CrossMark

Studies in Second Language Acquisition **42** (2020), 489–497
doi:10.1017/S0272263120000212

Research Report

INTRODUCTION TO THE SPECIAL ISSUE

LANGUAGE LEARNING FROM MULTIMODAL INPUT

Elke Peters *

KU Leuven

Carmen Muñoz

University of Barcelona

INTRODUCTION

This special issue is concerned with an area that has only recently attracted the attention of SLA researchers: the effects of multimodal input (image + voice + subtitles/captions; picture + voice + text) on second language (L2) learning. As stated by Ellis and Shintani (2014, p. 7), "learning can only take place when learners are exposed to input," but compared to the bulk of studies into written input, only few studies have investigated the potential of multimodal input. Recent research has shown that watching (subtitled) television and exposure to (and interaction with) audiovisual material enhances learners' L2 skills (e.g., Peters & Webb, 2018; Rodgers & Webb, 2017; Sockett, 2014). Studies on multimodal input are generally in line with theories of multimodal input (e.g., Mayer, 2009). However, these studies also raise many questions as to how and when learning is promoted and the individual differences that influence such learning.

MULTIMODAL INPUT

Drawing on Paivio's (1986) Dual Coding Theory, Mayer's (2014) cognitive theory of multimedia learning states that learning is better when information is processed in spoken as well as written mode because learners make mental connections between the aural and visual information provided there is temporal proximity. Examples in the domain of language learning are storybooks with pictures read aloud (e.g., Tragant & Pellicer-Sánchez, 2019), audiovisual input (e.g., Peters & Webb, 2018), subtitled audiovisual input (e.g., Peters, 2019; Pujadas & Muñoz, 2019), captioned audiovisual input (e.g., Montero Perez et al., 2013; Peters, 2019; Pujadas & Muñoz, 2019), or glossed audiovisual input (e.g., Montero Perez et al., 2018). What these types of input have in common is the

* Correspondence concerning this article should be addressed to Elke Peters, KU Leuven, Faculty of Arts, Sint-Andriesstraat 2, 2000 Antwerp, Belgium. elke.peters@kuleuven.be

© *The Author(s), 2020. Published by Cambridge University Press.*

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

combination of pictorial information (static or dynamic) and verbal input (spoken and/or written). In studies on multimodal input, learning gains are often explained in terms of the visual support (Peters, 2019; Peters & Webb, 2018; Rodgers, 2018; Rodgers & Webb, 2019; Wong & Samudra, 2019). However, most of these input types combine not two but three sources of input: (1) pictorial information, (2) written verbal information in captions or subtitles, or in written text, and (3) aural verbal input. In light of Baddeley's (2007) dual-processing theory of working memory, it could be argued that language learners might experience cognitive overload when engaging with both pictorial and written information in addition to aural input. As a result of processing difficulties, language learners might divide their attention between the pictorial information and the written verbal information (Ayres & Sweller, 2014; Mayer & Moreno, 1998). However, eye-tracking research has demonstrated that language learners are able to process both pictorial and written verbal information (Bisson et al., 2012), on the condition that they are familiar with the script of the foreign language (Winke et al., 2013).

In addition to imagery, there are other advantages inherent in multimodal input and audiovisual input in particular. Learners need fewer words to understand TV programs compared to books. Webb and Rodgers (2009a, 2009b) have put forward knowledge of the 3,000 most frequent word families and proper nouns to reach 95% coverage of the input. However, the lexical coverage figures for TV viewing have recently been found to be lower (Durbahn et al., 2020), so the lexical demands are not as high as for reading (knowledge of the 4,000 most frequent word families for adequate comprehension and 8,000 word families for detailed comprehension; Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006). Rodgers and Webb (2011) also established that words are repeated more often in TV programs than in reading, especially in related TV programs, which is beneficial for vocabulary learning. Another advantage is the wide availability of audiovisual input using the Internet and streaming platforms. It can, thus, easily provide language learners with large amounts of authentic language input (Webb, 2015). Finally, language learners are motivated to watch L2 television, as has been well documented in surveys on language learners' engagement with the L2 outside of the school (Lindgren & Muñoz, 2013; Peters, 2018; Peters et al., 2019).

LANGUAGE LEARNING FROM MULTIMODAL INPUT

Previous research into language learning from multimodal input has focused on three main areas: comprehension, vocabulary learning, and, to a lesser extent, grammar learning.

A consistent finding in this area is that audiovisual input is beneficial for comprehension, in particular when learners have access to captions (e.g., Gass et al., 2019; Rodgers & Webb, 2017; Winke et al., 2010; see also the meta-analysis by Montero Perez et al., 2013). Captions assist comprehension by helping to break down speech into words and thus facilitating listening and reading comprehension. Crucially, a unique support offered to learners' comprehension by multimodal input is imagery, as noted in the preceding text (e.g., Durbahn et al., 2020; Jones & Plass, 2002). Research into audiovisual input has shown that it can work as a compensatory mechanism especially for low-proficiency learners (e.g., Rodgers, 2016).

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

The bulk of research into multimodal input has focused on vocabulary learning. A seminal study on the effect of TV viewing on vocabulary learning is Neuman and Koskinen's 1992 study. They were among the first to stress the potential of audiovisual input for vocabulary learning. Other early examples of research into TV viewing include Vanderplank (1988), Koolstra and Beentjes (1999), and the work by d'Ydewalle and colleagues (d'Ydewalle & Van de Poel, 1999; Pavakanun & d'Ydewalle, 1992). Although there were a considerable number of studies on TV viewing in the 1990s, audiovisual input was largely neglected by SLA researchers. It was not until 2009 that the field of SLA started to pay more attention to audiovisual input. Two key studies were the corpus studies by Webb and Rodgers (2009a, 2009b), which showed the lexical demands of different types of audiovisual input. They argued that in addition to reading, audiovisual input may also be a valuable source of input for language learners. Since then, the field of SLA has witnessed a steady increase in the number of studies investigating vocabulary learning from audiovisual input (e.g., Sydorenko, 2010; Winke et al., 2010). A great deal of previous research into audiovisual input has focused on the efficacy of captions (= subtitles in the L2) on vocabulary learning (see Montero Perez et al.'s, 2013, meta-analysis). Surprisingly, fewer studies have focused on noncaptioned and nonsubtitled audiovisual input (see Peters & Webb, 2018, and Rodgers & Webb, 2019, for two exceptions). Even though the majority of studies have been conducted with university students (e.g., Montero Perez et al., 2014; Peters & Webb, 2018; Rodgers & Webb, 2019; Winke et al., 2010), there has been an increasing number of studies with learners in secondary schools (e.g., Peters, 2019; Peters et al., 2016; Pujadas & Muñoz, 2019; Suárez & Gesa, 2019), primary schools (e.g., Muñoz, 2017), and even preschool children (e.g., Samudra et al., 2019). Research has also moved from using short, educational clips to using full-length TV programs (e.g., Peters & Webb, 2018) and even complete TV shows (e.g., Pujadas & Muñoz, 2019; Rodgers & Webb, 2019). Finally, in addition to studying the effectiveness of multimodal input for vocabulary learning, research has also started to study language learners' processing of multimodal input (e.g., looking patterns of captions or pictures) by means of eye-tracking (Bisson et al., 2012; Montero Perez et al., 2015; Serrano & Pellicer-Sánchez, 2019; Winke et al., 2013; Wong & Samudra, 2019). Together, there seems to be robust evidence that language learners can indeed pick up unfamiliar words from multimodal input and that the provision of captions has the potential to increase the learning gains.

Research into grammar learning through multimodal input is very scarce. Early studies (d'Ydewalle & Van de Poel, 1999; Van Lommel et al., 2006) found only small gains in young learners' grammar learning under different subtitling conditions: reversed (L1 audio, L2 subtitles) and standard (L2 audio, L1 subtitles). In the study by Van Lommel and colleagues grammar learning gains were observed only when rules were presented in advance, which led these authors to conclude that, in contrast to vocabulary acquisition, grammar might be too complex to learn incidentally from a short video presentation. More recent studies involving captions and grammar in longer treatments have provided evidence of positive benefits for L2 grammar development in adults, especially when captions are textually enhanced (Lee & Révész, 2018). However, results have not been similarly positive for all target structures, suggesting the influence of other factors such as the structure-specific saliency of a grammar token (Cintrón-Valentín et al., 2019).

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

RATIONALE FOR AND OVERVIEW OF THE SPECIAL ISSUE

Research into the potential of multimodal input has been gaining traction, but the number of studies is still limited and mainly confined to vocabulary learning. Now that research into multimodal input is starting to broaden its focus to different aspects of learning as well as its research techniques, the present issue provides an up-to-date account of research in this area with a view to include innovative work and a range of approaches.

The special issue pursues new avenues in research into multimodal input by focusing on pronunciation, perception and segmentation skills, grammar, multiword units, and comprehension. In addition, it extends previous eye-tracking research (e.g., Winke et al., 2013) by investigating the effects of underresearched pedagogic interventions on learners' processing of target items, target structures, and text. The studies nicely complement each other in their research methodologies and participant profiles. The special issue comprises six empirical studies and one concluding commentary.

The six empirical papers in the special issue explore:

- Different types of *input* (TV viewing with and without L1 or L2 subtitles, reading-while-listening, reading, listening);
- Different types of *captioning* (unenhanced, enhanced, no captioning);
- Different *components* of language learning (single words, formulaic sequences, comprehension, grammar, pronunciation);
- Different mediating *learner-* and *item-related* factors (e.g., working memory, prior vocabulary knowledge, frequency of occurrence);
- Different *learning conditions* (incidental learning, intentional learning, experimental and classroom-based) and time conditions (short video clips, full-length TV programs, extensive viewing);
- Participants of different *ages* (children, adolescents, adults); and
- Different *research tools* (eye-tracking, productive and receptive vocabulary tests, comprehension tests).

Two studies in this special issue address vocabulary learning. They broaden the scope of existing research into multimodal input by comparing different types of input (written, spoken, audiovisual input in Feng & Webb) and by focusing on learning formulaic sequences (in Puimège & Peters).

Because several researchers (e.g., Webb, 2015) have recommended the use of audiovisual input for enlarging learners' vocabulary, Feng and Webb's study compared the relative effectiveness of three types of input (audiovisual, spoken, written) for vocabulary learning. It is the first study that compares reading, listening, and TV viewing in one experiment. To date, it is not clear whether audiovisual input is more effective than other types of input for language learning. To answer that question, Feng and Webb set up a quasi-experimental study, in which Chinese university students were assigned to one of four conditions: viewing condition, listening condition, reading condition, and control condition (= no input). The authors used a full-length TV program and tested EFL learners before and after the viewing on a set of target words. Participants in each experimental condition were exposed to the same content of the TV program, but in different ways depending on the condition to which they were assigned. The viewing group watched the TV program (audio + imagery); the reading group read the script of the TV program without having access to the audio; and the listening group listened to the audio of the TV

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

program without having access to the visual support. Feng and Webb's findings showed that all three groups learned vocabulary to the same extent. Regardless of the experimental group, there was incidental vocabulary learning from meaningful input. The authors conclude that in addition to extensive reading programs, there should "also be programs developed to encourage extensive listening and extensive viewing" (p. 520).

The study by Puimège and Peters explores the incidental learning of formulaic sequences from audiovisual input and factors affecting the learning of formulaic sequences. There is some evidence that formulaic sequences can be learned incidentally through reading (e.g., Pellicer-Sánchez, 2017; Toomer & Elgort, 2019) and reading-while-listening (Webb et al., 2013). However, very little is known about learning formulaic sequences from audiovisual input, as the few studies into the effect of TV viewing so far have tended to focus on learning single words (Peters & Webb, 2018). This is surprising given that corpus research has shown that the oral discourse typical of audiovisual input could provide sufficient exposure to formulaic sequences for incidental learning to occur (Lin, 2014). Like Feng and Webb, Puimège and Peters used a full-length episode as input. By using a pretest-posttest, within-participant design, the authors found that Dutch-speaking English-as-a-foreign-language learners learned formulaic sequences both at the level of form recall and meaning recall after watching a 1-hour English-language documentary without subtitles. In line with previous research (e.g., Peters, 2019; Peters & Webb, 2018), they also found that learners' prior vocabulary knowledge was positively related to learning (Matthew effect). Interestingly, the study shows the incremental nature of learning formulaic sequences, as learners' preknowledge of the meaning of the formulaic sequences prior to the TV viewing was the strongest predictor of learning the form of formulaic sequences.

Two studies in this special issue focus on comprehension from multimodal input. Pujadas and Muñoz compared the differential effect of subtitles and captions on comprehension in a longitudinal, classroom-based study, while Pellicer-Sánchez et al. explore processing differences in reading-only and reading-while-listening.

The study by Pujadas and Muñoz contributes to the area by addressing the effects of captions and subtitles on TV viewing comprehension in a classroom intervention that extends over 8 months, which affords a unique longitudinal view with great ecological validity and pedagogical implications. Once a week four groups of adolescent beginner to low-intermediate learners of English viewed an episode of a sitcom series, two with captions and two with subtitles; one group in each language condition received explicit instruction of the target vocabulary prior to viewing the episode. The study also aimed at investigating how learners' comprehension was affected by factors related to the learner, the episodes, and the testing instruments. The study found that, with learners at this age and proficiency level, viewing with subtitles leads to higher content comprehension than viewing with captions, and that having explicit instruction on target vocabulary had a small negative effect on comprehension (possibly because of the difficulty of splitting attention between the two demands). Other effects of on-screen text language were observed in the influence of vocabulary size only in the captions condition, and in differences in the way in which learners processed textually explicit information and inferential information, while the lexical coverage of the episodes was a significant predictor of comprehension independently of the language group.

Pellicer-Sánchez et al. make a unique contribution to research into multimodality by investigating young EFL learners' processing of multimodal reading materials by means of

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

eye-tracking. Compared to the number of studies focusing on university students, there is still little research into young language learners. The authors used a counterbalanced, within-participants design to explore processing differences in storybook reading with or without audio by analyzing learners' looking patterns at the text and the pictures. All participants were exposed to both the reading-only and reading-while-listening condition. The authors found that learners spent proportionally more time on the pictures in the reading-while-listening condition than in the reading-only condition. However, it should be noted that this was not at the expense of text comprehension. Their study does not provide evidence for Mayer's (2009, 2014) redundancy principle, which states that learning is negatively affected when verbal information is presented in both spoken and written form. On the contrary, their findings suggest that there was better integration of pictorial and verbal information when the storybook also had audio (reading-while-listening).

The final two studies both investigate the effect of captions, but each addresses a different aspect of learning, that is pronunciation development (Wisniewska & Mora) and grammar (Lee & Révész).

The study by Wisnieska and Mora is the first to examine the benefits of extended exposure to multimodal input through captioned and uncaptioned video on L2 pronunciation development. The scarce research in this area found benefits for L2 speech segmentation from captioned viewing (e.g., Charles & Trenkic, 2015) and evidence that captions help retune L2 learners' phonetic categories (Birulés-Muntané & Soto-Faraco, 2016; Mitterer & McQueen, 2009). Wisnieska and Mora extend previous research by examining processing skills and perceptual and productive accuracy in learners' L2 phonology. Participants were randomly assigned to four experimental conditions involving two viewing modes (captioned or uncaptioned) and two task focus conditions (focus on phonetic form or focus on meaning). Benefits for L2 pronunciation were assessed in a pretest/posttest design after an 8-week exposure treatment. Unlike previous research, captioned video in this study was not superior to uncaptioned video in effecting gains in speech perception, which may reflect differences in testing. An intriguing result concerns differences in the benefits obtained for phonological accuracy under the two viewing modes and the two different attention-directing conditions. Captioned viewing led to pronunciation gains as long as there was no focus on phonetic form and learners' attention was directed to meaning, whereas uncaptioned viewing led to pronunciation gains only when there was a focus on phonetic form. The authors interpret these findings as a case of cognitive overload that made it difficult to synchronize reading with online auditory speech processing. Like in Pujadas and Muñoz's study (this issue), overloading may have resulted from asking learners to split their attention between competing tasks.

As noted in the preceding text, there is as yet little available evidence as to whether captioning may promote development in L2 grammatical knowledge. Lee and Révész's pioneering study contributes to expand this area in multiple directions. Their study is novel in employing eye-tracking methodology to investigate whether attention allocation to target grammatical features is moderated by type of captioning and related to their development. The study followed a pretest-posttest-delayed posttest experimental design, and learners were assigned to an enhanced captions group, an unenhanced captions group, and a no captions group. The findings of the study show that captioning, with or without textual enhancement, facilitated the acquisition of grammatical knowledge, and that the enhanced captions group obtained greater gains. The authors suggest that in comparison with reading

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

studies, where weaker effects are found (Lee & Huang, 2008), textual enhancement and captioning might have increased the salience of the target features leading to a greater depth of processing. In addition, learners in the enhanced captioning condition paid more attention to the target grammatical construction and showed stronger association between attention and target construction development. This study has direct pedagogical implications illustrating how videos can be used effectively in meaningful tasks that are directly relevant for the language classroom (following a task-based language teaching approach) and exploring the effectiveness of directing learners' attention to the target construction.

We are confident that this special issue will make a significant contribution to theories of language learning from multimodal input and will help us gain a deeper understanding of the mechanisms underlying language learning. The research questions addressed in this proposed special issue are also concerned with essential topics in the field, such as the effects of learner proficiency and individual differences (attention, working memory) or the effects of textual enhancement.

The research presented in this special issue demonstrates how multimodal materials can be used to increase and enrich learner exposure to language. Such research carries important educational implications. It speaks to language teachers indicating ways of enriching classroom input. It also speaks to language policy makers and society in general, increasing their awareness of the potential of audiovisual input for L2 learning outside the classroom, a learning context that is gaining prominence in the current globalized world.

REFERENCES

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 206–226). Cambridge University Press.

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford University Press.

Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS One*, *11*, e0158409.

Bisson, M.-J., Van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2012). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*, 1–20. https://doi.org/10.1017/S0142716412000434.

Charles, T. J., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173–198). Peter Lang.

Cintrón-Valentín, M., Garcia-Amaya, L., & Ellis, N. (2019). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal*, *47*, 1–21. https://doi.org/10.1080/09571736.2019.1615978.

Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System*, *88*, 1–13. https://doi.org/10.1016/j.system.2019.102166.

d'Ydewalle, G., & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, *28*, 227–245.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.

Gass, S., Winke, P., Isbell, D., & Ahn, J. (2019). How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning & Technology*, *23*, 84–104. https://doi.org/10125/44684.

Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, *86*, 546–561. https://doi.org/10.1111/1540-4781.00160.

Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*, 51–60. https://doi.org/10.1007/BF02299476.

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*, 15–30.

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal*, *102*, 557–577. https://doi.org/10.1111/modl.12503.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition*, *30*, 307–331. https://doi.org/10.1017/S0272263108080479.

Lin, P. M. S. (2014). Investigating the validity of Internet television as a resource for acquiring L2 formulaic sequences. *System*, *42*, 164–176. https://doi.org/10.1016/j.system.2013.11.010.

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*, 105–129. https://doi.org/10.1080/14790718.2012.679275.

Mayer, R. (2014). Introduction to multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–24). Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320. https://doi.org/10.1037/0022-0663.90.2.312.

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS One*, *4*, 1–5.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, *18*, 118–141. http://dx.doi.org/10125/44357.

Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, *99*, 308–328. https://doi.org/10.1111/modl.12215.

Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, *31*, 1–26. https://doi.org/10.1080/09588221.2017.1375960.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*, 720–739. https://doi.org/10.1016/j.system.2013.07.013.

Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System*, *67*, 77–86. https://doi.org/10.1016/j.system.2017.04.015.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*, 59–82. https://doi.org/10.1353/cml.2006.0049.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*, 94–106. https://doi.org/https://www.jstor.org/stable/747835.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Pavakanun, U., & d'Ydewalle, G. (1992). Watching foreign television programs and language learning. In F. L. Engel, D. G. Bouwhuis, T. Bosser, & G. d'Ydewalle (Eds.), *Cognitive modelling and interactive environments in language learning* (pp. 193–198). Springer.

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research*, *21*, 381–402. https://doi.org/10.1177/1362168815618428.

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners ' vocabulary knowledge. *ITL—International Journal of Applied Linguistics*, *169*, 142–168. https://doi.org/10.1075/itl.00010.pet.

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*, *53*, 1008–1032. https://doi.org/10.1002/tesq.531.

Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, *63*, 134–148. https://doi.org/10.1016/j.system.2016.10.002.

Peters, E., Noreillie, A.-S., Heylen, K., Bulté, B., & Desmet, P. (2019). The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning*, *69*, 747–782. https://doi.org/10.1111/lang.12351.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*, 1–27. https://doi.org/10.1017/S0272263117000407.

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: An exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal*, *47*, 424–438. https://doi.org/10.1080/09571736.2019.1638630.

Pujadas, G., & Muñoz, C. (2019). Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by adolescents. *The Language Learning Journal*, *47*, 479–496. https://doi.org/10.1080/09571736.2019.1616806.

Rodgers, M. P. H. (2016). Extensive listening and viewing: The benefits of audiobooks and television. *The European Journal of Applied Linguistics and TEFL*, *5*, 43–57.

Rodgers, M. P. H., & Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, *45*, 689–717. https://doi.org/10.5054/tq.2011.268062.

Rodgers, M. P. H., & Webb, S. (2017). The effects of captions on EFL learners ' comprehension of English-language television programs. *Calico Journal*, *34*, 20–38. https://doi.org/10.1558/cj.29522.

Rodgers, M. P. H. (2018). The images in television programs and the potential for learning unknown words The relationship between on-screen imagery and vocabulary. *ITL-International Journal of Applied Linguistics*, *169*, 191–211. https://doi.org/10.1075/itl.00012.rod

Rodgers, M. P. H., & Webb, S. (2019). Incidental, vocabulary learning through viewing television. *ITL—International Journal of Applied Linguistics*. Advance online publication. https://doi.org/10.1075/itl.18034.rod.

Samudra, P. G., Wong, K. M., & Neuman, S. B. (2019). Promoting low-income preschoolers' vocabulary learning from educational media: Does repetition support memory for learned word knowledge? *Journal of Cognitive Education and Psychology*, *18*, 160–173. http://dx.doi.org/10.1891/1945-8959.18.2.160.

Serrano, R., & Pellicer-Sánchez, A. (2019). Young L2 learners' online processing of information in a graded reader during reading-only and reading-while-listening conditions: A study of eye movements. *Applied Linguistics Review*. Advance online publication. https://doi.org/10.1515/applirev-2018-0102.

Sockett, G. (2014). *The online informal learning of English*. Palgrave MacMillan.

Suárez, M. d. M., & Gesa, F. (2019). Learning vocabulary with the support of sustained exposure to captioned video: Do proficiency and aptitude make a difference? *The Language Learning Journal*, *47*, 497–517. https://doi.org/10.1080/09571736.2019.1617768.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, *14*, 50–73. http://dx.doi.org/10125/44214.

Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, *69*, 405–439. https://doi.org/10.1111/lang.12335.

Tragant, E., & Pellicer-Sánchez, A. (2019). Young EFL learners ' processing of multimodal input: Examining learners' eye movements. *System*, *80*, 212–223. https://doi.org/10.1016/j.system.2018.12.002.

Vanderplank, R. (1988). The value of teletext subtitles in language learning. *ELT Journal*, *42*, 272–281. https://doi.org/10.1093/elt/42.4.272.

Van Lommel, S., Laenen, A., & D'Ydewalle, G. (2006). Foreign-grammar acquisition while watching subtitled television programmes. *British Journal of Educational Psychology*, *76*, 243–258. https://doi.org/10.1348/000709905X38946.

Webb, S. (2015). Extensive viewing: Language learning through watching television. In D. Nunan, & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159–168). Routledge.

Webb, S., Newton, J., & Chang, A. C. S. (2013). Incidental learning of collocation. *Language Learning*, *63*, 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x.

Webb, S., & Rodgers, M. P. H. (2009a). The lexical coverage of movies. *Applied Linguistics*, *30*, 407–427. https://doi.org/10.1093/applin/amp010.

Webb, S., & Rodgers, M. P. H. (2009b). Vocabulary demands of television programs. *Language Learning*, *59*, 335–366. https://doi.org/10.1111/j.1467-9922.2009.00509.x.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*, 65–86. http://dx.doi.org/10125/44214.

Winke, P., Sydorenko, T., & Gass, S. (2013). Factors influencing the use of captions by foreign language learners, an eye-tracking study. *The Modern Language Journal*, *97*, 254–275. https://doi.org/10.1111/j.1540-4781.2013.01432.x.

Wong, K. M., & Samudra, P. G. (2019). L2 vocabulary learning from educational media: Extending dual-coding theory to dual-language learners. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2019.1666150.

https://doi.org/10.1017/S0272263120000212 Published online by Cambridge University Press