

The Development of Multi-modal Corpus Tool For Teaching English to Young Learners

Eri Hirata

Abstract

In recent years, children have been exposed to various types of technological tools and devices at a very early stage of their lives. In educational settings, different types of educational tools are emerging as a result of the advancement of information technology. While the analytical methods developed in corpus linguistics are applied in different areas of research, the applications to classroom education are somewhat limited, especially those aimed at young learners. Corpus-based language learning and teaching is still limited to adult learners due to the authenticity of the corpus data and the difficulties involved with using available concordance software tools. This paper discusses an ongoing research project to develop a multi-modal corpus application tool for young learners aimed at assisting their English learning in context.

Keywords: multi-modal corpus tool, EFL, young learners,
corpus-based approach to language learning
Teaching English to Young Learners (TEYL)

Introduction

With the development of corpus linguistics and advances in the related information technology including software developments, it is possible to access large electronic text databases of naturally occurring spoken and written language and to analyse them both quantitatively and qualitatively (Adolphs and Carter, 2013). The analytical methodologies developed in corpus linguistics are often applied to other areas of language study. Nevertheless, the mainstream has been “limited to the textual dimension of communication” (Carter & Adolphs, 2008: 275). This is largely due to the availability of data and technology to allow

presentation of semiotic information in spoken corpora.

Along with such initiatives, multi-media corpora or multi-modal corpora are being developed (e.g. Blache et al. 2009; Carter & Adolphs, 2008) and several attempts have been made to provide the framework and coding of the multimodal aspects of the recorded spoken discourse (Allwood et al. 2007; Allwood & Ahlsen, 2009; Colletta et al. 2009; Carter and Adolphs, 2008, Ekman et al. 2002; Matsuoka et al. 2009). Depending on the research aims, the coding features differ from study to study. For instance, Colletta et al. (2009) attempted to annotate the multi-modal behaviour of children by tracking language and gestures they produced. Other studies include Allwood and Ahlsen (2009) who conducted a study investigating the intercultural differences of multi-modal features for “the parameters to be taken into account in designing and evaluating a system for multimodal intercultural ICT” (Allwood and Ahlsen, 2009: 173). Blache, et al. (2009) reported a project (i.e. The ToMa project) on the creation of multimodal corpora and discussed ways to exploit such corpora. Difficulties still remain in terms of the automatic or manual annotation of the data, which has delayed further expansion of this area of corpus linguistics.

While such initiatives exist, none of them are aimed directly at the teaching of young learners. This would seem to indicate the need for further research into how such technology can benefit EFL education for young learners. This paper proposes a framework for a multi-modal corpus application designed for teaching young learners in an EFL setting. The first part of this paper reviews the literature on the relationship between young learners and digital media, and how the technological environment around children is changing its implication to their learning. This is followed by a discussion of corpus-based language learning with a special focus on data-driven learning (DDL) (Johns, 1997). The framework of a Multi-modal Corpus Tool (MmCT) is described with details of the software application. The final section presents a summary of this paper with possible future directions of this research project.

Young learners and digital media

In recent years, children have been exposed to different kinds of digital technologies from a very early stage of their lives. There is a tendency for many young children to have opportunities to encounter and engage with various technological devices and digital media (Marsh, 2010; McPake et al 2013). This

recent change in the environment around children both at home and in learning environments has implications for how young children experience discerning meaning (Levy, 2009; Yamada-Rice, 2011) and how they develop their early literacy skills (Burnett & Daniels, 2015).

The mainstream of research on children and ICT has tended to focus on the general effect of children's exposure to and their use of technological devices at home. The focus is often on a debate on whether such exposure poses detrimental effects on children, such as lack of sleep (Kondo et al. 2012), obesity and vision problems (Alper, 2011; Staiano and Calvert, 2012).

However, regarding the use of computers by young children, Plowman and Stephen (2003: 151) state that "there does not appear currently to be any clear evidence on the deleterious effects of exposure to ICT". Therefore, the use of digital technology does not seem to be hazardous, as long as sufficient precautions are taken (Endepohls-Ulpe et al. 2015). As Endepohls-Ulpe et al. (2015) suggest, it is necessary to focus on what children can learn from such technology. It is important to distinguish the usage of ICT at home for games or long-hours of watching TV from the use of such technology in an educational setting. Research concerning multi-modal learning in various digital contexts has focused primarily on how various media intersect to present information, specifically on the development of children's capacity to "read, create and share digital information" (Shuker & Terreni, 2013, cited in Binder et al. 2015: 114).

Adam and Wild (1997) report on a study which investigated the effect of multimedia CD-ROM storybooks on primary students' attitudes to reading. They (ibid: 122) describe CD-ROM interactive storybooks as a useful addition to digital technology as they employ several features which can potentially improve a reading process. Using such storybooks readers have the capacity to interact with the characters and surroundings, both audibly and visually through watching animation (Anderson, 1992), which is not possible in traditional paper-based media. Such storybooks also give readers greater control of the reading process, as they can choose whether to read the story or have it read to them with highlighted subtexts. They can also gain further understanding of language by clicking on unknown words and hearing them spoken (Anderson, 1992: 65). The study conducted by Adam and Wild (1997) indicated that the use of technology has encouraged unwilling readers to read more after a technology-mediated classroom.

Binder et al. (2015: 90) report similar results by incorporating the use of blogging with classroom-based teaching which "afforded children the opportunity to practice making sense of multimodal texts and also to create texts of their own through expressing and responding to a receptive audience". When engaging in

activities with digital media, children need to “draw on their traditional literacy practices to decode not only print, but to negotiate diverse interactive and participatory modes of representation that frequently include images and videos and are authored by individuals from a broad range of backgrounds, locales and expertise” (ibid, 2015: 94).

Since children’s increasing experience with digital media and technology is embedded in the “textual landscape” (Kress, 2013: 166) as well as from other experiences, it is impractical to neglect this influence. Rather, it is important for educators to utilize their experience of understanding meanings with the use of technological devices in their teaching. As noted earlier, young learners’ increasing exposure and their engagement with digital media (McPike et al. 2013) affects how they frame early literacy (Burnett & Daniels, 2015).

Understanding the process of meaning-making through digital media needs further research. Balajthy (1989) reports that computer-based learning has a positive effect on students who had previously difficulty or failure. As was earlier the case when CD-ROM was introduced into TEYL (Anderson, 1992; Adam and Wild, 1997), the use of multi-modal text contributes to children’s overall comprehension (Binder et al, 2015).

Corpus-based teaching and teaching English to young learners

Corpora have influenced various aspects of ELT, in terms of dictionary production, material and syllabus design, and even in classroom use. The application of corpora appears to be employed especially in the field of teaching English for Specific Purposes (ESP) or English for Academic Purposes (EAP) (e.g. Jabbour, 1998) and in the application of ‘specialised’ corpora in teaching (e.g. Flowerdew, 2001; Johns 1989, 1991; Stevens, 1991; Tribble, 2001). While the corpus used for a general ELT course needs to ensure that the corpus represents “a wider cross-section of registers and genres” (Flowerdew, 2001: 72), the corpus used in teaching English for special purposes (such as ESP and EAP) needs to represent language related mainly to the discipline in focus, taking learners’ specific needs into consideration.

While some methods in applying corpus-based research findings are available for adults (e.g. Tribble & Jones, 1997: 36; Scott & Tribble, 2006), very little research has been conducted on how the electrically-stored texts can be utilized in pedagogy from the perspectives of ‘applied’ corpus linguistics. One of the

applications of corpora to the ELT classroom pedagogy was termed by Johns (1991) as “data-driven learning” (DDL). This approach is rather inductive, having the characteristic of relying mostly on “learners’ intelligence to find answers” (Johns, 1991: 12) by consulting evidence from corpora (i.e., concordance lines in the Key-Word-In-Context format). On these grounds, it was assumed to be appropriate for adult learners (Johns, 1991: 12). Some research reports the effectiveness of DDL in EFL settings in facilitating learners’ consciousness of patterns and lexical items (Lee, 2006; Sripicharn, 2002). Sripicharn (2002: 408) suggests future research to test DDL with students at a lower level of proficiency, with a suggestion that the DDL materials should be rewritten to suit the level of the students. Moreover, Lee (2006) suggests that the use of literature corpora along with DDL or other methods such as CALL (Computer Assisted Language Learning) can promote students’ understanding of vocabulary and extended texts, and thus improve their reading skills. However, the direct use of corpus data in the form of DDL, is often limited to adult learners (e.g. Johns, 1991).

In the context of first language (L1) teaching with children, Sealey and Thompson (2004; 2006; 2007) conducted a study aimed at identifying the possibility of using a corpus with British primary school children as a part of their literacy curriculum, and investigated how young learners respond to corpus-based activities. They used the CLLIP (Corpus-based Learning of Language In Primary school) corpus which comprises of texts written for children extracted from the British National Corpus (BNC). They employed concordances produced from their CLLIP corpus, which was used to promote metalinguistic awareness among children in the L1 context, in order to lead children to discover and identify patterns in English. The concordances were initially colour-coded according to word classes by means of the concordance software, SARA (Dodd, 2001), integrated with the BNC. Sealey and Thompson (2004: 88) report that colour seemed to help children to notice the grammatical words and describe the difference between lexical words in concordance lines.

While corpus-based language learning is being conducted for L1 English young learners (Sealey and Thompson, 2004, 2006, 2007; Sealey, 2009, 2011), there are very few attempts at data-driven learning being used in the case of teaching young learners in L2 pedagogy. One reason is that it is traditionally believed that the language learning with young learners takes place in interaction and therefore L2 young learners’ teaching is focused very much on teaching in the classroom with a predominant emphasis on speaking and listening. Drawing on the evidence from the positive outcome in corpus-based L1 learning (Sealey and Thompson, 2004, 2006, 2007; Sealey, 2009, 2011), it seems to indicate some

prospects for introducing a corpus-based approach to the teaching of young EFL learners. For instance, the study by Sealey and Thompson (2004) clearly indicates how much additional support there was from the layout (i.e., the colour-coded concordance lines), which made the concordances 'accessible' to children. As Thompson and Sealey (2007) point out, the analysis of a corpus should help to highlight the kinds of patterns in language which learners may need to familiarise themselves with as they learn. Tsui (2004: 40) suggests that teachers use corpus findings concentrating on high frequency words, i.e., which are 'usual' rather than 'exceptional' usages, especially for elementary and intermediate levels.

With a different environment from first language education and the characteristics of teaching young learners, it is important to take into consideration the aspects of 'learnability', 'usefulness' and the 'accessibility', as well as the choice of data when considering a pedagogical application of corpus findings. Considerations such as those mentioned above are particularly important when designing a corpus for teaching English to young learners. Moreover, in the case of young learners who are learning English as a foreign language, it may be too demanding to learn through data-driven learning with the use of a traditional textual corpus.

Although the understanding of children's meaning-making process through digital media needs to be clarified more, as noted earlier, the meaning can be addressed multi-modally in real life. For these reasons, it can be argued that having the multi-modal corpus for young learners would benefit especially L2 young learners as the multi-modal information (i.e. visual and textual presentation with audio) assists young learners' comprehension of meaning. This is somewhat different from the CD-ROM based storybooks described earlier as it allows the identification of keywords from different scenes, if any, and corresponding multi-modal information at the same time to be presented on the screen.

It should be noted that the importance of interactional elements in children's language learning needs to be accounted for, and it is not the intention of this study to suggest the use of this application alone in a classroom. Interaction is an important element of language learning. For instance, it has been suggested that learners are more likely to notice grammatical information in recasts (i.e. 'corrective reformations of children's utterances that preserve the child's intended meaning') provided by teachers (Long & Robinson, 1998: 25). Studies conducted with L1 learners (Baker & Nelson, 1984; Farrar, 1992) and L2 learners (Oliver, 1995; Ortega & Long, 1997) all suggest that the use of recasts is more effective for the acquisition of grammatical information than the models presented to the learners. Therefore, corpus-based activities may provide one aspect of

interactional activities in the classroom.

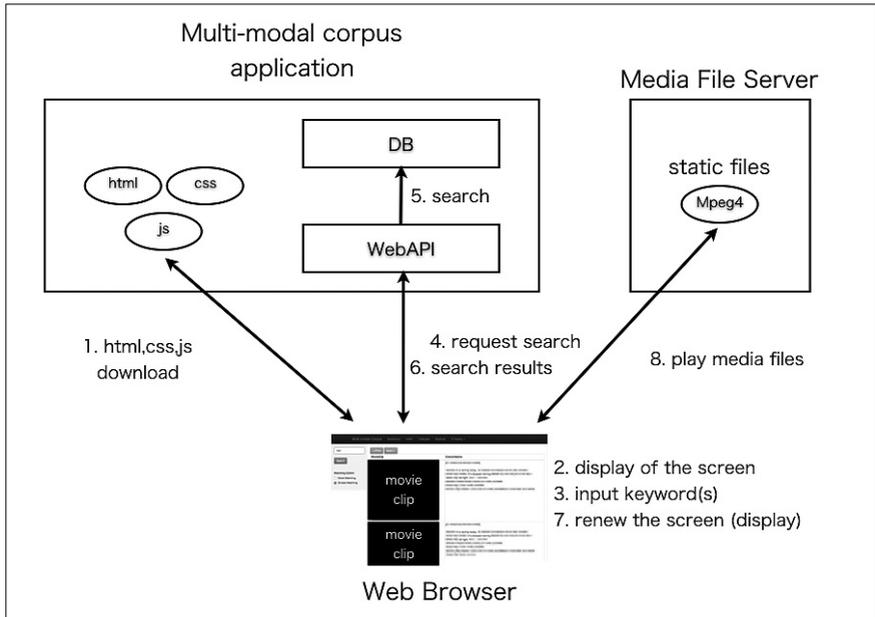
Framework of the Multi-modal Corpus Tool (MmCT)

As discussed earlier, having a multi-modal corpus tool designed for young learners can be useful in terms of assisting young learners meaning-making while learning and interacting through the use of application software. In this section, the framework of the multi-modal corpus application (MmCT) is described in detail.

Generally, the application software can be largely divided into two types: one is a web application, and the other is a desktop application. The web application is a program that receives the request through a web browser (e.g. Internet Explorer, Chrome, Safari, FireFox) and responds on Web UI (Web User Interface). The desktop application is software installed on a certain device (e.g. PC, tablet computer). Each type has advantages and disadvantages in its own right. For instance, a desktop application does not need to be connected to the internet, but the software needs to be developed separately according to the different operating systems (e.g. Linux, Microsoft Windows, MacOS) (Mikowski and Powell, 2013). In addition, operating systems are updated regularly, which requires adjustments to the desktop application accordingly. On the other hand, with a web application, users can utilize the application independently of the operating system of their device, simply with the requirement of a web browser and internet access (Charland and Leroux, 2011). With the web application, it is also possible to limit the availability within the devices which share the same WiFi access (Mikowski and Powell, 2013).

Considering the advantages and disadvantages described, it can be argued that it is more flexible to use a web-based application, as it would work with any kind of operating system, as long as the device has a web browser installed and enabled with access to the internet. Moreover, access can be limited within the devices by use of a local area connection (e.g. <http://localhost/xxx>) through a web browser without internet access, as long as the web application is installed on the devices (Barr, 1996; Oppliger, 2001). For the reasons mentioned above, in this current study, the development of a web application version of the tool is suggested. Figure 1 shows the overall framework of the Multi-modal Corpus Tool (MmCT) for this study.

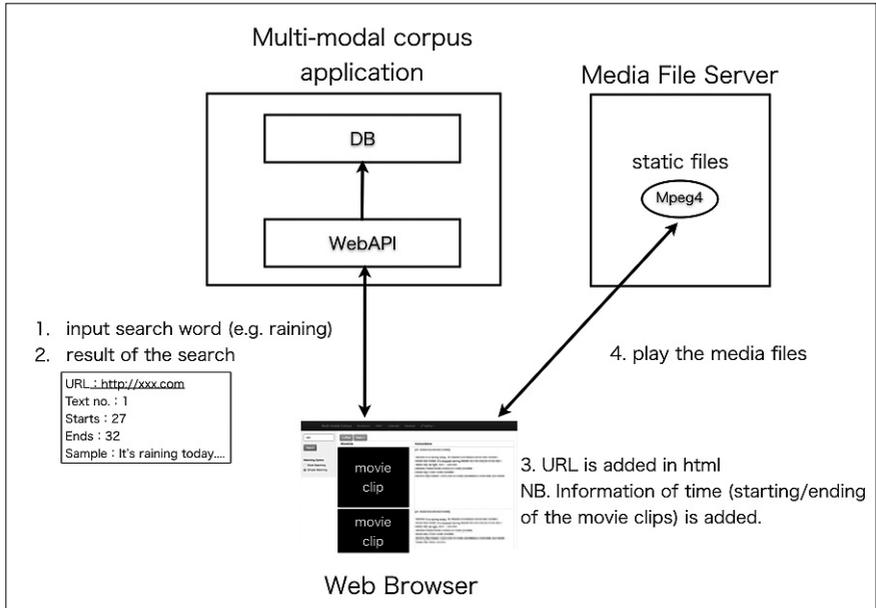
Figure 1. The framework of the MmCT web application



Once the application is installed on the device, the search flow takes the following steps before displaying the appropriate concordances and the location of the video/audio information based on the search:

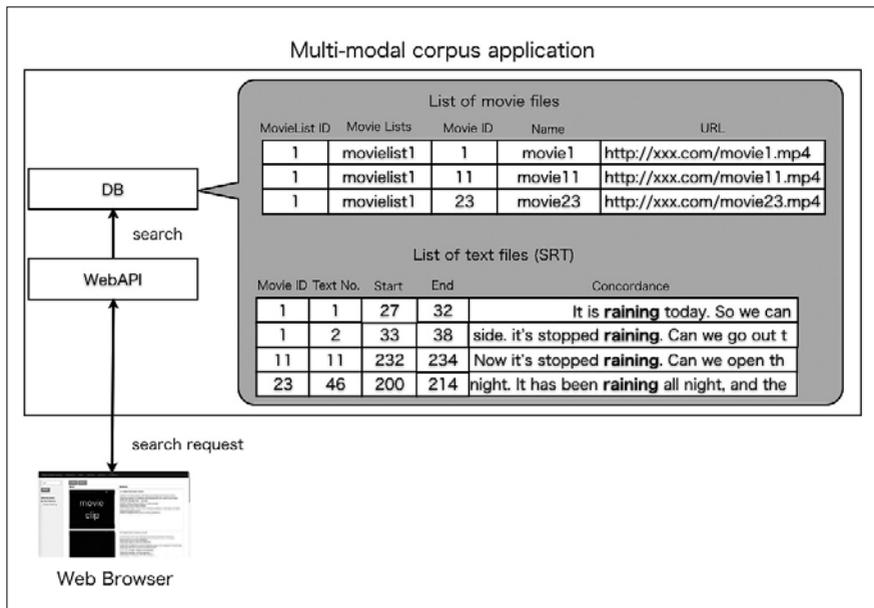
1. Once the URL of the multi-modal application (i.e. Multi-modal Corpus Tool (MmCT)) is entered in the browser, the browser sends the request to the MmCT to retrieve and display WebUI (HTML, CSS, JavaScript), and the search screen is displayed;
2. A user inputs the search keyword(s) in WebUI on the browser, and presses the [search] button;
3. JavaScript identifies the search keyword(s) being entered;
4. JavaScript sends requests to the MmCT to search the keyword(s);
5. MmCT searches the keyword(s) entered in the search box, and retrieves them from the database (DB) (e.g. search word: raining; for more details, see Figure 2 and 3);

Figure 2: The process of the display of media files



6. MmCT sends back the results identified in the DB to the web-browser;
7. WebUI displays the results with the information received from MmCT, which include the concordance output (text) and the URLs of the media files, and these are added in the HTML in the browser; and
8. The web browser obtains the media files from the URLs and a user is able to play the files by clicking the [play] button.

Figure 3: Identification of the textual and visual data from the database



Not only displaying the textual outputs of a certain word or phrase in the corpus, MmCT enables the viewing of movie clips in which the identified keyword(s) appear in different scenes. In addition, the concordance output and the words occurring closely to the node (i.e. the search word(s)) can be sorted alphabetically so that it helps the patterns of the language become visible. Sealey and Thompson (2006) report that the function of colour highlights assisted children to identify the characteristics of the language and thus promoted their meta-linguistic awareness. The function of sorting the concordance and adding the colour depending on the position to the node is also employed. With the use of MmCT, it is possible to conduct DDL activities with children and provide them with opportunities to practise making sense of multi-modal texts through the intersection between the text, images and audio.

Summary and Future research

Young learners' aptitude for digital media means that this is an area which can no longer be excluded from TEYL. Corpus linguistics is one area for further

study in order to make the most of both current technology, and children's growing ability with it. This paper reports on an ongoing project, aiming to create a multi-modal corpus application tool for teaching young learners, and a framework of the suggested multi-modal corpus tool is presented. Further research needs to be conducted on how to maximise opportunities for children's learning. Such challenges include:

- 1) further specification of data to be included in a multi-modal corpus for TEYL, and clearance of copyright issues involving the use of data;
- 2) specification of coding multi-modal features and methods;
- 3) further development of child-friendly interfaces by drawing on existing research on how children interact with ICT (e.g. haptic interactions such as pointing and dragging, Joiner et al. 1998) in order to make corpus-based activities more accessible to children; and
- 4) further investigation into the application of corpus linguistics in TEYL.

References

- Adam, N. and Wild, M. (1997). Applying CD-ROM interactive storybooks to learning to read. *Journal of Computer Assisted Learning* 13(2): 119-132.
- Adolphs, S. and Carter, R. (2013). *Spoken corpus linguistics: From monomodal to multimodal*. London: Routledge.
- Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. & Paggio, P. (2007): The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Proceedings of the International Language Resources and Evaluation Conference (LREC) 41(3-4)*, 273-287.
- Allwood, J. and Ahlsen, E. (2009). Multimodal intercultural information and communication technology: A framework for designing and evaluating multimodal intercultural communicators. In M. Kipp, J.-C. Martin, P. Paggio, D. Heylen (Eds.), *Multimodal Corpora: From models of natural interaction to systems and applications*. (pp. 160-175). Berlin: Springer.
- Alper, M. (2011). Developmentally appropriate new media literacies: supporting cultural competencies and social skills in early childhood education. *Journal of Early Childhood Literacy*, 13: 175-97.
- Anderson, J. (1992). Living books and other books without pages. *Unicorn* 18(3): 64-67.
- Baker, N.D. & Nelson, K. E. (1984). Recasting and related conversational techniques for triggering syntactic advances by young children. *First Language* 5: 3-21.
- Balajthy, E. (1989). *Computers and reading: Lesson from the past and the technologies of the future*. New Jersey: Prentice Hall.
- Barr, D. (1996). Common DNS operational and configuration errors. Retrieved 20th December, 2015 from <http://tools.ietf.org/pdf/rfc1912.pdf>.

- Binder, M. J., Sorin, R., Nolan, J., and Chu, S. (2015). Multimodal meaning-making for young children: Partnerships through blogging. In S. Garvis and N. Lemon (Eds.). *Understanding digital technologies and young children*. Oxon: Routledge. (pp. 92-111).
- Blache, P., Bertrand, R., and Ferrè, G. (2009) Creating and exploiting multimodal annotated corpora: The ToMa project. In M. Kipp, J.-C. Martin, P. Paggio, D. Heylen (Eds.). *Multimodal Corpora: From models of natural interaction to systems and applications*. (pp. 38-53). Berlin: Springer.
- Burnett, C. and Daniels, K. (2015). Technology and literacy in the early years: Framing young children's meaning-making with new technologies. In S. Garvis and N. Lemon (Eds.). *Understanding digital technologies and young children*. Oxon: Routledge. (pp. 18-27).
- Carter, R. and Adolphs, S. (2008). Linking the verbal and visual: new directions for corpus linguistics. In A. Gerbig & M. Mason (Eds.). *Language, people, numbers: Corpus linguistics and society*. Amsterdam: Rodopi. (pp.275-291).
- Charland, A., & Leroux, B. (2011). Mobile application development: web vs. native. *Communications of the ACM*, 54(5): 49-53.
- Colletta, J.-M., Kunene, R. N., Venouil, A., Kaufmann, V., & Simon, J.-P. (2009). Multi-track annotation of child language and gestures. In M. Kipp, J.-C. Martin, P. Paggio, D. Heylen (Eds.), *Multimodal Corpora: From models of natural interaction to systems and applications*. (pp. 54-72). Berlin: Springer.
- Dodd, T. (2001). *SARA: Technical Manual*. Oxford University Computing Services. Available from: <http://info.ox.ac.uk/bnc/sara/TechMan/>.
- Ekman, P., Friesen, W.V. & Hager, J.C. (2002): *Facial Action Coding System: The Manual on CD ROM*. Salt Lake City: A Human Face.
- Endepohls-Ulpe, M. Quaiser-Pohl, C. and Deckers, C. (2015) Availability and use of personal computers in German kindergartens: preconditions and influences. In S. Garvis and N. Lemon (Eds.). *Understanding digital technologies and young children*. Oxon: Routledge. (pp. 112-121).
- Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology* 28: 90-98.
- Flowerdew, J. (2001). Concordancing as a tool in course design. In M. Ghadsessy, A. Henry & R. L. Roseberry (eds.). *Small corpus studies in ELT: theory and practice*. Amsterdam: John Benjamins, pp. 71-92.
- Garvis, S. (2015). Digital technology and young children's narratives. In S. Garvis and N. Lemon (Eds.). *Understanding digital technologies and young children*. Oxon: Routledge. (pp. 18-27).
- Gerbig, A. & Mason, M. (2008) (Eds.). *Language, people, numbers: Corpus linguistics and society*. Amsterdam: Rodopi.
- Ho, C. M. L., Anderson, K. T., Leong, A. P. (2012). (Eds.). *Transforming literacies and language: Multimodality and literacy in the new media age*.
- Joiner, R, Messer, D. Light, P and Littleton, K. (1998). It is best to point for young children: a comparison of children's pointing and dragging. *Computers in Human Behavior* 14(3): 513-529.
- Jabbour, G. (1998). *Corpus Linguistics, contextual collocation and ESP syllabus creation: A text analysis approach to the study of medical research articles*. Unpublished PhD Thesis, University of Birmingham, Birmingham, UK.
- Johns, T. (1989). Whence and whither classroom concordancing? In T. Bongaerts, P. de-Haan,

- S. Lobbe & H. Wekker (eds.). *Computer applications in language learning*. Dordrecht: Foris Publications, pp. 9-33.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *ELR Journal* 4: 1-16.
- Kondo, Y., Tanabe, T., Kobayashi-Miura, M., Amano, H., Yamaguchi, N., Kamura, M., Fujita, Y. (2012). Association between feeling upon awakening and use of information technology devices in Japanese children. *Journal of Epidemiology*. 22(1): 12-20.
- Kress, G. (2003). *Literacy in the new media age*. London: Routledge.
- Lee, H. (2006). *Using English children's literature in particular Arthur Ransome's Swallows and Amazons for teaching EFL in Taiwan*. Unpublished PhD Thesis University of Birmingham, Birmingham, UK
- Levy, R. (2009). "You have to understand words...but not to read them": young children becoming readers in a digital age. *Journal of Research in Reading*, 32(1): 75-91.
- Long, M. H. & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams. (eds.). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, pp.12-41.
- Matsuoka, Y. Katagiri, Y. Ishizaki, M. & Enomoto, M. (2009). Unsupervised clustering in multimodal multiparty meeting analysis. In M. Kipp, J.-C. Martin, P. Paggio, D. Heylen (Eds.). *Multimodal Corpora: From models of natural interaction to systems and applications*. (pp. 93-108). Berlin: Springer.
- McLean, K. & Edwards, S. (2015). Beginning the conversations about young children's engagement with technology in contemporary times. In S. Garvis and N. Lemon (Eds.). *Understanding digital technologies and young children*. Oxon: Routledge. (pp. 155-165).
- Marsh, J. (2010) Childhood culture and creativity: a literature review. Newcastle: Creativity, Culture and Education. Retrieved 15th December, 2015 from <http://www.creativitycultureeducation.org/wp-content/uploads/CCE-childhood-culture-and-creativity-a-literature-review.pdf>
- McPake, J., Plowman, L., and Stephen, C. (2013). Pre-school children creating and communicating with digital technologies in the home. *British Journal of Educational Technology*, 44 (3): 421-31.
- Mikowski, M. S. & Powell, J. C. (2013). *Single page web applications: JavaScript end-to-end*. New York: Manning Publications.
- Oliver, R. (1995). Negative feedback in Child NS/NNS conversation. *Studies in Second Language Acquisition* 17: 459-482.
- Oppliger, S. (2001). *Internet and intranet security*. Second edition. MA: Artech House.
- Ortega, L., & Long, M. H. (1997). The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish. *Spanish Applied Linguistics* 1(1): 65-86.
- Plowman, L. and Stephen, C. (2003). A "benign addiction?" Research on ICT and pre-school children. Research on ICT and pre-school children. *Journal of Computer Assisted Learning*. 19:149-64.
- Robinson, L. (2003). Technology as a scaffold for emergent literacy: interactive storybooks for toddlers. *Young Children*. 58: 42-48.
- Scott, M., & Tribble, C. (2006). *Textural patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

- Sealey, A. (2009). Exploring vocabulary with young L1 learners: the contribution of a corpus. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (eds.). *Vocabulary studies in first and second language acquisition: The interface between theory and application*. Basingstoke: Palgrave Macmillan, pp. 39-58.
- Sealey, A. (2011). The use of corpus-based approaches in children's knowledge about language. In S. Ellis, and E. McCartney (eds.) *Applied linguistics and primary school teaching*. Cambridge University Press, pp. 93-106.
- Sealey, A., & Thompson, P. (2004). "What do you call the dull words?": Primary school children using corpus-based approaches to learn about language. *English in Education* 38(1): 80-91.
- Sealey, A., & Thompson, P. (2006). "Nice things get said": Corpus evidence and the National Literacy Strategy. *Literacy* 40(1): 22-28.
- Sealey, A. and Thompson, P. (2007). Corpus, concordance, classification: young learners in the L1 classroom. *Language Awareness* 16(3): 208-223.
- Segars, E. and Verhoeven, L. (2002). Multimedia support of early literacy learning. *Computers & Education* 39: 207-221.
- Shuker, M.-J. and Terreni, L. (2013). Self-authored e-books: expanding young children's literacy experiences and skills. *Australasian Journal of Early Childhood*, 38(3): 17-24.
- Sripicharn, P. (2002). *Evaluating Data-driven Learning: the use of classroom concordancing by Thai learners of English*. Unpublished PhD Thesis. University of Birmingham, Birmingham, UK.
- Staiano, A. E. & Calvert, S. L. (2012). Digital gaming and pediatric obesity: At the intersection of science and social policy. *Social Issues and Policy Review*, 6(1): 54-81.
- Stevens, V. (1991). Classroom concordancing: Vocabulary materials derived from relevant, authentic text. *English for Specific Purposes* 10: 35-46.
- Thompson, P., & Sealey, A. (2007). Through children's eyes? Corpus evidence of the features of children's literature. *International Journal of Corpus Linguistics* 12(1): 1-23.
- Tribble, C. (2001). Small corpora and teaching writing. In M. Ghadsessy, A. Henry & R. L. Roseberry (eds.). *Small corpus studies in ELT: theory and practice*. Amsterdam: John Benjamins, pp. 71-92.
- Tribble, C. & Jones, G. (1997). *Concordances in the classroom*. Houston, Tx: Athelstan.
- Tsui, A. M. (2004). What teachers always wanted to know-how corpora can help. In Sinclair, J. (ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, pp. 39-61.
- Yamada-Rice, D. (2011). New media, evolving multimodal literacy practices and the potential impact of increased use of the visual model in the urban environment of young children's learning. *Literacy*, 45(1): 32-43.