Spellchecking for Children in Web Search: a Natural Language Interface Case-study

Casey Kennington
Computer Science
Boise State University
caseykennington@boisestate.edu

Jerry Alan Fails
Computer Science
Boise State University
jerryfails@boisestate.edu

Katherine Landau Wright
Literacy, Language and Culture
Boise State University
katherinewright@boisestate.edu

Maria Soledad Pera
Computer Science
Boise State University
solepera@boisestate.edu

Abstract

Given the more widespread nature of natural language interfaces, it is increasingly important to understand who are accessing those interfaces, and how those interfaces are being used. In this paper, we explore spellchecking in the context of web search with children as the target audience. In particular, via a systematic literature review of work that illustrate challenges and limitations, we show that, while widely used, popular search tools are ill-designed for children. We then use spellcheckers as a case study to highlight the need for an interdisciplinary approach that brings together natural language processing, education, human-computer interaction to address a known information retrieval problem: query misspelling. We conclude that it is imperative that those for whom the interfaces are designed have a voice in the design process.

1 Introduction

The most popular web search engine, Google, logs roughly 3.5 billion searches per day.[1] Like many other commercial search engines, Google's primary interface is very simple: a search box (i.e., a single line text-box) and two buttons. The expected input is a natural language query using text as the primary modality; after the search is initiated, Google provides a ranked list of websites, presented also in text form. But how well does this well-used, simple natural language interface (NLI) serve everyone, including children whose natural language and user interface understandings is qualitatively different from that of adults? Gutnick and Green and Holloway (2014) documented that by age 8 two-thirds of children will use the Web on a daily basis, but mainstream search engines offer weak support for children's inquiry approaches. While these *emergent searchers* quickly develop the technical ability to use search engines and may frequently conduct online inquiry tasks, they still might not find relevant information – or information that matches their search intent (Anuyah et al., 2019; Bilal and Huang, 2019). These challenges go beyond just applying "safe search" filters to results; children struggle with the natural language *input* aspect of the interface for several well-studied reasons:[2]

Keyword Choice. Children find it difficult to formulate succinct, meaningful keyword queries (Broch, 2000; Duarte Torres et al., 2012). Search terms are generally nouns, a search convention that children have not yet grasped (Bilal and Ellis, 2011), which causes children to struggle to find keywords that express their information needs (Aspray and Bernat, 2000). This sometimes leads children to favor longer, grammatical natural language queries despite the difficulty they have in typing (Druin et al., 2009).

Over/under-specified Queries. Young children favor thematic categorizations, which may lead them to over/under-specify queries. Though typing is challenging, some children may over express their search intent and enter a complete sentence to initiate the search process or enter multiple keywords, not all related to their desired result (Kammerer and Bohnacker, 2012; Schacter et al., 1998).

Spelling. Emergent searchers are generally in either the Within Word Pattern or Syllables & Affixes stages of spelling development, meaning they tend to rely on their knowledge of letter names (rather than letter sounds), struggle to appropriately represent vowel sounds (Bear et al., 2016), and spell words phonetically. All of this can lead to queries that search engines do not process correctly. For example, as Spink et al. (2010) found, a

[1] https://www.internetlivestats.com/google-search-statistics/

[2] Note that few of these studies are recent. This is partially due to the fact that public data sets and benchmarks are not available (children are a protected population), so there is less research in the area of child-oriented information retrieval.

Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing, pages 8–13
April 20, 2021.

child searching for information about endangered animals may type "ndanged anemls." In this example, the child uses the letter name "N" to represent "en", struggles to represent the short /i/ sound, and omits other vowels (Bear et al., 2016). Moreover, Druin et al. (2010), who investigated children's use of keyword search in homes, shared an example of a child slowly spelling D-O-L-F and pressing the enter key as he searched for "dolphins." In these examples, search engines could assume users meant to type "deranged" or the acronym or first name DOLF, missing completely the intent of the query.

Homonyms & Taxonomies. Children are only becoming aware that written representations sometimes have different senses (Large et al., 2008). For example, a jaguar could be an animal or a car. While it might be easy for children to use related words ("dog" – "leash"), they can struggle with abstract categorizations ("dog" – "mammal") because of incomplete understanding of taxonomies and hierarchies (Nguyen and Murphy, 2003). For example, they may enter "lion," when looking for books about animals, or may enter "fairies," when they are interested in fantasy.

In addition to the challenges with input, children also struggle in different ways with the search *results*, which are displayed as text in a ranked order:

Query Reformulation and Prompts. Children have difficulty reformulating a query if their initial attempt does not lead to relevant information (Gossen et al., 2014; Kammerer and Bohnacker, 2012). In fact, they are more likely to input one query and read through all retrieved results instead of optimizing their time using new queries (Gossen et al., 2014). Additionally, children do not always take advantage of the query suggestions that search engines provide (Druin et al., 2009).

Connection of Results with Query. Children are not trained to identify suitable resources from among those that their favored search engines retrieve (Bilal and Ellis, 2011), nor are they trained to understand the direct connection between their queries and retrieved results (Karatassis, 2017).

Quality & Comprehension of Results. Children step through retrieved resources, though children rarely judge those resources for quality (Graham and Metaxas, 2003; Walker, 2013). The average readability level of the (snippets of) resources retrieved by search engines is the 10th-grade level (typically children 15-16 in the United States) (Azpiazu et al., 2016; Anuyah et al., 2019). This level is far beyond that appropriate for emergent searchers. As such, children may not be able to comprehend retrieved resources, thus making these results irrelevant (Bilal and Ellis, 2011; Graham and Metaxas, 2003; Walker, 2013).

As indicated above, the *input* and *results* elements of the search task for children have components that relate directly to natural language processing. However, these elements cannot be addressed individually and purely from a traditional natural language processing perspective. Additionally, the obvious must be stated: *children are not simply small adults*, which has implications for all interfaces that children use that frequently are designed for (and by) adults. Addressing these search challenges for children requires an interdisciplinary approach that sits at the crossroads of information retrieval (IR–the search task) and human-computer interaction (HCI–the interface), both with focus on children as users, as well as natural language processing (NLP–the medium) and education (Ed–developmental understanding of the user).

In this paper, we use Spelling as a case study for applying this interdisciplinary approach to illustrate that an interface as commonplace as web search and a basic operation such as automatic spellchecking are not yet effective tools when children are the users, and applying an interdisciplinary approach can result in important research and technological advances that impact users.

2 Case Study: Spellcheckers for Children in Web Search Settings

Children make many errors as they learn how to spell, which need to be addressed if search tools are to effectively respond to children's information needs. Spelling words incorrectly has potentially harmful implications for children because a word that is auto-corrected by an adult-designed spellchecker can completely change the meaning of the search query, thus affecting search success. In this section, we focus our attention on spellchecking. We first discuss assumptions about spellcheckers and sketch an interdisciplinary approach to resolving issues with spellcheckers for children. Lastly, we give examples of recent work on an effective spellchecker for children.

Assumptions Inherent from the Task Query spelling is a "crucial component of modern search engines" (Li, 2020), as misspelled queries can lead to the retrieval of results that do not respond to the

nformation needs of the users, or even no results at all if the misspelling cannot be mapped to common vocabulary. Within this context, IR and HCI researchers have allocated resources to understanding and addressing the issue of spelling in search (Spink et al., 2010; Druin et al., 2010). Still, the majority of the literature in this area responds to the needs and interactions of traditional search engine users, i.e., adults (Chang and Deng; Duan and Hsu, 2011; Li et al., 2012; Gupta et al., 2019). In this case, there is an implicit assumption that these user groups would understand the conventions of using the interface of a search system where text is the primary modality of expressing a user's search intent. Strategies adopted by IR researchers to address spelling problems – often users' typos or misused words/phrases (Li, 2020) – depend on lookup dictionaries or existing adult query-logs. These corrections happen after the query is complete (e.g., "Did you mean..." feature in Google). These assumptions and challenges are confounded when the user is a child; they do not usually understand the interface conventions and may be confused by corrected spelling mistakes that use a dictionary lookup only after submitting a query. This is a problem as children may not realize that a correction occurred, nor identify that the results are skewed because they do not reflect the original intent of their queries. For example, if a child begins to type "ndanged" but the spell-checker corrects to "danger" instead of "endangered," the reformulated query clearly does not reflect the child's original intent. This poses a concern, especially in light of the fact that young searchers can have misconceptions about the search tools themselves – children expect results, but do not question the relevance of those results (Halavais, 2017; Anuyah et al., 2019). Of note, in the previous example, the child relies upon knowledge of the name of the letter "N" and not the sound the letter makes (Bear et al., 2016) a perspective that spellcheckers (designed for adults) generally miss.

2.1 Sketch of an Interdisciplinary Approach

As evidenced from our discussion above, tackling misspellings in children's query is not something that the IR community can undertake *alone*. Spellcheckers are not just the underlying technical mechanism that identifies the spelling mistakes, they are at the forefront, fostering interactions with the interface. Moreover, spelling correction and learning is an educational endeavor that can be addressed by combining NLP, HCI, and Ed. NLP is needed because the primary medium is text, albeit often not grammatical, and searching often happens across multiple turns, which is similar to interactive, spoken dialogue – a sub-field of NLP. We should treat spellchecking with children as an interactive, multi-turn process. Ed is needed because the interactive, multi-turn process and the spellchecker itself should apply a more principled model of spelling correction based upon educational research in spelling development so spelling mistakes can inform the system how to act, which is where HCI is needed. For example, spelling suggestions could be displayed, highlighting spelling differences, to draw the child's attention to the specific spelling mistake made, and providing an option to alter their input (i.e., in potential query reformulations). This naturally makes the process more interactive, drawing the query over several turns between the user and the system, and informs the child that query reformulation is a part of the search process. HCI expertise leverages researched principles and guidelines in addition to participatory design practices (Fails et al., 2013) to help design intuitive displays which can provide scaffolding in the form of visual cues to guide child users to describe their search intent accurately.

2.2 Recent work on Interdisciplinary Modeling and Evaluation of Child-directed Spellcheckers

In Downs et al. (2019), we examined what children actually expect from spellcheckers in search tools. We did so via *participatory design* (specifically using the Cooperative Inquiry method (Druin, 1999; Fails et al., 2013)) in order to involve children in designing the spellchecker. The main take-away from this research was the need for visual cues to act as scaffolding for the children to use as they learn to navigate the spellchecker's interface to identify and correct spelling mistakes. What otherwise appears to be a solved task for NLP is far from solved when children are the users. These efforts informed the modeling of the spellchecker *KidSpell* we introduced in Downs et al. (2020a). KidSpell is highly influenced by established Educational research in spelling development which led to the design of a phonetic-based spellchecker because children tend to make mistakes with the phonetic mappings of sounds to alphabet symbols. The authors evaluated

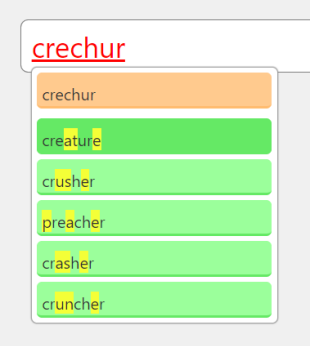

Figure 1: Example of the KidSpell child-directed spellchecker from Downs et al. (2020a).

their spellchecker on mistakes made in the context of handwritten writing samples (that were digitally transcribed) as well as spelling mistakes made during web search tasks. The phonetic spellchecker worked better for children than the current state-of-the-art spellcheckers used by adults (including Bing's spellchecker[3], and Aspell[4]), despite being a more parsimonious, rule-base approach. An example of the spellchecker in use is depicted in Figure 1, with visual cues showing the parts of the word that differ phonetically from what the child spelled.

The development and advancement of KidSpell included a multi-faceted approach that included algorithmic advances and user-interface enhancements (Downs et al., 2019, 2020a,b). Algorithm advances (applying NLP research) included evaluating input, identifying misspellings, and generating ranked spelling suggestions. With regards to the user interface, research included multiple participatory design sessions with an intergenerational design team consisting of children ages 6-11 and adult researchers (applying HCI research). The team collaboratively designed various approaches that could convey to children when a word was misspelled and enable them to better select the intended word from the suggested spelling corrections.[5] We then investigated the effectiveness of various cues within a web-search task (applying HCI and IR research). One cue was images that prototypically represented the word in question (e.g., if a child typed *dinsoer* and the spellchecker suggested *dinosaur*, then hovering over *dinosaur* with the mouse would display an image of a dinosaur, as depicted in Figure 2. Another cue we evaluated was speech synthesis; for the *dinosaur* example, hovering over it would play a synthesized voice reading the word over the speakers. We also looked at the combination of the two and showed that both cues had impact, but synthesized speech especially helped guide the children to select the correctly spelled word.

Figure 2: Example of the KidSpell child-directed spellchecker using both visual and synthesized voice cues. Taken from Downs et al. (2020b).

3 Conclusions: Lessons for NLP & HCI

Spellchecking (for English) appears to be low-hanging fruit in NLP, but it is clearly not a solved problem for children. In the context of search within IR, spellchecking is necessary for the best search results, but the search task and interface are not well-suited for children. In this paper, we sketched an interdisciplinary approach that combines knowledge from the fields of Education, NLP, and HCI in an IR task and reviewed recent literature that shows how critical the interdisciplinary approach is for solving spellchecking for children in web search settings.

As is the case with spellcheckers, search for children cannot be assumed to be solved because it is already widely adopted and useful for adults. The interface, the input medium, and the audience all need to be taken into account. We can learn an important lesson here from the field of HCI that has far-reaching implications: the best way to design an interface that is suitable to a target audience is to involve them in the design and evaluation process. This is becoming more critical as research from the NLP field turns into practical products that face users; NLIs are not always *natural*; they need to be designed with and for the users that they will serve.

[3] https://www.microsoft.com/en-us/bing/apis/bing-spell-check-api

[4] http://aspell.net/

[5] The spellchecker is active on http://cast.boisestate.edu.

Acknowledgements

This research was partially supported by NSF Award 1565937 as well as NSF Award 1763649. We appreciate the anonymous reviewers' comments and feedback.

References

Oghenemaro Anuyah, Ashlee Milton, Michael Green, and Maria Soledad Pera. 2019. An empirical analysis of search engines' response to web search queries associated with the classroom setting. *Aslib Journal of Information Management*.

William Aspray and Andrew Bernat. 2000. Recruitment and retention of underrepresented minority graduate students in computer science. In *Report on a Workshop by the Coalition to Diversity Computing*.

Ion Madrazo Azpiazu, Nevena Dragovic, and Maria Soledad Pera. 2016. Finding, understanding and learning: Making information discovery tasks useful for children and teachers. In *ACM SIGIR Workshop on Search as Learning*.

DR Bear, M Invernizzi, S Templeton, and F Johnston. 2016. Words their way: Word study for phonics, vocabulary, and spelling instruction, 2nd.

Dania Bilal and Rebekah Ellis. 2011. Evaluating leading web search engines on children's queries. In *International Conference on Human-Computer Interaction*, pages 549–558. Springer.

Dania Bilal and Li-Min Huang. 2019. Readability and word complexity of serps snippets and web pages on children's search queries. *Aslib Journal of Information Management*.

Elana Broch. 2000. Children's search engines from an information search process perspective. *School Library Media Research*, 3.

Yi Chang and Hongbo Deng. *Query Understanding for Search Engines*. Springer.

Brody Downs, Oghenemaro Anuyah, Aprajita Shukla, Jerry Alan Fails, Maria Soledad Pera, Katherine Wright, and Casey Kennington. 2020a. Kidspell: A child-oriented, rule-based, phonetic spellchecker. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 6937–6946.

Brody Downs, Tyler French, Katherine Landau Wright, Maria Soledad Pera, Casey Kennington, and Jerry Alan Fails. 2019. Children and search tools: Evaluation remains unclear. In *KidRec Workshop co-located with ACM IDC 2019*.

Brody Downs, Aprajita Shukla, Mikey Krentz, Maria Soledad Pera, Katherine Landau Wright, Casey Kennington, and Jerry Fails. 2020b. Guiding the selection of child spellchecker suggestions using audio and visual cues. In *Proceedings of the Interaction Design and Children Conference*, pages 398–408.

Allison Druin. 1999. Cooperative Inquiry: New Technologies for Children. page 8.

Allison Druin, Elizabeth Foss, Leshell Hatley, Evan Golub, Mona Leigh Guha, Jerry Fails, and Hilary Hutchinson. 2009. How Children Search the Internet with Keyword Interfaces. In *Proceedings of the 8th International Conference on Interaction Design and Children*, IDC '09, pages 89–96, New York, NY, USA. ACM. Event-place: Como, Italy.

Allison Druin, Elizabeth Foss, Hilary Hutchinson, Evan Golub, and Leshell Hatley. 2010. Children's Roles Using Keyword Search Interfaces at Home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 413–422, New York, NY, USA. ACM. Event-place: Atlanta, Georgia, USA.

Huizhong Duan and Bo-June Hsu. 2011. Online spelling correction for query completion. In *Proceedings of the 20th international conference on World wide web*, pages 117–126.

Sergio Duarte Torres, Djoerd Hiemstra, Ingmar Weber, and Pavel Serdyukov. 2012. Query recommendation for children. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 2010–2014.

Jerry Alan Fails, Mona Leigh Guha, and Allison Druin. 2013. Methods and Techniques for Involving Children in the Design of New Technology for Children. *Foundations and Trends® in Human–Computer Interaction*, 6(2):85–166.

Tatiana Gossen, Juliane Höbel, and Andreas Nürnberger. 2014. A Comparative Study About Children's and Adults' Perception of Targeted Web Search Engines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 1821–1824, New York, NY, USA. ACM.

Leah Graham and Panagiotis Takis Metaxas. 2003. " of course it's true; i saw it on the internet!" critical thinking in the internet era. *Communications of the ACM*, 46(5):70–75.

Lelia Green and Donell Holloway. 2014. Zero to eight: Very young children and the domestication of touch screen technologies in australia.

Jai Gupta, Zhen Qin, Michael Bendersky, and Donald Metzler. 2019. Personalized online spell correction for personal search. In *The World Wide Web Conference*, pages 2785–2791.

Aviva Lucas Gutnick. Always connected: The new digital media habits of young children. In *In New York: The Joan Ganz Cooney Center at Sesame Workshop.*

Alexander Halavais. 2017. *Search engine society*. John Wiley & Sons.

Yvonne Kammerer and Maja Bohnacker. 2012. Children's web search with Google: the effectiveness of natural language queries. In *Proceedings of the 11th International Conference on Interaction Design and Children*, IDC '12, pages 184–187, Bremen, Germany. Association for Computing Machinery.

Ioannis Karatassis. 2017. Websail: Computer-based methods for enhancing web search literacy. In *Proceedings of the Conference on Conference Human Information Interaction and Retrieval, CHIIR*, pages 403–405.

Andrew Large, Valerie Nesset, and Jamshid Beheshti. 2008. Children as information seekers: what researchers tell us. *New Review of Children's Literature and Librarianship*, 14(2):121–140.

Yanen Li. 2020. Query spelling correction. In *Query Understanding for Search Engines*, pages 103–127. Springer.

Yanen Li, Huizhong Duan, and ChengXiang Zhai. 2012. A generalized hidden markov model with discriminative training for query spelling correction. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 611–620.

Simone P Nguyen and Gregory L Murphy. 2003. An apple is more than just a fruit: Cross-classification in children's concepts. *Child development*, 74(6):1783–1806.

John Schacter, Gregory KWK Chung, and Aimée Dorr. 1998. Children's internet searching on complex problems: Performance and process analyses. *Journal of the American society for Information Science*, 49(9):840–849.

Amanda Spink, Susan Danby, Kerry Mallan, and Carly Butler. 2010. Exploring young children's web searching and technoliteracy. *Journal of documentation*, 66(2):191–206.

Henry M Walker. 2013. Homework assignments and internet sources. *ACM Inroads*, 4(4):16–17.