



Jan Kochanowski University Press

This is a contribution from *Token: A Journal of English Linguistics*
Volume 15/2022.

Edited by John G. Newman, Marina Dossena.

© 2022 Jan Kochanowski University Press.

Kids in the House: How the U.S. House of Representatives addresses youngsters

Gianmarco Vignozzi

University of Pisa

ABSTRACT

Educational websites for children are becoming more and more popular as sophisticated sources of specialized content dissemination. *Kids in the House*, the website for children created and run by the U.S. House of Representatives, is an interesting example as it is distributed in four different versions targeting preschoolers, grade schoolers, middle schoolers and high schoolers. This contribution aims to analyze quantitatively, qualitatively and comparatively some of the most recurring linguistic features that characterize the four website versions, by concentrating on verbal, and occasionally nonverbal, popularizing strategies. The analysis reveals that there are marked differences among the website versions both in terms of lexical complexity and discourse. Also, verbal and nonverbal popularizing devices seem to closely reflect the needs and tastes of the target audiences.

Keywords: educational websites for children, popularization for children, verbal and nonverbal language, corpus analysis.

1. Introduction: Educational websites for children

This article intends to inquire into how the U.S. House of Representatives explains its principal duties and how it functions to youngsters of different age groups who have reached specific educational stages. In particular, the object of the analysis is the official website for children created and run by the Office of the Clerk, namely *Kids in the House*.¹

¹ As regards the website “[t]he Office of the Clerk provides this website as a public service. The information on this site is considered public information and may be

As the UNICEF website states (UNICEF 2019), in the digital, media-saturated environment we are living in, “like adults, children use the Internet for their right to information”. Indeed, in the last two decades different types of educational website especially designed to convey sectoral information to children (e.g., in the fields of science, sociology, ecology, politics) have started to flourish, with English-speaking countries playing a pioneering role (Buckingham – Scanlon 2004; Diani 2020; Diani – Sezzi 2019, 2020). In line with what has been described in seminal studies that considered specialized knowledge dissemination in different types of texts targeting adult laypeople (cf. *inter alia* Calsamiglia 2003; Calsamiglia – van Dijk 2004; Garzone 2020), specialized knowledge in child-oriented websites is carefully moderated and tends to undergo a meticulous process of recontextualization and mediation in order to be accessible and appreciated by this special group of receivers. Children, in fact, are a particularly demanding target audience as they represent the quintessence of the lay reader (or better lay user) for whom the capacity to decode opaque expert information is further complicated by their not fully completed cognitive development and their partial encyclopaedic knowledge (Diani 2015; Bianchi 2018; Bruti – Manca 2019; Cappelli – Masi 2019; Sezzi 2019). That is why, both verbal and nonverbal knowledge dissemination strategies need to be tactically deployed to translate and “digest” the package of information and create a multimodal environment suited for children’s requirements, which should be entertaining as much as it is educationally stimulating.

Unlike grown-ups, youngsters are generally not merely looking for informational materials on a website. They are looking for a virtual space where they can have fun and, possibly, discover new content that could also be useful at school. In other words, when looking for specialized information about something, they most often seek edutainment contents that are more than just informative, but combine entertainment, engagement and education in a hybrid multi-layered text type (Buckingham – Scanlon 2004). Indeed, if an educational website is not able to grab and hold their attention, the risk is that children will just move on to other web content, which is nowadays very easy, given the large number of options available in the digital(ized) world.

A recent survey on children’s perception and consumption of websites in the USA (Sherwin – Nielsen 2019) suggested that, in order to be successful,

distributed or copied unless otherwise specified. Images on this site are provided as a contribution to education and scholarship”. All the images used throughout this paper are in the public domain (<https://kids-clerk.house.gov/privacy-policy.html>, accessed September 2021).

educational websites are required to follow certain usability approaches (e.g., the use of guided and interactive learning paths) and narrowly targeted content. This is because children, as digital native literates, are becoming more and more web-proficient, demanding, and aware of age differences in content design. Thus, they will not engage with something they perceive as too babyish, and they will be discouraged from interacting with a website they find too challenging and thus unappealing.

In line with these ideas, *Kids in the House*, was released in four different versions each of which targeting children at different stages in the U.S. educational system, i.e., young learners (3 to 6 years), grade schoolers (6 to 11 years), middle schoolers (11 to 14 years), and high schoolers (14 to 18 years). In the present study, I therefore aim to examine quantitatively and qualitatively (through comparison) some of the most recurrent linguistic features in each of these website versions, by concentrating in more detail on lexical features and on the popularizing strategies and devices, both at a verbal and nonverbal level, selected to convey specialized information and cater for the needs of children at different educational stages.

The article is organized as follows: Section 2 describes in detail the corpus under analysis and outlines the methods used for the analysis discussed in Section 3 and Section 4. Section 3 presents a corpus analysis of the most salient lexical features of the four website versions, and Section 4 expands on the popularizing tools selected to effectively reach the different target audiences. Some final remarks and suggestions for further research are gathered in Section 5.

2. Corpus description and methodology

The website analyzed here, i.e., *Kids in the House*, was launched in 2009 by the Office of the Clerk of the U.S. House of Representatives as a public service to promote and disseminate knowledge about the legislative branch of the United States Government among schoolchildren of all ages. The website is the result of an agenda set by President Bill Clinton in 1997 for the World Wide Web to be in every classroom by the year 2000. In fact, he wanted those websites sponsored by United States government agencies to be enriched and become a tool for students (and teachers) for retrieving reliable and accessible information on the principal institutions. Therefore, he suggested children's pages as a possible improvement (Clarke 2006).

As stated in the mission statement present on the main landing page, all information is meant to be both educational and entertaining

for the target audience, making the website a typical example of an edutainment or infotainment text (Djonov 2008). As previously suggested, the distinctiveness of this page as compared to other similar institutional websites for children is that from the navigation menu of the main page the user is given the opportunity to choose between four distinct versions aimed at schoolchildren of different age groups. The contents adapted for each website version are very similar and they are sorted into the same thematic macrosections that revolve around i) the functioning of the Congress, ii) how laws are made, iii) the works of art that can be found in the Capitol building and some interesting historical facts, iv) a guided tour of the Capitol, and v) some information about the Clerk of the House. This repetitive structure, which enhances thematic consistency for the benefit of the recipients who may want to engage with different website versions, makes the texts and the paratexts easily comparable. In general, the website has a drill-and-practice approach, i.e., it is based on systematic structural and content repetitions, as has also been noted by Diani and Sezzi (2019) for educational websites for children about the EU.

In order to carry out a quantitative linguistic analysis of the verbal language across the four website versions, all the textual content retrievable on the website was used to assemble a corpus (henceforth *KidsHouseCorpus*) with four sub-components. Table 1 offers an overview of the composition of *KidsHouseCorpus*.

Table 1. *KidsHouseCorpus* composition

Corpus components	Word tokens	Word types
<i>Young Learners</i>	911	281
<i>Grade School</i>	7,351	1,445
<i>Middle School</i>	10,385	1,951
<i>High School</i>	11,288	2,126
TOTAL <i>KidsHouseCorpus</i>	29,935	2,664

All the webpages from which the texts were retrieved were also saved and stored as PDF files so as to have access to nonverbal data as well.

Young Learners targets preschoolers from 3 to 6 years old who have not yet developed reading skills. As shown in Figure 1, the menu page is very cheerful, featuring an outdoor setting with a cartoon-like blue sky as the background and animated graphics below on the grass (an anthropomorphic smiling scroll of parchment and the capitol building) that not only adorn

the page to appeal to very young children, but they also frame, introduce, and meaningfully mirror the contents. The four macrosections into which the website is organized are presented by animated visual icons that again hint at the contents (e.g., *Around the Capitol* is introduced by a camera taking shots with the Capitol building in the background) and are accompanied by introductive short sentences.



Figure 1. *Young Learners* main page

As is evident from the low number of word tokens and types, texts within the thematic macrosections are kept to a minimum and rely more heavily on graphics. Interestingly, apart from short textual paragraphs, each section features colourful icons depicting either coloured pencils or scissors that are anchors sending the user to printable educational content. The website also contains a *Little-Known Fact* box that tells a different item of interest about the House of Representatives every time you refresh the webpage and a link to a glossary section that lists the definitions of seven technical words (e.g., *Bill*, *Capitol*, *Impeach*) found in the texts.

Grade School is designed for children in primary school aged from 6 to 11 years old (1st to 5th grade) who have just started to read texts on their own. The menu page is very different from the *Young Learners'*. Here all the graphic elements evoke the school setting: the banner headline is a blackboard sketch of the Capitol building and the Stars and Stripes Flag, the plain white background appears to be stapled to resemble a piece of paper, and the thematic sections are organized by four colourful pieces of paper

attached to the background with thumb tacks. The iconic illustrations for *How Laws Are Made* and *Around the Capitol* are the same as those for *Young Learners*, but here they are not animated. *What is Congress?* and *Art & History*, instead, are extensible sections that enclose different thematic subsections all accompanied by photos related to the contents (Figure 2).

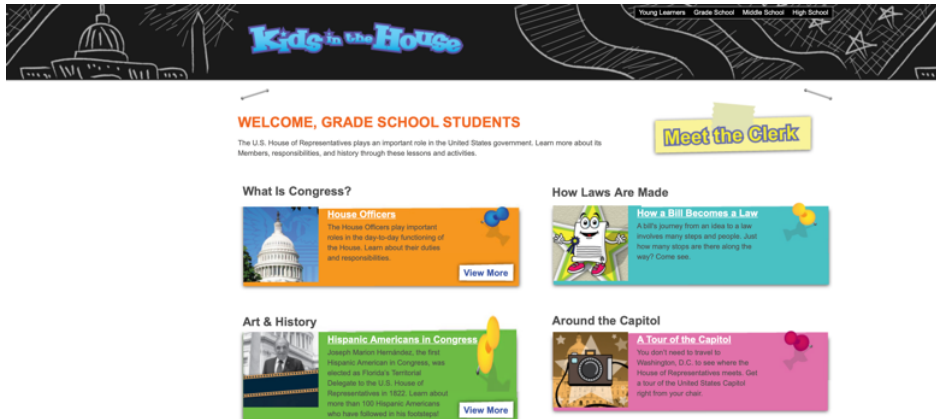


Figure 2. *Grade School* main page

Within the sections, the amount of written text is much greater, with numerous paragraphs, though short, for each topic. In this case, sectorial and potentially opaque words are highlighted and when placing the mouse over them their definitions automatically pop up. There is also a glossary page where all twenty-six terms are located. Educational activities are also embedded in the text and are linked to specific terms (e.g., the *Capitol Maze* is accompanied by a colouring picture).

Middle School addresses children within the age range of 11-14 (6th to 8th grade). The page has a more traditional layout with a thematic background representing stars and stripes (recalling the American flag). However, it is rendered more engaging for teenagers by the option *choose your style* allowing the user to select a different colour and style for the background. The macrosections are still sorted in colour bands, but some of the cartoon-like illustrations that adorn *Young Learners* and *Grade School* are substituted by more realistic thematic icons. As Figure 3 illustrates, the landing page also contains a *search* band that enables the user to search for specific contents on the website, a link to the *Glossary* page (here containing sixty-nine items), and a link to the *For Teachers* section. This last page gathers educational resources and activities specifically created on the basis of the contents of the site.

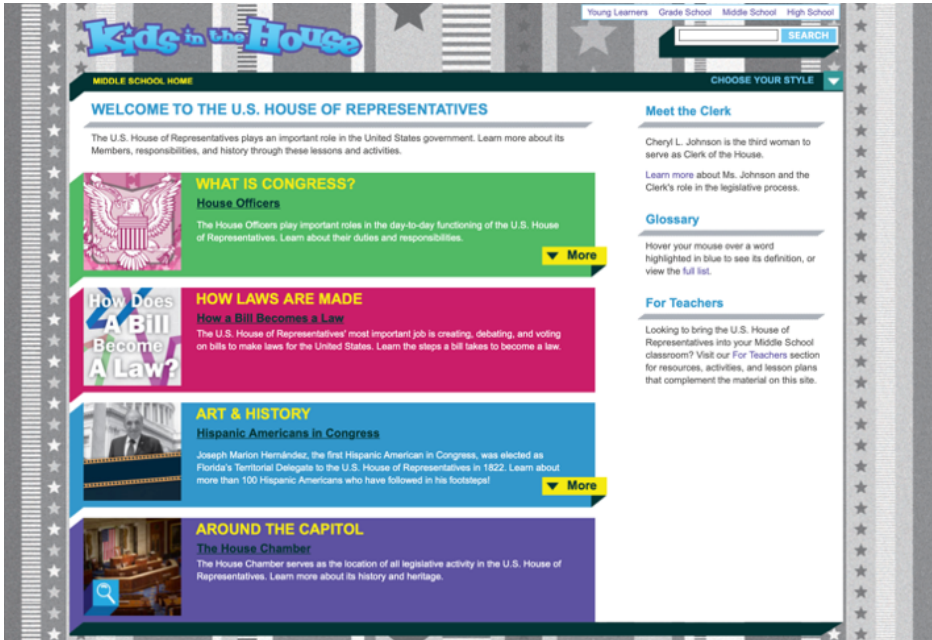


Figure 3. Middle School main page

The sections are structured in paragraphs with longer texts often arranged in bulleted lists. Besides the highlighted terms accompanied by the corresponding glossary definitions, other words in bold are direct links to different internal sections of the website. A new resource that is added to this website version is an apparatus of photos embedded within the text and accompanied by explanatory captions. For example, when talking about the *House of Chamber*, a picture of the House and a brief description of its functions are embedded within the text. A last noteworthy characteristic of all section subpages is the *Did You Know?* box, which is a more grown-up version, in a question/answer format, of the *Little-known Fact* present in the *Young Learners'* page.

The last website version is *High School*, a page aimed at and designed for adolescents from 14 to 18 years old (9th to 12th grade). The layout of the page is akin to the *Middle School* one, with a plain background with the American bald eagle symbol foregrounded at the top of the page and four monochromatic dark bands framing the thematic macrosections. Specific to this website version are the *What's New* box, which lists the latest contents published, and the link to a *Social Media* webpage, i.e., the official YouTube Channel of the Clerk of the House (Figure 4).

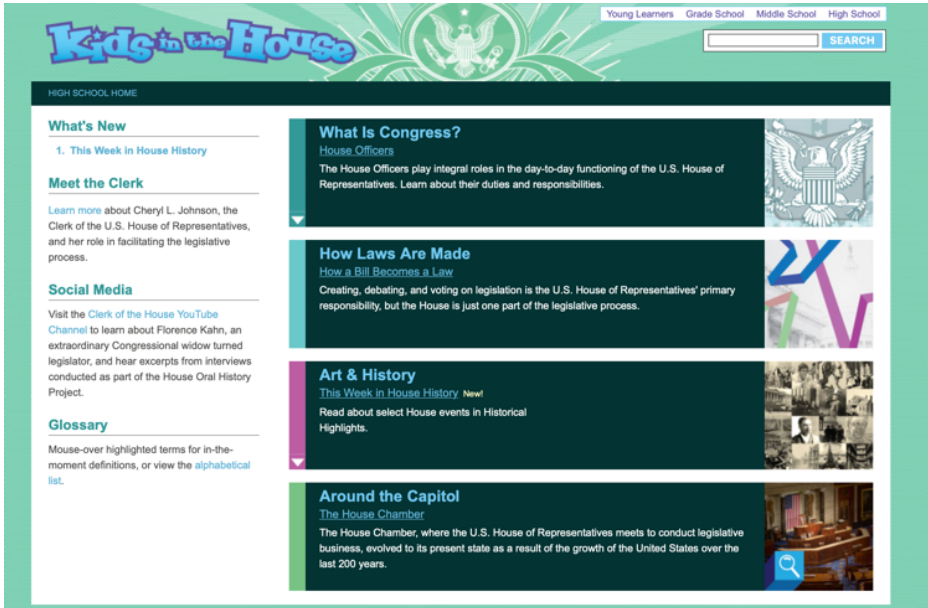


Figure 4. *High School* main page

The internal organization of the thematic sections is not dissimilar to *Middle School*, with the addition of a greater number of captioned photos, glossary terms, and a series of words highlighted in grey that are links to the official website of the House of Representatives and to other governmental websites, so that the user can examine certain topics in more depth. Accordingly, at the end of the texts in each macrosection there is an *Additional Sources* box with different references to external websites.

Having outlined the most relevant organizational features of the webpages, the following sections focus on the verbal contents, in particular the lexical features, of the *KidsHouseCorpus*. These are analyzed with some of the tools of corpus-assisted discourse analysis (cf. Baker 2006; Partington – Duguid – Taylor 2013 for an introduction to corpus-assisted methods) which will be introduced when relevant throughout the analysis. The corpus software used to process the data are: *Word Smith Tools* (Scott 2016), for calculating some measures of lexical complexity, i.e., the lexical variation and the lexical density, and *Sketch Engine* (Kilgarriff et al. 2014), for retrieving keywords and assessing concordances. The subsequent examination of the popularizing strategies draws on Calsamiglia and van Dijk's (2004) classification, which brings to the fore the rhetorical means used to turn "specialized knowledge into 'everyday' or 'lay' knowledge" (Calsamiglia – van Dijk 2004: 370). These include i) definition, to introduce unknown

words or concepts by briefly outlining the salient traits of the referent, ii) denomination, to assign specialized labels to specialized concepts, iii) description, through which additional information on the subject matter is added so as to enrich the reader's knowledge, iv) reformulation, to offer a more accessible paraphrase of a selected discourse fragment, v) exemplification, to provide the addressee with familiar examples to support the decoding and understanding of specific situations, vi) generalization, to exploit wide-ranging and general conclusions in order not to burden the receiver with technical information, vii) analogy, to resort to metaphors, similes, or other tropes to bridge old and new knowledge by means of figuration.

Relevant comments on significant nonverbal features associated with knowledge popularization are inspired by the tenets of multimodal discourse analysis according to Kress and van Leeuwen (2001).

3. Quantitative analysis of lexical features across website versions

3.1 Descriptive statistics

The first exploratory statistics carried out to describe the lexical complexity of *KidsHouseCorpus* and its components was the standardized type/token ratio (STTR), which is a standardized index computed by *Word Smith Tools* to measure the level of lexical variation of a text. The type/token ratio is often used in discourse analysis when describing the specificity of a particular text type or register (Baker 2006). Its value is obtained by dividing the number of unique words (i.e., types) by the number of running words (i.e., tokens), thus turning the result into a standardized and comparable percentage². The higher the result, the wider the variety of different words in a corpus.

The other measure that was computed to estimate lexical complexity across the corpus components is lexical density, a parameter that calculates the proportion of lexical words to grammatical words (Ure 1971). Hence, a high lexical density value indicates a large number of information-carrying words, whereas a low value denotes relatively few of them. The results of these analyses are displayed in Table 2.

² For an introduction to the operations carried out by *Word Smith Tools* to standardize the values, cf. https://lexically.net/downloads/version5/HTML/index.html?type_token_ratio_proc.htm.

Table 2. Lexical variation and lexical density in *KidsHouseCorpus*

<i>KidsHouseCorpus</i> components	Lexical variation (STTR)	Lexical density
<i>Young Learners</i>	34.35%	55%
<i>Grade School</i>	38.73%	61%
<i>Middle School</i>	41.09%	62%
<i>High School</i>	43.67%	65%

As clearly emerges from the table above, both lexical variation and lexical density progressively increase from *Young Learners* to *High School*, thus indicating the different degrees of complexity in terms of lexical diversity and vocabulary richness across the four website versions.

If we compare the percentages found for lexical variation, ranging from 34.35% to 43.67%, with the written and the spoken sections of the British National Corpus (BNC)³ that have a STTR of respectively 57.81% and 32.45%, it emerges that all *KidsHouseCorpus* subcomponents have values more similar to the spoken register, generally known for its repetitiveness and constrained vocabulary range. A comparable, but less marked, trend was noted by Bruti and Manca (2019), who computed STTRs values ranging from 43% to 48% for a collection of specialized corpora of British, American, and Italian science magazines for teenagers. By contrast, in terms of information content (i.e., the lexical density) the percentages of the *KidsHouseCorpus* components are decidedly on the high end of the 40-65% span for written non-fiction texts signalled by Stubbs (1996) on the basis of a large-scale study that also assessed that conversation hardly ever exceeds 40% of content carrying words.

Therefore, these statistics point to the fact that the texts in *KidsHouseCorpus* contain a rather restricted vocabulary whose range steadily increases from young learners to high schoolers. Notwithstanding this, they seem to be quite dense in informational content as testified by the significant presence of lexical words. These results could be telling of the nature of popularizing texts for children, for which the limited (and repetitive) amount of different lexis may be interpreted as a strategy to enhance textual coherence and cohesion. In fact, repetitiveness can ease decoding especially for children and can help them cope with the complexity, both at a linguistic and cognitive level, brought about by the specialized information load (Cappelli – Masi 2019).

³ The BNC is intended here as a general reference corpus to compare *KidsHouseCorpus* with spoken and written English.

3.2 Keywords

With the aim of continuing the analysis with an inductive approach, the four components of *KidsHouseCorpus* were compared through a keyword analysis. Keyword lists, i.e., statistically overused words in a corpus as compared to a reference corpus (cf. *inter alia* Scott – Tribble 2006), were created by using *Sketch Engine*. In this case, the non-queried components of the corpus were set as the benchmark. In other words, when extracting the keywords for *Young Learners* the other parts of *KidsHouseCorpus* were set as the reference corpus, so as to extract some particularly significant words for each subcomponent as compared to the others.

The analysis yielded numerous keywords for each corpus component, and it would be beyond the scope of this paper to consider them all. The following observations focus on some interesting cases, established through manual concordance reading, among the top-twenty most significant words for each website version. Starting from *Young Learners*, the most relevant among the function words, i.e., highly frequent grammatical words telling of text style (Flowerdew 2003), is ‘you’ that occurs as the third most significant element in the corpus. As amply suggested in the literature (cf. *inter alia* Diani – Sezzi 2019; Bruti – Manca 2019; Vignozzi 2020), the second person pronoun is often used in monologic discourse to directly address and involve the non-present audience by setting a dialogic, symmetrical, and friendly atmosphere (1).

- (1) In the House Chamber, where the Members of the U.S. House of Representatives meet, you will see eagles on special items like [...].

In this case, the direct receivers are children who are not able to read and navigate the internet on their own, consequently adults are inevitably addressed as well. They are what can be called the covert audience. In fact, various other lexical keywords (e.g., ‘help’, ‘print’, ‘grown-up’) point to the active role adults have in making the information presented accessible and enjoyable to a young audience. The interactional style of this website version is also testified by the keyword ‘hello’, an opening greeting which is repeatedly used at the beginning of paragraphs, before the explanations, to build rapport and establish an informal friendly tone. By contrast, the second person pronoun and the greeting rarely occur in the other website versions. As for lexical keywords, which are indicators of the aboutness (i.e., subject matter) of the text, the most salient one is ‘lawmaker’. This term

is specific to this corpus component as in the others the more formal and technical synonym 'legislator' is employed. 'Most' also emerged as key, as it is repeatedly used to generalize about some potentially specific and too detailed passages (2).

(2) Most laws in the United States begin as bills.

Moving on to grade schoolers, the first functional word in terms of keyness is the first plural person pronoun 'we', which often occurs when introducing denominations, perhaps as a strategy to reduce the distance between the expert/narrator and the receiver, and to turn the specialist information into everyday shared knowledge (3,4).

(3) In 1791, George Washington chose the 10 square miles offered by Maryland to become the capital. We know this area as Washington, D.C.

(4) When Congress first moved to the Capitol, the House of Representatives met in the room we now call Statuary Hall.

As for 'you' in the *Young Learners* subcorpus, the pronoun 'we' is specific to this corpus component. A further group of keywords is represented by 'stuff' and 'cool', two words pointing to a very informal tone reminiscent of the language used by teens. In terms of knowledge recontextualization, the keyword 'like' is particularly interesting as it is used in similes that help non expert readers correlate new knowledge with known experiences, which, in the case of the example below, parallels the functioning of the House Chamber with the school setting (5).

(5) First the Members elect the Speaker of the House. The Speaker is a lot like the class president.

Taking *Middle School* as the focus corpus, the most striking result is that the first item that was yielded in the keyword list is the punctuation mark '?', which points to the high occurrence of direct questions. Accordingly, the concordance lines confirm that on different occasions information is presented in a question/answer format (6), which is clearly a strategy used to arouse curiosity in the reader.

(6) So what happens on Congress's first day? The Constitution states that each new Congress must convene, or assemble, for the first time at noon on January 3 [...].

In addition, imperative verb forms seem to be particularly relevant in this corpus section, as testified by the keywords ‘see’, ‘learn’ and ‘get’, which are usually employed to stimulate users to acquire new information (7) and engage in educational activities.

(7) See how Representatives use committees to organize their work!

Some lexical keywords specific of the aboutness of the texts are also relevant, for example ‘delegates’ and ‘caucus’, both followed by their definitions, as described in more detail in the next section.

To conclude, the most striking result when assessing the keywords for *High School* is that the only function words occurring among the top-twenty most key items are the personal pronoun ‘they’, which especially occurs in longer and more complex sentences as an anaphoric referent (8), and the conjunction ‘and’, again signalling an additive style (i.e., hypotaxis).

(8) The committee members will seek expert input, hold “mark-up” sessions to make any changes or updates deemed important, and, if necessary, send the bill to a subcommittee for further analysis through research and hearings.

Indeed, most keywords in *High School* are sectorial nouns conveying specific contents (e.g., ‘committees’, ‘injunction’), adjectives employed in detail descriptions (e.g., ‘historical’, ‘pending’) and more formal verbs (e.g., ‘highlight’, ‘preserve’), all signals of a more academic style.

4. Exploring strategies for popularization

As declared in the mission statement and as confirmed by some of the keywords evaluated in the previous section, the intent to present knowledge in a simplified, and possibly amusing, way to children of all ages is the main endeavour of the website. Therefore, in what follows I concentrate on the analysis of the popularizing devices that are deployed to disseminate specialized knowledge to readers with different needs and tastes. By relying on the model put forward by Calsamiglia and van Dijk (2004), all verbal techniques of popularization in the four corpus components were singled out and classified. In order to be able to draw comparisons among corpora of different sizes, the quantitative results of this scrutiny into popularizing strategies are collected in Table 3 as percentages.

Table 3. Popularizing strategies in *KidsHouseCorpus*

Popularizing strategies	Young Learners	Grade School	Middle School	High School
<i>Definition</i>	13%	29%	30%	31%
<i>Denomination</i>	–	7%	16%	21%
<i>Description</i>	45%	34%	20%	16%
<i>Reformulation</i>	–	4%	21%	25%
<i>Exemplification</i>	33%	12%	9%	7%
<i>Generalization</i>	9%	5%	3%	–
<i>Analogy</i>	–	9%	1%	–

4.1 Young Learners

Starting from *Young Learners*, it is evident that popularization in texts is mainly based on (paratactic) descriptions (45%), which depict objects in a more extensive, but less specialized, way than definitions. Descriptions contain details considered unnecessary for experts but that help the lay reader (or in this case the hearer as pre-school children cannot read alone) picture potentially missing items in their encyclopaedic knowledge. Example (9) shows how a description is used to define the artistic and historic value of the United States Capitol building.

- (9) The United States Capitol is more than 200 years old. Many important people have walked its halls, including John Quincy Adams and Abraham Lincoln. It is also home to many beautiful and important works of art, such as paintings and statues.

Here the description first gives an idea of the epoch and of the importance of the building by concretely presenting its approximative age and two notable people related to the place, names which are very likely familiar to American kids. It then continues with some words on its artistic contents, whose reputation is rendered through positive everyday descriptive adjectives (e.g., ‘beautiful’, ‘important’) and is linked to the concrete sphere by means of two generic examples of categories of works of art. Furthermore, the description is also complemented by a ludic activity. Children are asked to open a link and print (with the aid of a grown-up explicitly encouraged) a colouring sheet representing a comic-like version of the most iconic painting in the Capitol: *The Apotheosis of Washington*, where angels are substituted by smiling

kids with George Washington in the middle, easily recognisable by his stern facial expression (Figure 5).

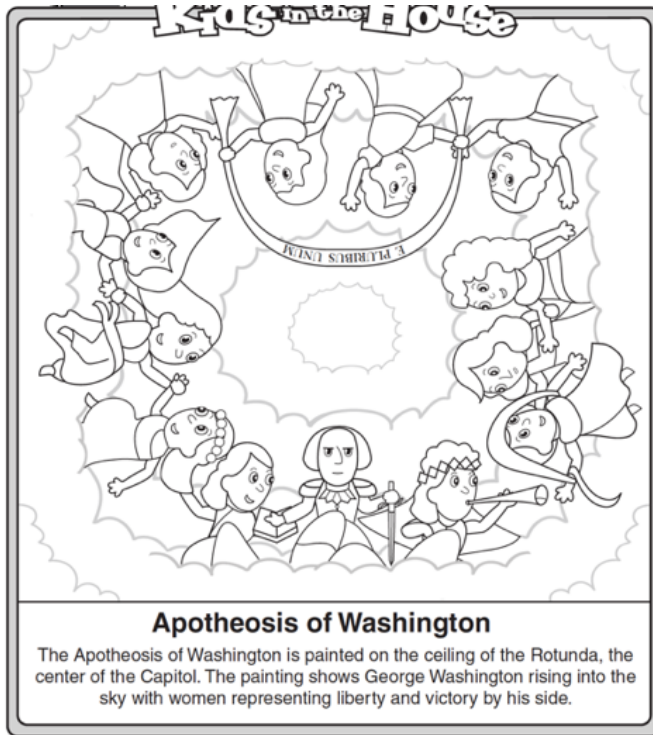


Figure 5. Colouring sheet for *The Apotheosis of Washington*

Hence, the colouring activity in this case not only offers a moment of relaxation and play but also illustrates and further clarifies the contents of the generic description.

Exemplifications (33%) are frequently used to enhance the comprehension of difficult concepts by providing down to earth and more specific illustrations of general and vague concepts that might be difficult for young children to grasp and concretize. The passage below (10) illustrates the examples provided in the text to explain what laws are.

- (10) The laws Congress makes help Americans. There are laws that say kids have to go to school, laws that protect animals and nature [...].

The scenarios selected (school and nature) are meant to be familiar and easily distinguished by children in as much as they are already part of their still developing knowledge of the world.

To a lesser extent, definitions (13%) occur to explicate some selected unknown words key to understanding the topic being examined. It is noteworthy that in many instances they are not expressed in the text but are contained in colouring sheets (Figure 6) linked to specific words, as example (11) shows for the term ‘hopper’.

- (11) Once the bill is written, it is placed in the hopper, and introduced to the rest of the Members of the U.S. House of Representatives.

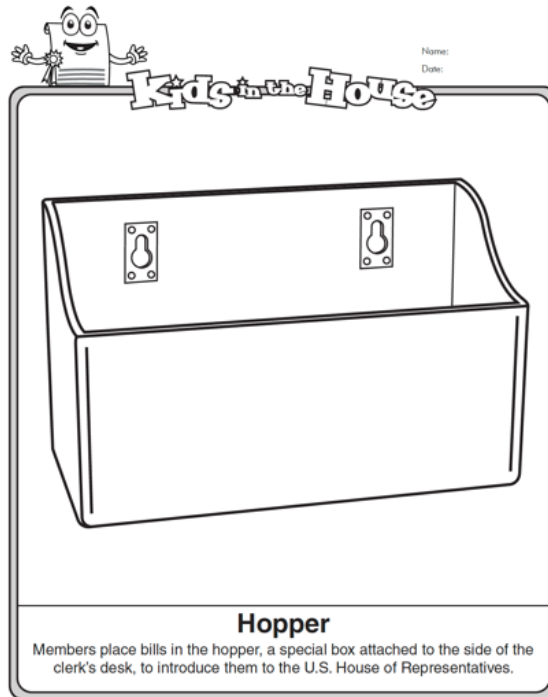


Figure 6. Colouring sheet for *Hopper*

The colouring sheet presents an illustration of the object to be coloured as well as a defining verbal caption.

The only other popularizing strategy found in *Young Learners* is generalization (9%), which is resorted to in cases where a wide-ranging description of a definite concept can ease the comprehension without loading the recipient with unnecessary details. The excerpt that follows (12) exemplifies how the long list of duties and responsibilities of the Clerk of the House is cut short by using the vague quantifier ‘lots of’.

- (12) The Clerk has lots of jobs.

Finally, it is worth pointing out that metaphors, denominations and reformulations are absent in this corpus component perhaps because it was felt they are ineffective and unnecessary for the cognitive capacities and needs of preschoolers.

4.2 *Grade School*

In *Grade School* the use of simple and not too detailed descriptions (34%) is still the most common explanatory device to introduce unknown concepts, as in the account of the ‘fresco technique’ showcased in the example below (13).

- (13) Constantino Brumidi [...] used a fresco technique in which water-based paints are applied to freshly spread plaster.

This description is enriched by various captioned photos which display sections of the described work of art.

Definitions (29%) are used far more in this website version than in *Young Learners*, as the recipients are children of reading age who can understand verbal clarifications. Some of these definitions are embedded in the text as periphrases (14); others, instead, are introduced by the verb ‘mean’.

- (14) Subcommittees are smaller groups of the committee’s members who are experts in a specific part of the committee’s area of public policy.

However, in most cases they are linked to the words that need to be defined as self-opening boxes that appear when the mouse is placed over the highlighted term (Figure 7). In this way, the detailed definition is accessed only if necessary and if the reader is interested. This allows the displayed cognitive load, which may present too much specialized information, to be reduced.

Each Representative, Delegate, and Resident Commissioner serves on two standing assignments are given at the start of each Congress. Delegate A nonvoting representative in the U.S. House of Representatives from American Samoa, the Commonwealth of the Northern Mariana Islands (CNMI), Guam, the Virgin Islands, or the District of Columbia. and returning Representatives usually have expertise and seniority. To be a Delegate, one must be a resident of the District of Columbia. Resident Commissioner Congress.

Figure 7. Pop up definition of *Delegate*

Cases of exemplifications (12%) are present, though to a slightly lesser extent than in *Young Learners*. In many contexts they are functional in adding

information or introducing curious items rather than explaining something unknown, as displayed in (15), where the name of a famous citizen whose portrait hangs in the Capitol's Rotunda is added as a parenthetical element introduced by 'such as'.

- (15) Also, distinguished citizens, such as Rosa Parks, are honored in the Rotunda after they die.

Differently from what was found for preschoolers, analogies in the form of metaphors and similes (9%) are repeatedly employed in this corpus section. It is relevant to observe that the term of comparison introduced by these figures of speech is very often taken from the school environment (16), a domain familiar and mastered by elementary schoolers whose comprehension of metaphoric language could otherwise be limited.

- (16) Just like your teacher, the Clerk of the House takes attendance to be sure everyone is there so they can get to work.

Denominations (7%) are also sometimes used to enhance popularization for grade schoolers. Hence, some specific labels are attached to concepts which are generally introduced by metadiscursive verbs such as 'called' or 'named' (17).

- (17) To be sure that one branch does not become more powerful than the others, the Government has a system called checks and balances.

As for generalization (5%), it can be observed that it is slightly less common in this corpus component compared to *Young Learners* (but still higher than in *Middle* and *High School*). It is principally used to recapitulate by using informal vague quantifiers such as 'lots of' (18), 'a lot of', or 'most'.

- (18) With the rules decided, the first day activities are over, but Congress has lots of work to do making laws to keep Americans safe.

Lastly, a few cases of reformulation (4%) could also be singled out, as in (19) that is taken from a passage where the different methods for voting on a bill are listed.

- (19) – Viva Voce (corresponding to voice vote).

Here the technical term *Viva Voce*, a Latin phrase, is paraphrased with its meaning given in this particular context and placed between round brackets in a simpler and more straightforward way.

4.3 Middle School

In contrast to the two previously discussed corpus sections, *Middle School's* most employed popularizing device is definition (30%), which surmounts description (20%). Definitions are both embedded in the text or linked to some specific disciplinary terms. However, a peculiarity of definitions in *Middle School* is that they are very often connected to one another through partial repetition (20).

- (20) Like the majority leader, the minority leader is not mentioned in the U.S. Constitution.

In the example above, and in many other cases, the defined object (in this case the 'minority leader') is introduced by referring to a previous definition through the preposition 'like' that works as a textual anaphora. Some defined words are also accompanied by photos that can be clicked on and enlarged in order to read the descriptive caption that adds extra information and peculiar items of interest. For example, 'hopper' is defined in the text as "a wooden box on the side of the Clerk's desk", and is complemented by the photo and the caption in Figure 8.



Representatives introduce bills by placing them in the bill hopper that is attached to the side of the Clerk's desk. The term comes from a funnel-shaped storage bin filled from the top and emptied from the bottom. The storage bins usually hold grain or coal.

CLOSE

Figure 8. Photo of the *hopper* with its descriptive caption

The second most used strategy is reformulation (21%), which is much more frequent here than in the previously discussed website versions. This is principally because the texts contain more numerous terms that need to be abridged for students to fully grasp their meaning. In (21) we can notice how an apposition (the most recurrent type of reformulation in this corpus section) is used to explain the sense of ‘Representatives’.

- (21) Representatives, the title given to Members of the U.S. House of Representatives, are elected to serve a two-year term.

Descriptions (20%) are employed to a similar extent to reformulations and typically occur in self-opening in-depth information boxes such as the one in Figure 9, which explains what the ‘Great Depression’ is and expands on some details, such as its time span and location.

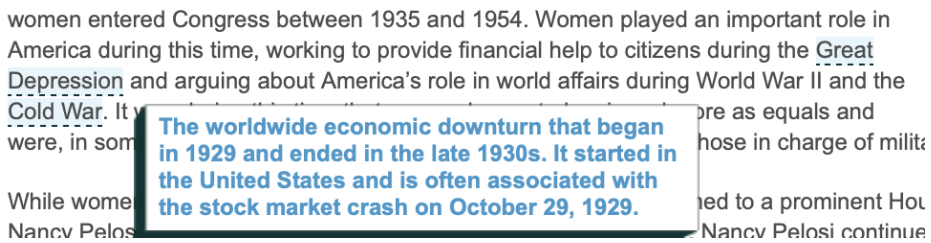


Figure 9. Pop up box for the *Great Depression*

Moving on to the practice of denomination (16%), we can note that specific concepts are designated by technical terms more than twice as much in this corpus section than in *Grade School*. Denominations are usually triggered by some specific verb phrases such as ‘meaning’, ‘called’ and ‘known as’ (22), and the technical term is regularly put between inverted commas so as to highlight its status as a sectorial expression.

- (22) The Member introducing the bill is known as its “primary sponsor”.

It is interesting to notice that, embedded in the heading picture of the thematic page where most denominations can be found (i.e., ‘The Role of the House’), there are two buttons that send users to some educational printable activities related to technical words. More in detail, there is a word search where users must find hidden words in a puzzle, and a crossword where there are definitions of words used in the text (Figure 10).



Name: _____

Date: _____

CAPITOL CROSSWORD

The House of Representatives has to juggle many different activities on the first day. Juggle the answers to the questions below by filling them in the indicated boxes on the puzzle grid.

DOWN

1. The document that outlines the United States government and its functions.
4. A person designated to act for or represent another or others; a deputy; representative, as in a political convention.
5. A proposed law requiring the approval of both Houses and the signature of the President to enact.
7. A rule of conduct established and enforced of by the authority, legislation, or custom

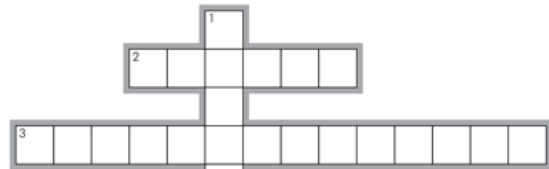


Figure 10. Capitol crossword activity

The use of exemplifications (9%) plays an even more marginal role in the adaptation for middle schoolers. Some of the examples that were provided in *Young Learners* and *Grade School* are simply omitted, probably because they were found to be unnecessary for comprehension as well as too obvious.

Also, generalizations (3%) and analogies (1%) are kept to a minimum, perhaps because it is felt they are not required to explain content.

4.4 High School

Popularization in *High School* principally rests on definitions (31%), reformulations (25%) and denominations (21%), which score the highest results in this website version. Definitions are very similar, both in quality and quantity, to those found for middle schoolers. Reformulations, instead, are far more numerous. Some of the technical terms which were left unexplained in *Middle School* are explained here especially through paraphrases, as in the example that follows (23) where the concept of ‘Committee of the Whole’ is paraphrased through juxtaposition.

- (23) The Committee of the Whole, a committee on which all Members of the U.S. House of Representatives serve, is a mechanism for quickly moving legislation through to the House floor for a vote.

Regarding denomination, abbreviation is the only type singled out in this website version. In some cases, acronyms and initialisms (24) are offered so as to provide the user with the abbreviated label that is commonly found in texts for experts and that, therefore, is worth learning.

- (24) Quickly surrounded by well-wishers, the first Asian Pacific American (APA) Member of Congress [...].

Another feature of *High School* is that the images and photos that often accompany the explained terms in the other corpus sections are significantly reduced here. They are very often substituted by smaller captioned images (Figure 11) that, if enlarged, send the user to the official website of the U.S. House of Representatives, where in-depth information sheets can be found.



**Historical Highlight: The first
Speaker of the House,
Frederick A.C. Muhlenberg
of Pennsylvania
June 04, 1801**

Figure 11. Captioned picture in *High School*

Also, educational activities are omitted and substituted by direct links placed on technical words or phrases that point to the higher degree of hypertextuality of *High School*. These links lead to the official page of the House of Representatives or to other official governmental pages (e.g., the webpage of the U.S. Senate), where the highlighted item is dealt with in more detail. This continuous reference to external sources is also the reason why descriptions (16%) are minimized.

Although the percentage of explaining examples (7%) is further reduced compared to the other website versions, they are sometimes used in passages in which not only do they help to picture a concept, but they also add information, as in (25) where an example of a Hispanic American Member of the Congress is provided.

- (25) Over the years, Hispanic American Members, such as Representative Henry B. González who chaired the powerful House Banking and Currency Committee [...].

Finally, the analogies and generalizations which were exploited in the other website versions to ease comprehension are not used in *High School*. In most cases, the concepts that for young learners, grade schoolers and middle schoolers are explained through an analogy, or a generalization, are left to the user's interpretation, or they are clarified through a link to the pop-up glossary.

5. Concluding remarks

The study of *Kids in the House* was an attempt to describe how differently the four specific versions of this educational website address the respective designated audiences. The quantitative and qualitative analysis principally focused on verbal language with occasional comments on the interaction with the nonverbal apparatus. More precisely, the attention was on the lexical contents of the four website versions and on the popularizing structures that are exploited to mediate and translate specialized contents (e.g., the functioning of the U.S. House of Representatives) according to the accessibility needs and tastes of young learners, grade schoolers, middle schoolers and high schoolers (cf. Sherwin – Nielsen 2019). On the whole, the analysis brought to light the specificities of each website version, showing how linguistic resources and edutainment contents in *Kids in the House* are carefully adapted for each group of addressees, making it an example of a highly accessible educational website.

The inductive quantitative analysis showed that both indicators of lexical richness, i.e., lexical variation and lexical density, increase steadily from young learners to high schoolers, thus clearly showing how lexical complexity is gradually amplified according to the target audience. Nevertheless, if lexical variation remains within the reference value ranges

that point to repetitiveness, lexical density is overall quite high, thereby testifying to the lexical complexity of specialized contents. The exploratory keyword analysis helped to disclose the most salient words distinguishing each corpus component and revealed interesting discourse-bearing trends. For example, *Young Learners* emerged to be overtly oriented towards the receivers and their involvement (e.g., the occurrence of the pronoun 'you' and the greeting 'hello'). In *Grade School* a more inclusive perspective is usually adopted (e.g., the usage of the pronoun 'we') so as to share knowledge with the users. *Middle School* is distinguished by the abundance of questions and of the imperative mood, which are other tools used to encourage the audience to get involved and engage with the website contents. *High School*, instead, presents elements pointing to a more complex and elaborate style (e.g., 'and' and 'they'), more suitable for teenagers already familiar with a hypotactic essay-like style. A group of keywords across the four website versions point to their popularizing intent (e.g., 'most' indicating generalizations in *Young Learners*, or 'like' used for similes in *Grade School*), thus confirming the mission statement of the website landing page, according to which information should be presented and explained to schoolchildren of all ages.

The more detailed qualitative analysis on the popularizing devices selected to facilitate the understanding of specialized content revealed marked differences in the way distinct target audiences are catered to. For preschoolers, paratactic descriptions and targeted examples are the most exploited popularizing strategies, with occasional recourse to definitions and generalizations. Descriptions and definitions are often accompanied by educational activities (e.g., printable colouring sheets), which involve the nonverbal apparatus and help the user disambiguate and memorize new specialized concepts. Grade schoolers are children who have started elementary school and thus are learning how to read and surf the internet on their own. Popularization for them chiefly rests on descriptions and definitions, but also other strategies not used for young learners could be located here, for example denominations, reformulations, and analogies in the form of similes and metaphors. The presence of captioned photos and hypertextual elements (e.g., pop up definitions) play a crucial role in supporting verbal popularizing strategies, which make them more interesting and engaging. When looking at texts for middle schoolers and high schoolers, it is evident how reformulations and denominations, which are quite marginal in grade school, take on an important role, whereas analogies, examples and generalizations are often omitted. In *Middle School* a wide range of educational activities accompany and

elucidate denominations and, in general, specialized terms. *High School*, instead, more frequently relies on external supplementary worksheets thought to be for grown-ups (e.g., the official webpage of the House of Representatives).

Experimental research would be needed to further interpret the results and assess the actual usability and accessibility feedback of target users. In particular, it would be interesting to verify whether the verbal and nonverbal popularizing devices adopted by the author(s) are recognized as useful and adequate to disseminate knowledge, and if the explained concepts are successfully assimilated.

REFERENCES

Primary sources

British National Corpus (BNC)

<http://hdl.handle.net/20.500.12024/2554>, accessed September 2021.

Kids in the House

<https://kids-clerk.house.gov/>, accessed September 2021.

Secondary sources

Baker, P.

2006 *Using Corpora in Discourse Analysis*. London: Bloomsbury.

Bianchi, F.

2018 "Rewriting *Romeo and Juliet* for a young audience. A case study of adaptation techniques", *Lingue e Linguaggi* 27, 43-65.

Bruti, S. – E. Manca

2019 "The popularization of environmental issues in children's magazines: A cross-cultural corpus analysis". In: M. Ji (ed.) *Translating and Communicating Environmental Cultures*. London/New York: Routledge, 182-201.

Buckingham, D. – M. Scanlon

2004 "Connecting the family? 'Edutainment' web sites and learning in the home", *Education, Communication & Information* 4 (2), 271-291.

Calsamiglia, H.

2003 "Popularization Discourse", *Discourse Studies* 5 (2), 139-146.

Calsamiglia, H. – T. A. van Dijk

2004 "Popularization discourse and knowledge about the genome", *Discourse & Society*, 15 (4), 369-389.

- Cappelli, G. – S. Masi
 2019 “Knowledge dissemination through tourist guidebooks: Popularization strategies in English and Italian guidebooks for adults and for children”. In: M. Bondi – S. Cacchiani – S. Cavalieri (eds.) *Communicating Specialized Knowledge: Old Genres and New Media*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 124-161.
- Clarke, A. M.
 2006 *ETSI White Paper No. 2. Young children and ICT: Current Issues in the Provision of ICT Technologies and Services for Young Children*. Sophia–Antipolis: European Telecommunications Standards Institute.
- Diani, G.
 2015 “Exploring knowledge dissemination strategies in English and Italian newspaper articles for children: A focus on legal issues”, *Textus* XXVIII (2), 109-126.
 2020 “‘Health for Kids’. Multimodal resources for popularising health knowledge on websites for children”, *Lingue e Linguaggi* 40, 67-93.
- Diani, G. – A. Sezzi
 2019 “The EU for children: A case of web-mediated knowledge dissemination”. In: S. Maci – M. Sala (eds.) *Representing and Redefining Specialized Knowledge: Variety in LSP*. Bergamo: CERLIS Series, 203-221.
 2020 “Scientific websites for children: Nurturing children’s scientific literacy through the conflation of multiple semiotic resources”, *Journal of Visual Literacy* 39 (3-4), 273-291.
- Djonov, E. N.
 2008 “Children’s website structure and navigation”. In: L. Unsworth (ed.) *Multimodal Semiotics: Functional Analysis in Contexts of Education*. London/New York: Continuum, 216-236.
- Flowerdew, L.
 2003 “A combined corpus and systemic-functional analysis of the problem-solution pattern on a student and professional corpus of technical writing”, *TESOL Quarterly* 37 (3), 488-511.
- Garzone, G.
 2020 *Specialized Communication and Popularization in English*. Roma: Carocci.
- Kress, G. – T. van Leeuwen
 2001 *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- Kilgarriff, A. et al.
 2014 “The Sketch Engine: Ten years on”, *Lexicography* 1, 7-36.
- Partington, A. – A. Duguid – C. Taylor
 2013 *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam/Philadelphia: John Benjamins Publishing.
- Scott, M.
 2016 *WordSmith Tools Version 7*. Stroud: Lexical Analysis Software.

- Scott, M. – C. Tribble
2006 *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins Publishing.
- Sezzi, A.
2019 “Go on an art adventure. Popularizing art for children through museum websites”. In: M. Bondi – S. Cacchiani – S. Cavalieri (eds.) *Communicating Specialized Knowledge: Old Genres and New Media*. Newcastle upon Tyne: Cambridge Scholars Publishing, 162-177.
- Sherwin, J. – K. Nielsen
2019 <https://www.nngroup.com/articles/childrens-websites-usability-issues/>, accessed September 2021.
- Stubbs, M.
1996 *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Oxford: Blackwell.
- UNICEF
2019 “Growing UP in a connected world”, <https://www.unicef-irc.org/growing-up-connected>, accessed September 2021.
- Ure, J.
1971 “Lexical density and register differentiation”. In: G. Perren – J. L. M. Trim (eds.) *Applications of Linguistics*. London: Cambridge University Press, 443-452.
- Vignozzi, G.
2020 “Involvement markers in American broadcast politics: A case study on TV political interviews”, *Lingue e Linguaggi* 36, 315-330.

Address: GIANMARCO VIGNOZZI, Università di Pisa, Dipartimento di Filologia, Letteratura e Linguistica, Piazza Torricelli, 2 - 56126 Pisa, Italy.
ORCID code: orcid.org/0000-0003-3131-0258