This work has sought to systematically review research on the applications of captioning and revoicing in language learning carried out in the last 20 years in order to offer an overview of the state of the art and encourage further research. The review of research on the use of the most widely studied AVT mode in language learning, namely captioning, illustrates how a number of publications have focussed on standard and reverse interlingual subtitling. Scholars argue for the integration of AVT tasks within a communicative approach to language learning and teaching. Interlingual subtitling is seen as an innovative and motivating exercise that allows learners to notice the language and negotiate meaning, thus enhancing listening comprehension and vocabulary acquisition (Díaz Cintas, 1995, 1997, 2001; Talaván, 2006a). Apart from the benefits deriving from translation, standard subtitling can prove particularly beneficial for improving listening skills, while reverse subtitling can enhance writing skills. In both cases, peer-to-peer review is encouraged to develop metalinguistic reflections. In research considering a teacher’s perspective, examples of methodology-based subtitling models are also suggested for language learning and intercultural education (Borghetti, 2011; Incalcaterra McLoughlin & Lertola, 2011), as well as an adaptation of professional subtitling norms and assessment criteria for pedagogical purposes (Lertola, 2015). Finally, interlingual subtitling can also enrich language for specific purposes education (Kantz, 2015; Talaván, 2006b).

The greatest number of experimental studies focus on standard interlingual subtitling as a pedagogical tool. The benefits of interlingual subtitling have been identified with regards to listening comprehension (Talaván, 2010, 2011; Talaván & Rodríguez-Arencón, 2014a), vocabulary acquisition (Lertola, 2012, forthcoming), idiomatic expression retention (Bravo, 2008), intercultural education (Borghetti & Lertola, 2014), and pragmatic awareness (Incalcaterra
McLoughlin, 2009a; Incalcaterra McLoughlin & Lertola, 2016; Lopriore & Ceruti, 2015). In all the studies considered, English is one of the two languages involved in the subtitling task. In almost half of cases identified and examined, English is the L2; Italian is the L2 in all the other studies found. The great majority of these studies have been carried out in face-to-face university contexts involving a minimum of ten participants whose language levels ranged from A1 to C1. The AV material generally used is a movie or sitcom excerpt to be subtitled with the software Learning via Subtitling, ClipFlair, or Subtitle Workshop. The type of analysis applied is both qualitative and quantitative. Concerning reverse interlingual subtitling, the research focus is the development of writing and translation skills (Burczyńska, 2015; Talaván et al., 2016a; Talaván & Rodríguez-Arancón, 2014b). In the three experimental studies available, a qualitative and quantitative analysis is carried out on reverse subtitling of movie excerpts (with Aegisub or Subtitle Workshop) with English as the L2. Two of the studies involve a great number of university students as participants in online learning settings (Talaván et al., 2016a; Talaván & Rodríguez-Arancón, 2014b), whereas investigation on the application of intralingual subtitling tasks in language learning is quite limited. To this regard, only two studies were identified, one of which combines the use of intralingual subtitling and dubbing. The study on intralingual subtitling alone reports encouraging results on writing skills enhancement in undergraduate distance learning students of English, who subtitled a sitcom excerpt using ClipFlair (Talaván et al., 2016b).

Revoicing as a language learning tool is gaining more scholarly attention. Within revoicing, dubbing has enjoyed the highest consideration as an AVT technique; however, audio description is starting to be viewed as another effective task in the foreign language classroom. The literature review shows that dubbing has been suggested by many authors among other communicative language activities, either with or without translation (Bilbrough, 2007; Duff, 1989; Kumai, 1996; Zohrevandi, 1994). For instance, González Davies (2004) makes a general dubbing task proposal, without specifying whether to use standard or reverse dubbing in which the core exercise is translation, since no actual revoicing is required. Two options in dubbing tasks not involving translation have been identified: ‘simple video dubbing’, which consists in substituting the original
soundtrack with students’ voices; and ‘scenario creation’, suitable for more advanced students who can prepare their own storyline and script for a muted video (Burston, 2005). While the first option mainly focuses on the improvement of listening and speaking skills, the second option can also enrich reading and writing skills and boost grammar and vocabulary knowledge. Dubbing thus provides similar pedagogical benefits to video making while saving classroom time and being less complex in terms of logistics. It still requires preparation for both teachers and learners, however, especially in technical terms. Peer-to-peer collaboration and public presentation of the dubbed product also prove to be strong motivators (Burston, 2005; Danan, 2010; Duff, 1989; Kumai, 1996).

Experimental research has mainly focused on intralingual dubbing; only one experimental study on reverse interlingual dubbing was found. A limited number of experimental studies that combine reverse interlingual subtitling and dubbing were identified. As with experimental research on captioning, the target language in most intralingual-dubbing experimental studies is English, followed by Spanish. These studies focus on the enhancement of speaking skills, mainly in face-to-face secondary-school and university settings. The number of participants varies greatly and so do their proficiency levels. In most studies, qualitative and quantitative data are both gathered. Movie or TV series excerpts are the most preferred AV materials, and ClipFlair is the dubbing software selected both in face-to-face and distance learning (Navarrete, 2013; Talaván & Costal, 2017). The research focus of the qualitative study on reverse interlingual dubbing is the development of speaking and writing skills, since army students in the United States are required to dub a number of AV excerpts from English into Dari, Pashto, or Farsi (Danan, 2010).

Scholars have carried out a considerable amount of research on the application of AD in foreign language learning in the last decade. Again, as is the case for dubbing and subtitling, English is the target language in most AD experimental studies, followed by Spanish. In general, the focus of AD tasks is the development of speaking skills (Talaván & Lertola, 2016). However, lexical and phraseological competence, integrated learning skills, and writing skills have also been considered (Calduch & Talaván, 2018; Ibáñez Moreno & Vermeulen,
Interestingly, a number of studies focus on the development of speaking skills in Chinese learners of English (Chiu, 2012; Florente, 2016; He & Wasuntarasophit, 2015). However, videos used in some of these studies exceed the recommended length of video clips in language education. In other studies with students of Spanish as a foreign language, the content of clips selected for AD tasks is related to blindness in order to introduce AD as an accessibility mode and sensitise learners on visually-impaired conditions (Herrero & Escobar, 2018; Ibáñez Moreno & Vermeulen, 2013, 2014). All the experimental studies on AD were carried out with university students (mainly in face-to-face contexts) whose language levels ranged from B1 to B2. Participants carried out the AD either with ClipFlair or the VISP mobile app. In half of the cases, only qualitative data was gathered.

The application of combined captioning and revoicing has also gained scholars’ attention in the last few years. A limited number of experimental studies on combined reverse interlingual subtitling and dubbing have been identified. Although focussing on different aspects – speaking skills, pragmatic awareness, and collaborative language learning – all studies are carried out with a considerable number of university students of English, mainly in distance education settings (Talaván et al., 2014, 2015; Talaván & Ávila-Cabrera, 2015). Quantitative and qualitative collection techniques were used in almost all the studies. The captioning software selected was either Aegisub, DivXLand Media Subtitler, or ClipFlair; while the revoicing software consisted of Audacity, Windows Movie Maker, or ClipFlair. As previously mentioned, two studies on combined intralingual captioning and revoicing were identified. The first qualitative study by López Cirugeda and Sánchez Ruiz (2013) reports on an English subtitling and revoicing task within teacher training in a face-to-face university context. The AV material was created by each participant who performed the task using one subtitling software programme among those suggested (AVI Subtitler, DivXLand Media Subtitler, and Subtitle workshop) and Virtualdub as revoicing software. The second qualitative study by Herrero et al. (2017) combines film analysis, AD, and subtitles for the deaf and hard-of-hearing. Although the study aims at developing an ad hoc teaching model, these combined AVT tasks could foster writing and speaking skills.
All AVT tasks examined provide learners with the opportunity to benefit from authentic multimodal input and produce a tangible output, which helps develop a sense of communicative achievement. Teachers play an important role both in revoicing and captioning tasks. For instance, Danan (2010) reckons that, although more successful in terms of motivation, entirely student-led projects present more language-related mistakes. She suggests finding the right balance between teachers’ control and learners’ initiative through periodical guidance and feedback. Borghetti and Lertola (2014) also acknowledge that more opportunities for intercultural learning arise with teachers’ input. The main limitations of employing AVT tasks in foreign language learning are the lack of ready-made activities, which is in part overcome by the advent of ClipFlair, and technology-related issues. The creation of tasks can be time consuming. In particular, selecting videos with suitable language and content for AVT tasks can prove challenging. Furthermore, learners should be computer-literate and have access to a computer (or a mobile in the case of AVT apps) with an internet connection. The employment of apps is still quite limited. Experimental studies often report technological issues related to the software used encountered by learners while performing captioning and revoicing tasks.

Overall, the literature review on captioning and revoicing in language learning has revealed that a fast-growing body of research supporting the integration of AVT tasks in the foreign language classroom, especially in English as a foreign language learning, is currently available. Nevertheless, the literature review provides suggestions for future research. Besides developing current lines of research, future investigation should bridge the current gaps according to several of the authors cited. In particular, it should foster research on reverse interlingual subtitling and dubbing, intralingual subtitling, and combined intralingual captioning and revoicing. Moreover, scholars urge future research involving different language combinations. Participants should also come from a variety of learning contexts other than higher education, and be in online or blended settings. Finally, although they provide encouraging outcomes, a very limited number of studies is currently available on the benefits of AVT tasks in language for specific purposes education and teacher training, calling for future research in these regards.