Chapter 1

Introduction

Rapid technological changes have given rise to new methods and opportunities in language learning. In the past two decades, interest has been growing in the integration of Audiovisual Translation (AVT) with a communicative approach to language learning and teaching. AVT indicates the transfer of verbal language in audiovisual media and it is usually used as an umbrella term which refers to screen-translation, film translation, multimedia translation, or multimodal translation (Bollettieri, Di Giovanni, & Rossato, 2014; Chiaro, 2009; Perego, 2005). AVT modes¹ can be divided into two main groups: captioning² (written language transfer procedures) and revoicing (oral language transfer procedures). These language transfer procedures can take place between two languages (interlingual) or within the same language (intralingual). Both captioning and revoicing tasks can be used effectively in second language acquisition³.

Captioning allows learners to add the written translation (interlingual subtitling) or a condensed transcription (intralingual subtitling) of the original spoken language to an Audiovisual (AV) product. Captioning can foster listening,

---

¹ The terms mode and type are used as synonyms in AVT research (Gambier, 2003). Other terms like procedure, method, and technique are employed as synonyms of mode and type (Pérez-González, 2014).

² In AVT research and practice, the term subtitling is usually employed to indicate the AVT written type (Pérez González, 2009; Pérez-González, 2014). In this work, the term captioning is used – instead of subtitling – as hyponym to refer to “ANY thing done in terms of […] writing on the screen” as proposed in the ClipFlair theoretical framework (Zabalbeascoa, Sokoli, & Torres, 2012, p. 18). As stated by Sokoli (2018), “[a]dding written words includes processes such as standard interlingual subtitling but it can also refer to inserting captions for the deaf and hard-of-hearing, intertitles, annotations or speech bubbles” (p. 79).

³ The term second language acquisition is employed as an umbrella term encompassing both second and foreign language acquisition. Unless otherwise specified, the terms learning and acquisition are used interchangeably. The term second language (L2) generally refers to another language, acquired in some measure after one’s native language (L1), and this term can also indicate the study of a third or even fourth language. The defining condition is that an L1 has already been learned, in formal or informal educational settings. The term foreign language also indicates another language acquired to some measure, but it is differentiated from L2 since here learning usually takes place in a classroom context located outside the country or community where that language is spoken. Here within, the term L2 will be used to indicate non-native language learning in the country where the L2 is spoken. In the past few years, however, the term second language acquisition has been more widely used to indicate the study of another language regardless of the environment in which it takes place, perhaps because “[t]he term ‘second’ is more neutral and it is totally free of the negative nuances that might be associated to ‘foreign’” (Ma, 2009, p. 20).
reading, and writing as well as improve transferable skills. Revoicing offers learners the opportunity to enhance their speaking skills through dubbing, Audio Description (AD), and other voice-over tasks. Dubbing allows learners to replace the voice of someone in an AV product with their own voice. Dubbing – also known in the professional practice as *lip-synchronised dubbing* – can prove to be more challenging for learners due to the synchronisation between spoken language and onscreen images. AD – which aims to assist visually impaired people – requires learners to provide a spoken account of relevant visual aspects of the AV product between stretches of the original spoken dialogue. Depending on the task undertaken, revoicing can also promote listening, reading, and writing skills. AVT activities are learner-centred tasks which can be carried out both individually or in groups, and thus have the potential to promote autonomous and cooperative learning.

Captioning and revoicing in language learning can be either standard or reverse (Table 1). Standard interlingual subtitling refers to spoken second language (L2) text translated into written first language (L1), while interlingual reverse subtitling refers to spoken L1 text translated into written L2. Intralingual subtitling is a condensed transcription in L2 of the original spoken language as in the professional practice. With regard to revoicing, standard interlingual dubbing refers to spoken L2 text translated into spoken L1; while reverse dubbing refers to L1 spoken text translated into spoken L2. Although involving translation, standard interlingual dubbing is not frequently used in language learning. This may be because, contrary to standard interlingual subtitling, L2 input disappears completely from the final dubbed product. Finally, intralingual dubbing is a voice repetition of the original spoken language. While subtitling

---

4. The professional practice distinguishes a number of revoicing transfer methods: lip-synchronised dubbing, voice-over, narration, free commentary, and simultaneous interpreting (Pérez-González, 2014). This work follows the ClipFlair framework, which employs “the term revoicing to include all kinds of recording speech on video, including dubbing, voice over, audio description for the blind and visually impaired, free commentary, karaoke singing, or reciting” (Sokoli, 2018, p. 79).

5. In the professional practice, intralingual subtitling, also known as *captions* (Caimi, 2015) and referred to as bimodal, same language, or unilingual subtitling as well, “traditionally addressed at minority audiences, such as immigrants wishing to develop their proficiency in the language of the host community or viewers requiring written support to fully understand certain audiovisual texts [...] has now become almost synonymous with subtitling for the deaf and hard of hearing” (Pérez González, 2009, p. 15). However, subtitling for the deaf and hard-of-hearing also contains paralinguistic information otherwise not accessible to deaf people, while intralingual subtitling in language learning does not include paralinguistic information.
and dubbing remain the two most studied AVT techniques in language learning, audio description is gradually gaining scholars’ attention.\(^6\)

Table 1.1. Main aspects of subtitling and dubbing in language learning

<table>
<thead>
<tr>
<th>AVT Modes</th>
<th>Standard</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Captioning</strong></td>
<td>L2 =&gt; L1</td>
<td>L1 =&gt; L2</td>
</tr>
<tr>
<td>Interlingual Subtitling</td>
<td>L2 =&gt; L1/(another L2)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Is the written translation of the original spoken language.</td>
<td></td>
</tr>
<tr>
<td><strong>Intralingual Subtitling</strong></td>
<td>L2 =&gt; L2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Involves one language: L2.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Is a condensed transcription of the original spoken language.</td>
<td></td>
</tr>
<tr>
<td><strong>Revoicing</strong></td>
<td>L2 =&gt; L1</td>
<td>L1 =&gt; L2</td>
</tr>
<tr>
<td>Interlingual Dubbing</td>
<td>Involves two languages: L2 and L1/(another L2).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Is the oral translation of the original spoken language.</td>
<td></td>
</tr>
<tr>
<td><strong>Intralingual Dubbing</strong></td>
<td>L2 =&gt; L2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Involves one language: L2.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Is a voice repetition of the original spoken language.</td>
<td></td>
</tr>
</tbody>
</table>

The AVT practice in language learning has raised interest among scholars and teachers. European institutions have also recognised the potential of AVT in language education by funding research-led projects like *Learning Via Subtitling* (LeViS)\(^7\) (2006-2008) and ClipFlair\(^8\) (2011-2014). LeViS developed a free subtitling software and activities specifically designed for language learning. These activities promote a ‘hands-on’ approach where multimedia represents

---

6. In addition, there has been an increasing interest in using AVT tasks in translator (and subtitler) training (Barbe, 1996; Incalcaterra McLoughlin, 2009b; Martí Ferriol & Martí Marco, 2016; Rundle, 2000; Williams & Thorne, 2000; to name a few). The analysis of this phenomenon goes beyond the scope of the present work.


8. [http://clipflair.net/](http://clipflair.net/)
the central aspect of an activity, rather than a marginal feature (Hadzilacos, Papadakis, & Sokoli, 2004). A final evaluation of the project through questionnaires has revealed that learners found the subtitling task motivating and suitable for the development of different skills; most of the respondents expressed that they would like to have these activities in their regular foreign language class (Sokoli, Zabalbeascoa, & Fountana, 2011). ClipFlair – based on the LeViS experience – aims at promoting language learning through interactive clip captioning and revoicing. Hence, learners have a wider range of activities at their disposal (Sokoli, 2015).

AVT modes differ greatly from one another and do not necessarily involve translation, as in the case of intralingual subtitling and dubbing (Orrego Carmona, 2013). In the broader terms of captioning and revoicing, ClipFlair includes any activity in which learners can interact with (and thus modify) a video by adding their writing or speech respectively. Three types of verbal learner responses can be identified in ClipFlair activities: repeating, rephrasing, or reacting.

“Repeating refers to verbatim rendering of the verbal elements of the clip as literally as possible. Rephrasing means free rendering or rewording the text, and includes concepts such as “loose” paraphrase, gist, and summary. Reacting has to do with producing a new communicative contribution in response to a previous one” (Sokoli, 2018, p. 80, emphasis in the original).

ClipFlair activities can be employed successfully in face-to-face (Lertola & Mariotti, 2017), online (Incalcatera McLoughlin & Lertola, 2015), or blended learning contexts (Talaván & Rodríguez-Arancón, 2014a). Furthermore, learners’ involvement in learning activities is seen on a continuum from teacher-driven learners – who attend a face-to-face, online, or blended course in which the teacher decides how learners can best use ClipFlair – to independent learners, who select and organise their own language learning (Sokoli, 2015; Zabalbeascoa et al., 2012). AVT tasks allow learners to interact with multimodal material that combines verbal (oral and written speech) and non-verbal elements (sound and image) in an innovative way. According to Sokoli (2018), “[m]ultimodal literacy
Jennifer Lertola

[…] requires further skills: being able to interpret the multimodal text in its totality, as a complex communication act, and make sense from a combination of verbal and non-verbal sign elements” (p. 80), contrary to the unimodal text approach. In this context, ClipFlair has attempted to develop the traditional four-skill model by proposing the concept of AV skills: AV speaking, AV writing, AV listening and AV reading (Zabalbeascoa et al., 2012). These four AV-specific skills can be defined as follows:

“[AV speaking and AV writing] refer to the ability to produce speech and writing, respectively, in combination with the video, taking into consideration and adapting to its other elements, such as speed, voice quality, performance, shot transition. […] Similarly, AV reading and AV listening refer to oral and written comprehension with the combined effect of the elements of the multimodal material” (Sokoli, 2018, p. 80, emphasis in the original).

In order to foster the integration of captioning and revoicing activities in the FL classroom, ClipFlair has created a single online platform – the ClipFlair Studio – where teachers, learners, and activity creators can interact (Baños & Sokoli, 2015). All users can create and access (or modify) ready-made AVT activities in the ClipFlair Studio. Over 400 activities were created within the project framework to help better acquisition of the 15 target languages. To ensure quality, activities were peer-reviewed and tested by project members at different stages and in a number of learning settings. The one-year ClipFlair pilot phase involved more than 1,200 learners, 37 tutors, and 12 languages (Baños & Sokoli, 2015, p. 211). Overall results demonstrate how ClipFlair is appreciated by both learners and teachers. More than 80% of the learners involved in the pilot phase found the activities useful for language learning, and reported that they would like to have such activities in their regular foreign language classroom. The LeViS and ClipFlair projects have created a fruitful environment for teaching-
oriented research, national, and international collaborations\textsuperscript{10}, publications and a specific conference on the topic\textsuperscript{11}. However, after the end of the ClipFlair project, related research has gradually decreased.

This work focusses on the foreign language classroom, presenting a systematic review of studies carried out in the last twenty years on the applications of captioning and revoicing in this context in order to offer an overview of the state of the art and encourage further research. The literature review will present research on the topic, paying particular attention to relevant experimental studies (i.e. empirical research that involves data collection), which will be reviewed in terms of research focus, target languages, participants, learning settings, audiovisual materials, and captioning/revoicing software and type of analysis (i.e. qualitative and/or quantitative)\textsuperscript{12}. The first part will consider studies on captioning, namely on standard interlingual subtitling, reverse interlingual subtitling, and intralingual subtitling. The second part will focus on the research currently available on revoicing: reverse interlingual dubbing, intralingual dubbing, and audio description. Finally, the third part will present the studies that combine captioning and revoicing; in particular, reverse interlingual and intralingual subtitling and dubbing.

\textsuperscript{10} At a national level, in 2010, the National Academy for the Integration of Research, Teaching and Learning (NAIRTL), Ireland, funded the open-ended collaborative website project called Sub2Learn (\url{http://www.sub2learn.ie/}) that offers guidelines on how to subtitle, examples of subtitling activities in English, French, and Italian, a list of subtitling software, evaluation guidelines, and a bibliography on subtitles and subtitling in language learning.

\textsuperscript{11} ClipFlair Conference, Barcelona, 18-19 June 2014 (\url{http://clipflair.net/conference2014/}).

\textsuperscript{12} In quantitative research, data collection procedures provide principally numerical data which is analysed through statistical methods, while in qualitative research data collection procedures mainly aim to present open-ended, non-numerical data which is analysed through non-statistical methods (Dörnyei, 2007). Qualitative and quantitative approaches have become complementary in education research since their combination often results in “findings that may […] provide a more complete explanation of the research problem than either method alone could provide” (Ary, Cheser Jacobs, & Sorensen, 2009, p. 23).