but the other way around; they have to type out the phonetic version by using the provided IPA keyboard. Similarly, to KEPHAM’s tasks and as suggested by Lintunen (2005) phonemic transcription could make learners more aware of their use of language, and therefore benefit from it in terms of pronunciation learning. Especially if the learners are already accustomed to orthography. The listening part produces a word by sound, and the user has to type out the word they just heard by using the IPA keyboard provided by the application.

The topics of the application were divided into categories, which provided an easier to understand way of describing the phonetic alphabet. If the whole IPA chart was without these categories it might seem intimidating to the learner. The pictures could be a bit clearer in showing exactly how the sound is made, but the pronunciation tips make it easy to follow. Similar to KEPHAM’s application, TFlat could offer a more precise way of dealing with the sounds, giving them the correct terms that are used, such as voiceless velar plosive for the sound /k/. It is unclear how the application recognizes the correctly pronounced words and what algorithms the rating system uses. This could lead into misunderstandings on how certain words are pronounced, even though the specific sound that is being practiced is correct. The other tasks provided by the application seem to be featured less prominently and they act as variations from the main tasks. For example, in the listening tasks the audio quality can sometimes make it hard to hear what the application wants the learner to type, which might frustrate the learner.

As the TFlat's application functioned and was structured very similarly to the KEPHAM's application, a comparison in terms of application features is relevant. TFlat provides a similar imitation and shadowing practice collection with a scoring system to provide goals for the learner. It also offered the possibility for recording and listening to the user's own utterances, increasing its effectiveness as a self-monitoring tool (Ma et al., 2011). The other notable difference compared to KEPHAM, it had a more visual progress marking method in addition to the scoring method after each individual practice. As the user completes practices under a specific sound, a meter begins filling up. This gives the learner clear goal for filling the meter completely, and can produce a sense of achievement, which is important for motivation in games (Peterson, 2013). Unlike KEPHAM, TFlat had no clear presentation of English stress rules or any other element of pronunciation, except the instructions for producing specific sounds. As the application provides otherwise virtually the same features as KEPHAM's, it can also be described as a tool that provides the learner with an independent and interactive learning environment outside of the classroom.

### 5.2.3 Thomson's English Accent Coach

Similarly, to the other two applications, the Thomson's English Accent Coach (henceforth, TEAC) concentrates on specific topics, vowels and consonants. The topics are not explored further, the tasks/practices instead highlight one vowel or consonant at a time, with the help of the IPA symbols. In addition to the tasks, the application has "learn" sections for instructions on pronouncing vowels or consonants. The main idea of the application seems to be to teach the learner how to recognize and use the IPA symbols for pronunciation. The vowel and consonant symbols are placed within the practices in accordance to how easily they are confused with each other. The user clicks the syllables or words that match with the sound they hear in the tasks. In addition to individual sound-based tasks, the application also has "Echo" game to memorize IPA symbols, that plays four different sounds in alternating sequence, and the user must remember the sequence.

Beth Sheppard (2016) provides some insight and evaluation for Thomson's English Accent Coach. He also briefly explores some of the strengths of the application, such as the use of multiple speakers' voices in the application. At least when looking at the early studies, using multiple speakers' voices in phonetic contexts have been found to improve phonetic training and memorization of the sounds significantly more than using the same speaker voice in all contexts (Thomson, 2012: 1235). When playing with the application, users also receive immediate feedback on their success. If the user presses a wrong phonetic alphabet, the application produces a sound

indicating that it was wrong. The phonetic alphabet also turns red during it to more clearly point out the error. Similarly, a sound indicating getting the answer correct also plays when the user chooses the correct option. HVPT studies also point out, that more learning seems to occur when the users are given this immediate feedback (Thomson, 2012: 1235). In fact, the research on HVPT is exactly what resulted in the making of this application for both, learning purposes and using it for further studies in the computer assisted pronunciation teaching (CAPT) field (Thomson, 2015: 1254)

The application also has a well-functioning interface and the includes all of ten vowels at once in default vowel game practices. The benefits of an engaging interface, especially paired with immediate feedback has been found to better orient the learner towards necessary phonetic information (Thomson, 2012: 1251). This reduces learners' distortion in the perception of vowels (Logan, Lively & Pisoni, 1991, as cited by Sheppard, 2016). Sheppard also mentions its unavailability for Android devices and the applications incapability of reacting to the user's proficiency in different areas as notable weaknesses.

TEAC offers many opportunities for learning pronunciation for specific sounds. However, it does not provide a similar opportunity for voice recording and evaluating these recordings in regard of quality of pronunciation. This greatly decreases its effectiveness as a developing environment for self-regulatory and self-monitoring (Kucirkova & Falloon, 2017; Ma et al. 2011). However, as Sheppard (2016) points out, the application offers a tool for developing the learner's self-monitoring competence via its tasks. TEAC provides the user with progress tracking features, that will highlight problem areas for one's pronunciation during the use of the application. It offers progress reports for each vowel or consonant game played individually, and overall progress reports from the time spent doing the application's practices. This is clear reflection of rule and goal-oriented game-like features that have a positive effect on the user's sense of achievement and motivation (Peterson, 2013). The application also offers a chance to use the color vowel chart, which transforms the vowels into colors and thus provides a multisensory tool that can help the learner perceive and identify the used sounds in a different manner. Furthermore, TEAC is made as a game-like application. This is most explicitly shown with the vowel chart practice, which is basically a game with very simple rules. Another example of it is how the application makes the learner acknowledge their errors by forcing them to click on the correct symbol when an error occurs, in order to progress (Sheppard, 2016).

## 6. Discussion and conclusion

Our research aims were to find out what kind of English pronunciation teaching materials are available on YouTube and as mobile applications, as well as how suitable they are for learning English pronunciation. According to the background material we found, there were a number of studies regarding use of technology in language teaching, but studies and sources reviewing YouTube materials on this topic were nonexistent.

### 6.1. Videos

Considering the possibilities for innovative and different approaches made possible by YouTube, we were slightly disappointed by how many of the videos failed to utilize the YouTube platform and concept of video lessons to be more interesting. This was shown by how the speakers had opted to use props and aids, such as whiteboard in their lessons, instead of using enhanced visual and auditory methods, such as animations to do something different than classroom-based teaching. The five categories for YouTube videos were word stress, sentence stress, rhythm, intonation and connected speech. The videos were divided into these themes according to the prevalence of a specific theme in a video, for example if word stress was the main topic of the video that video would be categorized under the word stress theme, even if the actual content in the video turned out to not correctly reflect the title. Several videos were similarly constructed and utilized tools, such as whiteboards and animations while teaching, for example videos 7, 8 and 9 in sentence stress - category. Furthermore, many videos also clarified themselves via post editing. The differences between the videos mostly had to do with overall quality, methods used, and theoretical background presented. Additionally, one clear difference in the videos was the use of visual aids, such as highlighting text and animations.

As shown in the data review and analysis parts, some of the videos had different terms and explanations for the same phenomenon in pronunciation, such as the “partial fall” in video 17 or the working backwards - method in video 3. This raises the question whether it is more confusing to use own terminology for the topics, or if the importance is not great enough to affect usefulness for the learner. On one hand, it could be argued that as long as the primary information is correct, the



video functions as a valuable learning aid. On the other hand, learners who would want to explore the specific topics more deeply, may find it confusing when they cannot find information on the issue with the same terms the videos use. In addition, many of the videos also emphasized sounding native-like as an important goal for their viewers, a view which is not supported by our literature. For example, Lane (2010) claims that native-like speech is an unrealistic goal, and should be replaced with more achievable goals, such as intelligibility and comprehensibility. It should also be noted that the videos offered almost exclusively one type of practice method; repeating after the speaker.

Another concern in many of the videos is whether the information they offer is relevant, correct and properly presented. This was shown, for example, with videos that were only barely or not at all reflecting the theoretical background that we had for the study, as was the case when video 20 mistook one topic for another, when the video was actually addressing sentence stress, despite the video title having intonation in it. Some of the videos also tended to deal with a large amount of information, for example video 1, that could be confusing for the learner to deal with all at once. However, as mentioned before in the study, the videos should not be thought of as the main learning source, but instead of extra supporting sources for the teachers or as sources for independent learning from the learner's point of view. We suggest that reliable sources are needed to balance the unpredictability of quality and reliability in YouTube videos, stemming from the fact that anyone could make them.

For example: as per literature, teachers are encouraged to teach how native speakers highlight stressed syllables, such as by using pitch, volume and length. Focus is also directed on producing unstressed syllables that is often achieved by using vowel reduction (Celce-Murcia et al., 2010: 198). Naturally, the students also need to know what the main levels of stress are. The selected videos seemed to disregard vowel reduction and multiple levels of stress altogether. Furthermore, none of them specifically taught the three main levels of stress. This was strange, because the videos did not seem to assume that the learners were aware of these concepts beforehand and started lessons from clean slates. This type of discrepancy supports the expectation that the videos could not be utilized fully in educational environment. However, the videos succeeded in addressing the problem of which syllable to place the stress on, which is one of the challenging issues for the students considering this topic (Lane, 2010: 18).

There are many suggested ways of visualizing the use of word stress, such as using capital letters, accent marks, bubbles or underlines (Celce-Murcia et al., 2010: 199). Most prevalent of these methods were the use of capital letters or the use of different colors to mark the stressed syllables. It

would be beneficial if the methods were consistent with each other as to not confuse the learner. If the materials were provided by ESL teachers, this could be achieved by searching for materials with similarly illustrated stress patterns to be presented for the students. However, it could also be helpful for the students to learn these different lexical patterns to help them understand all of the possible methods used in the online materials, such as these videos. Individual preferences also play a part in this matter. It could also be beneficial to teach how stress is displayed in dictionaries and other similar contexts, as was done in case of videos 3 and 4. This would offer great self-study possibilities. When it comes to sentence stress, the videos have great emphasis on teaching how the meaning of the sentence changes if different words are being stressed. There was a discrepancy between our selected literature and our selected videos in this matter. In literature, sentence stress is explored through rhythmic patterns shared by words and sentences (Celce-Murcia et al., 2010: 209) or included into rhythm category altogether (Lane, 2010: 45). It is important for the learners to realize how this placed emphasis on certain words in the sentence can completely change its meaning.

While we expected that YouTube videos might not always reach the same level of solidity as official information and teaching material in terms of reliable information, we did expect that they could find creative and easy to understand ways to present and teach their topics. This assumption was made based on the chosen videos being the first ones to appear with the specific keywords, which in YouTube indicates higher visibility to the viewers. Some of the videos succeeded in this, in the sense that they utilized features, such as animations, to highlight important parts of their lessons to the viewers. Many auditory and kinesthetic modalities are also used, such as clapping to emphasize the stress patterns, which can help the learner to place the stress in a word better. Another aspect that requires attention is the presenters themselves. Even though most of the videos were presented in a classical classroom manner where the teacher presents the topic at hand, some videos differed from this route by being more engaging with the audience than others.

This draws attention to how humor, for example, is utilized in videos. We suggest that the use of humor in teaching makes the lesson more interesting and holds the viewer's attention. Equally, if not even more prevalent than the content of the lessons themselves were how the lesson was taught by the presenters. Whereas some of the videos went straight to the point, some included humor and encouragement for interactions with the speaker. For example, examples drawn from popular culture could be more easily remembered than typical textbook examples. The use of snippets of songs, such as in the video 20 where the pronunciation of the word "police" was demonstrated by the song "Fuck tha Police" by N.W.A or displaying flashing lights and the sound of sirens at later

part when addressing strong and weak forms in the sentence “The police are coming”. YouTube videos enable these kind additional methods to teaching, which were almost entirely absent in most of the reviewed videos.

YouTube as a platform allows the content creators to address one issue at a time, as the posted videos are accessible at all time, and complex subjects can be divided into many videos. This was specifically illustrated by the speaker in video 19, where she suggested the viewers to go see her other videos for information on the closely related topics. Before, we mentioned examples of videos not reflecting our background literature or reflecting it loosely. However, there were also multiple cases where the videos and literature reflected one another well, such as the videos in connected speech category, showing that YouTube videos have merits as well. In addition, the search seemed to bring up multiple videos that were ran by professionals, such as accent trainer (video 18), teachers (video 11) and university channels (video 3), which increased their credibility.

Obviously, the videos succeed in some aspects and fail in other. However, the important question here is whether the good aspects found in the video materials outweigh the poor aspects. When approaching this question, we looked at it from the perspective of using the materials to support teachers and offer chances for independent/extra learning for the learners. As mentioned previously, this suggestion is based on the unpredictability of the YouTube material that the study has showcased. In an ideal situation, teaching can be realized through multiple modalities, such as auditory, visual, kinesthetic etc. Sometimes these modalities can even be invoked together (Odisho, 2003: 95). YouTube offers a platform where this can be done.

## 6.2. Applications

A trend in applications was prominent utilization and teaching of IPA-symbols. As shown in the data, it was one of the main features in the applications, along with voice recognition supported repetition and imitation practices. Because these applications do not offer much possibility for the communicative approach of language learning, they could be reviewed by using other methods provided by literature. As seen in Table 1, listening and imitation practices were one of the suggested methods for pronunciation teaching methods, although not quite as popular recently because of the prevalence in communicative approaches. The listening and imitation method also gives direct feedback when it comes to applications such as TFlat and KEPHAM, further benefiting the learning process. Furthermore, these two applications also provide phonetic training, which is

realized through articulatory descriptions, diagrams and the IPA in teaching. The applications provide all of these jointly with the pronunciation tasks, offering great opportunities for learning.

The applications also provide visual aids, a method that was demonstrated in Table 1 to be used with the teaching. Especially KEPHAM's application provided a lot of visualization - not only drawn charts of the vocal tract and mouth but also videos with front and side view. Recording learners' production is also a big part of KEPHAM and TFlat's functionality. It gives immediate feedback using a rating system and enables self-assessment. However, because the algorithm is not entirely trustworthy, it should be taken as such, a method to hear and compare the learners' own recorded voice to examples given by the application. Finally, the possibility to focus on stress shifts can also be seen with KEPHAM application. However, it does not offer context for these words in many cases and instead leaves them as separate entities. This is something the applications could expand on. Referring back to the table 1, the lack in use of minimal pair drills is quite questionable, as they would clearly demonstrate how the sounds can be just slightly distinctive. The applications also provided a wide array of information for the topics they addressed, which was in line with the literature; for example, phonetic sounds and their display was described similarly as shown in the descriptions of how and where these sounds are made in sections of the mouth (e.g. Celce-Murcia et al., 2010: 116). However, these mapping pictures could be displayed more clearly and some of them are missing the exact location of the sound, but they do have a description of where they are made.

HVPT methods, such as was on focus especially in Thomson's English Accent Coach, but to some extent on the other two applications as well and their utility of computer assisted pronunciation training could be very beneficial to the learner, as already suggested by studies (Thomson, 2012). As more emphasis is placed on different techniques of pronunciation teaching in digital environments and applications, studies such as this can produce even more powerful and innovative applications. Especially since techniques such as this cannot be easily produced in classroom teaching they truly stand out in the field of pronunciation teaching and language learning as individual learning tools.

As we expected, the applications were limited in this sense, as they offered only a certain type of activity. However, the applications offered different realizations for the repetition type task, where writing or memory games were utilized for learning the IPA-symbols, in addition to learning the pronunciation of specific words or phrases. For example, the Thomson's English Accent Coach included a color-based memory game to memorize different IPA-symbols, where the sounds could be heard while the symbol was pressed. KEPHAM and TFlat offered transcription tasks, that were engaging and more innovative than we first expected to find. It would be very difficult to

incorporate these tasks easily in classroom teaching as the keyboard of phonetic symbols was an unique take on learning. As Lintunen (2005) showed in his study on university students, the phonemic transcription can have a clear positive impact on a learner's pronunciation skill development. However, as Lintunen's study included only advanced English learners with considerable knowledge and skill in the language, it may not reflect the impact on less proficient learners' development. Further studies on this particular matter in conjunction with the e-materials that use the phonetics would bring valuable insight into the issue.

The applications had some features that could be compared to game-like concepts, which have been shown to have a positive effect on learning (Peterson, 2013; Ma et al., 2011), and for developing learner's self-monitoring skills. Especially Thomson's English Accent Coach provided a good example of this with the vowel chart game. We suggest that this type of interactive practice and example combination concept is what applications should be able to do. As the applications also provided the progress tracking features, they further provided more motivating features that can be compared to ones observed in games. KEPHAM also provided an option for the user to place reminders to practice specific topics. As explored by Kucirkova & Falloon (2017), it is important to develop self-regulatory abilities in learner, in order for them to be able to learn independently and reflect on their own learning goals with the help of materials not provided by a teacher. The applications offer opportunities for autonomous and more personalized learning. However, not only should the applications be used as materials for learning independently, they should be integrated into the practices in teaching in schools (Kucirkova & Falloon, 2017: 164).

As was briefly pointed out in the analysis for KEPHAM's application, it is also important to point out that when using these applications, the user is also likely to have access to internet as a source for information and more material. This was listed as one of the main reasons children find devices useful in learning according to a study (Kucirkova & Falloon, 2017: 108), along with access and portability, engagement and comparison to other tools. The children in that particular study compared mobile devices to computers and book, describing them as easier to use than computers and much more versatile than textbooks.

As criticism for the applications, we explored their functionality, the user interface and how well they presented the topics. The emphasis on shadowing and imitation practices has been criticized as being boring and demotivating for students (Derwing & Munro, 2015:123), which is definitely a downside of the applications. However, as mentioned before, this game-like nature of these applications might help the learners to remain interested in these tasks, especially since they offer progress bars for these different tasks.

KEPHAM differed from the two other applications in a sense of also putting emphasis on word stress. For an application that otherwise focuses on segmentals, such as the phonetic alphabet, it was interesting to see one suprasegmental feature in the application. For the most part, this feature offered some further insight into pronouncing the words that had already been emphasized via the main tasks. This kind of approach can be more user-friendly than the videos. The word stress section is always readily available, and the user can open it and go through the rules presented in the application to check the correct form of pronunciation and stress. The section contains more information than what the videos offered, and because of this availability it could be viewed beneficial for the application. However, some criticism could be directed towards the presentation of this section, as it was not very straightforward with its content. Furthermore, it is questionable if this sort of solid theoretical content is relevant to the nature of an application and how useful it is to have the content available in a mobile device, even if it could be accessed as an offline source.

The user interface of the applications was also confusing sometimes. For example, KEPHAM's application had the practice function as a small yellow button in the top right corner of the screen once a specific voice was selected to be learned. In addition, some of the features of the application had inconsistencies in accessibility and functionality. For example, the *reading* and *writing* features in TFlat's application could no longer be accessed after a certain amount of time using it. This could be major critique regarding the application's integrity and credibility as learning material. These problems might occur depending on the user's device or the updates - or the lack of updates from the developer's side. We suggest that the applications should be addressed similarly to the videos, and be considered supports for teaching and independent learning, as they are geared for providing information and material for specific topics and offering opportunities for some practicing them.

### 6.3. Conclusion

We have explored YouTube videos and mobile applications for English pronunciation teaching. Our findings are that there is a wide variety of pronunciation learning materials available on YouTube in the area of teaching suprasegmental features of pronunciation. The applications also provide useful learning material for specific issues. However, we argue that due to the general unreliability of them, they should not be used as the main source for learning, but instead as supportive elements parallel to reliable material, such as official teaching materials obtained from classroom. They offer opportunities for independent learning without concern for time and place, and also authentic materials that the classroom may lack. However, this point must be paired with

the development of self-regulatory and self-monitoring skills of the learners, in addition to the integration into classroom practices. This will provide more meaning to the use of applications and also provide some safety, in terms of the credibility of the material available to the learners. They also have potential to offer different modalities and methodologies for teaching, as well as offer interesting or humorous approaches for language and pronunciation learning. Since Finnish schools are mostly focused on textbook materials and tasks even when considering teaching pronunciation, suprasegmental elements could be incorporated in teaching via the use of applications and YouTube materials. However, their suitability should be reviewed by the teacher beforehand.

Some limitations for the study should be noted. The amount of data, and especially the lack of quantitative data, was limited. This resulted in the findings and analysis of the data becoming partly speculative, instead of solid and affirmative. The findings on videos reflected this issue more clearly. The categorization of the videos resulted in a minor discrepancy, in terms of whether the videos selected would actually be the ones the learners would find independently. The algorithms used by YouTube's search and video display functions play a part on what kind of videos the learners would be exposed to. However, generally the search function does bring up more popular videos when searched with key terms of pronunciation and each suprasegmental element. It might be more probable that the learners would simply search for materials using more general keywords such as "English pronunciation", instead of specific elements such as "intonation".

Thus, further studies could include more quantitative approach, to map trends and phenomenon in the data more accurately and provide a deeper review of materials, and therefore provide a more affirmative view of their usefulness. An interview-based study of learners using the YouTube materials would also provide interesting angle to the issue, as it could then be pointed out how the learners view these materials.

## REFERENCES

## Primary sources

NO.	Channel	Theme	Name (URL)
1	Jeff Roy	Word stress	<a href="https://www.youtube.com/watch?v=65AgbiwQ6ko">https://www.youtube.com/watch?v=65AgbiwQ6ko</a>
2	Your English Web: Weekly English video lessons	Word stress	<a href="https://www.youtube.com/watch?v=6pvWiAh528">https://www.youtube.com/watch?v=6pvWiAh528</a>
3	Academic Skills, The University of Melbourne	Word stress	<a href="https://www.youtube.com/watch?v=SNmEeNmIxNI">https://www.youtube.com/watch?v=SNmEeNmIxNI</a>
4	Rachel's English	Word stress	<a href="https://www.youtube.com/watch?v=08qBN29mIBs">https://www.youtube.com/watch?v=08qBN29mIBs</a>
5	Billie Haase	Word stress	<a href="https://www.youtube.com/watch?v=bX-_YSDM7ic">https://www.youtube.com/watch?v=bX-_YSDM7ic</a>
6	Learn English with Emma [engVid]	Sentence Stress	<a href="https://www.youtube.com/watch?v=L4u0vINDZ_g">https://www.youtube.com/watch?v=L4u0vINDZ_g</a>
7	StGeorgeInternation	Sentence Stress	<a href="https://www.youtube.com/watch?v=-0G_yZfXJUQ">https://www.youtube.com/watch?v=-0G_yZfXJUQ</a>
8	Learn English with Benjamin [engVid]	Sentence Stress	<a href="https://www.youtube.com/watch?v=xcjrgkGhB4">https://www.youtube.com/watch?v=xcjrgkGhB4</a>
9	English Lessons with Adam - Learn English [engVid]	Sentence Stress	<a href="https://www.youtube.com/watch?v=4Iqrm82LED4">https://www.youtube.com/watch?v=4Iqrm82LED4</a>
10	eltraining	Sentence Stress	<a href="https://www.youtube.com/watch?v=btz0_MbKyaw">https://www.youtube.com/watch?v=btz0_MbKyaw</a>
11	QuickESL	Rhythm	<a href="https://www.youtube.com/watch?v=cLdyLqxQ9qU">https://www.youtube.com/watch?v=cLdyLqxQ9qU</a>
12	Jill Diamond	Rhythm	<a href="https://www.youtube.com/watch?v=Neskv8MqCXc">https://www.youtube.com/watch?v=Neskv8MqCXc</a>
13	ElementalEnglish	Rhythm	<a href="https://www.youtube.com/watch?v=OTZV3zHohdc">https://www.youtube.com/watch?v=OTZV3zHohdc</a>



14	English Jade - Learn English (engVid)	Rhythm	<a href="https://www.youtube.com/watch?v=8XVeMLYiNM0">https://www.youtube.com/watch?v=8XVeMLYiNM0</a>
15	Pronunciation Pro	Rhythm	<a href="https://www.youtube.com/watch?v=G3NB19i949E">https://www.youtube.com/watch?v=G3NB19i949E</a>
16	JenniferESL	Intonation	<a href="https://www.youtube.com/watch?v=BRLG2FAFR6w">https://www.youtube.com/watch?v=BRLG2FAFR6w</a>
17	ElementalEnglish	Intonation	<a href="https://www.youtube.com/watch?v=tzh3Owutf5Y">https://www.youtube.com/watch?v=tzh3Owutf5Y</a>
18	Jill Diamond	Intonation	<a href="https://www.youtube.com/watch?v=3JCTywIDrVk">https://www.youtube.com/watch?v=3JCTywIDrVk</a>
19	Rachel's English	Intonation	<a href="https://www.youtube.com/watch?v=p8DJFNjZiIM">https://www.youtube.com/watch?v=p8DJFNjZiIM</a>
20	Learn English with Papa Teach Me	Intonation	<a href="https://www.youtube.com/watch?v=cu0FE31oAg0">https://www.youtube.com/watch?v=cu0FE31oAg0</a>
21	Gabby Wallace - Go Natural English	Connected speech	<a href="https://www.youtube.com/watch?v=F_lee1xlp4">https://www.youtube.com/watch?v=F_lee1xlp4</a>
22	Oxford Online English	Connected speech	<a href="https://www.youtube.com/watch?v=nbm3_QKU">https://www.youtube.com/watch?v=nbm3_QKU</a>
23	Learn English with Papa Teach Me	Connected speech	<a href="https://www.youtube.com/watch?v=-P-Bk-Sm9dY">https://www.youtube.com/watch?v=-P-Bk-Sm9dY</a>
24	English with Lucy	Connected speech	<a href="https://www.youtube.com/watch?v=ZZTV0Gwyo_Q">https://www.youtube.com/watch?v=ZZTV0Gwyo_Q</a>
25	ElementalEnglish	Connected speech	<a href="https://www.youtube.com/watch?v=4VnSAr9No6k&amp;t=72s">https://www.youtube.com/watch?v=4VnSAr9No6k&amp;t=72s</a>

KEPHAM. (2014). English Pronunciation (Version 5.5.5) [Mobile application software]. Retrieved from <http://play.google.com/store>

TFlat. (2014). Speak English Pronunciation (Version 5.5.2) [Mobile application software]. Retrieved from <http://play.google.com/store>

Thomson, R. (2012). English Accent Coach (Version 2.3) [online]

<https://www.englishaccentcoach.com/> (24 May, 2018)

### Secondary sources

Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. *Recent research in second language phonetics/phonology: Perception and production*. Barcelona: University of Barcelona.

[https://www.researchgate.net/profile/Joan\\_C\\_Mora/publication/229021613\\_Assessing\\_the\\_effects\\_of\\_phonetic\\_training\\_on\\_L2\\_sound\\_perception\\_and\\_production/links/551bc4dc0cf2fe6cbf75e64a/Assessing-the-effects-of-phonetic-training-on-L2-sound-perception-and-production.pdf](https://www.researchgate.net/profile/Joan_C_Mora/publication/229021613_Assessing_the_effects_of_phonetic_training_on_L2_sound_perception_and_production/links/551bc4dc0cf2fe6cbf75e64a/Assessing-the-effects-of-phonetic-training-on-L2-sound-perception-and-production.pdf)

Antonio, A., Tuffley, D. (2015). YouTube a valuable educational tool, not just cat videos. [online].

n. pag. <https://theconversation.com/youtube-a-valuable-educational-tool-not-just-cat-videos-34863>. (23 May, 2018)

Baker, A., Murphy, J. (2011), Knowledge Base of Pronunciation Teaching: Staking Out the Territory. *TESL Canada Journal* [online] 28(2), 29-50

<http://www.teslcanadajournal.ca/index.php/tesl/article/view/1071>

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation. A course book and reference guide*. Cambridge University Press.

Chan, W. M. (2011). *Media in foreign language teaching and learning*. Boston: De Gruyter Mouton.

Dalton, C. & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford University Press.

Derwing, T. M. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam, Netherlands ; Philadelphia, Pennsylvania: John Benjamins Publishing Company.

Heselwood, B. (2013). *Phonetic Transcription in Theory and Practice*. Edinburgh: Edinburgh University Press.

- Hişmanoğlu, M. (2006) Current Perspectives on Pronunciation Learning and Teaching. *Journal of Language and Linguistic Studies*. 10 (2), 101-110.  
<http://www.jlls.org/index.php/jlls/article/view/26>
- Jenkins, H., Purushotma, R., Weigel, M., Clinton, K. and Robinson, A. (2006). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. John D. and Catherine T. MacArthur Foundation, October 2006.
- Kucirkova, N., and Falloon, G. (2017). *Apps, Technology and Younger Learners: International Evidence for Teaching*. Abingdon, Oxon: Routledge.
- Lane, L. (2010). *Tips for Teaching Pronunciation: A Practical Approach*. Pearson Education, 10 Bank Street, White Plains, NY 10606
- Lintunen, P. (2005). *Phonemic Transcription and its Effect on Learning*. Paper presented at Phonetics Teaching and Learning conference, London, 1-4 (February, 2005)  
[http://www.academia.edu/12808547/Phonemic\\_Transcription\\_and\\_its\\_Effect\\_on\\_Learning](http://www.academia.edu/12808547/Phonemic_Transcription_and_its_Effect_on_Learning).  
 (25. May, 2018)
- Jain, L. C. (2002). *Virtual environments for teaching & learning*. Singapore; River Edge, NJ: World Scientific.
- Ma, M., Oikonomou, A. and Jain, L. (2011). *Serious Games and Edutainment Applications*. London: Springer London.
- McMahon, A. S. (2016). *An Introduction to English Phonology*. Edinburgh: EUP.
- Mompean, J. A. & Fouz-González, J. (2015). *Investigating English pronunciation: Trends and directions*. Houndmills, Basingstoke Hampshire; New York: Palgrave Macmillan.
- Munday, P. (2016).: The case for using DUOLINGO as part of the language classroom experience. RIED. Revista Iberoamericana de Educación a Distancia, 19(1), 83-101.  
[http://digitalcommons.sacredheart.edu/lang\\_fac/40/](http://digitalcommons.sacredheart.edu/lang_fac/40/)
- Odisho, E. Y. (2003). *Techniques of teaching pronunciation in ESL, bilingual & foreign language classes*. München: Lincom Europa.

- Peterson, M. (2013). *Computer games and language learning* (First edition.). New York, NY: Palgrave Macmillan.
- Reinders, H. (2012). *Digital games in language learning and teaching*. Houndmills, Basingstoke, Hampshire ; New York, NY: Palgrave Macmillan.
- Seidel, J., (1998) Qualitative Data Analysis, *The Ethnograph v4.0: A Users Guide, Appendix E*. Colorado Springs, Colorado: Qualis Research. [www.qualisresearch.com](http://www.qualisresearch.com)
- Sheppard, B. (2016). English Accent Coach. *The Electronic Journal for English as a Second Language*. 20(1) n. pg. <http://www.teslej.org/wordpress/issues/volume20/ej77/ej77m2/> (25. May, 2018)
- Smith, C. (2018). 17 Amazing Duolingo Facts and Statistics [online]. <https://expandedramblings.com/index.php/duolingo-facts-statistics/> (23 May, 2018)
- Statista. The Statistics Portal (2016). Mobile Gaming - Statistics & Facts. <https://www.statista.com/topics/1906/mobile-gaming/> (23 May, 2018)
- Tergujeff, E. (2013). *English pronunciation teaching in Finland*. Jyväskylä: University of Jyväskylä.
- Thomson, R. (2012). Improving L2 Listeners' Perception of English Vowels: A Computer-Mediated Approach. *Language Learning* 62: 1231-1258. doi:[10.1111/j.1467-9922.2012.00724.x](https://doi.org/10.1111/j.1467-9922.2012.00724.x)