THESIS

THE CORPUS, WORD LIST, AND USABILITY ANALYSIS FOR THE CORPUS OF EXTENDED REALITY

Submitted by

Lauren Anne Mangus

Department of English

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Spring 2023

Master's Committee:

Advisor**:** Tatiana Nekrasova-Beker

Anthony Becker
Francisco Raul Ortega

ABSTRACT

THE CORPUS, WORD LIST, AND USABILITY ANALYSIS FOR THE CORPUS OF EXTENDED REALITY

There are multiple uses for word lists and corpora inside of linguistics-based research and studies; however, their primary function is the evaluation of authentic language for pedagogical application. This study presents both current and novel methods of approaching word list curation, word list evaluation by experts before publishing, and a collaborative analysis of pedagogical applications by focus group participants. The curation of the Corpus of Extended Reality (CoXR), a discipline within Computer Science, and its word list, the CoXR Word List, capture the language structures that are specific to the discipline of Extended Reality (XR). The CoXR Word List additionally includes hyphenated word forms that are specific to the discipline of XR. To test the efficacy of the corpus it was evaluated by graduate students in the discipline of XR to gain insight to whether traditional methods of discipline-specific corpus curation actually generate discipline specific words and whether the generated examples are meaningful. Firstly, the study found that verification of a word list is useful for determining the efficacy of the corpus itself and the word list with its coordinating examples. Secondly, the study found that (1) students from XR and outside of XR (English Language studies) found the CoXR Word List to be meaningful. Thirdly, all participants found the CoXR Word List as a tool they might use in the classroom as a teacher or student. Lastly, participants indicated that if they had access to

discipline-specific word lists such as the CoXR Word List they would be incentivized to engage in cross-disciplinary research.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Tatiana Nekrasova-Beker for her consistent support and encouragement throughout the development of this thesis and my time in the program. Despite the numerous challenges you always held us up and kept us moving forward; in this I learned tremendous life lessons. Thank you for believing in me and my wild ideas, and always finding a way to channel them into something tangible. With wholehearted sincerity, thank you for your unwavering belief in me.

I would also like to acknowledge my committee members Dr. Anthony Becker and Dr. Francisco R. Ortega. Dr. Anthony Becker, you also always supported my ideas in and out of the classroom that allowed me to explore the study contained herein in ways that were both novel and revolutionary for me, thank you. Dr. Francisco R. Ortega, you have been more than just a committee member to me these past years, but also a cherished mentor. Thank you for always believing in me, guiding me, and listening to me. It is with your support and the NUILab that I was able to put the study portion of this thesis together in a limited time span. I will always be grateful for your never-ending support for me, my ideas, and constant need for your patience.

Most importantly, I would like to take this opportunity to thank my family, especially my parents, for supporting me through this graduate program. I truly could not have made it this far without your unwavering faith in me, and second chances. Thank you. To Dr. Dhruva Patil, thank you for being my watchman at night, a guiding hand in the dark, my sword in the day, and always a heart of light. Without you I could not be here today writing this, and I will never forget how much you sacrificed to help me be here, thank you my love.

DEDICATION

To all who extended their hand as I walked from the fire, as others more brilliant than myself were left behind; I will never forget that this opportunity is a gift, a hammer, to shatter the ceiling.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

The classroom across the world generally has one look, a teacher, a board, and students who sit together facing the first two entities. This format has been used for centuries as the consistent modality for instruction (excluding the current development of the Communicative method and other non-Global North modalities of instruction). With this in-person format the structure of the classroom has also remained largely the same. Students come to class, leave, do homework, take exams and write essays. Yet, with the change in how education is accessed and how students engage with access to online materials, the classroom has been developing into new modalities across the world as the result of access to technology. With this technology has come a large shift in what language and educational areas we have access to, as well as the language we use.

The change in language used across the world is arguably undoubtable as this global economy increases in its influence with one another. Further, as English becomes the *lingua franca* of a world that didn't ask for one, how language is used is rapidly being re-shaped. With this re-shaping comes many challenges, especially those who are not high-level speakers of the English Language (L2+ [denoted as L2+ to decolonize the practice that all English Language Learners are only proficient at one language after English]). These language users are faced with numerous challenges in regard to communicating in a general manner with first language speakers (L1) of that language and in the business world. The current study is concerned with the language acquisition of graduate students that are both L1 and L2+ speakers of the English language. In graduate school, many people immigrate to new places for greater educational opportunities, but with this opportunity there is a cost.

This cost is often at the expense of one's language abilities, which is to say that many graduate students come to a field having enough prior knowledge to be accepted, but not entirely the discipline-specific language required for success in their given field (Groves & Mundt, 2015). The same could also be said of native speaker of the dominant language. This, therefore, points to a common problem. Language creates barriers in education, especially at the graduate level and beyond which creates numerous challenges for the student and those around them (Lin & Morrison, 2021; Huang, 2010). Furthermore, if one wishes to collaborate and/or work outside of their own specific domain there are challenges in joining the specific discourse community, because often a new language system must be created (Huang, 2010). To evaluate this gap, the research contained within this master's thesis evaluates a field of computer science, Human Computer Interaction, and a rapidly growing field under its umbrella, Extended Reality (which encompasses Mixed Reality (MR), Augmented Reality (AR), and Virtual Reality(VR)).

This project includes a corpus, curated for the purpose of developing a word list of relevant discipline-specific terms in the fields of HCI and XR, as well as a two-part survey analysis to examine whether or not having access to the language of the discipline encourages access and communication in and across disciplines within academic spaces. Additionally, the corpus itself contains a novel introduction of hyphenated words to reproduce a more intentionally realistic representation of the words in the discipline of computer science. The word list generated by the corpus of documents and its following evaluation aim to understand (1) if the language used within HCI and XR is specific to their respective disciplines, or if the language used is specific to only computer science, (2) if graduate students and professionals in the field of XR found the word list to be both representative of the domain and meaningful as a learning tool, and (3) whether or not graduate student teachers who are unassociated with

computer science find that having access to discipline specific language impacts their pedagogical repertoire when teaching CS students and then with this access, their potential to collaborate across disciplines.

Chapter 2: Literature Review

Computer Science and Language

In the field of Computer Science (CS) there are many disciplines within the field that cover a vast expanse of computing technologies. Unlike the natural sciences and mechanical sciences, which have branches that hold as their own fields, CS has kept its very different fields under one umbrella. As a result of this, there is a significant amount of mystique that surrounds the field of CS and this image of the field has led to a significant amount of internal research about the field, rather than being done by researchers in other fields. For example, the natural sciences have been touched by philosophy (Kallio-Tamminen, 2020), physics and language studies (Jalilifar & Memari, 2017), the arts and environmental science (Spagnolo et al., 2018), and other fields outside of its own. Engineering, for example, is a field often studied by linguists in the field of English for Specific Purposes (ESP), yet is not solely reliant on internal language research. On the other hand, CS is internally studied regarding its pedagogical applications, Computer Science Education (CSE), and its language which is targeted within Natural Language Processing (NLP).

One large reason for this internal study might be that computer scientists can study their own field with more efficiency through CS-based research methods. Another reason may be that CS is an insular field with strongly constructed barriers of entry from the earliest levels to the highest levels within the field (Cheryan et al., 2016). As a result of these characteristics, the field is not well studied outside of its own walls and this is increasingly becoming a problem. As Artificial Intelligence applications such as Chat-GPT (OpenAI, 2023) merge into our daily lives it is mystique that causes a public relations issue and as the field dreams of, seamless integration.

Yet, as there are limited studies on the language used within CS, and therefore limiting the access for external populations, there grows the potential for public-based pedagogy, collaboration across fields to dismantle the archaic power structures of secrecy, and internal-based pedagogy to build masters of the field's specific languages in order for more straightforward and unambiguous publications.

One way to open and expand the field of CS is to study the language within it. Understanding a particular language invites opportunities for access and social mobility (Csata & Marácz, 2021). In terms of writing, the field of CS is unique in that it does not generally challenge the differences between "good writing" among its diverse writers but instead focuses on correctness and clarity in its work (Hynninen & Kuteeva, 2017). Furthermore, as Hynninen & Kuteeva (2017) demonstrated, CS writing allowed for more "universally recognizable" versions of English so that "a text that conveys the meaning in an understandable way for the international computer science research community" (p. 63). While CS writing may have systems built internally to withstand the diversity of Englishes housed under its roof, the discipline-specific language is worthy of evaluation, maybe even more so due to its freedom for linguistic development to outsiders as CS is the field that will determine our futures. Additionally, if the English within CS has its own global-standard, this then could potentially be the early years of developing a new English language in its own right, which may then further limit access to its materials for external consumption and comprehension.

To expand then the language within the field of CS for those within the field, and those outside the field, more methodologies need to be constructed through cross-disciplinary research such as linguistic evaluations of the field. Research in a field to define its language practices opens opportunities for access to populations that may struggle with the language (Csata &

Marácz, 2021). While fields such as NLP and CSE can evaluate based on their own methodologies, re-inventing the wheel after almost 50 years of ESP studies (Huang, 2010) seems academically irresponsible on the basis that instead of opening doors for cross-disciplinary methodologies and research, the field has instead been self-reliant. To better understand why studying the language of CS is important outside of the discipline, noting the demographics of the field paints a clearer picture.

The Demographics in Computer Science

In the field of CS, the demographics of student bodies have been changing over time. According to the Open Doors Report by the Institute of International Education, in 2012, international students that studied CS were a student body of 53,953 (with a growth rate of 5.1%). Whereas in 2022, this has increased to a student population of 157,896 students with a 13.1% (Open Doors, 2022). In 2012, Business Management had the highest student population, whereas in 2022, Mathematics and CS held the trophy for the highest international student population (Open Doors, 2022). This demonstrates that in ten years, the field of CS has grown by almost one hundred and four thousand students. Of these students and United States of America (US) citizens, according to the Department of Education, the number of graduate degrees awarded in CS annually has increased from 20,925 in 2012 to 51,521 in 2020 (Department of Education, 2021). The graduate degree awardee level for graduate students is important because graduate students make up 89% of the international student body, 844,235 students in 2022 (Open Doors, 2022). We can therefore infer, based on the data, that international students consist of a large portion of the student population in the field of CS. If then, there is a large portion of L2+ speakers within a field then there are likely challenges of language acquisition and use within the field of not only the English language, which is the dominant publishing language in

CS (Hynninen & Kuteeva, 2017), and the field of CS. While CSE strives towards building more inclusive resources for its students (Yates & Plagnol, 2021), I would like to argue that cross-disciplinary relationships, especially from fields that have long studied how to serve international and L2+ students in the education system, are important for long-term success and may provide more immediate results.

Extended Reality – A Growing Novel Field in CS

One of the smaller but growing fields within CS is the discipline of Extended Reality. This field incorporates MR, AR, and VR. What makes this field unique and a discipline of its own, is that it is under the field of HCI and thus studies how humans (at this point in time) can use devices (such as glasses) to add to reality (AR) or escape into a different reality (VR) or maybe a mix of both (MR) to shape our day-to-day lives. This field is unique in that the sole purpose is to enhance our current reality by some form of a human-centered approach. Each part of the field, AR, MR, VR are related in the production of materials and language just enough to be considered and evaluated by this study as a singular body. This is largely because, due to the size of the field and its current interaction rates, those who work in AR are likely to work with those in VR at the university level by some degree.

This field was chosen for this study due to the researcher's close proximity to the field in terms of language access and that at the university being evaluated, the student body of degree seeking graduate students in XR have a majority L2+ student body. By the researcher's observations, the student population being evaluated demonstrated a grasp of the language, in that they could use it to publish often successfully, but not always a mastery (fluidity with clear and concise explanations) of the language they were using. As a result, due to the proximity and age of the field of study, XR and its graduate student population will be evaluated in the current

project as a chance to glimpse the discipline, but to also indicate challenges potentially within the field of CS itself. To study the language, the researcher employed the general technique of the ESP field by constructing an authentic language-based corpus in XR and then subsequently a word list from the corpus that was curated.

A History of Corpus Linguistics

Corpus linguistics is one of the multiple areas of study within a larger field of Applied Linguistics . Corpus linguistics is the study and curation of corpora which are "large, systematic, computerized collections of texts" (Römer, 2021, p. 233). The primary function of corpora is to give detailed insights into the construction, use, and function of language in a particular context (Römer, 2021). These corpora serve as gateways into disciplines and areas of authentic language production across languages. These gateways have allowed for new perspectives on language (Römer, 2021; Herbst et al, 2004). Corpora and their analysis have a long history, but it is how they came into being and have been shaped across linguistic disciplines that is vital for understanding their modern purpose.

With the advent of computers much has been shaped in the past few decades as portable devices and ease of access to computing power is accessed by the common person. The field of linguistics has not been excused from this growth. Rather, computational linguistics and the use of computer programs has accelerated the evaluation and study of language through Machine Translation (MT) (Groves & Mundt, 2015; Hirst, 2013). Yet, as MT continued to show failure from the 1940's to 1970's it was eventually replaced by the now more common studies within computational linguistics such as natural language and sub-processing techniques like corpus linguistics (Hirst, 2013). Within this time frame standards were set in corpus linguistics that are still used today. These standards are analysis by frequency, concordance for analysis, sampling

in a manner that allows for even dispersion of materials, and statistics that analyzed dispersion as much as raw frequencies (McEnery & Hardie, 2013).

The standards that are still present in corpus-based analysis today have survived decades of exploration and the rejection of corpus linguistics by the notable, and large figure within linguistics, Noam Chomsky. He argued that language can not be categorized based on its occurrences, and solely an external process because it reduces language to generalizations that are unquantifiable (McEnery & Hardie, 2013). As McEnery and Hardie state "it is undoubtedly true that no corpus can completely represent a language; Chomsky's observation that the set of possible sentences is infinite is very valuable" (pp. 730-731). While still arguably true, corpora have in these more modern times moved from trying to explain how to understand and limit language to expanding on how language is being used and developing. Furthermore, with corpus-based systems such as the Corpus of Contemporary American English (COCA), it is possible to represent the infinite amount of occurrences of language online – it just might not be humanly possible for an individual to personally evaluate them. It is here, that modern abilities of computational linguistics start to push the boundaries and pave new paths.

Corpus Linguistics in the Modern Age

The modern corpus is used to study multiple areas of language development and usage such as vocabulary and grammar. With the creation of corpora that can contain spoken language, such as the Santa Barbara Corpus of Spoken American English, the ability for linguists to use corpora to analyze grammar grew (McEnery & Hardie, 2013). Evaluating grammar and larger sections of language allowed for the opportunity to study collocates, lexical bundles, and other important aspects of language use. A foundational method for these novel uses of a corpus in large part came from the work done by Randolph Quirk and the Survey of English Usage (SEU)

in 1959 (McEnery & Hardie, 2013). This modern development was due to the bridge that was built between linguistics and computer science (McEnery & Hardie, 2013). As a result, by the end of the 1900's corpus-based methodologies had been used throughout linguistics-based fields (McEnery & Hardie, 2013).

Neo-Firthian Corpora Studies

In *The History of Corpus Linguistics* by McEnery and Hardie (2013) they describe a sub-set of study within corpus linguistics, *neo-Firthian* corpus linguistics after John Sinclair's research on the notable linguist J.R. Firth (p. 742). This perspective of study in corpus linguistics expanded on the word development to seek more than just grammar, but the structures within language itself. McEnery and Hardie first discussed their concept of neo-Firthian corpus linguistics in 1965 in "The Oxford Handbook of the History of Linguistics". In their 2013 description, this method of corpus analysis involves examining the collocate-based structures within the evaluated corpus rather than primarily the grammatical structures as was previously done (McEnery & Hardie, 2013) (where collocate is defined as a series of specific word groupings that have a frequency of occurrence greater than one). Additionally, Sinclair's writings describe meaning as being centralized in collocates and preconstructed phrases as much as the writing as a whole (McEnery & Hardie, 2013; Römer, 2011). This has been expanded upon in the past few decades as by Wolfgang Teubert in 2005, where "for Tuebert, collocation – and in fact corpus linguistics in general – is not a window into a mind-internal phenomenon; rather it is a tool for exploring meaning in discourse" (McEnery & Hardie, 2013, p. 743). It is by this evaluation of collocates that one might speculate on the relationship between these collocates and the user who produces them.

As writing becomes a process that can increasingly be analyzed and created by artificial intelligence, one might speculate that the way in which language is produced could become more formulaic than structures of grammar. Yet, as many can relate, lexical bundles are a part of our daily language production. These lexical bundles, or "multi-word prefabricated expressions" (Biber, 2004, p. 372), are a fundamental linguistic element that are brought to status and a concrete place within language by their frequency of use (Biber, 2004). With this then, how might a corpus introduce, share, and analyze lexical bundles and word forms in individual fields so as to determine at what level those entering the field must be able to produce? One such answer is the examination, and then comparison, of discipline specific corpora and their resulting word lists containing lexical forms such as hyphenated words.

The Standards of Corpus Analysis

Frequency

As has been demonstrated, the curation and analysis of corpora has been occurring for over half a century. With the advent of "The General Service List" in 1953, many criteria are used to evaluate words in a corpus, one of which is the frequency of occurrence (Coxhead, 2000). The standard of using frequency allows for words that are used frequently in a corpus to be evaluated on their position amongst other words and elements being evaluated. Once these word lists are established and set aside from the materials they came from, further disambiguation and evaluation can occur. Avril Coxhead writes in her "A New Academic Word List" (2000) that "the ideal word list should be divided into smaller, frequency-based sublists to aid in the sequencing of teaching and in materials development" (p. 214). Frequency is additionally important in terms of understanding one's lexical coverage.

Lexical coverage is defined as "the percentage of running words in the text known by the reader" (Nation, 2006, p. 61). Readers of anything must be able to understand the words within every sentence and then the structures that expand upon that. In areas where there is a significant amount of jargon there becomes an increase in difficulty to achieve this task. To evaluate the percentage of words that one must be able to understand throughout the text, lexical coverage is used. According to the research done by Hsu (2014), which evaluated the lexical coverage required for university freshman to understand their engineering textbooks a 95% lexical coverage was needed. Hsu's research demonstrated that there were different levels of word families for each area of research within her Engineering English Word List but that the 95% lexical coverage held true across these disciplines within the field. This is important to this research because CS, as previously mentioned, is a vast field with many independent and dependent offshoots into other disciplines while being held as one field. Despite this vastness, it may be assumed that regardless of varying levels of discipline-specific knowledge for every discipline within the field of CS, they will require, as Hsu (2014) found in Engineering, a 95% lexical coverage for comprehension of texts within the field.

The Size of a Corpus's Resulting Word List

The size of a corpus that is specific to a particular research area, such as a discipline or genre of speech, is to have according to John McHardy Sinclair (1991) "millions of running words (tokens) to ensure that a very large sample of language is available" (Coxhead, 2000, p. 216). While this is the standard size of a corpus taking from a discipline itself, the sub-disciplines that are evaluated within most overarching corpora have a much smaller number of tokens. The General Service Word List and one of the industry standard word lists had 7,822 types (West, 1953). The Academic Word List, which is the industry standard for an academic word list

contains 3,082 types (Coxhead, 2000). The Computer Science Word List built by Minshall (2013) contained 1,919 types. These counts of types demonstrate that there are fluctuations in what can be considered a "standard" size of a corpus, but that as the language is narrowed from general language to academic to a specific discipline or field the count of types within a word list is exponentially reduced. It is then expected that a word list that focuses on a particular field would result somewhere between the academic level of about three thousand to the field specific size of about one-thousand-word types. As the size of word lists dramatically reduces in size depending on how concentrated the fields get, there is an opportunity for an in-depth analysis of this language use for application in everyday life and for pedagogical purposes. These potential uses for word lists are concentrated in the field of English for Specific Purposes.

Corpus Analysis in English for Specific Purposes

The field of English for Specific Purposes (ESP) is a vast field but is defined as "the teaching and learning of English as a second or foreign language where the goal of the learners is to use English in a particular domain" (Paltridge & Starfield, 2012, p. 2). In Coxhead's Academic Word List, the individual sections such as the arts or sciences had hundreds of thousands of tokens, whereas the word list itself held into the millions of tokens (i.e., Table 2) (Coxhead, 2000, p. 220). If this standard setting corpus demonstrates that sub-disciplinary tokens can contribute to a meaningful whole despite being lower in number than as proposed by Sinclair (1991), then one can infer that corpora that are sub-disciplinary in nature are likely to result in a fewer number of tokens. Yet, a corpus should aim towards a sizeable amount of tokens in order to produce a sample size that is representative of that which is being evaluated.

Word Selection

There are multiple ways of establishing how a word is selected for a word list such as frequency, but additionally the range of the word and its keyness value. The range of a word is the amount of documents that use a particular word in a corpus (Heatley et al., 2002). The keyness value evaluates one corpus against another to evaluate the unique value of frequency from the corpus being evaluated to its comparison (Nekrasova-Becker et al., 2019). Once frequency has been established and the words have been separated from their host documents further analysis is required to understand their role within the language area being studied. It is common that a keyword analysis is to be conducted, for establishing the keyness value of a word asserts whether it is specific to the subject matter being evaluated or not (Motschenbacher, 2016; Nekrasova-Becker et al., 2019). Corpora that are generally used for comparison to in a keyword analysis are the Brown Corpus of Standard American English (1961), "the British National Corpus (2007) or the Corpus of Contemporary American English" (Nekrasova-Becker et al., 2019).

Expanding the Technicality of Language

Once a word list has been established by some criterion, a word list must provide meaning in some fashion (McEnery & Hardie, 2013, p. 743). To provide meaning in a field such as Computer Science, the language used must be explicit and precise in nature so that abstract concepts such as artificial intelligence and virtual worlds can be communicated across language boundaries. To evaluate language such as this, Nekrasova-Becker et al. (2019) categorized engineering words by their level of technicality. These levels were: general academic, semi-technical, and technical words (Nekrasova-Becker et al., 2019; Mudraya, 2006; Ward, 2009). While Nekrasova-Becker et al. (2019) generated these levels by categorizing a word based on its

activity of use, the ambiguity of fields within computer science made this process challenging to accomplish with a reasonable degree of certainty in this present study.

Word Lists Across Disciplines

Word lists across disciplines can differentiate between size, how they are analyzed, and what the results could mean for the field it is in. The breadth of research on corpora for word list curation is expansive. Spanning from Durovic's word list of Marine Engineering (2021) to student Biology papers by Swales (2014), a Medical Academic Word List (Wang et al., 2008), Agriculture (Martinez et al., 2009), Nursing (Yang, 2015), and Computer Science (Minshall, 2013). In each of these word lists there were a variety of methods and results. In Martinez et al. (2009) their agricultural corpus was a 826,416 word corpus and was based on academic journals and resulted in having 126,437 tokens. Martinez used the GSL and AWL for cross-comparison to differentiate between the discipline-specific words within their corpus. In Yang's Nursing corpus they used research articles to build a 1+ million-word corpus with 676 word families within. Yang (2015) used the AWL for cross-comparison pulling out the data from their corpus to construct the word list. These word levels and comparison methods are consistent across most, if not almost all, published word lists.

Pedagogy and English for Specific Purposes: The Now and Current Gaps

Graduate students across disciplines face many hardships while writing in an academic setting. This issue is heightened for those who are foreign-language learners of the dominant language. In disciplines, such as computer science, there is a notable stigma that surrounds the area of study and shrouds it in mystery as much as barriers of entry. To break down this issue and to understand how language access relates to the acquisition of language, many pedagogical

tools must be explored. One such tool for exploration is sub-disciplinary word lists that are regularly updated with recent literature to provide students of the discipline at all levels with access to materials that not only provide examples of use, but also potentially definitions.

Word lists are not new pedagogical tools, in the way that dictionaries are not new pedagogical tools for language acquisition. Yet, the issue with dictionaries in this modern age is trifold. Firstly, dictionaries are expensive if purchased in hard copy. Secondly, dictionaries are not accessible to many in the world whether it be online or lugging around a 5-pound book. Lastly, dictionaries are impossible to truly keep up to date. It is these challenges that limit access to potential students across the globe. With the growth in technology and the ease of compiling corpora, and subsequently word lists, there is now an opportunity to establish whether word lists that are discipline specific could prove useful to students in and outside of this discipline. Furthermore, using technology in the classroom has proven thus far to be an asset, rather than a burden (Baba, 2014; Mahini et al., 2012; Firmin & Genesi, 2013; Vrasidas & Glass, 2005).

To the instructor inside of a discipline, giving students access to external tools for self-development as well as quick guides to in-class materials (source) improves student success. Furthermore, for those who are entering joint masters and doctoral programs, often there is a large discourse community bridge that must be crossed. In this time students could get lost and start to fall behind. This gap could arguably be larger for students that are not “native” speakers of the language being used to discuss their area of study. In terms of XR, within the next decade it is likely that there will be an increase in integrating XR technology into the classroom (Meccawy, 2022). This integration would likely find that it could be seamlessly integrated into the classroom if the general public had access to the language to both understand and discuss the

technology without fear of shame in sounding ill-informed, or even learning false information regarding the tools due to a lack of linguistic access to the truth.

A second challenge of heavy, costly, and out of date discipline specific dictionaries is that instructors who wish learn more about a different discipline are often greeted at the gates of that discipline, shamed by their understandable ignorance, and quickly sent back from whence they came. Collaboration starts where there is common ground. To increase access to collaboration opportunities between students and their instructors, between students and students, instructors and instructors, and researchers to researchers, there must be common ground (Ābeltiņa & Rizhamadze, 2021). Language has long been the first intermediary for two different entities to meet and bond in collaboration. If dictionaries, as they are today, are inaccessible then creating the foundation for language is further inaccessible to the common person. It is because of this that word lists provide an opportunity for quicker and up-to-date access to discipline-specific languages. Currently, such resources do not exist. This research aims to evaluate this gap and determine whether potentially meeting these needs is both feasible and worthy of investment.

Computer Science and English for Specific Purposes

In "When a bug is not a bug: An Introduction to the Computer Science Academic Vocabulary List" by David Roesler (2021), a modernized word list was created for fields of computer science in response to expanding on Minshall's (2013) Computer Science Word List. In the time since 2013, Roesler expanded upon the work by Minshall (2013) by filling the gap of not including textbooks as part of the corpus for analysis, as well as distinguishing between the role of headwords and their derivational forms as potentially representing different discipline specific meanings. Which is to say, that in other disciplines often a headword and its derivational

form can present with two very different meanings, whereas in an alternative context the opposite may be true (Roesler, 2021).

Roesler's analysis of the corpus of computer science demonstrates that there is always room for improvement upon word lists. This is furthered by his call to the field for a word list to be constructed in the discipline of computer science with hyphenated word forms. Roesler (2021) writes "a comprehensive multi-word CS vocabulary list that also includes hyphenated forms and a list of common collocations may be a useful project for future research and could provide learners with an additional level of description of the language of academic CS" (p. 10). The research within this thesis answers this call by providing hyphenated word forms in a discipline of computer science, Human-Computer Interaction as well as Extended Reality.

This research uses a frequency, keyness, and range-based analysis of a corpus to construct a word list. This word list is in the sub-discipline of Extended Reality (XR), which constitutes the more widely known fields of Augmented Reality, Virtual Reality, and Mixed Reality. Due to the sub-disciplinary nature of the corpus, the token level keeps in line with Sinclair's (1991) suggestion of a minimum of 1 million tokens whilst providing a smaller amount of types, or single occurrences of a single word. It is with these types that once disambiguated from non-specific words to the discipline of XR, it was necessary to make meaning from the word list to engage in a further categorization process. This categorization of reducing the words to a technical to non-technical scale was done with human participants that are both learners and experts in the field of XR. It is with this insight that pedagogical applications of sub-disciplinary word lists as the result of a corpus could be further evaluated.

Research Questions

The research questions for this study are multifold, in that they ask if particular pieces of words lists are important, how they are interpreted by users, and if they may be long-term useful pedagogical and collaborative tools across academia.

1. Is a word list based on a corpus in Extended Reality possible to create? To what extent do frequency, keyness, and range contribute to the creation of a word list in Extended Reality? With a word list in Extended Reality what role do hyphenated word forms play in the accuracy and representation of the field?
2. Does soliciting feedback from graduate students and/or professionals in the field on a discipline-specific word list help shape or reshape the of the efficacy and utility of the word list?
3. When computer science and non-computer science graduate students are presented with a word list in Extended Reality, what are their perceptions of that word list as an educational tool?
4. Do graduate students see a use for discipline-specific word lists as a potential tool for incentivizing cross-disciplinary conversations and/or collaboration?

Chapter 3: Method

This study required three phases before completion. Phase One of the study required the construction and analysis of the new corpus of XR. Phase Two consisted of a survey that asked graduate students of XR to evaluate the resulting discipline specific words generated by the corpus to determine their efficacy. The final phase, Phase Three involved new questions that were surveyed to the same graduate students of XR and additional graduate student teachers in the department of English. The word lists in Phase Three answered the research questions by confirming the usability and evaluating how the users interacted with the contents of the word list. These three phases construct the entirety of the research, and each will be presented individually despite their minimal overlaps to account for all elements of the study (each phase can be seen in Abstract F).

Phase One

The first phase of the research involved curating the corpus in the computer science sub-discipline of HCI, which then has the sub-sub-discipline of XR. Due to the intertwined nature of the umbrella, HCI and XR, HCI was considered in the curation and analysis of the corpus. To account for the texts that graduate students would read the most, academic journals and conferences proceedings were the two sources for this corpus. The leading associations in computer science are the Associations for Computers and Machinery (ACM) and Institute of Electrical and Electronics Engineers (IEEE). From these associations, the conferences of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE Annual International Symposium on Virtual Reality (IVR), and the ACM Virtual Reality Software and Technology (VRST) were selected for the corpus document compilation. The journals that were

selected for the corpus curation are the IEEE Transactions on Visualization and Computer Graphics (TVCG) and the ACM Transactions on Computer-Human Interaction (TOCHI). Each of these venues were identified via personal communication with a professional and academic researcher in the field (F. Ortega, Personal Communication, February 15, 2022).

The ACM and IEEE conference and journal sources were accessed through the database access provided by the institution. To ensure that the corpus represented the development of language in the disciplines of HCI and XR, files from January 2012 – December 2022, were collected where possible. The conference IEEE ISMAR had not yet occurred by the date of collection, and therefore 2022 was not collected. Additionally, the last issue of TVCG 2022 had not been published by the time of collection, and therefore was also not included to be a part of the corpus. To ensure that there was a similar size in the data collection of journals and conferences, the top 9 by citation publications were collected from ISMAR, IVR, and VRST for each year. Similarly, the top 12 papers by citation in journals from TVCG and TOCHI were acquired for each publication produced in a single year. This selection process resulted in the following number of downloads:

Table 1
Corpus Downloads by Publication

Conference Proceedings		Downloaded Files by Count
	ISMAR	90
	IVR	99
	VRST	91
	Total Conference Proceedings	280
Journal Articles		
	TVCG	390
	TOCHI	182
	Total Journal Articles	572

When downloading these documents, there were a few issues in obtaining the documents themselves. Firstly, the ACM database limits the amount of daily downloads to one hundred files. Furthermore, the ACM database does not list the published articles/proceedings by values such as citation count, but rather by keywords or discipline-specific sub-categories as seen in the rectangle box of the ACM Spatial User Interaction example in Figure 1. While this is likely useful for those in the field who would be locating by keyword, as a non-expert, I did not have the capability to discern which categories would be appropriate to include for this corpus as there were no clear boundaries for maintaining the consistency of the data.

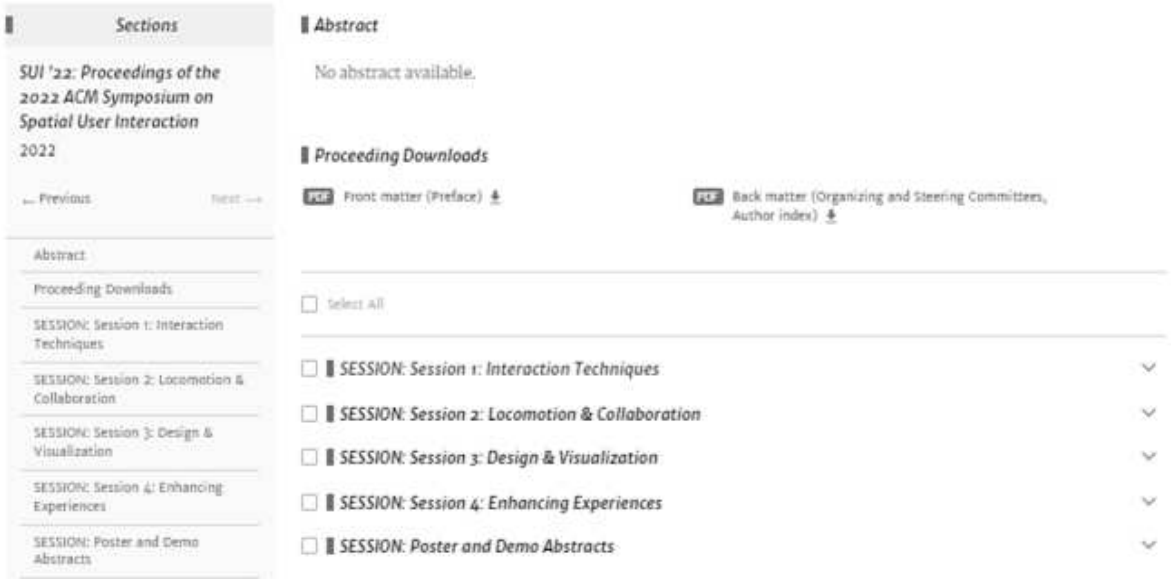


Figure 1 - Conference Page Example for ACM Download

Therefore, this limited the size of the corpus and the materials used to obtain the results. Other incidences of note in the download process were as follows:

- When there was a tie for top citations, I chose the paper with the most downloads to citations out of those with matching citation counts.

- All IEEE rankings were based on number of paper citations as given by the IEEE website. If more than one paper had identical number of citations, the papers were downloaded in the order provided by IEEE (there was no observed reasoning for the order, not alphabetical, not number of patents – possibly random).
- Regarding special issues, for TOCHI, volume 21 was not included twice, despite being shown as published twice (once in late 2015 and then again in early 2016).
- Lastly, for TOCHI, the special issues were cancelled in 2017 and were not collected further.

Each of these decisions were to ensure that the data being collected were as representative of the domain being evaluated and that there was continuity between the different conferences and journals.

Data Processing for the Word List (non-hyphenated)

Antconc

To start the Corpus of Extended Reality (CoXR) compilation process, first the saved pdf files were converted to text files through the AntFileConverter program (AntFileConverter, 2022). Each text file was saved by its original categorization and retained in a folder for the conference or journal (henceforth publication venue) from which it originated. The text files are necessary so that they may be analyzed by the AntConc corpus analysis software (AntConc, 2022). The AntConc software was used to extract a word list from each publication venue. This word list was then copied into a Microsoft Excel sheet for each publication venue. The individual sheets allowed the analysis of the word lists to occur independently before the collective analysis.

Removing Irrelevant Components from the Data

To remove irrelevant components from the data, such as symbols and characters, tables, and visuals from the text files used to generate the word list the AntFile Converter (2022) was used to transform publications from .pdf to .txt form. Then to remove symbols, tables, and visuals from the .txt files (where they remained) the word lists were generated in AntConc (2023) which by default removes these items. To then narrow this process further, each publication venue was reduced to contain words with a frequency of at least 10 words. The number of 10 words was chosen because by an observable degree a majority of the words within the corpus that were misspelled words, symbols, or broken hyphens (i.e. the algorithmic discards from the process AntConc uses) had a frequency of ten or less.

To start the process of determining whether a word should be considered discipline specific or not, the word lists were limited to the frequency of occurrence of 20, 25, 50, 100, 200, and 250. This was done to ascertain the exponentially declining rate of frequency of occurrence to the words within the corpus because this is a frequency-based word list. Since this is a small level corpus that focuses on a sub-discipline the words that would be available in each word list occur less frequently the standard academic corpus that would be available in more commonly published work. Therefore, to understand where the rate of decline limited discipline-specific words it was crucial to identify the frequency of occurrences for all of the publication venue word lists. The number of occurrences that resulted was 50 occurrences within each individual word list. This constraint allowed for words such as *haptics*, *cybersickness*, and *avatar* which would not have been visible if the limit was set to a more traditional 100-250 level frequency of occurrence. Since the goal of the corpus is to isolate the discipline specific words to XR it was

essential to keep to the purpose rather than what is traditionally done so that a thorough analysis could be completed for the following phases.

With this limitation of frequency all other words were deleted from the individual publication venue word lists resulting in only words that had the frequency of at least 50 occurrences in their respective word lists. With this limitation, the word lists were combed through to delete remaining proper nouns that managed to get through such as *IEEE*, *Occulus*, and *Meta*. Furthermore, names that were remaining were removed such as *Xi*, *Zhou*, and a collection of other surnames. Lastly, any individual letters were removed and symbols that made it through the frequency elimination process. These word lists were then ready to be broken down into discipline specific words.

Isolating the Discipline Specific Words

To isolate the discipline specific words, the individual publication venue word lists were compared to the MIT 10,000 word list (MIT, 2020) and the Coxhead Academic Word list (Academic Word List, 2000). The MIT top 10,000 words in use in the English language was used as it is a current word list developed by Dr. Eric Price in his time at the Massachusetts Institute of Technology (MIT). This word list is additionally open-source, of which this factor and the more current nature of its curation is different than the generally used General Service Word List (1953) and therefore deemed more appropriate for a more modern word list in its curation. The Academic Word List (AWL) was used as it is the only leading standard academic word list in this discipline at this point in time. These current word lists allowed for a discipline-specific word list to be constructed using what is believed to be the best resources currently available.

The isolation of discipline specific words was done in Microsoft Excel using the VLOOKUP function. Utilization of this function allowed the cross-referencing of the CoXR Word List with the external word lists. Once the words were appropriately designated as being present in an external word list or not, the words that were present in both the MIT 10,000 word list and the Coxhead Academic Word list could be removed as they were likely not discipline specific words if they have a high frequency of occurrence. The following word count for each publication venue is as follows:

Table 2
Discipline-Specific Words by Publication

Conference Proceedings		Word (tokens) by Count
	ISMAR	172
	IVR	134
	VRST	139
	Total Conference Proceedings	445
Journal Articles		
	TVCG	927
	TOCHI	726
	Total Journal Articles	1,653

Each of these lists were then combined into one larger word list, and a word list by their publication type, conference proceeding or journal article, and then further into specifically the publication venue and disciplinary location. This was done to isolate the discipline specific words by their genre.

To evaluate the number of shared words between the publication venues, the VLOOKUP function was used to cross-reference individual publication venue documents. From this cross-referencing, a word list for having all five venues sharing a word was created. To further examine the shared words between publication venues, they were additionally limited to the occurrence of appearing in at least 4, 3, and 2 publication venues. This information was used to

further extrapolate whether discipline specific words can be discovered through this analytical process, or if it neglected the inclusion of important discipline specific words. The word list that included an occurrence in all five publication venues was analyzed in AntConc for its keyness and range to further understand if it is specific to the discipline or not. This process was to analyze whether frequency was enough for justification to be discipline specific, or if keyness and range also played a significant role.

Keyness

With AntConc, the Brown Corpus List (Brown Corpus of Standard American English, 1961) was used as the base corpus for a keyness analysis of each word in the CoXR Word Lists. The keyness analysis is important because it demonstrates which words are considered less frequent in everyday language use versus their use within the CoXR corpus. Following the keyness entry for each word that occurred in the CoXR Word List the range was also given. This value was obtained by using the AntConc Concordance analysis. The range was determined by imputing each word into the Concordance search line of AntConc (2022), and then every file entry accessed was copied into a Microsoft Excel worksheet. Using the *Remove Duplicates* function in Microsoft Excel, the files that were duplicated in the concordance results were removed, and therefore only the file names of each word remained. Once the frequency, keyness, and range were calculated for each of the final resulting words for all five publication venues as a whole entity, the examples for the word list could be accessed.

Obtaining Examples

To choose examples of each word in the CoXR corpus, each publication venue was evaluated for the highest keyness value for each word. The highest keyness value was selected

because it demonstrates that the most unique and frequent uses of the word lay within a particular publication venue. Once the highest keyness value has been chosen, the example was selected using the Concordance Plot option through the search function on AntConc. Each example was selected near the beginning of the publication chosen because it was often the clearest use of the word without other technical jargon and often explained or elaborated upon the word in some way. Each of these examples was then compiled alongside the frequency of occurrence, keyness, and range for each publication venue. These results constructed the CoXR corpus is the materials used in Phase Two of this study.

word	ISMAR	keyismar	rangeismar	TVCG	keytvcg	rangetvcg	VRST	keyvrst	rangevrst	TOCHI	keytochi	rangetochi	IVR	keyivr	rangeivr
annotations	82	146.34	13	219	92.25	45	90	171.7	11	78	44.95	27	87	158.91	5
augmented	1377	2359.82	87	1731	646.47	60	457	791.89	56	242	86.82	64	527	880.62	73
dataset	108	192.75	27	1985	836.32	6	63	120.19	13	385	221.89	51	125	228.33	16
discomfort	80	101.45	10	77	0	26	62	79.79	20	119	33.9	26	54	62.34	17
explored	53	47.45	33	383	97.47	163	89	111.19	39	403	161.21	116	119	153.27	47
find	152	271.28	54	1062	447.4	185	82	156.44	53	439	253.02	144	125	228.33	61
first	509	908.57	86	3467	1460.95	119	242	461.72	90	1453	837.56	64	441	805.66	87
gaze	181	245.78	20	754	233.94	34	338	552.04	24	1073	519.74	25	295	449.87	32
gesture	207	214.98	20	179	0	38	141	134.46	29	422	100.4	42	131	110.65	28
immersive	310	553.3	43	1008	424.65	96	643	1227.04	66	104	59.94	20	564	1030.43	74
interacting	56	90.93	28	231	87.74	106	69	122.13	33	283	152.58	101	77	130.97	31
manipulation	138	197.58	23	478	151.38	128	292	497.54	33	118	33.44	43	150	223.95	29
metrics	68	121.36	24	962	405.27	141	63	120.19	19	332	191.35	42	54	98.63	17
multimodal	50	89.23	13	220	92.67	48	66	125.91	24	145	83.57	42	69	126.03	14
occlusion	214	361.34	36	629	244.61	103	82	139.48	20	66	25.53	9	71	113.4	19
perceptual	136	194.21	34	692	236.39	38	140	217.62	31	146	46.67	30	64	78.36	30
physiological	141	145.83	12	56	0	32	128	140.54	18	200	32.74	38	140	149.57	11
questionnaires	79	103.94	20	84	0	42	165	268.65	49	118	36.58	50	76	102.04	54
randomly	61	93.24	35	313	114.27	79	55	89.54	26	160	76.2	54	52	79.92	37
realism	59	30.76	24	244	0	62	164	192.64	40	84	0	10	153	163.59	42
sensory	59	53.43	16	178	25.33	44	73	81.74	28	139	29.55	37	51	44.11	22
significant	468	835.37	71	1502	632.79	100	121	230.84	63	865	498.57	135	552	1008.5	73
specific	145	258.78	51	1379	580.97	242	86	164.07	84	932	537.2	153	115	210.06	61
statistically	76	108.09	17	269	84.88	11	87	137.05	33	206	90.19	48	95	144.12	25
stimuli	56	70.62	14	334	105.22	72	82	123.03	25	230	98.01	28	64	86.15	25
subjective	124	132.3	30	457	99.18	24	147	184.24	42	291	80.29	58	131	147.93	41
usability	154	274.85	30	414	174.4	100	160	305.25	30	501	288.75	73	205	374.47	38
visualization	484	850.62	66	13923	5855.54	224	318	594.19	63	558	309.73	62	528	951.12	72
visually	100	148.77	41	892	337.79	242	74	113.51	36	209	91.81	44	63	88.87	31
wearable	66	117.79	34	109	45.92	38	112	213.67	26	293	168.87	48	51	93.15	21

Figure 2 - Main CoXR Word Llist with Frequency, Keyness, and Range

Hyphenated Word List

To build the hyphenated word list, code in Python was developed. The pseudocode shown in Table 3 describes the sequence of steps to count the number of hyphenated words in text. A directory with the name of the conference contains all files with extension '.txt' and are obtained after processing a paper through AntConc. A table entry, denoted WordCountLog in the pseudo code is initialized to log (register) the occurrences and count for the respective hyphenate

word. A loop is initialized to iterate over all files in the directory and shown by the syntax "for…do". Each text file is read with an encoding of "UTF-8'', a character encoding system that considers each individual element as an ASCII character that can be read by a computer program. The entire text read from one text file is available in the variable denoted as AllText. This variable is then iterated over using another for…do loop where every entry corresponds to a line in the text.

Table 3:
Algorithm 1: Counting hyphenated words in text

```
Data: Directory containing text files of papers from a conference
Result: Log of all hyphenated words with their word count

ListOfFiles:= Files of extension '.txt' in a conference directory
WordCountLog:= Empty table entry to log respective word and its count
begin
   for TextFile in ListOfFiles do:  #Iterating over each file in the list of files
AllText:= Text data in TextFile
      begin
         for Line in AllText do:  #iterating over each line in the text
               WordList:= List of all words in Line
               begin
               for Word in WordList do:  #iterating over every word in the list of words
               if '—' in Word then #hyphen detected in word
               if Word present in WordCountLog then  #hyphenated word belong to word count
               log
               WordCountLog[Word] += 1  #Increase the count of the word by 1
               else: #hyphenated word does not belong to word count log
               WordCountLog[Word] = 1  #Add the word to the word count log, and initialize
               the count as 1
               end
      end
end
```

Each line in the text is denoted as a collection of words and denoted as WordList. Since it is a collection, the variable can be iterated over using a "for…do" loop to get the individual word in the list. Once a word is available, a conditional block is applied in it to determine whether it is

hyphenated or not. If the word in question contains a hyphen (‘’—‘’) then it is added to the WordCountLog table. If the word exists in the table, then the count for it is increased by 1, otherwise it is added to the table with an initial value of 1. As a result, each word in the table will have a lower bound of 1, and no specific upper bound on the value. Once all the text in every file in the directory is processed, the WordCountLog table is returned by the function.

The results of the code were copied into csv files. These csv files contained the frequency of occurrence, the hyphenated word, and the documents the word occurs within. It was from this list that the range for each hyphenated word could also be found. The hyphenated word list was categorized by frequency, and like the main word list limited to 50 occurrences in frequency as a cut off line for acceptance into the general word list.

Phase Two

The Qualtrics Study, Compilation

To build the survey needed for expert evaluation of the word list, a completely anonymous survey structure was needed. The Qualtrics system (2023) was used in its anonymous mode which means that no identifiable information such as an IP address, email address, or name were collected. This was done to ensure that all participants could remain anonymous and the study could be based on collected demographics alone. The questions asked respondents their degree of study, year of study, and whether the English Language was their first language. Following this, each word of the corpus was presented along with its coordinating example. The questions for each word and its example are as indicated in Figure 3.

Of these questions, questions one and three were given on a yes/no answer option . This was determined as appropriate as there are no other potential responses. Questions 2 and 4 were

text response questions that allow the participants to write freely. Question 5 was multiple choice with the options of *discipline specific*, *somewhat discipline specific*, and *general academic language*. The purpose of this question was to test the efficacy of the corpus itself on whether it lies more heavily towards being discipline specific or general in academic language according to graduate students currently studying in the field of XR.

Delivering the survey

Once the survey had been completed, the survey was sent via a webpage link to an attached standard practice email to graduate students in the Computer Science Department at a large, public, research one university in the western portion of the United States of America. Specifically, the students recruited are graduate students of an XR specific laboratory, of which only focuses within the realm of HCI and more specifically, XR. These students are the primary audience for this research and are, therefore, the optimum audience to test the efficacy of the CoXR corpus list. This trial phase was to test the words used and was only a small part of the study as a whole. To further understand which examples are best used, if only strict data analysis can provide a discipline specific corpus, and whether the corpus is useful for pedagogical applications outside of the field of XR, phase three was immediately implemented following the results of phase two.

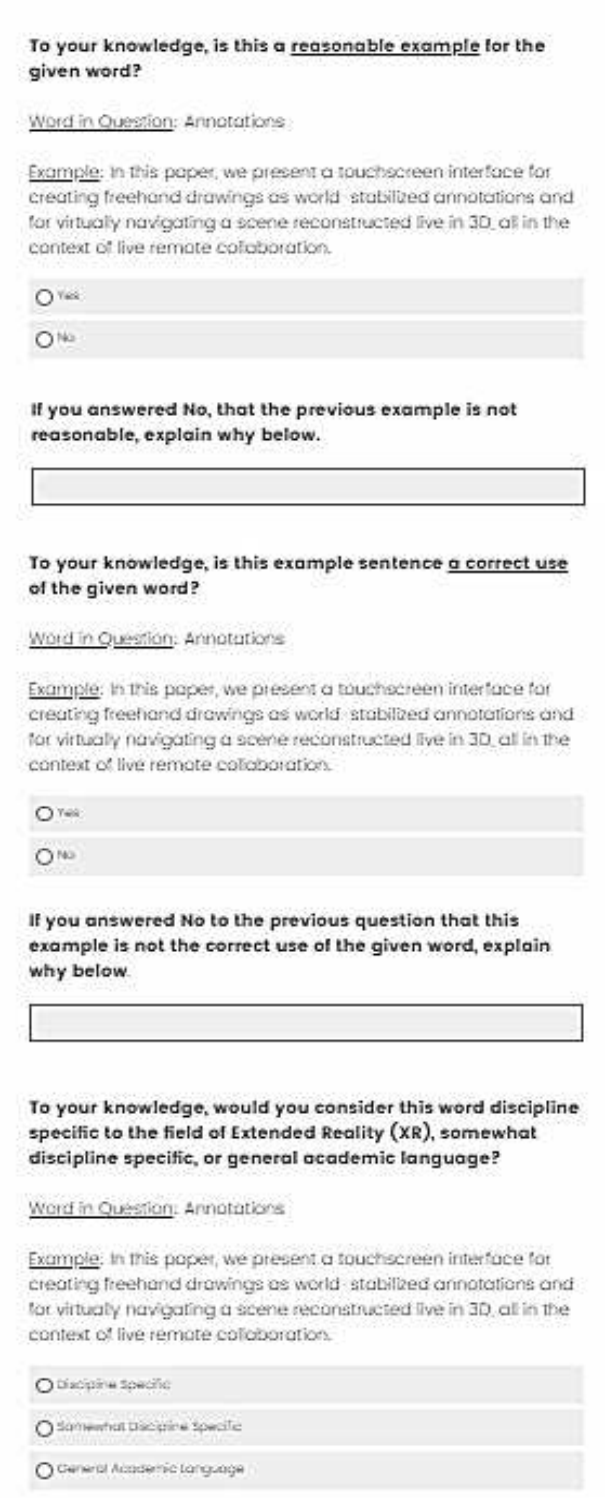


Figure 3 - Phase Two Qualtrics Questions Format and Example

Phase Three

In the third and final phase, the results for Phase Two were analyzed. From these results, the words that were analyzed in Phase Two informed which words would be present in Phase Three, as well as the examples used. The results of Phase Two indicated that some of the words were discipline specific and pointed towards a certain keyness value (see Results for more information). Since the keyness value seemed to inform whether or not a word might be considered discipline specific, the Phase Three word list included words that had a total document frequency of occurrence of in at least three documents. Each word's example was curated in the same manner as the word list used in Phase Two of this study.

Once the words that would be used in Phase Three had been gathered, the survey was constructed. The surveys used in Phase Three are in two parts, (1) the survey that would be used to evaluate non-XR discipline specific graduate students, all of whom have teaching experience, (2) the original word list of Phase Two with updated examples and including the word list that was gathered post-addition of the total document frequency of three or more occurrences. The second survey was delivered to graduate students and professionals (1 postdoc, and 1 assistant professor). Of the all of the surveys, the questions remained the same. The questions asked started with demographic questions that asked (1) the person's level of study or professional experience, (2) what year of study they are in (if a graduate student), (3) if the English language was their first language, and (4) what their current program of study was. The fourth question was used to survey the non-XR discipline specific population only.

Post-demographic questions, the words within the survey and their examples were examined individually with each of the following questions for every word: (1) Do you understand this example, (2) Do you understand this word when it is used in the given sentence,

(3) With the example provided, could you use this word in a sentence, and (4) If you encountered this word later, might you be able to understand its meaning. For all of the questions for the words and their examples, the answers were in a multiple-choice format with only “yes” and “no” as the options for answering the question.

Following the survey for the non-XR discipline specific population the participants were asked, and audio-recorded, answering the following questions immediately after completing the survey:

- After completing this survey, do you see a use for a detailed word list that is discipline specific in your classrooms?
- In your opinion, what are your questions or concerns about using this word list in your classroom?
- If this word list were published with each word and its following example, what would you, based on your perspective and experience, do to improve the word list in terms of its accessibility and user friendliness?
- If you were to use this word list for personal use to communicate with someone in Extended Reality, how comfortable would you feel? Explain why or why not?
- Did you find the hyphenated words different in difficulty to the single word examples? Why or why not?
- Would the use of a discipline specific word list such as this provide you with incentive to collaborate and/or explore across disciplines if you had access to it?
- As someone encountering some of these words for the first time, what was your experience like in terms of inferring a definition based on the example given?

These questions were answered by every participant in the non-XR study group via an audio recording by Zoom (2023). The recordings were saved in a protected file on the co-Investigator's computer and later transcribed in a manner that left each participant anonymous.

The survey for the XR discipline specific population in Phase Three was sent as a link via Qualtrics to the participants. This study population was evaluated to evaluate whether this word list could be used as a potential tool in CS Education as well as a guide for graduate students for their acquisition of discipline-specific language. The second survey for this portion of the study included the same demographic questions in a pre-survey and the same question regarding each word and its example. The difference of the post-word list survey questions are the exclusion of questions relating to pedagogy as this participant population did not have unanimous experience in teaching previously. As such, in an open text box question format, the second survey asked in the post-word list evaluation survey the following questions:

- After completing this survey, do you see a use for a detailed word list that is discipline specific in your classrooms as a student?
- In your opinion, what are your questions or concerns about using this word list in your classroom as a student?
- If this word list were published with each word and its following example, what would you, based on your perspective and experience, do improve the word list in terms of its accessibility and user friendliness?
- Did you find the hyphenated words different in difficulty to the single word examples? Why or why not?
- Would the use of a discipline specific word list such as this provide you with incentive to collaborate and/or explore across disciplines if you had access to it?

The questions resulting from the two surveys within Phase Three were coded individually and the results were published individually for collective analysis.

Transcription

To transcribe the Zoom (2023) recordings, that were audio recording only, for Phase Three all of the in-person interviews were conducted by recording with video off on Zoom. This source was chosen for its ease of access, availability to the researcher, and that it stored the recorded files in a secure location on the researcher's personal computer. Once all of the files were uploaded as MP4 files, audio only, to the researcher's personal computer they were played to Microsoft Word using the "dictate" function. The "dictate" function is an audio to text function provided by the Microsoft Word software. This ability allowed the text to be auto-generated, and then to ensure accuracy the researcher reviewed the transcript and modified incorrect words where necessary. Since one of the sessions had a recording that failed to properly record and be saved on the researcher's personal computer, Participants two through five were not recorded. However, in the case of this issue, the researcher took notes of the exact words used by respondents in the in-person interview for Participants two through five. These results were transcribed verbatim in response to each question. Finally, to compile the results the answers to the questions on their "yes, somewhat, or no" response was recorded in a spreadsheet for evaluation in the results.

Chapter 4: Results

The results of this research are presented by their phases because each phase is an individual portion of research and while the phases built upon one another to answer final research questions, they are also independently answering either field calls and/or research questions in their own right. Phase One answered the call in Roesler (2021) for a hyphenated word list in CS and developed answers towards research question one. Phase Two introduced novel methodology for testing the efficacy of a word list before its introduction into application and was foundationally important for answering the final research questions in Phase Three. In Phase Three the two-part survey process evaluated not just the CoXR Word List but additionally the hyphenated word list. In addition, Phase Three was developed to answer pedagogical questions and the final three research questions of this study.

Phase One

In Phase One the documents that were downloaded from ISMAR, IVR, TOCHI, TVCG, and VRST resulted in some differences of the number of downloaded documents. The top nine 9 papers for every given year in conference proceedings resulted in 90 downloaded documents for ISMAR, 99 for IVR, and 91 for VRST. The top 12 papers by citation for every year collected in journal articles resulted in 182 downloaded documents for TOCHI and 390 for TVCG. The citation count for each publication venue was the main criterion for being downloaded. The resulting average of citations for the conference proceeding publications is $M = 25.75$ ($SD = 5.8$), whereas the resulting average of citations for journal articles is $M = 49.35$ ($SD = 10.75$). This resulted for the entire corpus the average citation count of all combined documents to be $M = 35.19$ ($SD = 14.59$). With these results, there might be some assumption as to potential

outliers. This concern is valid, and demonstrated by both minimum and maximum values by publication in the following table:

Table 4
Minimum and Maximum Citation Values for CoXR Corpus Files

Publication Venue	Minimum # of Citations per paper	Maximum # of Citations per paper	Average Citation Count
ISMAR	0	248	30.1
IVR	0	146	28.03
VRST	0	104	19.12
TOCHI	1	389	41.75
TVCG	1	943	56.95

The entire corpus itself resulted in 130,249 types and 9,476,100 tokens. By publication venue these numbers vary and are best represented by Table 5:

Table 5
Types and Tokens for CoXR Corpus Files

Publication Venue	Types	Tokens
ISMAR	24,352	711,015
IVR	22,348	686,386
VRST	25,840	641,974
TOCHI	55,738	3,067,285
TVCG	88,487	4,369,440

These results show that based on the download count there was a comparable type value of words in each publication venue. The average number of types per publication venue was 261 (SD = ± 34.1).

Table 6
Types and Tokens to Downloaded Papers for CoXR Corpus Files (rounded up to nearest whole number)

Publication Venue	Types	Tokens
ISMAR	271	7,900
IVR	226	6,933
VRST	273	7,055
TOCHI	306	16,853
TVCG	227	11,204

As a result of these similarities, when the word list was constructed, there were quite a few shared words, but very little shared exactly across all publication venues. All five publication venues were evaluated here to understand what words specifically were being used across all publication venue types. Furthermore, evaluating all of the publication venues together allowed for HCI, AR, and VR dominant publications to be evaluated together. This was problematic, in that since the type count was so similar, the word lists for each publication venue easily cancelled out other words from the word list. This is to say that if one word list had 500 types to compare against one with 100 types, there are 400 extra word types that could be used for cross comparison. Whereas, in the case of this word list, when there are 226 words being cross compared with 227 types, the resulting word list is almost at a 1:1 ratio, resulting in significantly less potentially shared words (despite being within the same discipline) by being given a broader array of words to choose from.

In total, the word list resulted in 30 shared types across all publication venues. The shared types all have a frequency of at least 50 occurrences within their respective publication venue. These types are as seen in Table 7.

The results of Table 7 demonstrate that despite having 130,249 types in the corpus, the word lists themselves once reduced to not include general everyday language and academic language result in a much lower number of types. To ensure that the types were being looked at accurately, the occurrences across publication venues were evaluated. The result was, that with four out of the five publication venues being counted there are 35 shared types. For three publication venues being evaluated, there were 69 shared types and for two publication venues being evaluated there were 259 shared types. This growth did not result in a clear pattern. For example, the cross comparison of two venues alone did not display a pattern as to why certain

types were shared across publication venues. As a result, the word list generated for this phase included only the 5 shared venues of the CoXR Word List.

Table 7
Shared Types for All Publication Venues & their Frequency of Occurrence Normalized to One Hundred Thousand Words

Word	ISMAR	IVR	VRST	TOCHI	TVCG
Annotations	12	13	14	3	5
Augmented	194	77	71	8	40
Dataset	15	18	10	13	45
Discomfort	11	8	10	4	2
Explored	7	17	14	13	9
Find	21	18	13	14	24
First	72	64	38	47	79
Gaze	25	43	53	35	17
Gesture	29	19	22	14	4
Immersive	44	82	100	3	23
Interacting	8	11	11	9	5
Manipulation	19	22	45	4	11
Metrics	10	8	10	11	22
Multimodal	7	10	10	5	5
Occlusion	30	10	13	2	14
Perceptual	19	9	22	5	16
Physiological	20	20	20	7	1
Questionnaires	11	11	26	4	2
Randomly	9	8	9	5	7
Realism	8	22	26	3	6
Sensory	8	7	11	5	4
Significant	66	80	19	28	34
Specific	20	17	13	30	32
Statistically	11	14	14	7	6
Stimuli	8	9	13	7	8
Subjective	17	19	23	9	10
Usability	22	30	25	16	9
Visualization	68	77	50	18	319
Visually	14	9	12	7	20
Wearable	9	7	17	10	2

The hyphenated word list resulted in both whole and spliced values which reinforces a similar challenge faced in Roesler (2021). Despite this, due to the provided frequency of use and the range of the words they showed a clear picture of what the hyphenated word forms may look

like within the publication venues being evaluated. To maintain consistency across the main word list to the hyphenated word list, the hyphenated word forms with a frequency of occurrence of 50 or above were evaluated. This resulted in 274 results. The highest of these results was the frequency of 5,035 occurrences; however, this occurrence was the result of being heavily cited and therefore could not count directly as a representation of the language within the texts being evaluated.

Following the first and largest hyphenated form of occurrence is “real-time” which was used 1,944 times across a range of 406 documents. The large range of documents where “real-time” is used demonstrates its importance within the literature of the discipline; however, it is not specifically discipline specific. Other words of high frequency that may be considered specific to the discipline of CS and not necessarily XR, based on the author’s experience, are: (hyphenated word, frequency) “human-computer, 806”, “multi-user, 242”, and “self-avatar, 162”. While there are cross-disciplinary hyphenated word forms high in the frequency of occurrence, they are also specifically used within the discipline of XR to convey particular points relating to the machinery used in the discipline such as “see-through, 771”, “head-mounted, 727”, “image-based, 365”, and “full-body, 359”. What is notable about these words is their common collocates which limit them specific to the discipline such as a “see-through optical lens” and a “head-mounted display”. It is therefore then that the hyphenated forms must be evaluated not just on their frequency or their range, but by their top collocates. An example of this can be seen in Figure 4) below which is a screenshot of the AntConc Software as it demonstrates the collocates of the word “see-through”.

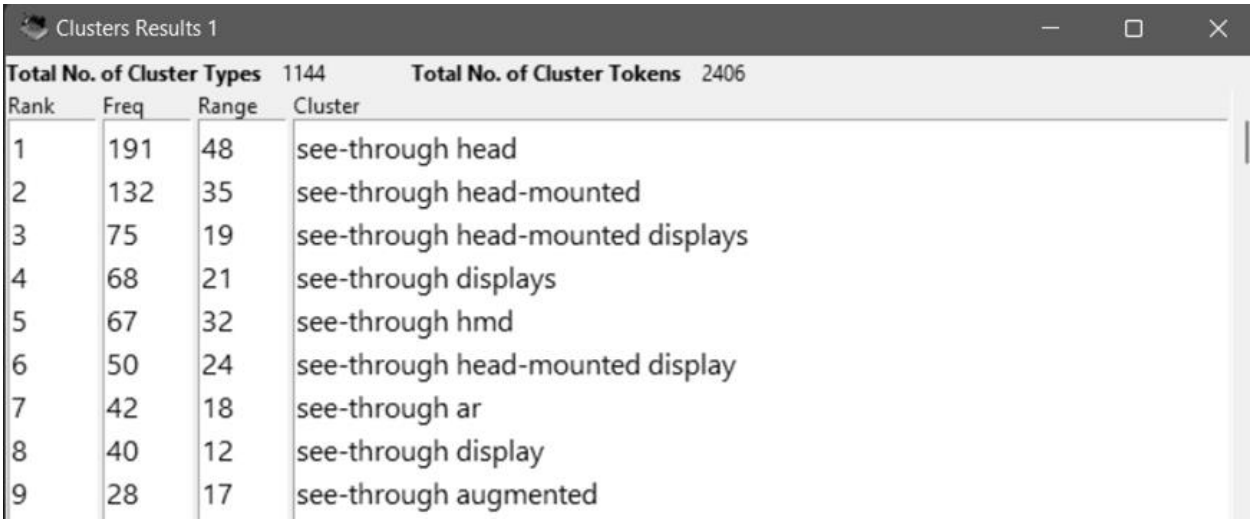


Figure 4: AntConc Results Example for Hyphenated Words and their Collocates

Phase Two

The participants of the survey were asked a series of demographic questions. Two students responding to the survey, one a doctoral track student and the other a masters' track student. All participants were either in their second or third year of study. Importantly, however, the two respondents to the survey were split on being L1 or L2+ speakers of the English language. With these results we are shown in Table 8 that the demographics are fairly similar in their years by program of study and they represent the two demographic types, L1 and L2+ speakers.

Table 8
Phase Two Survey Demographics

Participant Number	Degree Type	Year of Study	L1 or L2+ Status
1	Doctoral	2	L2+
2	Masters	2	L1

The reactions to the word list were mixed. Participant Two provided more written feedback, generally in a negative response to the examples, but scored higher (by count) on

describing more words as *discipline specific*. The differences between the allocation of general academic language to discipline specific can be seen in Table 9.

Table 9
Phase Two Participant Allocation of Word List Specificity

Generalization	Participant 1	Participant 2
Discipline Specific	2	4
Somewhat Discipline Specific	11	12
General Academic Language	12	9

Participant One answered that *Augmented* and *Visualization* are discipline specific words, whereas Participant Two responded that *Augmented, Interacting, Multimodal, and Perceptual* are discipline specific words. All of the words that are different between Participant One and Two regarding what is a *discipline specific* word, were listed as *somewhat discipline specific* by Participant One. Of the differences between participants one and two regarding *general academic language*, Participant One responded that *Usability, Visually,* and *Wearable* are "General Academic Language". In all of the cases of Participant One allocating the general category to these words, Participant Two described them as "Somewhat Discipline Specific". These differences are further seen in the free-responses.

In the free-responses, Participant Two responded six times and Participant One responded four times. In the four times that Participant One responded, it was to state that the example contained too much jargon and in one case had difficulty understanding the example altogether because "…my first language is not English, if I don't know this word, just read this example, I can't understand it". In the example on "interacting" both Participant One and Two found the example to be too full of jargon to be understandable. The example given writes:

> "For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research)" (TOCHI – 80).

Specifically, Participant Two wrote that this example did not seem as a typical usage of the word "interacting" in regards to the discipline. These results aided in guiding Phase Three and its following results.

Phase Three

In Phase Three, there were two different participant groups. First, there were the English-major degree seeking students and a second survey population of XR-major seeking students and professionals in the field. Both survey groups were asked a series of questions to determine if they found the word list useful, how it might play into their role in the classroom as either a student or teacher, and if they found the word list helpful as a potential example of a tool that could be used for collaboration with other disciplines. The purpose of seeking this information was to determine if the students themselves in XR found the word list as potentially helpful for their own academic journey. An additional goal was to analyze whether it might be helpful for others with limited knowledge of the field to get a glimpse at the language in the field and whether this access might incentivize cross-disciplinary collaboration. This question of access is set on the author's own belief that cross-disciplinary conversation does not occur as often as it could because of language barriers across fields. Therefore creating conversational power structures where one person in the conversation holds more power than the other participant in terms of language ability and subsequently changes the dynamic of conversation.

English Major Graduate Students

In Part A of Phase Three, the participant linguistic-based demographics vary in discipline of the English Language studies in that not all students are L1 speakers of the English language; however, only one student represented the L2+ category. All of the respondents were masters' level students with three participants in their second year of study and three participants in the first year of graduate studies. Of these students, only one responded "I feel somewhat comfortable" to the question "How comfortable do you feel using computer science specific language?". The student that responded to this question was the L2+ student.

The participants in Part A responded to the words with more "No" responses, indicating that there were more words they did not understand in context and/or could not figure out in the context given. The words that 2 or more participants did not understand for at least two of the questions each for that word are: *Annotations, Dataset, Metrics, Occlusion, Visualization, and Haptic*. As these participants received 24 words to evaluate, the results imply that they could only comprehend about 75% of the words presented. The word with the greatest unanimous difficult was "occlusion" as presented in Table 10 below.

Table 10
Phase Three – Part A: "Occlusion" Response Sample per Participant

Question	1	2	3	4	5	6
Do you understand this example?	N	Y	N	Y	N	Y
Do you understand this word when it is used in the given sentence?	N	N	N	N	N	N
With the example provided, could you use this word in a sentence?	N	N	N	N	N	N
If you encountered this word later, might you be able to understand its meaning?	N	Y	N	Y	N	N

The participant responses are in response to the following example for occlusion from ISMAR: "This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture" (ISMAR-33). Another word that

received not quite as much “no” results was Haptic, but the word itself is discipline-specific to computer science and engineering, which could have been the cause for its 54% “no” answer response rate.

The XR Major Participants

The Part B responses are the inverse in English language proficiency in comparison to the Part A respondents. In Part B there were six participants all of whom are graduate students. Of these students the respondents included six doctoral track students varying from three to five years of study in their respective program. Finally, there were five L2+ participants and one L1 participant. Beyond these demographics, however, the Part B participants were asked two additional questions regarding their discipline specific language acquisition. The first question was: “How did you learn words specific to your discipline of Extended Reality? (Please select all answers that apply)” and the second was “What would you say is the primary source of your discipline-specific language acquisition?”. The participants responded in a very similar manner, with a greater emphasis on how their primary source of their language acquisition was from academic journal articles. The percentage of respondents that answered about their sources of acquiring language are heavily influenced by self-directed content such as watching YouTube videos instead of classroom-based materials, as seen in Table 11 below.

The Table 11 demonstrates that only one student considered classroom-based materials to be their primary source of language acquisition. The demographic categories of this participant is a second-year doctoral student whose first language is not the English language. All other participants responded that their primary source of language acquisition was from reading academic journal articles – which is also the primary source of the examples curated for the word list in this research.

Table 11
Percentage of Part B Participants Who Responded to their Influences and Primary Influence of Language Acquisition

Source	Of Many Influences	Primary Influence
Dictionaries	0	0
An Online Corpus	0	0
YouTube and Online Videos	83	0
From Reading Textbooks	100	0
From Reading Academic Journal Articles	100	83
From Attending Conferences	50	0
From Classroom-based Materials and Instruction	83	17

Of the questions in the word list, the respondents in Part B had two different word list sets. In addition to the words used by Part A participants, Part B participants additionally had words that were taken from the reduced coverage across all five publication venues. The words used in Version A and Version B are listed in Appendix (letter). Of the words that were similar with Part A participants, the participants in Part B responded at least twice per individual to the words: *Annotations, Metrics, Multimodal,* and *Perceptual.* Of these words that the participants responded a "no" value, and therefore a low response to understanding the word and its example, the highest word to receive the "no" response was *Perceptual*. The example used for Perceptual comes from TVCG and reads "In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity" (TVCG – 173). This result demonstrates that there is a particular understanding between the participants and how they are encountering the word. Importantly, the two participants who gave "no" values for *Perceptual* are both L2+ speakers of the English Language, have three years each in their doctoral program, and learned their discipline-specific vocabulary from academic journals.

In the final portions of the surveys for Phase Three the participants were asked a series of questions relating to how they might use the CoXR Word List, their opinion on the hyphenated words, and if they would be incentivized to collaborate across disciplines if they had access to that discipline's language through a word list such as the CoXR Word List. The results of the two surveys were mostly consistent in answers with all of the English-major participants giving approximately the same answers for each question. The participant answers for the CS students were more mixed, but leaned in the same direction as the English-major participants.

The English major participants were overwhelmingly supportive of the word list and were engaged in seeing it as a tool. For question one "After completing this survey, do you see a use for a detailed word list that is discipline specific in your classrooms?" the participants all answered "yes". Participant One responded "yes, absolutely" and participants two through six responded that they found the word list to "encourage interaction". For question two "In your opinion, what are your questions or concerns about using this word list in your classroom?" the participants responded that there were challenges with:

- How might we make the word list applicable to our classrooms?
- Who would be in charge of distribution and development of the word lists?
- How much access would they, as educators, have to the word list?
- Would the word list include definitions if it was used in the classroom?

The responses to question three "If this word list were published with each word and its following example, what would you, based on your perspective and experience, do to improve the word list in terms of its accessibility and user friendliness?" reinforced some of the considerations in question two. Participants two through five responded that the word list was in-

accessible completely in that it lacked audio recordings for each word and its example, there is not a pronunciation chart, there are no images, and there weren't any direct definitions.

In response to question four "If you were to use this word list for personal use to communicate with someone in Extended Reality, how comfortable would you feel? Explain why or why not?" the participants were split. Participant one responded that they were about 25% comfortable using the word list if discussed on a scale of zero (not comfortable) to 100 (comfortable). Participant six said that they would need more exposure and time with the word list before giving a final answer. Participants two through five answered question four by discussing that they, like participant six, would want more time with the word list, have a chance to become better acquainted with the word list, and practice the contents before interaction.

For question five "Did you find the hyphenated words different in difficulty to the single word examples? Why or why not?" the participants all responded "yes" to the question. In question six it is asked "Would the use of a discipline specific word list such as this provide you with incentive to collaborate and/or explore across disciplines if you had access to it?". The participants responded to question six that they would, in two cases, "absolutely" be incentivized to explore across disciplines if they had access to the language of that discipline. Participant six responded that: "I honestly like definitely wholeheartedly want to say yes because I think it would give me some like comfort some peace of mind knowing that I have access to things easily to help improve my understanding without having to be like *what was that, what was that,* over and over again because it does become like exhausting not understanding".

The CS participants had slightly different questions since the nature of their experience was varied and not all had taught previously. To question one "After completing this survey, do you see a use for a detailed word list that is discipline specific in your classrooms as a student?"

all participants except for one L2+ student responded "yes" to the question. The one participant that had an alternative answer wrote that they were "unsure". For question two "In your opinion, what are your questions or concerns about using this word list in your classroom as a student?" the participants had mixed responses. The responses are as follows:

Table 12
CS Participant Responses – Question 2 Phase Three

Participant	Participant Response
A1	Not enough examples, or no one give further explanations about the misunderstanding of the word.
A2	I feel like it might be hard to cover all the words well enough without prior knowledge of some things - so some vocabulary might require other vocabulary
A3	None.
B1	Understanding the need for why certain words qualify as discipline specific when they have the same meaning in all regards.
B2	Limited concerns, the disclaimer of the specific field of use should be stated, some terminology is defined differently depending on the field it is used in.
B3	if a professor says something you didn't understand you would rarely have the time to look it up right away

In response to question three "If this word list were published with each word and its following example, what would you, based on your perspective and experience, do improve the word list in terms of its accessibility and user friendliness?" the participants replied that including visual media would be beneficial, adding examples from new publications, adding audio, and having a definition would be beneficial.

In response to question four "Did you find the hyphenated words different in difficulty to the single word examples? Why or why not?" the participants responded with mixed responses. Two participants, both L2+ said "no", one L2+ said "yes", and the other three participants had ambiguous answers that surmounted to a "somewhat" answer. For example, participant B3 responded with just a hyphen itself and participant B2, a L1 participant wrote "When to use hyphens is always a mystery-to-me". The participants continued their mixed responses in

response to the last question, question five "Would the use of a discipline specific word list such as this provide you with incentive to collaborate and/or explore across disciplines if you had access to it?". The participants that had a "yes" answer were three in number with one writing "Definitely. Since the word list helps me understand the use of the word in my domain, I am able to relate to how it is used across disciplines. E.g morphology". Three participants wrote that they would not use a word list, with the only L1 participant writing "No, but the ability to search for key words from other fields as the relate to a topic would speed up my paper searches when I am working outside of my domain".

The shared responses between both participant groups can only be calculated for questions one, four/five (English/CS), and five/six. The response rates are shown in Table 13 below:

Table 13
Phase Three – Shared Free Responses for CS and English Participants by Percent

Question (English/CS)	Yes	Somewhat	No	N/A
1	92	8	0	0
4/5	58	8	17	17
5/6	67	25	8	0

These responses demonstrate that there is a majority response towards using discipline-specific word lists such as the CoXR Word List in their classrooms, that hyphenated words are trickier to get a grasp on, and that there is a potential for increased encouragement to use word lists for cross-disciplinary collaboration.

Chapter 5: Discussion

Phase One

In Phase One the successful curation of documents developed a sizable corpus that provided a high level of types, 130,249 from conference proceedings and academic journals. Unlike other corpora in the same field (Roesler, 2021; Minshall, 2013), the CoXR Word List did not use textbooks. This was done because based on the experience of the author, most graduate level students and professionals in XR and CS in general do not utilize books to the same degree in which they use conference proceedings and journal articles to acquire their information. This reason is likely because the materials can be quickly published and distributed in a rapidly developing field, such as on the digital archive system by Cornell University, arXiv. To better understand and address the current language within the discipline of XR the only viable route to capture the language was to use journal articles and conference proceedings in the corpus's development. This was a challenge, however, because not every database is built the same and not all have the same access ability.

This corpus was limited in that it attempted to collect more articles from locations such as ACM's User Interface Software and Technology, Intelligent User Interfaces, and Multimodal Interfaces and Machine Learning for Multimodal Interaction sources in addition to the Spatial User Interaction articles. This collection could not occur because ACM categorized each of these conferences by the key words for which they presentation was housed under (see Figure 5) instead of by general placement in the program and therefore allowing ease of access. Due to this, the researcher would have been forced to "cherry pick" the materials to meet the

requirements of the corpus, which would have negatively moved the data towards a non-expert's bias.

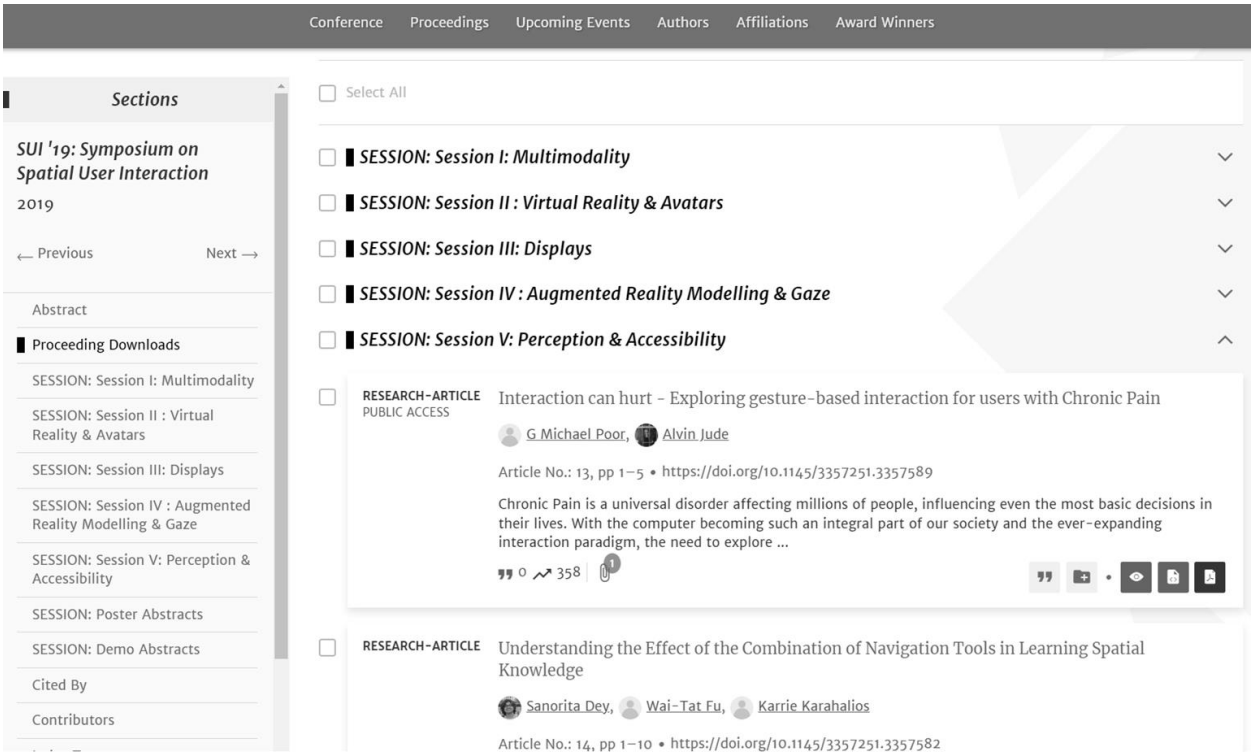


Figure 5 - Screenshot of the ACM Database for the SUI Conference to Demonstrate Sections

In addition to this challenge, curating the corpus from ACM was time consuming since there is a built-in download limit from the database. After a certain amount of downloads the researcher encountered a message that blocked downloads over 200MB in a 15-minute span and would thus be blocked from accessing the site for at least two hours. Challenges such as this might encourage those build corpora in the Computing or Engineering realms to face multiple complications if using ACM as a source website. Despite this, not every location contained the same challenges, and the corpus does include one journal-based source, TOCHI and one conference-based source VRST from ACM as they did not require the same acquisition technique (as seen in Figure 5).

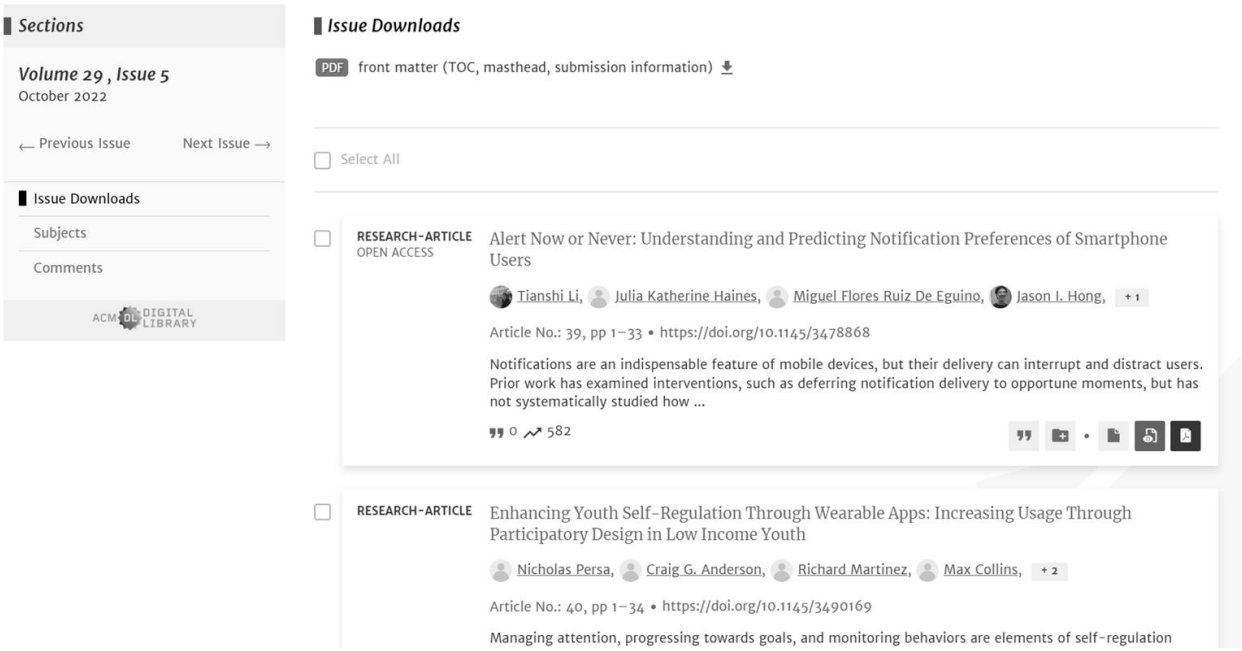


Figure 6 - Screenshot of IEEE Database to demonstrate lack of sections

Due to this lack of picking based on words that stood out (as seen in Figure 6) the corpus itself is believed to be a close representation of the field because of several reasons. Firstly, it evaluates ten years of the discipline to capture not just the language in use today, but partial evidence of the development of the language. This area is not as thoroughly discussed but is a point for future research. Secondly, the corpus contains documents that would be encountered by the participants of this study. Graduate students and professionals in the field generally consider a paper or proceeding to be outdated after two years, and historically important from five to ten years (F. Ortega, Personal Communication, October 2021). Therefore then, the audience should be familiar with the top papers from the past decade in the journals and conferences of their field, such as those presented in this corpus.

One area of improvement for the corpus to truly capture the needs of the audience and population would be the inclusion of company-based publications that are posted on arXiv (Cornell University, 2023) because they represent a large portion of the research in the field.

Companies such as Meta have a large investment in XR and therefore spend a significant amount of money to get the latest literature to the field. Since they do not require the same notoriety that a university might, due to their marketability, they often publish on the discipline-specific hosting site of arXiv for public consumption. This is a place where more documents should be curated, but as in the case of the ACM documents, it would require a level of picking specific documents based on keywords instead of the choosing documents in a central area altogether.

A tool that might help with this process is access to a web crawler system (such as COCA) that pulls documents from across the web based on keywords specifically. This would allow for a similar baseline, but unfortunately is not something that can be produced by a major university and web crawlers are banned on most university internet systems. Therefore, to expand upon this corpus one might need to have the support of a company or private enterprise with a small team of experts who could weigh in on the keywords necessary for accurate representation of the discipline. This then poses interesting quandaries and questions for the future of corpus compilation, because if it requires experts and private capital to capture the authentic language of the discipline it invites a large shift from the ESP academic-centric space into a new realm entirely.

The Word List

The CoXR Word List was curated to represent the authentic language of the field. Authentic language in the context of this research is the language produced by professionals and individuals within a field to communicate in as close to optimal levels of productivity and efficiency with one another so that meaningful communication can occur. To maintain the exploration of only using authentic language, the references were kept in this corpus and thus their words are represented in the word list. Although this is a practice not commonly done, for

the references can be reproduced more often across documents, and therefore in a frequency-based word list dominate spots that could be held by words within the text itself they play a vital role in demonstrating authentic language in the fields of CS. I believe that including references is essential to demonstrating authentic language in CS because the titles of papers within the field can be jargon filled and therefore require a high lexical coverage (Hsu, 2014) to understand.

With this then, if a student and/or professional in the field were to encounter a paper in their discipline, their first line of reading the paper is to understand the title. As many generally do, judge a book by its cover, a student exploring and learning a discipline might encounter a title, have low lexical coverage of the title's contents, and entertain a different perspective of the material within the paper based on the intelligibility of the title. This dissuasion is common for many students (as sourced through conversations by the author) and therefore including the words within allows for representation of the authentic language of the discipline as a whole.

In addition to the issue of first understanding the titles of papers and within references, the words within top papers must also be well known in order to communicate within the discipline. Communication within XR, based on the types given for the word list may be difficult to learn altogether. As seen in Table 5: Types and Tokens for CoXR Corpus Files, despite the consistent number of types per category (depending on the publication venue type), not many words are shared across the publication venues. This is further demonstrated by the 130,249 types across the corpus itself despite the average conference venue having 24,180 types and the journal venues of the corpus having an average of 72,112 types. When these numbers are compared to the final word list results of all five publications having only 24 words in common that are discipline specific raises many questions on what then might be considered pertinent for practitioners in the field to know.

One question that might be raised regarding what words are discipline specific across the field versus a particular venue is that: are words specific to a discipline alone, or populations within a discipline? If yes, what does this signal about the potential growth of a language and those entering the community? Another question that may arise is: how narrow should linguistic research go regarding discipline-specific corpora? Finally, one might consider how, if students gain access to their discipline-specific lexicons through primarily reading conference proceedings and journal articles (see Table 12) then how can word lists be used to scaffold the community language levels within a discipline? This is to say, that if a field uses specific words, yet under this umbrella each publication venue has its own discipline-specific lexicon, how might students access and learn these nuances for not only their understanding of the materials, but as researchers themselves. These questions pose many challenges towards the study of disciplines within a field but raise many opportunities for further research in ESP, especially in a rapidly changing field such as CS.

Hyphens and the Future of Coding for Word List Results

In the world today many people are surrounded by an ever growing number of novel and innovative technologies. With new entities, comes new names and ways to describe them. These changes generally lead to the compounding of words, using hyphens to demonstrate new meanings, and using language to demonstrate the breadth of freedom that the English language (among many others) allows. In CS hyphenated words are frequently used to indicate a state of being, such as *state-of-the-art*, *real-time*, *real-world*, and many more. With these forms comes exclusive language to disciplines within CS, such as XR. Since XR deals with both the real-world and new and novel world forms here within our space (AR) and immersed into another

(VR) or a combination thereof (MR) the language used in the field is no exception to this. As the word list of this research shows, there are 274 hyphenated word forms across the CoXR corpus.

While the results from the CS participants indicated that they understood the words given that were hyphenated, the English major participants indicated that they struggled with the word and example for "dataset" because it was used with and without a hyphen in the example: "We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values". This sentence excerpt in its own proves, for this instance, that the writers did not understand the difference between the hyphenated and non-hyphenated form. This is an important area of research – especially in relation to understanding how L1 and L2+ student confidence interacts with compound understanding and production in fields such as CS where the rate from hyphen to compound might be occurring at a high speed.

Based on the high range of words such as "real-time" which has a range of 406 documents, looking at hyphenated words may be important for understanding discipline-specific language as well as the potential evolution of the language. The hyphenated words may present new entities and/or describe new structures in the field, such as XR where something that is *head-mounted* is an important differentiation when placed with the collocate "optical see-through" in that an *optical see-through head-mounted displays* could be a particularly designed MR structure that utilizes VR head-mounted devices. Whereas if in an AR-based structure "optical see-through" could indicate a lens of any kind in use to be seen through with the eyes by glasses are not often referred to as *head-mounted* in the same way as VR devices.

This is important in CS because hyphenated words, like any compound structure, start to explain a feature and then evolve into a compound itself. CS has a quick turnover rate for publications and therefore the language must, by the law of nature of language itself be evolving

with equal speed. This leaves hesitancy when writing publications and applying for grants because the process of accepting the compound is not an overnight process. In the researcher's own time involved in writing and editing in the field, there was confusion for some time on when to use hyphens on some word forms such as "data-set/dataset, virtual-reality/virtual reality". Future work might examine the progression of hyphen use over time to see if there is a common rate of development between the time a hyphenated word is formed and the time it becomes compounded into a single word. Furthermore, hyphens that are specific to the discipline might have specific collocates that they are used worth. Future research might evaluate the relationship between collocates and single discipline-specific words and hyphenated word forms. The purpose of this research would be to expand upon the word lists that have been constructed to evaluate if there are more ways that word lists need to be improved upon to meet modern needs.

Phase Two

Word Lists: The Trend of Non-Expert Experts in ESP

While Phase Two is only a part of the entire research presented in most cases the construction of a word list is the only requirement for publication in ESP. This novel introduction of evaluating the word list's efficacy was conducted for several reasons. Firstly, in academic work it is required that one's every thought, regardless of your notoriety be backed up by a previous scholar, otherwise it is generally considered (1) invalid and (2) demonstrates the author's awareness of the field or lack thereof. If all academic writing that is considered with a high degree of scientific respect requires validation by a source, then why is it that the research, development, and publication of word lists (especially those who are discipline specific) do not require similar validation?

The reasons for a lack of research could be multifold, in that there is generally little to no funding for studies in language research, accessing experts could prove difficult depending on ones situation due to financial reasons, bias across academic disciplines, and the general trend of insularity in academia which restricts cross-disciplinary interaction. The validation and surveys conducted in this research were made possible by the author's own connection to cross-disciplinary research and the willingness of a professor to lend their students' time. Studies in today's quick moving world often pay up to twenty dollars and hour for participants, require long tedious manners of acquiring and signing up participants, and the lack of public space for more research to be conducted. Due to these limitations the study in Phase Two was limited to questions that took only half an hour to complete, was conducted anonymously online, and was unpaid – all in the interest of the participants time and the lack of academic infrastructure for non-celebrity fields with minimal resources. The results of the study were meaningful to the research questions they answered, but did have some limitations that require further research to be perfected as a potential route of establishing methodology for word list construction.

Human Participation for Word List Efficacy

In Phase Two the participants of the study were asked to evaluate the initial word list. This phase was done because the author of the study is a non-expert in the field of XR and therefore, was unqualified to determine whether a word is or is not discipline specific. Furthermore, the author of this study, as a non-expert, could not be assumed to have chosen the optimal examples for the words in the word list. To fill these gaps left by the non-expert role, graduate students that are currently enrolled in a XR based program surveyed the word list with its corresponding examples. The survey asked the participants demographic information as well

as if the example was appropriate for the given word, why or why not, and to examine if it was discipline specific or a general academic word.

The results of this phase indicated that some of the examples were not favored, but a majority of them were. The responses to this section by one participant in particular were what was most curious. In that one of the participants, a L1 master's track student critiqued the examples themselves in their writing. For example, for the word *interacting* they wrote:

> "I would not generally describe all objects as capable of "interacting." I could see discussion of people MANIPULATING these objects, but to say they INTERACT with them would require that the objects be able to take inputs and produce outputs."

This result then may not directly answer whether the question should be used or not, but it does demonstrate a few needs from the study. Firstly, participants needed further reinforcement in the knowledge that these examples were examples of authentic language. Secondly, the participants may have needed clearer instructions, so that they did not critique how the word is used instead of answering "is this a good example of the word". These would have more directly benefited the results of the survey. The other participant (L2+) did not respond in the same manner as the L1 participant and instead answered the questions, as the instructions asked, on whether they understood the example or not.

A large potential issue with this survey is that the participants know the author of the study and may have been biased towards the researcher for not being a member of their discipline. This rationale could explain the participant who believed that the examples were not correctly worded, but further research should be conducted to examine this issue in more detail. Research that could lend a perspective to this is giving members of the CS community a blind

survey to who the author of the survey is, and their respective backgrounds. By eliminating any potential bias on either side of the survey could lead to better results. Additionally, having more participants in this phase could have elaborated upon whether or not the questions of the survey were accurately represented or not.

Phase Three

Human Participation in two Disciplines: A Glance at Scarcity and Celebrity Fields

In the section on Phase Two it was discussed that there is a lack of resources in academia for cross-disciplinary work, especially for celebrity and non-celebrity fields (such as Artificial Intelligence which has billions of dollars allocated annually for research, whilst the arts such language studies generally receive only a small percentage of that sum). Having celebrity fields has been productive towards enhancements in the fields receiving the sums, however as demographics show, most of the incoming workers within these fields, such as CS, are L2+ speakers of English (Open Doors Report, 2022). As several experts in the field of CS with doctoral degrees or the experience equivalent, all at top major companies across the United States of America have confided in me (and thus must be represented anonymously), that themselves and/or their co-workers who have over hundreds of citations to their name still struggle to manage and grasp the language of their field when presented in the English language.

I would like to argue here that this creates an issue of scarcity – scarcity of resources that aid white collar careers due to the expectation of language osmosis (learning as you go), a scarcity of pedagogy at higher levels of the research field from academia to industry, and a scarcity of general awareness by management entities. When workers/researchers do not have a

good grasp, or lexical coverage, of the language they work and publish in this not only results in a deep inefficiency it is also therefore costly.

This issue could be costing the managing entity financially by having researchers that lack the skills to write and win grants, get papers accepted because the ability to describe them is limited, and in general on the daily basis because workers are inefficient with their communication. While CS is a field that in general allows more freedom for "good writing" (Hynninen & Kuteeva, 2017) ambiguity or un-clear wording can easily lead in this day in age to misinformation. This corpus, and its resulting word list, demonstrate that researchers in the field of XR must have a grasp on at least 130,249 types to have close to full lexical coverage. While most speakers of a language, and those who speak multiple languages have a lexicon into the millions of words, not having access to the way in which the discipline would like the speaker to perform with these 130,249 types could prove costly to not only the company or university with the researcher, but the researcher's success themselves. Due to what has been mentioned, more research needs to be done to understand the level of scarcity as the result of inadequate levels of research and pedagogical application to post-graduate education in industry and academia.

Non-Expert and Expert Evaluation of Word Lists

The purpose of Phase Three was to evaluate if (1) the word list might be useful to those in XR and as a tool beyond XR if created in other fields, (2) if educators in another field found the word list productive towards their understanding of the words themselves, and (3) if access to a word list could increase the desire for cross-disciplinary research and collaboration. The third question was asked with the assumption that collaborative research is impeded by the lack of access to technical language across fields, and therefore creates inefficient research.

The results of Phase Three showed that the participants that are in the field of English language studies had difficulty with six of the given words, whereas of the same words the XR graduate student respondents to the survey were challenged by four of the words used. The English major participants had difficulty with the words: *Annotations, Dataset, Metrics, Occlusion, Visualization, and Haptic*. Whereas the XR student participants had difficulty with the words: *Annotations, Metrics, Multimodal,* and *Perceptual*. These results seem to indicate challenges by both cohorts.

Firstly, the English-major respondent answers indicate that there are some discipline-specific words that require multiple iterations of the same word in different contexts for comprehension because one example is enough. This could be backed up by the XR-student participant response that the majority of them acquired their lexicon in the discipline by reading examples of the words in context in academic literature. However, this is problematic because doctoral students in the field could not use four of the words based on their use in context despite their years of commitment to the field. What this may demonstrate then, is that despite exposure to a word, if you do not have a sound lexical coverage (Hsu, 2014) of the context a word needed for acquisition lies within, there will be multiple challenges towards that acquisition. As a result, a word list might be useful for those in the field to use to explore direct uses of the word list, but it may need to contain several examples from all over the field to aid those with a low lexical coverage so that they can find the "in-context" description of the word around other words they are familiar with.

A limitation of this research is that there were not multiple examples paired alongside the words, but this was intentionally done. With one example the participants were forced to see if they understood the word or not based on a singular scenario. This was intentional because in

conversation and in classrooms, new words are not often discussed nor defined, they are thrown in with everything else. It is often that we encounter a word in context once, then build up more scenarios for acquisition. Additionally, when there are jargon filled sentences, students wanting to understand a word in context will have to ignore the noise surrounding the word itself. This is a skill that requires practice and is not re-enforced.

Future research might examine this further, to see how students evaluate in-context words when there are limited examples and how they might produce them. Additionally, in CS education, examining the vocabulary and the intentions behind the vocabulary being taught could be expanded upon. Finally, for cross-disciplinary research it might be worthwhile to examine at what point of lexical coverage someone might feel comfortable having to start discussing and exploring a new field with the intention of trial by error, rather than a smooth level of acquisition. This would be important because as new fields and ideas are generated every day there are also new ways of understanding them and therefore we are all learning to adapt in novel ways.

Hyphenated Word Responses

The students in XR that participated in the study the majority of them agreed to some level that the hyphenated word forms were trickier than the stand alone words. There were two participants that had no trouble with them with one writing “they convey information as a whole and can be considered a word in itself”. This, however, does show that there is opportunity for further research into understanding how hyphenated word forms and other groupings of words such as collocates and acronyms are used in the discipline of CS and potentially other sciences. For L2+ students, hyphenated words might cause further challenges, especially if one’s lexical coverage is low. One XR L2+ participant wrote that hyphenated words “… increase the chances

of different interpretation". This demonstrates that they do build some form of confusion because each word in the hyphenated word form has an individual meaning that lends itself to the whole meaning but does not always necessarily describe the entire meaning itself when put together.

Examples of more direct hyphens in the field might be "head-mounted" which literally means mounted to the head via some device such as a headset or helmet. Whereas "state-of-the-art" is more challenging because it requires the reader to understand that a state has multiple meanings, and that art here may be defined in reference to the Latin *magnum opus*, meaning the best of one's work (generally art like paintings and sculptures). Combined, state-of-the-art means the best there is and requires much context into not just the English language, but the cultures surrounding it.

With this, there is also potential for future research into the level of discipline-specific words and hyphenated word forms that require cultural understanding to be understood by the user. An example of this being necessary for single words is the word "artifact" which has multiple meanings across disciplines but is generally understood as either something to be held, or something that is a relic of some kind. In the case of XR it is used as a tangible entity, since the field works in the intangible. Yet, one doctoral track L2+ student struggled with the word artifact. This may be due to their low exposure to the word itself, the example it was presented in, or potentially (and this is where more research could be done) the challenge of understanding how the word is culturally used in different places with multiple meanings.

Pedagogical Applications of a Word List: Are they viable tools?

The English major participants, all who are currently teachers or have taught in the past professionally, were asked if they would use a discipline-specific word list in their classrooms.

They all responded, in general, that they might consider it if there was a specific reason or application for its use. With this, the XR participants responded, in general, that they might consider using a word list as a resource in their classes. With these results then they lean towards providing positive insight that word lists have a pedagogical application for both instructors and students. While word lists have been evaluated as potential tools for use in ESP to aid L2+ students, reinforcement of their potential applications is necessary, especially as technological resources continue to grow and diversify. One particular finding of the English-participant study responses was that a participant said they believed that the word list might "encourage interaction". This is curious because it opens up questions on how some students may or may not be engaging with courses, materials, and other students across disciplines due to issues with access to language.

With this then, future research might consider evaluating discipline-specific word lists as interaction tools between students of different disciplines, and potentially language status. If undergraduate students specifically could gain access to the language of other fields early on, might they be encouraged to explore across disciplines instead of locking themselves into one by their senior year? How might changing how we treat intradisciplinary work change how academia is done? Would removing barriers at the lowest levels in terms of language access re-shape how students view academic field boundaries? There are many questions that could follow this, but it can only be answered if students and beyond, find word lists useful for more than just education.

Are Word Lists Useful?

To the question "are word lists useful" it would depend on who you ask and what discipline you are in. The English-major participants, a group that is in a field where the research

requires the study or analysis of another discipline for existence (e.g. rhetoricians study Communications or Political Science, Non-Fiction researchers studying Natural Resources, or Linguists explore other disciplines such as Engineering). Areas for potential collaboration based on this, would be using the communicative abilities of the humanities to elevate how the sciences are communicated. In industry there are marketing and communications teams, economists, writers, and editors supporting scientific work; however, this is not done in the same way in academia. Arguably this has been detrimental to how the sciences in academia are perceived by the outside world (e.g. the misinformation spread surrounding COVID-19 vaccinations). Since English majors generally have a specific interest for analysis, I would like to argue that more often than not they take the initiative to reach out to another discipline for collaboration (this would be an interesting research topic). Whereas, in the author's experience, those in CS focus on their discipline and are generally solicited from outside their own field for collaboration. Therefore, based on the results, where all of the English-major participants responded that "Yes" they would be interested in using a word list for collaboration with another field in order to understand it better, the CS participants when not responding "yes" said they would only do so for publication-based research. This indicates several potential pieces for consideration.

Firstly, is this word list just about the word list, or is it asking bigger questions such as above? Do CS-based researchers lack accessible resources to language and are therefore less inclined to use something as a resource when they are used to trial by error? Or, are their challenges with how English majors might view themselves in comparison to other disciplines? How might the super star status of a field change the mentality surrounding access to collaboration and cross-disciplinary collaboration? How might the denigration of humanities fields in turn build inferiority complexes around the next generations, graduate students? How

also might the superstar status of a field build a superiority complex within its students so that they do not seek resources but instead expel energy seeking answers on their own? Secondly, if language is given for collaboration purposes will it be used based on one's exposure to utilizing resources or will it rely on personality to be utilized? Lastly, if disciplinary word lists are made accessible for collaboration purposes will it start to dismantle the mystique or strong barriers of entry to some fields for cross-academic scholarship and collaboration?

Chapter 6: Conclusion

While corpus linguistics and the analysis of word lists has been a research area for linguists for over fifty years there is still little research on how small of a field should be studied and evaluated for authentic language use, growth, and for application to broader scopes. In the field of CS, two corpora (Minshall, 2013; Roesler, 2021) evaluated CS as a whole, whereas this research narrowed the field to XR. The purpose of narrowing the field was to evaluate if the standard methodology for corpus linguistics on larger scales such as the GSL or AWL could be applied to small-scale corpora. While the results demonstrated some discipline specific words, in comparison the bigger lists a much smaller selection of final types that spread across all five publication venues were generated despite having a significantly large type value than common word lists at 130,249 types. In addition to these results, the study answered the call in Roesler (2021) for the generation and evaluation of a hyphen-based corpus analysis within CS.

Based on the word list of hyphens, research question one "Is a word list based on a corpus in Extended Reality possible to create? To what extent do frequency, keyness, and range contribute to the creation of a word list in Extended Reality? With a word list in Extended Reality what role do hyphenated word forms play in the accuracy and representation of the field?" was answerable. The results for question one indicated that a word list was possible to create in the field of Extended Reality. Furthermore, the results show that a word list based on frequency as the main indicator of a word being discipline-specific is not enough. As a result, more research needs to be conducted to discover the best parameters for smaller corpora that focus on disciplines rather than fields. Additionally, the results show that for those unfamiliar with the field, hyphenated words were tricky to grasp, but for those in the field it was a split

chance that it might interrupt their comprehension and/or reading process. As a result, it may be worth consideration that hyphenated words can create challenges for understanding discipline-specific language and due to their frequent use within the discipline they are worth further evaluation. Further research might consider the progression of compounding in the field, how L2+ or novice L1 learners interact and understand hyphenated word forms as they join a new discipline within CS, and how they can be pedagogically reinforced by educators so that students are aware of when and how to use them.

The second research question asked: "Does soliciting feedback from graduate students and/or professionals in the field on a discipline-specific word list help shape or reshape the of the efficacy and utility of the word list?" Based on the results gained in Phase Two soliciting feedback can prove helpful in the examination of the word list. Yet, with this, there needs to be considerable effort in having a relatively larger population of respondents so that more feedback can be provided. In terms of the efficacy of the wordlist, the results demonstrate that the feedback given in Phase Two can help shape which authentic language examples are used in the word list itself. Future research might consider creating a sturdy methodology for the incorporation of this technique as a tool for determining the efficacy and utility of a word list.

The student populations that evaluated this research had some varying perspectives on the word list but in general favored it as a tool for their education. In research question three, which asks: "When computer science and non-computer science graduate students are presented with a word list in Extended Reality, what are their perceptions of that word list as an educational tool?" is answered as a "yes" it can be used as an educational tool. While the English-majoring participants evaluated it as a pedagogical tool and the CS-majoring students as a student-based tool, both groups by at least a majority of the participants agreed they would use it in the

classroom. This indicates that word lists are a potential pedagogical tool and more research should be conducted across disciplines to evaluate this further.

The final research question asked: "Do graduate students see a use for discipline-specific word lists as a potential tool for incentivizing cross-disciplinary conversations and/or collaboration?" This question resulted in a majority response from participants that "yes" they would be more likely to interact with another discipline and/or collaborate if they had access to the language needed to effectively communicate and understand the discipline. The participants who did not agree to this were all CS students but one of the participants responded that they would use a word list, such as the CoXR Word List, as a way to look for keywords in other fields to more effectively conduct research and to write papers in or about that field. With this information, there is a significant amount of room for future research on using word lists as a cross-disciplinary collaboration tool.

With the sum of these findings, this study demonstrates that word lists are useful tools for pedagogical application whether that be in the classroom or teaching oneself more about a discipline specific language. Furthermore, the research opens the doors toward exploring what level of information must be present for a word list to be effective, such as hyphenated word forms (in disciplines where it is necessary or useful), images, audio, more examples, and/or definitions. With these elements and the access to the word list there is much to research further on how word lists can impact and improve the access, incentivization, and success of cross-disciplinary collaboration. These findings, in summary, ask the question what new ground can we break to build word lists that have stronger and novel methods for curation and verification so that there may be more widespread utilization of this language acquisition and research tool.

REFERENCES

Ābeltiņa, A., & Rizhamadze, K. (2021). The Complexity of Relationships within Business and the Importance of Common Ground. *SHS Web Conferences, 90*, 1-6. https://doi.org/10.1051/shsconf/20219001001

Anthony, L. (2022). AntConc (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Anthony, L. (2022). AntFileConverter (Version 2.0.2) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Baba, P. A. (2014). Technology in the classroom: A tool or drag. *Journal of Education and Practice, 5*(27), 75-78.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371-405.

Brigham Young University. (2023). Corpus of Contemporary American English. https://www.english-corpora.org/coca/

Cheryan, S., Ziegler., Montoya, A.K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than other? *American Psychological Association Psychological Bulletin, 143*(1), 1-35. https://psycnet.apa.org/doi/10.1037/bul0000052

Cornell University. (2023). *arXiv*. https://arxiv.org/

Coxhead, A. (2000). The New Academic Word List. *TESOL Quarterly, 34*(2), 213-238. https://www.wgtn.ac.nz/lals/resources/academicwordlist

Csata, Z., & Marácz, L. (2021). Social inclusion and multilingualism: Linguistic justice and language policy. *Social Inclusion, 9*(1), 1-4. DOI: 10.17645/si.v9i1.3941

Đurović, Z. (2021). Corpus Lingusitics Methods for Building ESP Word Lists, Glossaries and Dictionaries on the Example of a Marine Engineering Word List. *Lexikos, 31*(1), 259-282. https://doi.org/10.5788/31-1-1647

Firmin, M. W., & Genesi, D. J. (2013). History and Implementation of Classroom Technology. *Procedia Social and Behavioral Sciences, 93*, 1603-1617.

Francis, W. N., & Kučera. H. (1964). Brown Corpus Manual. http://icame.uib.no/brown/bcm.html

Groves, M. & Mundt, K. (2015). Friend or Foe? Google Translate in language for academic purposes. *English for Specific Purposes, 37*, 112-121. http://dx.doi.org/10.1016/j.esp.2014.09.001

Heatley, A., Nation, P., & Coxhead, A. (2002). Range computer program. Retrieved from: http://www.victoria.ac.nz/lals/about/staff/paul-nation

Herbst, T., Faulhaber, S., & Uhrig, P. (2004). *The Phraseological View of Language*. De Gruyter Mouton. https://doi-org.ezproxy2.library.colostate.edu/10.1515/9783110257014

Hirst, G. (2013). Computational Linguistics. (K. Allan, Ed.) Oxford University. https://doi.org/10.1093/oxfordhb/9780199585847.013.0033

Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes, 33*, 54–65. https://doi.org/10.1016/j.esp.2013.07.001

Huang, L. (2010). Seeing eye to eye? The academic writing needs of graduate and undergraduate students from students' and instructors' perspectives. *Language Teaching Research, 14*(4), 517-539. DOI: 10.1177/1362168810375372

Hynninen, N., & Kuteeva, M. (2017). "Good" and "acceptable" English in L2 research writing: Ideals and realities in history and computer science. *Journal of English for Academic Purposes, 30*, 53-65. https://doi.org/10.1016/j.jeap.2017.10.009

Jalilifar, A., & Memari, M. (2017). Exploring nominalization in physics and applied linguistics textbooks with different levels of difficulty: Implications for English for specific purposes. *Teaching English Language, 11*(2), 131-159.

Kallio-Tamminen, T. (2020). Dynamic Universe – natural science and philosophy in unison. *Journal of Physics: Conference Studies*, 1-13. doi:10.1088/1742-6596/1466/1/012005

Lee, J., Han., B & Choi, S. (2016). Motion effects synthesis for 4D films. *IEEE Transactions on Visualization and Computer Graphics, 22* (10), 2300-2314.

Lin, L. H. F., & Morrison, B. (2021). Challenges in academic writing: Perspectives in Engineering faculty and L2 postgraduate research students. *English for Specific Purposes, 63*, 59-70. https://doi.org/10.1016/j.esp.2021.03.004

Mahini, F., Forushan, Z., & Haghani, F. (2012). The importance of teacher's role in technology-based education. *Procedia Social and Behavioral Sciences, 46*, 1614-1618.

Martínez, I.A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes, 28*(3), 183-198. https://doi.org/10.1016/j.esp.2009.04.003\

McEnery, T., & Hardie, A. (2013). The History of Corpus Linguistics. (K. Allan, Ed.) Oxford University. https://doi.org/10.1093/oxfordhb/9780199585847.013.0033

Meccawy, M. (2022). Creating an immersive XR learning experience: A roadmap for educators. *Electronics, 11*, 1-18. https://doi.org/10.3390/electronics11213547

Minshall, D. E. (2013). A computer science word list (Master's thesis). Retrieved from https://www.baleap.org/wp-content/uploads/2016/03/Daniel-Minshall.pdf.

MIT. https://www.mit.edu/~ecprice/wordlist.10000

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes, 25*, 235-256. doi:10.1016/j.esp.2005.05.002

National Center for Education Statistics. (2022). Graduate Degree Fields. *Annual Reports and Information Staff*. https://nces.ed.gov/programs/coe/indicator/cta/undergrad-degree-fields

Nation, P. (2006). How large a vocabulary is needed for reading and listening? The Canadian Modern Language Review, 63(1), 59–81

Nekrasova-Beker, T., Becker, A., & Sharpe, A. (2019). Identifying and teaching target vocabulary in an ESP course. *TESOL Journal*, 1-27. https://doi.org/10.1002/tesj.365

OpenAI. (2023). *ChatGPT*. https://openai.com/blog/chatgpt/

Open Doors Report. (2022). *International Students*. https://educationusa.state.gov/us-higher-education-professionals/recruitment-resources/open-doors-report

Paltridge, B. and Starfield , S. (2007). Thesis and dissertation writing in a second language: a handbook for supervisors. Routledge.

Qualtrics. (2023). *Qualtrics XM*. https://www.qualtrics.com/

Roesler, D. (2021). When a bug is not a bug: An introduction to the computer science academic vocabulary list. *Journal of English for Academic Purposes, 54*, 1-11. https://doi.org/10.1016/j.jeap.2021.101044

Rolim, C., Schmalstieg, D., Kalkofen, D., & Teichrieb, V. (2015). Design Guidelines for Generating Augmented Reality Instructions. *IEEE International Symposium*. DOI: 10.1109/ISMAR.2015.36

Römer, U. (2021). Applied corpus linguistics for language acquisition, pedagogy, and beyond. *Language Teaching, 55*, 233-244. doi:10.1017/S0261444821000392

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., & Dillenbourg, P. (2016). Using mobile eye-trackers to unpack the perceptual benefits of a tangible user interface for collaborative learning. *ACM Transactions on Computer-Human Interaction, 23*(6), 1-23.

Spagnolo, S., Gonella, F., Viglia, S., & Ulgiati, S. (2018). Venice artistic glass: Linking art, chemistry and environment – a comprehensive energy analysis. *Journal of Cleaner Production, 171*, 1638-1649. https://doi.org/10.1016/j.jclepro.2017.10.074

Swales, J. M. (2014). Variation in Citational Practice in a Corpus of Student Biology Papers. *Written Communication, 31*(1), 3-141. https://doi-org.ezproxy2.library.colostate.edu/10.1177/0741088313515166

Vrasidas, C. & Glass, G. V. (2005). *Achieving technology integration in classroom teaching* (C. Vrasidas & G.V. Glass, Eds.). Information Age Publishing.

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes, 27*, 442-458. doi:10.1016/j.esp.2008.05.003

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes, 28*, 170-182. doi:10.1016/j.esp.2009.04.001

West, M. (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Longman.

Yang, M. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27-38. http://dx.doi.org/10.1016/j.esp.2014.05.003

Yates, J., & Plagnol, A. (2022). Female computer science students: A qualitative exploration of women's experiences studying computer science at university in the UK. *Education and Information Technologies, 27*, 3079-3105.

Zoom. (2023). *Zoom*. https://zoom.us/

Appendix A

Phase Two Email to Participants

Dear Participant,

My name is Lauren Mangus, and I am a researcher from Colorado State University in the English department. We are conducting a research study on a corpus of Extended Reality in the field of Computer Science that contains novel content such as hyphenated words, to evaluate the efficacy of a corpus before pedagogical (education) research, and then to briefly implement a small sample of pedagogical research to both the target and non-target audience to determine if corpora can not just be used to aid those in the field, but those potentially interested in collaboration between fields. The title of our project is The Participant Analysis Tool for the Corpus of Extended Reality. The Principal Investigator is Dr. Tatiana Nekrasova-Becker in the Department of English and the co-Investigator is Lauren Mangus.

We would like you to take an anonymous online survey. Participation will take approximately 60 minutes. Your participation in this research is voluntary. If you decide to participate in the study, you may withdraw your consent and stop participation at any time without penalty.

We will not collect your name or personal identifiers. When we report and share the data to others, we will combine the data from all participants. While there are no direct benefits to you, we hope to learn if discipline-specific corpora are helpful tools to influence greater collaboration between disciplines when access to language is made accessible to both parties. To do this, the investigators hope to learn that discipline-specific corpora increase user linguistic confidence when discipline-specific words are presented to the user along with their definition and an example of usage.

There are no risks to participating in this study. While it is not possible to identify all potential risks in research procedures, the researcher(s) have taken reasonable safeguards to minimize any known and potential (but unknown) risks.

To indicate your consent to participate in this research and to continue on to the survey, please click here: <insert link> (to be determined at a future date).

If you have any questions about the research, please contact Lauren Mangus at lamangus@colostate.edu or Dr. Tatiana Nekrasova-Becker at t.nekrasova_beker@colostate.edu. If you have any questions about your rights as a volunteer in this research, contact the CSU IRB at: RICRO_IRB@mail.colostate.edu; 970-491-1553.

PI: Dr. Tatiana Nekrasova-Becker
Associate Professor

Co-Investigator: Lauren Mangus
Graduate Student

Appendix B

Phase Three Email of Time Inquiry to Participants

Dear Participant,

My name is Lauren Mangus, and I am a researcher from Colorado State University in the English department. We are conducting a research study on a corpus of Extended Reality in the field of Computer Science that contains novel content such as hyphenated words, to evaluate the efficacy of a corpus before pedagogical (education) research, and then to briefly implement a small sample of pedagogical research to both the target and non-target audience to determine if corpora can not just be used to aid those in the field, but those potentially interested in collaboration between fields. The title of our project is The Participant Analysis Tool for the Corpus of Extended Reality. The Principal Investigator is Dr. Tatiana Nekrasova-Becker in the Department of English and the co-Investigator is Lauren Mangus.

We would like you to take an anonymous online survey that acts as an RSVP service for participation in a later study. The later study would require meeting in-person at a designated time with a group of others to participate in an anonymous survey and to discuss participant reactions to the survey. The survey will be audio-recorded. Participation will take approximately 60 minutes. Your participation in this research is voluntary. If you decide to participate in the study, you may withdraw your consent and stop participation at any time without penalty.

We will not collect your name or personal identifiers. When we report and share the data to others, we will combine the data from all participants. While there are no direct benefits to you, we hope to learn if discipline-specific corpora are helpful tools to influence greater collaboration between disciplines when access to language is made accessible to both parties. To do this, the investigators hope to learn that discipline-specific corpora increase user linguistic confidence when discipline-specific words are presented to the user along with their definition and an example of usage.

There are no risks to participating in this study. While it is not possible to identify all potential risks in research procedures, the researchers have taken reasonable safeguards to minimize any known and potential (but unknown) risks.

To indicate your consent to participate in this research and to continue on to the survey, please click here: <insert link> (an anonymous Google Form survey for responding yes/no to meeting in-person at a later set date and time).

If you have any questions about the research, please contact Lauren Mangus at lamangus@colostate.edu or Dr. Tatiana Nekrasova-Becker at t.nekrasova_beker@colostate.edu. If you have any questions about your rights as a volunteer in this research, contact the CSU IRB at: RICRO_IRB@mail.colostate.edu; 970-491-1553.

PI: Dr. Tatiana Nekrasova-Becker
Associate Professor

Co-Investigator: Lauren Mangus
Graduate Student

Appendix C

Phase Three Consent to be Recorded – Form for Participants

Colorado State University

Consent to Participate in Research

Participant Analysis Tool for the Corpus of Extended Reality (*Students)*

Introduction and Purpose

My name is Lauren Mangus (co-investigator). I am a graduate student at Colorado State University working with my faculty advisor, Dr. Tatiana Nekrasova-Becker (PI) in the School/Department of English. I would like to invite you to take part in my research study, which looks at a corpus of Extended Reality in the field of Computer Science that contains novel content such as hyphenated words, to evaluate the efficacy of a corpus before pedagogical (education) research, and then to briefly implement a small sample of pedagogical research to both the target and non-target audience to determine if corpora can not just be used to aid those in the field, but those potentially interested in collaboration between fields.

Procedures

If you agree to participate in my research, I will conduct an interview with you at a pre-determined time and location. The interview will involve questions about computer science specific words, a tool to learn them, and how that tool might be used in the classroom and for collaboration across academic disciplines. It should last about *75 minutes*. With your permission, I will audiotape and take notes during the interview. The recording is to accurately record the information you provide, and will be used for transcription purposes only. If you choose not to be audiotaped, I will take notes instead. If you agree to being audiotaped but feel uncomfortable or change your mind for any reason during the interview, I can turn off the recorder at your request. Or if you don't wish to continue, you can stop the interview at any time.

Benefits

There is no direct benefit to you from taking part in this study other than potentially learning a few new words. We hope to learn if discipline-specific corpora are helpful tools to influence

greater collaboration between disciplines when access to language is made accessible to both parties

Risks/Discomforts

The potential risk of this study is that you will be answering questions related to pedagogical applications that are based on your opinion in front of other participants within your own field. While the nature of the material is not sensitive and does not have any direct harm to the participants, you are free to decline to answer any questions you don't wish to, or to stop the interview at any time.

As with all research, there is a chance that confidentiality could be compromised; however, we are taking precautions to minimize this risk.

Confidentiality

Your study data will be handled as confidentially as possible. If results of this study are published or presented, individual names and other personally identifiable information will not be used.

To minimize the risks to confidentiality, we will maintain the materials on a password protected hard drive and laptop that is only accessible to the co-Investigator. All information from the study will only be in the hands of the co-investigator and remain anonymous throughout the study. Furthermore, any identifying information will be destroyed permanently by March of 2025, three years after the defense date of the co-Investigator.

The PI and co-investigator will transcribe the audio recordings as soon as possible after the interview, and then destroy the tapes. When the research is completed, the co-investigator will save the transcriptions and other study data for possible use in future research done by the co-investigator or others. The co-investigator will retain these records for up to 3 years after the study is over. The same measures described above will be taken to protect confidentiality of this study data. The PI and co-investigator may be asked to share the research files with the sponsor or the CSU Institutional Review Board ethics committee for auditing purposes.

Rights

Participation in research is completely voluntary. You are free to decline to take part in the project. You can decline to answer any questions and are free to stop taking part in the project at any time. Whether or not you choose to participate in the research and whether or not you choose to answer any questions or continue participating in the project, there will be no penalty to you or loss of benefits to which you are otherwise entitled.

Questions

If you have any questions about this research, please feel free to contact the co-investigator at 808-457-0641 or lamangus@colostate.edu.

If you have any questions about your rights or treatment as a research participant in this study, please contact the Colorado State University Institutional Review Board (IRB) at: 970-491-1381, or e-mail RICRO_IRB@mail.colostate.edu .

****

CONSENT

Do you consent for your interview to be audiotaped?

___Yes

___No

If you wish to participate in this study, please sign and date below. You will be given a copy of this consent form to keep for your own records.

_____

Participant's Name *(please print)*

_____ _____

Participant's Signature Date

Appendix D

Email Template for Anonymous Survey Phase Three

Dear Participant,

My name is Lauren Mangus, and I am a researcher from Colorado State University in the English department. We are conducting a research study on a corpus of Extended Reality in the field of Computer Science that contains novel content such as hyphenated words, to evaluate the efficacy of a corpus before pedagogical (education) research, and then to briefly implement a small sample of pedagogical research to both the target and non-target audience to determine if corpora can not just be used to aid those in the field, but those potentially interested in collaboration between fields. The title of our project is The Participant Analysis Tool for the Corpus of Extended Reality. The Principal Investigator is Dr. Tatiana Nekrasova-Becker in the Department of English and the co-Investigator is Lauren Mangus.

We would like you to take an anonymous online survey. Participation will take approximately 60 minutes. Your participation in this research is voluntary. If you decide to participate in the study, you may withdraw your consent and stop participation at any time without penalty.

We will not collect your name or personal identifiers. When we report and share the data to others, we will combine the data from all participants. While there are no direct benefits to you, we hope to learn if discipline-specific corpora are helpful tools to influence greater collaboration between disciplines when access to language is made accessible to both parties. To do this, the investigators hope to learn that discipline-specific corpora increase user linguistic confidence when discipline-specific words are presented to the user along with their definition and an example of usage.

There are no risks to participating in this study. While it is not possible to identify all potential risks in research procedures, =the researcher(s) have taken reasonable safeguards to minimize any known and potential (but unknown) risks.

To indicate your consent to participate in this research and to continue on to the survey, please click here: <insert link> (to be determined at a future date).

If you have any questions about the research, please contact Lauren Mangus at lamangus@colostate.edu or Dr. Tatiana Nekrasova-Becker at t.nekrasova_beker@colostate.edu. If you have any questions about your rights as a volunteer in this research, contact the CSU IRB at: RICRO_IRB@mail.colostate.edu; 970-491-1553.

PI: Dr. Tatiana Nekrasova-Becker
Associate Professor

Co-Investigator: Lauren Mangus
Graduate Student

Appendix E

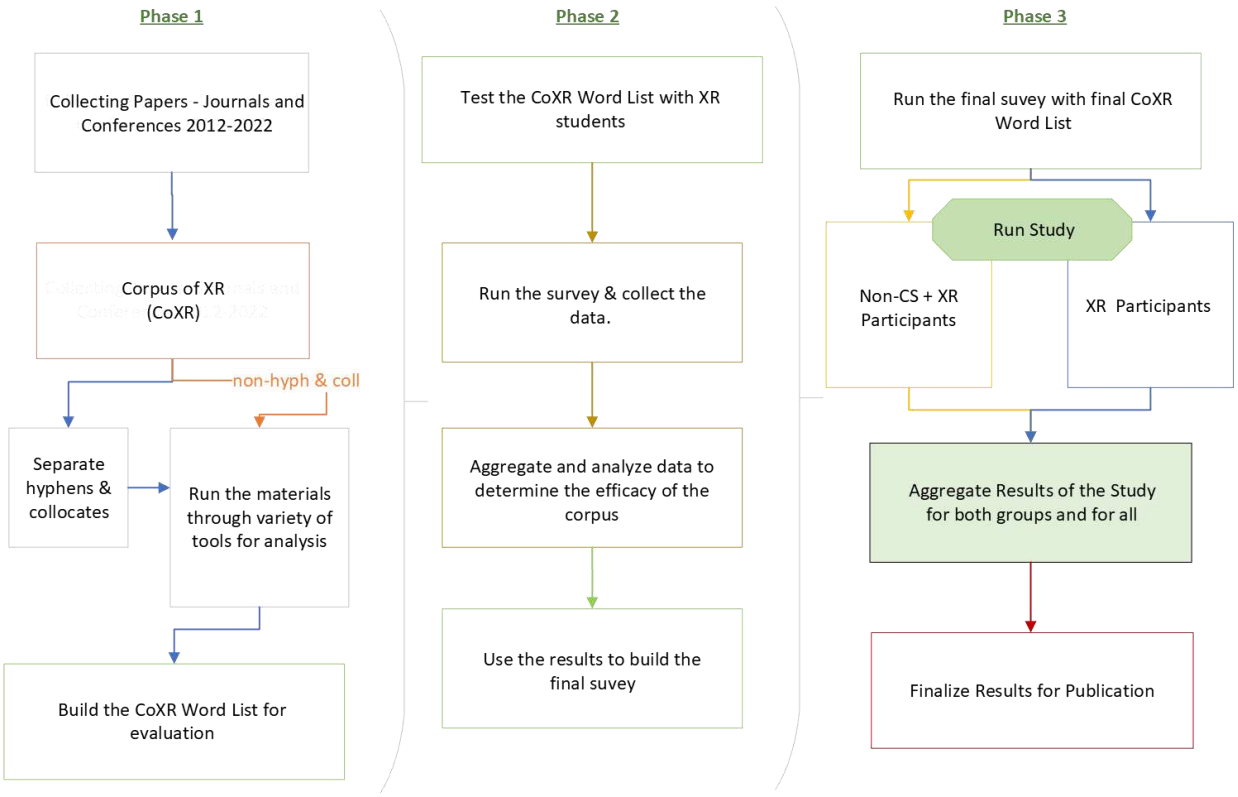
Phase Three Interview Questions

In the audio recording session, participants will be asked the following questions:

1) After completing the corpus-based survey, do you see a use for this tool in your classrooms?
2) In your opinion, what are you questions or concerns about using this tool in your classroom?
3) Is there any further insight you would like to offer regarding this tool from a pedagogical perspective?
4) If you were to use this tool for personal use to communicate with someone in Extended Reality, how comfortable would you feel? Explain why or why not?
5) Would the use of a sub-disciplinary corpus such as this provide you with incentive to collaborate across disciplines? Why or why not?

Appendix F

The Phases of this Study



Appendix G

Hyphenated Words by Frequency

Top 100 by Frequency (normalized) Hyphenated Words

Hyphenated Word	Frequency	Article for Range Example
computer-human	148088.24	109 - Personal Mobile Messaging in Context Chat Augmentations.txt
real-time	57176.47	194- Embedding Spatio-Temporal Information into Maps by Route-Zooming.txt
real-world	33352.94	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
human-computer	23705.88	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
see-through	22676.47	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
head-mounted	21382.35	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
large-scale	18794.12	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
long-term	17647.06	73- The Self-Avatar Follower Effect in Virtual Reality.txt
state-of-the-art	17323.53	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
high-dimensional	13411.76	387- Impact of Cognitive Biases on Progressive Visualization.txt
post-hoc	12882.35	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
high-level	12676.47	154- An â€˜In the Wildâ€™ Experiment on Presence and Embodiment using Consumer Virtual Reality Equipment.txt
image-based	10735.29	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
data-driven	10705.88	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
full-body	10558.82	65- Mixed Reality Office System Based on Maslowâ€™s Hierarchy of Needs Towards the Long-Term Immersion in Virtual Environments.txt
mid-air	10029.41	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
high-quality	9882.35	84- RNIN-VIO Robust Neural Inertial Navigation Aided Visual-Inertial Odometry in Challenging Scenes.txt
rgb-d	9382.35	139- Very High Frame Rate Volumetric Integration of Depth Images on Mobile Devices.txt
high-resolution	9352.94	12- MonoFusion Real-time 3D reconstruction of small scenes with a single web camera.txt
e-mail	9294.12	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
bas-relief	8970.59	116- Bas-Relief Generation and Shape Editing through Gradient-Based Mesh Deformation.txt
spatio-temporal	8852.94	73- The Self-Avatar Follower Effect in Virtual Reality.txt
non-linear	8382.35	26- Computer-Assisted Laparoscopic myomectomy by augmenting the uterus with pre-operative MRI data.txt

three-dimensional	7823.53	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
low-level	7500.00	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
human-centered	7411.76	73- The Self-Avatar Follower Effect in Virtual Reality.txt
multi-user	7117.65	109 - Personal Mobile Messaging in Context Chat Augmentations.txt
open-ended	6882.35	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
time-varying	6852.94	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
face-to-face	6647.06	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
near-eye	6558.82	196- Towards Pervasive Augmented Reality Context-Awareness in Augmented Reality.txt
context-aware	6441.18	47 - Sharing Domestic Life through Long-Term Video Connections.txt
co-located	6352.94	47 - Sharing Domestic Life through Long-Term Video Connections.txt
in-depth	6264.71	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
t-sne	6000.00	344- ProReveal Progressive Visual Analytics With Safeguards.txt
non-rigid	5823.53	40- Rich Intrinsic Image Decomposition of Outdoor Scenes from Multiple Views.txt
self-care	5764.71	143 - Disclosure
node-link	5735.29	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
web-based	5617.65	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
eye-tracking	5558.82	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
decision-making	5558.82	387- Impact of Cognitive Biases on Progressive Visualization.txt
hand-held	5529.41	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
follow-up	5500.00	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
well-known	5441.18	109 - Personal Mobile Messaging in Context Chat Augmentations.txt
end-user	5382.35	118 - It Is About What They Could Do with the Data.txt
two-way	5176.47	47 - Sharing Domestic Life through Long-Term Video Connections.txt
short-term	5117.65	45 - Ad Hoc Participation in Situation Assessment Supporting Mobile.txt
within-subjects	4911.76	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
self-avatar	4852.94	73- The Self-Avatar Follower Effect in Virtual Reality.txt
6-dof	4764.71	84- RNIN-VIO Robust Neural Inertial Navigation Aided Visual-Inertial Odometry in Challenging Scenes.txt
multi-view	4764.71	84 – Don’t Bother Me How to Handle Content-Irrelevant Objects in.txt

self-reported	4735.29	118- Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety.txt
model-based	4647.06	196- Towards Pervasive Augmented Reality Context-Awareness in Augmented Reality.txt
trade-off	4558.82	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
multi-scale	4441.18	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
in-situ	4323.53	196- Towards Pervasive Augmented Reality Context-Awareness in Augmented Reality.txt
patch-based	4205.88	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
t-test	4117.65	65- Mixed Reality Office System Based on Maslow's Hierarchy of Needs Towards the Long-Term Immersion in Virtual Environments.txt
3-d	4088.24	70- Distance Judgments to On- and Off-Ground Objects in Augmented Reality.txt
end-to-end	4058.82	84- RNIN-VIO Robust Neural Inertial Navigation Aided Visual-Inertial Odometry in Challenging Scenes.txt
co-occurrence	4000.00	273- Bridging Text Visualization and Mining A Task-Driven Survey.txt
leaning-based	4000.00	55 - Audio-Tactile Proximity Feedback for Enhancing 3D.txt
cohen-or	3970.59	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
post-test	3970.59	115 - Blending Human and Artificial Intelligence to Support.txt
low-resolution	3823.53	55 - Audio-Tactile Proximity Feedback for Enhancing 3D.txt
two-dimensional	3705.88	272- Task-Based Effectiveness of Basic Visualizations.txt
self-report	3705.88	154- An ˜In the Wild Experiment on Presence and Embodiment using Consumer Virtual Reality Equipment.txt
ground-truth	3588.24	26- Computer-Assisted Laparoscopic myomectomy by augmenting the uterus with pre-operative MRI data.txt
co-presence	3588.24	347- Mixed Reality Tabletop Gameplay Social Interaction with a Virtual Human Capable of Physical Influence.txt
top-down	3500.00	73- The Self-Avatar Follower Effect in Virtual Reality.txt
super-resolution	3500.00	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
self-motion	3470.59	77- Visual-Auditory Redirection Multimodal Integration of Incongruent Visual and Auditory Cues for Redirected Walking.txt
self-disclosure	3470.59	10 - All You Need is Love Current Strategies of Mediating Intimate.txt
first-person	3470.59	84 – Don't Bother Me How to Handle Content-Irrelevant Objects in.txt
third-person	3382.35	47 - Sharing Domestic Life through Long-Term Video Connections.txt
time-consuming	3352.94	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
hmd-based	3323.53	70- Distance Judgments to On- and Off-Ground Objects in Augmented Reality.txt
glyph-based	3294.12	378- Visual Cascade Analytics of Large-Scale Spatiotemporal Data.txt
small-scale	3294.12	186- Yarn-Level Cloth Simulation with Sliding Persistent Contacts.txt

signed-rank	3294.12	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
pre-test	3264.71	115 - Blending Human and Artificial Intelligence to Support.txt
low-rank	3264.71	370- Low Rank Matrix Approximation for 3D Geometry Filtering.txt
low-cost	3264.71	29- Elastic-Arm Human-scale passive haptic feedback for augmenting interaction and perception in virtual environments.txt
touch-based	3235.29	147- Reactive Vega A Streaming Dataflow Architecture for Declarative Interactive Visualization.txt
multi-dimensional	3235.29	159- Styling Evolution for Tight-Fitting Garments.txt
7-point	3205.88	347- Mixed Reality Tabletop Gameplay Social Interaction With a Virtual Human Capable of Physical Influence.txt
user-centered	3205.88	83 - How People Write Together Now Beginning the Investigation.txt
semi-structured	3205.88	47 - Sharing Domestic Life through Long-Term Video Connections.txt
k-means	3176.47	344- ProReveal Progressive Visual Analytics With Safeguards.txt
head-worn	3176.47	196- Towards Pervasive Augmented Reality Context-Awareness in Augmented Reality.txt
side-by-side	3176.47	363- Augmenting Sports Videos with VisCommentator.txt
rule-based	3147.06	66- Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw.txt
non-parametric	3147.06	193- PlenoPatch Patch-Based Plenoptic Image Manipulation.txt
so-called	3117.65	47 - Sharing Domestic Life through Long-Term Video Connections.txt
high-speed	3117.65	84- RNIN-VIO Robust Neural Inertial Navigation Aided Visual-Inertial Odometry in Challenging Scenes.txt
bottom-up	3058.82	202- Stereoscopic Thumbnail Creation via Efficient Stereo Saliency Detection.txt
well-being	3029.41	65- Mixed Reality Office System Based on Maslow's Hierarchy of Needs Towards the Long-Term Immersion in Virtual Environments.txt

Appendix H

Words Generated as Discipline Specific by Frequency (≥50) and Keyness (≥190)

Word Results based on Frequency and Keyness over 190 for Discipline Specific Inclusion: Listed in descending order of publication venues the word is present within from five to one

Word	Frequency of Occurrence	Keyness
visualization	484	3231.81
augmented	1377	802.62
immersive	310	539.79
gaze	181	444.6
metrics	68	303.65
occlusion	214	198.24
perceptual	136	188.63
manipulation	138	188.25
interacting	56	136.5
stimuli	56	120.18
annotations	82	114.15
multimodal	50	112.92
sensory	59	44.71
first	509	1255.15
significant	468	720.3
dataset	108	547.39
specific	145	545.54
find	152	381.88
usability	154	294.41
visually	100	238.81
explored	53	143.81
subjective	124	133.95
wearable	66	129.55
statistically	76	119.4
randomly	61	113.82
gesture	207	80.52
questionnaires	79	69.43
realism	59	45.25
physiological	141	44.22
discomfort	80	42.58
field	308	736.73
haptic	81	461.35
cues	299	396.38
defined	161	379.83
influence	128	261.97

difficult	121	249.24
efficient	131	241.44
final	100	238.36
latency	201	226.45
deformation	138	220.2
locomotion	51	220.09
classification	65	219.47
sickness	154	216.1
artifacts	73	211.84
fixed	90	196.48
identified	66	192.17
comput	65	1121.01
significantly	51	398.11
segmentation	145	302.42
clustering	1212	266.53
vertices	84	264.03
saliency	59	235.49
similarity	1100	227.55
behaviors	373	226.29
trajectories	549	220.09
filter	93	197.5
analytics	2154	485.79
flow	1881	425.62
five	678	262.18
findings	73	255.4
specifically	53	200.79
datasets	146	298.11
dementia	1093	224.61
fields	861	208.8

Appendix I

CoXR Survey Phase Two Participant Survey

Start of Block: Pre-Questions

Q12 As a graduate student, what level of study are you currently attending for?

o Master's Degree (1)

o Doctoral Degree (2)

Q13 As a graduate student, what year of study are you currently attending in

o 1 (1)

o 2 (2)

o 3 (3)

o 4 (4)

o 5 (5)

o 6+ (6)

Q14 Is the English Language your first language?

o Yes (1)

o No (2)

End of Block: Pre-Questions

Start of Block: Annotations

Q1 To your knowledge, is this a reasonable example for the given word?

Word in Question: Annotations

Example: In this paper, we present a touchscreen interface for creating freehand drawings as world-stabilized annotations and for virtually navigating a scene reconstructed live in 3D, all in the context of live remote collaboration.

o Yes (1)

o No (2)

Q3 If you answered No, that the previous example is not reasonable, explain why below.

Q2 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Annotations

Example: In this paper, we present a touchscreen interface for creating freehand drawings as world-stabilized annotations and for virtually navigating a scene reconstructed live in 3D, all in the context of live remote collaboration.

o Yes (1)

o No (2)

Q4 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q6 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Annotations

Example: In this paper, we present a touchscreen interface for creating freehand drawings as world-stabilized annotations and for virtually navigating a scene reconstructed live in 3D, all in the context of live remote collaboration.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Annotations

Start of Block: Augmented

Q7 To your knowledge, is this a reasonable example for the given word?

Word in Question: Augmented

Example: With the advent of more powerful hardware and improved methods, real-time approaches, known as 3D tracking and mapping have been demonstrated which allow new applications in augmented-reality.

o Yes (1)

o No (2)

Q8 If you answered No, that the previous example is not reasonable, explain why below.

Q9 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Augmented

Example: With the advent of more powerful hardware and improved methods, real-time approaches, known as 3D tracking and mapping have been demonstrated which allow new applications in augmented-reality.

o Yes (1)

o No (2)

Q10 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q11 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Augmented

Example: With the advent of more powerful hardware and improved methods, real-time approaches, known as 3D tracking and mapping have been demonstrated which allow new applications in augmented-reality.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Augmented

Start of Block: Dataset

Q15 To your knowledge, is this a reasonable example for the given word?

Word in Question: Dataset

Example: We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

o Yes (1)

o No (2)

Q16 If you answered No, that the previous example is not reasonable, explain why below.

Q17 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Dataset

Example: We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

o Yes (1)

o No (2)

Q18 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q19 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Dataset

Example: We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Dataset

Start of Block: Explored

Q20 To your knowledge, is this a reasonable example for the given word?

Word in Question: Explored

Example: In the early stages, advances in eye-tracking technology were primarily motivated by research in psychology and physiology, which explored the functioning of the human eye and its link with cognitive and perceptual processes.

o Yes (1)

o No (2)

Q21 We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

Q22 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Explored

Example: In the early stages, advances in eye-tracking technology were primarily motivated by research in psychology and physiology, which explored the functioning of the human eye and its link with cognitive and perceptual processes.

o Yes (1)

o No (2)

Q23 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q24 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Explored

Example: In the early stages, advances in eye-tracking technology were primarily motivated by research in psychology and physiology, which explored the functioning of the human eye and its link with cognitive and perceptual processes.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Explored

Start of Block: Find

Q30 To your knowledge, is this a reasonable example for the given word?

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

o Yes (1)

o No (2)

Q31 If you answered No, that the previous example is not reasonable, explain why below.

Q32 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

o Yes (1)

o No (2)

Q33 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q34 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Find

Start of Block: First

Q25 To your knowledge, is this a reasonable example for the given word?

Word in Question: First

Example: We first point out the connection, leading to the discussion of three core registration components in the light of the overfitting problem: Model selection, correspondences and constraints, and optimization.

o Yes (1)

o No (2)

Q26 If you answered No, that the previous example is not reasonable, explain why below.

Q27 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: First

Example: We first point out the connection, leading to the discussion of three core registration components in the light of the overfitting problem: Model selection, correspondences and constraints, and optimization.

o Yes (1)

o No (2)

Q28 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q29 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: First

Example: We first point out the connection, leading to the discussion of three core registration components in the light of the overfitting problem: Model selection, correspondences and constraints, and optimization.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: First

Start of Block: Gaze

Q35 To your knowledge, is this a reasonable example for the given word?

Word in Question: Gaze

Example: Finally, we need the ability to communicate with virtual agents using different cues such as gaze or eye contact.

o Yes (1)

o No (2)

Q36 If you answered No, that the previous example is not reasonable, explain why below.

Q37 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Gaze

Example: Finally, we need the ability to communicate with virtual agents using different cues such as gaze or eye contact.

o Yes (1)

o No (2)

Q38 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q39 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Gaze

Example: Finally, we need the ability to communicate with virtual agents using different cues such as gaze or eye contact.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Gaze

Start of Block: Immersive

Q40 To your knowledge, is this a reasonable example for the given word?

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

o Yes (1)

o No (2)

Q41 If you answered No, that the previous example is not reasonable, explain why below.

Q42 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

o Yes (1)

o No (2)

Q43 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q44 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Immersive

Start of Block: Interacting

Q45 To your knowledge, is this a reasonable example for the given word?

Word in Question: Interacting

Example: For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research).

o Yes (1)

o No (2)

Q46 If you answered No, that the previous example is not reasonable, explain why below.

Q47 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Interacting

Example: For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research).

o Yes (1)

o No (2)

Q48 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q49 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Interacting

Example: For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research).

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Interacting

Start of Block: Manipulation

Q50 To your knowledge, is this a reasonable example for the given word?

Word in Question: Manipulation

Example: The simultaneous manipulation (i.e.translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting.

o Yes (1)

o No (2)

Q51 If you answered No, that the previous example is not reasonable, explain why below.

Q52 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Manipulation

Example: The simultaneous manipulation (i.e.translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting.

o Yes (1)

o No (2)

Q53 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q54 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Manipulation

Example: The simultaneous manipulation (i.e.translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Manipulation

Start of Block: Metrics

Q55 To your knowledge, is this a reasonable example for the given word?

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

o Yes (1)

o No (2)

Q56 If you answered No, that the previous example is not reasonable, explain why below.

Q57 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

o Yes (1)

o No (2)

Q58 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q59 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Metrics

Start of Block: Multimodal

Q60 To your knowledge, is this a reasonable example for the given word?

Word in Question: Multimodal

Example: The current work presents design, development and a usability study of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.

o Yes (1)

o No (2)

Q61 If you answered No, that the previous example is not reasonable, explain why below.

Q62 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Multimodal

Example: The current work presents design, development and a usability study of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.

o Yes (1)

o No (2)

Q63 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q64 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Multimodal

Example: The current work presents design, development and a usability study of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Multimodal

Start of Block: Occlusion

Q65 To your knowledge, is this a reasonable example for the given word?

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

o Yes (1)

o No (2)

Q66 If you answered No, that the previous example is not reasonable, explain why below.

Q67 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

o Yes (1)

o No (2)

Q68 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q69 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Occlusion

Start of Block: Perceptual

Q70 To your knowledge, is this a reasonable example for the given word?

Word in Question: Perceptual

Example: In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity.

o Yes (1)

o No (2)

Q71 If you answered No, that the previous example is not reasonable, explain why below.

Q72 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Perceptual

Example: In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity.

o Yes (1)

o No (2)

Q73 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q74 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Perceptual

Example: In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Perceptual

Start of Block: Randomly

Q75 To your knowledge, is this a reasonable example for the given word?

Word in Question: Randomly

Example: It is often desirable to have a uniformly distributed yet randomly located point set.

o Yes (1)

o No (2)

Q76 If you answered No, that the previous example is not reasonable, explain why below.

Q77 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Randomly

Example: It is often desirable to have a uniformly distributed yet randomly located point set.

o Yes (1)

o No (2)

Q78 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q79 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Randomly

Example: It is often desirable to have a uniformly distributed yet randomly located point set.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Randomly

Start of Block: Sensory

Q80 To your knowledge, is this a reasonable example for the given word?

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

o Yes (1)

o No (2)

Q81 If you answered No, that the previous example is not reasonable, explain why below.

Q82 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

o Yes (1)

o No (2)

Q83 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q84 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Sensory

Start of Block: Significant

Q85 To your knowledge, is this a reasonable example for the given word?

Word in Question: Significant

Example: Prior research suggested a significant correlation between physiological signals and cybersickness severity, as measured by the simulator sickness questionnaire (SSQ).

o Yes (1)

o No (2)

Q86 If you answered No, that the previous example is not reasonable, explain why below.

Q87 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Significant

Example: Prior research suggested a significant correlation between physiological signals and cybersickness severity, as measured by the simulator sickness questionnaire (SSQ).

o Yes (1)

o No (2)

Q88 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q89 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Significant

Example: Prior research suggested a significant correlation between physiological signals and cybersickness severity, as measured by the simulator sickness questionnaire (SSQ).

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Significant

Start of Block: Specific

Q90 To your knowledge, is this a reasonable example for the given word?

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

o Yes (1)

o No (2)

Q91 If you answered No, that the previous example is not reasonable, explain why below.

Q92 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

o Yes (1)

o No (2)

Q93 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q94 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Specific

Start of Block: Statistically

Q95 To your knowledge, is this a reasonable example for the given word?

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

o Yes (1)

o No (2)

Q96 If you answered No, that the previous example is not reasonable, explain why below.

Q97 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

o Yes (1)

o No (2)

Q98 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q99 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Statistically

Start of Block: Stimuli

Q100 To your knowledge, is this a reasonable example for the given word?

Word in Question: Stimuli

Example: The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups.

o Yes (1)

o No (2)

Q101 If you answered No, that the previous example is not reasonable, explain why below.

Q102 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Stimuli

Example: The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups.

o Yes (1)

o No (2)

Q103 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q104 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Stimuli

Example: The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Stimuli

Start of Block: Subjective

Q105 To your knowledge, is this a reasonable example for the given word?

Word in Question: Subjective

Example: In this paper, we compare these two types of 6DoF navigation techniques in an immersive context, through an experiment using both objective and subjective measurements to assess user performance, the occurrence of cybersickness symptoms and the level of presence, when using either of these navigation paradigms.

o Yes (1)

o No (2)

Q106 If you answered No, that the previous example is not reasonable, explain why below.

Q107 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Subjective

Example: In this paper, we compare these two types of 6DoF navigation techniques in an immersive context, through an experiment using both objective and subjective measurements to assess user performance, the occurrence of cybersickness symptoms and the level of presence, when using either of these navigation paradigms.

o Yes (1)

o No (2)

Q108 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q109 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Subjective

Example: In this paper, we compare these two types of 6DoF navigation techniques in an immersive context, through an experiment using both objective and subjective measurements to assess user performance, the occurrence of cybersickness symptoms and the level of presence, when using either of these navigation paradigms.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Subjective

Start of Block: Usability

Q110 To your knowledge, is this a reasonable example for the given word?

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

o Yes (1)

o No (2)

Q111 If you answered No, that the previous example is not reasonable, explain why below.

Q112 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

o Yes (1)

o No (2)

Q113 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q114 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Usability

Start of Block: Visualization

Q115 To your knowledge, is this a reasonable example for the given word?

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

o Yes (1)

o No (2)

Q116 If you answered No, that the previous example is not reasonable, explain why below.

Q117 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

o Yes (1)

o No (2)

Q118 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q119 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Visualization

Start of Block: Visually

Q120 To your knowledge, is this a reasonable example for the given word?

Word in Question: Visually

Example: Large faces obscure much of a background map, are too visually dominant, and may suggest a higher density.

o Yes (1)

o No (2)

Q121 If you answered No, that the previous example is not reasonable, explain why below.

Q122 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Visually

Example: Large faces obscure much of a background map, are too visually dominant, and may suggest a higher density.

o Yes (1)

o No (2)

Q123 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q124 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Visually

Example: Large faces obscure much of a background map, are too visually dominant, and may suggest a higher density.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Visually

Start of Block: Wearable

Q125 To your knowledge, is this a reasonable example for the given word?

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

o Yes (1)

o No (2)

Q126 If you answered No, that the previous example is not reasonable, explain why below.

Q127 To your knowledge, is this example sentence a correct use of the given word?

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

o Yes (1)

o No (2)

Q128 If you answered No to the previous question that this example is not the correct use of the given word, explain why below.

Q129 To your knowledge, would you consider this word discipline specific to the field of Extended Reality (XR), somewhat discipline specific, or general academic language?

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

o Discipline Specific (1)

o Somewhat Discipline Specific (2)

o General Academic Language (3)

End of Block: Wearable

Appendix J

CoXR Survey Phase Three – English Major Participants

Start of Block: Pre-Questions

Q12 As a graduate student, what level of study are you currently attending for?

o Master's Degree (1)

o Doctoral Degree (2)

Q13 As a graduate student, what year of study are you currently attending in

o 1 (1)

o 2 (2)

o 3 (3)

o 4 (4)

o 5 (5)

o 6+ (6)

Q14 Is the English Language your first language?

o Yes (1)

o No (2)

Q134 What is your current program of study?

o Creative Writing (4)

o English Education (5)

o Literature (2)

o Rhetoric and Composition (3)

o TEFL/TESL (6)

Q135 How comfort do you feel using computer science specific language?

o I feel very comfortable (1)

o I feel somewhat comfortable (2)

o I do not feel very comfortable (3)

End of Block: Pre-Questions

Start of Block: Annotations

Q136 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Annotations

Example: In this paper, we present a touchscreen interface for creating freehand drawings as world-stabilized annotations and for virtually navigating a scene reconstructed live in 3D, all in the context of live remote collaboration.

Q1 Do you understand this example?

o Yes (1)

o No (2)

Q3 Do you understand this word when it is used in the given sentence?

o Yes (9)

o No (10)

Q137 Duplicated for ease of reference for the following questions.

Word in Question: Annotations

Example: In this paper, we present a touchscreen interface for creating freehand drawings as world-stabilized annotations and for virtually navigating a scene reconstructed live in 3D, all in the context of live remote collaboration.

Q2 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q6 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (4)

End of Block: Annotations

Start of Block: Augmented

Q140 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Augmented

Example: With the advent of more powerful hardware and improved methods, real-time approaches, known as 3D tracking and mapping have been demonstrated which allow new applications in augmented-reality.

Q7 Do you understand this example?

o Yes (1)

o No (2)

Q139 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q142 Duplicated for ease of reference for the following questions.

Word in Question: Augmented

Example: With the advent of more powerful hardware and improved methods, real-time approaches, known as 3D tracking and mapping have been demonstrated which allow new applications in augmented-reality.

Q9 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q11 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Augmented

Start of Block: Dataset

Q143 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Dataset

Example: We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

Q15 Do you understand this example?

o Yes (1)

o No (2)

Q145 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q144 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Dataset

Example: We use random graphs instead of real-world data-sets to remove the confounds of dataset size and the distribution of data attribute values.

Q17 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q19 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Dataset

Start of Block: Explored

Q146 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Explored

Example: In the early stages, advances in eye-tracking technology were primarily motivated by research in psychology and physiology, which explored the functioning of the human eye and its link with cognitive and perceptual processes.

Q20 Do you understand this example?

o Yes (1)

o No (2)

Q22 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q149 Duplicated for ease of reference for the following questions.

Word in Question: Explored

Example: In the early stages, advances in eye-tracking technology were primarily motivated by research in psychology and physiology, which explored the functioning of the human eye and its link with cognitive and perceptual processes.

Q148 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q147 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Explored

Start of Block: Find

Q150 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

Q30 Do you understand this example?

o Yes (1)

o No (2)

Q32 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q153 Duplicated for ease of reference for the following questions.

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

Q151 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q152 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Find

Start of Block: First

Q155 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: First

Example: We first point out the connection, leading to the discussion of three core registration components in the light of the overfitting problem: Model selection, correspondences and constraints, and optimization.

Q25 Do you understand this example?

o Yes (1)

o No (2)

Q27 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q158 Duplicated for ease of reference for the following questions.

Word in Question: First

Example: We first point out the connection, leading to the discussion of three core registration components in the light of the overfitting problem: Model selection, correspondences and constraints, and optimization.

Q156 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q157 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: First

Start of Block: Gaze

Q159 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Gaze

Example: Finally, we need the ability to communicate with virtual agents using different cues such as gaze or eye contact.

Q35 Do you understand this example?

o Yes (1)

o No (2)

Q37 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q162 Duplicated for ease of reference for the following questions.

Word in Question: Gaze

Example: Finally, we need the ability to communicate with virtual agents using different cues such as gaze or eye contact.

Q160 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q161 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Gaze

Start of Block: Immersive

Q164 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

Q40 Do you understand this example?

o Yes (1)

o No (2)

Q42 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q167 Duplicated for ease of reference for the following questions.

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

Q165 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q166 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Immersive

Start of Block: Interacting

Q168 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Interacting

Example: For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research).

Q45 Do you understand this example?

o Yes (1)

o No (2)

Q47 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q171 Duplicated for ease of reference for the following questions.

Word in Question: Interacting

Example: For more than a century (e.g., the pioneering work of James [1890]), psychologists have highlighted the benefits of interacting with physical objects for a variety of domains, such as memorization, mental imagery, problem-solving and, potentially, learning (though identifying which situations are most likely to benefit from the use of physical manipulatives is still an active area of research).

Q169 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q170 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Interacting

Start of Block: Manipulation

Q172 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Manipulation

Example: The simultaneous manipulation (i.e.translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting.

Q50 Do you understand this example?

o Yes (1)

o No (2)

Q52 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q173 Duplicated for ease of reference for the following questions.

Word in Question: Manipulation

Example: The simultaneous manipulation (i.e.translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting.

Q174 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q175 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Manipulation

Start of Block: Metrics

Q176 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

Q55 Do you understand this example?

o Yes (1)

o No (2)

Q57 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q179 Duplicated for ease of reference for the following questions.

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

Q177 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q178 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Metrics

Start of Block: Multimodal

Q180 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Multimodal

Example: The current work presents design, development and a usability study of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.

Q60 Do you understand this example?

o Yes (1)

o No (2)

Q62 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q183 Duplicated for ease of reference for the following questions.

Word in Question: Multimodal

Example: The current work presents design, development and a usability study of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.

Q181 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q182 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Multimodal

Start of Block: Occlusion

Q184 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

Q65 Do you understand this example?

o Yes (1)

o No (2)

Q67 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q187 Duplicated for ease of reference for the following questions.

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

Q185 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q186 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Occlusion

Start of Block: Perceptual

Q188 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Perceptual

Example: In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity.

Q70 Do you understand this example?

o Yes (1)

o No (2)

Q72 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q191 Duplicated for ease of reference for the following questions.

Word in Question: Perceptual

Example: In the meanwhile, many motion control algorithms have been proposed, but a general consensus is that a classical washout filter provides the best trade-off between perceptual quality and algorithmic simplicity.

Q190 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q189 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Perceptual

Start of Block: Randomly

Q192 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Randomly

Example: It is often desirable to have a uniformly distributed yet randomly located point set.

Q75 Do you understand this example?

o Yes (1)

o No (2)

Q77 To your knowledge, is this example sentence a correct use of the given word?

o Yes (1)

o No (2)

Q195 Duplicated for ease of reference for the following questions.

Word in Question: Randomly

Example: It is often desirable to have a uniformly distributed yet randomly located point set.

Q193 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q194 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Randomly

Start of Block: Sensory

Q196 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

Q80 Do you understand this example?

o Yes (1)

o No (2)

Q82 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q199 Duplicated for ease of reference for the following questions.

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

Q198 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q197 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Sensory

Start of Block: Significant

Q200 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Significant

Example: Prior research suggested a significant correlation between physiological signals and cybersickness severity, as measured by the simulator sickness questionnaire (SSQ).

Q85 Do you understand this example?

o Yes (1)

o No (2)

Q87 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q203 Duplicated for ease of reference for the following questions.

Word in Question: Significant

Example: Prior research suggested a significant correlation between physiological signals and cybersickness severity, as measured by the simulator sickness questionnaire (SSQ).

Q202 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q201 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Significant

Start of Block: Specific

Q204 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

Q90 Do you understand this example?

o Yes (1)

o No (2)

Q92 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q207 Duplicated for ease of reference for the following questions.

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

Q205 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q206 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Specific

Start of Block: Statistically

Q208 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

Q95 Do you understand this example?

o Yes (1)

o No (2)

Q97 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q211 Duplicated for ease of reference for the following questions.

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

Q210 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q209 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Statistically

Start of Block: Stimuli

Q212 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Stimuli

Example: The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups.

Q100 Do you understand this example?

o Yes (1)

o No (2)

Q102 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q215 Duplicated for ease of reference for the following questions.

Word in Question: Stimuli

Example: The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups.

Q214 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q213 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Stimuli

Start of Block: Subjective

Q216 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Subjective

Example: In this paper, we compare these two types of 6DoF navigation techniques in an immersive context, through an experiment using both objective and subjective measurements to assess user performance, the occurrence of cybersickness symptoms and the level of presence, when using either of these navigation paradigms.

Q105 Do you understand this example?

o Yes (1)

o No (2)

Q107 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q219 Duplicated for ease of reference for the following questions.

Word in Question: Subjective

Example: In this paper, we compare these two types of 6DoF navigation techniques in an immersive context, through an experiment using both objective and subjective measurements to assess user performance, the occurrence of cybersickness symptoms and the level of presence, when using either of these navigation paradigms.

Q217 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q218 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Subjective

Start of Block: Usability

Q220 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

Q110 Do you understand this example?

o Yes (1)

o No (2)

Q112 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q223 Duplicated for ease of reference for the following questions.

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

Q222 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q221 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Usability

Start of Block: Visualization

Q224 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

Q115 Do you understand this example?

o Yes (1)

o No (2)

Q117 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q227 Duplicated for ease of reference for the following questions.

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

Q225 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q226 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Visualization

Start of Block: Visually

Q228 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Visually

Example: Large faces obscure much of a background map, are too visually dominant, and may suggest a higher density.

Q120 Do you understand this example?

o Yes (1)

o No (2)

Q122 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q231 Duplicated for ease of reference for the following questions.

Word in Question: Visually

Example: Large faces obscure much of a background map, are too visually dominant, and may suggest a higher density.

Q230 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q229 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Visually

Start of Block: Wearable

Q232 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

Q125 Do you understand this example?

o Yes (1)

o No (2)

Q127 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q235 Duplicated for ease of reference for the following questions.

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

Q234 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q233 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Wearable

Start of Block: Haptic

Q242 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Haptic

Example: Researchers have started to combine augmented reality (AR) and haptic interaction to enable users to see and touch digital information that is embedded in the real world.

Q243 Do you understand this example?

o Yes (1)

o No (2)

Q244 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q245 Duplicated for ease of reference for the following questions.

Word in Question: Haptic

Example: Researchers have started to combine augmented reality (AR) and haptic interaction to enable users to see and touch digital information that is embedded in the real world.

Q246 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q247 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Haptic

Appendix K

CoXR Survey Phase Three – CS Students

Start of Block: Pre-Questions

Q12 As a graduate student, what level of study are you currently attending for?

o Master's Degree (1)

o Doctoral Degree (2)

Q13 As a graduate student, what year of study are you currently attending in?

o 1 (1)

o 2 (2)

o 3 (3)

o 4 (4)

o 5 (5)

o 6+ (6)

Q14 Is the English Language your first language?

o Yes (1)

o No (2)

Q375 How did you learn words specific to your discipline of Extended Reality? (Please select all answers that apply)

▢ Dictionaries (1)

▢ An Online Corpus (2)

▢ YouTube and Online Videos (8)

▢ From Reading Text Books (3)

▢ From Reading Academic Journal Articles (4)

▢ From Attending Conferences (6)

▢ From Classroom-Based Materials and Instruction (7)

Q377 What would you say is the primary source of your discipline-specific language acquisition?

o Dictionaries (1)

o An Online Corpus (2)

o YouTube and Online Videos (7)

o From Reading Text Books (3)

o From Reading Academic Journal Articles (4)

o From Attending Conferences (6)

o From Classroom-Based Materials and Instruction (5)

End of Block: Pre-Questions

Start of Block: Clustering

Q254 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Clustering

Example: In visual analytics, visualizations are often based on automatic models, for instance, clustering models are used to group data visually.

Q255 Do you understand this example?

o Yes (1)

o No (2)

Q256 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q257 Duplicated for ease of reference for the following questions.

Word in Question: Clustering

Example: In visual analytics, visualizations are often based on automatic models, for instance, clustering models are used to group data visually.

Q374 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q259 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Clustering

Start of Block: Cues

Q278 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Cues

Example: To address this gap, we conducted a mixed-method experiment to compare four levels of translational cues and control: none (using the trackpad of the HTC Vive controller to translate), upper-body leaning (sitting on a “NaviChair”, leaning the upper-body to locomote), whole-body leaning/stepping (standing on a platform called NaviBoard, leaning the whole body or stepping one foot off the center to navigate), and full translation (physically walking). Results showed that translational cues and control had significant effects on various measures including task performance, task load, and simulator sickness.

Q279 Do you understand this example?

o Yes (1)

o No (2)

Q280 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q281 Duplicated for ease of reference for the following questions.

Word in Question: Cues

Example: To address this gap, we conducted a mixed-method experiment to compare four levels of translational cues and control: none (using the trackpad of the HTC Vive controller to translate), upper-body leaning (sitting on a “NaviChair”, leaning the upper-body to locomote), whole-body leaning/stepping (standing on a platform called NaviBoard, leaning the whole body or stepping one foot off the center to navigate), and full translation (physically walking). Results showed that translational cues and control had significant effects on various measures including task performance, task load, and simulator sickness.

Q282 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q283 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Cues

Start of Block: Deformation

Q284 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Deformation

Example: The idea of data-driven shape deformation is to provide explicit examples of how the input shape should look like under some example deformations (example poses) and then to interpolate between these poses in order to obtain a specific shape/pose instead of using synthetic basis functions or variational principles to drive the deformation.

Q285 Do you understand this example?

o Yes (1)

o No (2)

Q286 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q287 Duplicated for ease of reference for the following questions.

Word in Question: Deformation

Example: The idea of data-driven shape deformation is to provide explicit examples of how the input shape should look like under some example deformations (example poses) and then to interpolate between these poses in order to obtain a specific shape/pose instead of using synthetic basis functions or variational principles to drive the deformation.

Q288 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q289 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Deformation

Start of Block: Find

Q150 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

Q30 Do you understand this example?

o Yes (1)

o No (2)

Q32 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q153 Duplicated for ease of reference for the following questions.

Word in Question: Find

Example: This study draws inspiration and methods from Gabbard's experiments and extends them to find: 1) how the device technology affects the text style readability and 2) the best way to convey a color coded information in AR regardless of the specific device.

Q151 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q152 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Find

Start of Block: Fixed

Q362 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Fixed

Example: One way to reduce VIMS is through the use of rest frames–portions of the virtual environment that remain fixed in relation to the real world and do not move as the user virtually moves.

Q363 Do you understand this example?

o Yes (1)

o No (2)

Q364 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q365 Duplicated for ease of reference for the following questions.

Word in Question: Fixed

Example: One way to reduce VIMS is through the use of rest frames–portions of the virtual environment that remain fixed in relation to the real world and do not move as the user virtually moves.

Q366 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q367 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Fixed

Start of Block: Gesture

Q260 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Gesture

Example: Recent research has explored free-hand gesture interaction with AR interfaces, but there have been few formal evaluations conducted with such systems.

Q261 Do you understand this example?

o Yes (1)

o No (2)

Q262 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q263 Duplicated for ease of reference for the following questions.

Word in Question: Gesture

Example: Recent research has explored free-hand gesture interaction with AR interfaces, but there have been few formal evaluations conducted with such systems.

Q264 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q265 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Gesture

Start of Block: Immersive

Q164 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

Q40 Do you understand this example?

o Yes (1)

o No (2)

Q42 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q167 Duplicated for ease of reference for the following questions.

Word in Question: Immersive

Example: In the immersive virtual jump experience, we embed a story using rich multimedia content, such as images and sound.

Q165 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q166 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Immersive

Start of Block: Locomotion

Q272 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Locomotion

Example: Particularly, natural walking experiences are crucial since locomotion is one of the most common activities connected to user interaction within 3D environments.

Q273 Do you understand this example?

o Yes (1)

o No (2)

Q274 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q275 Duplicated for ease of reference for the following questions.

Word in Question: Locomotion

Example: Particularly, natural walking experiences are crucial since locomotion is one of the most common activities connected to user interaction within 3D environments.

Q276 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q277 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Locomotion

Start of Block: Metrics

Q176 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

Q55 Do you understand this example?

o Yes (1)

o No (2)

Q57 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q179 Duplicated for ease of reference for the following questions.

Word in Question: Metrics

Example: The collected data allow (1) to determine the best set of parameters to use for this image-based quality assessment approach and (2) to compare this approach to the best performing model-based metrics and determine for which use-case they are respectively adapted.

Q177 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q178 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Metrics

Start of Block: Occlusion

Q184 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

Q65 Do you understand this example?

o Yes (1)

o No (2)

Q67 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q187 Duplicated for ease of reference for the following questions.

Word in Question: Occlusion

Example: This category contains techniques that deal with the correct placement and identification of virtual objects, by relying on size, occlusion and texture.

Q185 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q186 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Occlusion

Start of Block: Realism

Q248 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Realism

Example: For example, the effect of avatar appearance on our behavior has been confirmed for a variety of properties including gender, posture, figure, skin color, age and size, or degree of realism and anthropomorphism.

Q249 Do you understand this example?

o Yes (1)

o No (2)

Q250 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q251 Duplicated for ease of reference for the following questions.

Word in Question: Realism

Example: For example, the effect of avatar appearance on our behavior has been confirmed for a variety of properties including gender, posture, figure, skin color, age and size, or degree of realism and anthropomorphism.

Q252 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q253 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Realism

Start of Block: real-world

Q320 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Real-world

Example: TransforMR is designed for use in the real-world, supporting the substitution of humans and vehicles in everyday scenes, and runs on mobile devices using just their monocular RGB camera feed as input.

Q321 Do you understand this example?

o Yes (1)

o No (2)

Q322 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q323 Duplicated for ease of reference for the following questions.

Word in Question: real-world

Example: TransforMR is designed for use in the real-world, supporting the substitution of humans and vehicles in everyday scenes, and runs on mobile devices using just their monocular RGB camera feed as input.

Q324 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q325 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: real-world

Start of Block: see-through

Q326 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: see-through

Example: These displays optically combine the computer-generated image with the user's direct view of the surroundings ("optical see-through"), in contrast to smartphone- and tablet-based AR applications, which combine the computer-generated image with video imagery ("video see-through").

Q327 Do you understand this example?

o Yes (1)

o No (2)

Q328 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q329 Duplicated for ease of reference for the following questions.

Word in Question: see-through

Example: These displays optically combine the computer-generated image with the user's direct view of the surroundings ("optical see-through"), in contrast to smartphone- and tablet-based AR applications, which combine the computer-generated image with video imagery ("video see-through")

Q330 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q331 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: see-through

Start of Block: Sensory

Q368 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

Q369 Do you understand this example?

o Yes (1)

o No (2)

Q370 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q371 Duplicated for ease of reference for the following questions.

Word in Question: Sensory

Example: For people who are blind, then, the wealth of visual information encoded in graphs is likely lost unless tools using another sensory modality, such as sound, touch or a combination thereof, are developed to compensate for vision.

Q372 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q373 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Sensory

Start of Block: Specific

Q204 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

Q90 Do you understand this example?

o Yes (1)

o No (2)

Q92 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q207 Duplicated for ease of reference for the following questions.

Word in Question: Specific

Example: Research on remeshing implicitly account for surface mesh sampling, but most of the proposed techniques focus on the specific task of remeshing and are not general.

Q205 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q206 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Specific

Start of Block: Statistically

Q208 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

Q95 Do you understand this example?

o Yes (1)

o No (2)

Q97 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q211 Duplicated for ease of reference for the following questions.

Word in Question: Statistically

Example: Results indicate that the wobbly group had higher presence and social presence with the virtual human in general, with statistically significant increases in presence, co-presence, and attentional allocation.

Q210 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q209 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Statistically

Start of Block: Usability

Q220 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

Q110 Do you understand this example?

o Yes (1)

o No (2)

Q112 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q223 Duplicated for ease of reference for the following questions.

Word in Question: Usability

Example: At last, a comparing user study evaluation is proposed to demonstrate the usability of the touch-less approach, as well as the impact on user's emotion, running on a wearable framework or Google Glass.

Q222 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q221 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Usability

Start of Block: Visualization

Q224 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term "scientific visualization".

Q115 Do you understand this example?

o Yes (1)

o No (2)

Q117 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q227 Duplicated for ease of reference for the following questions.

Word in Question: Visualization

Example: The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization”.

Q225 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q226 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Visualization

Start of Block: Wearable

Q232 For the following questions on this page, refer to this word and its following example of usage.

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

Q125 Do you understand this example?

o Yes (1)

o No (2)

Q127 Do you understand this word when it is used in the given sentence?

o Yes (1)

o No (2)

Q235 Duplicated for ease of reference for the following questions.

Word in Question: Wearable

Example: As head-mounted wearable cameras allow us to capture a first-person view and share the experience, wearable cameras are one of the primary approaches for sharing first-person visual information.

Q234 With the example provided, could you use this word in a sentence?

o Yes (1)

o No (2)

Q233 If you encountered this word later, might you be able to understand its meaning?

o Yes (1)

o No (2)

End of Block: Wearable

Start of Block: post-questions

Q355 The following questions ask you to respond in the text boxes below the question to the word list you have just encountered in this survey and asks you to consider other potential word lists in and out of the field of XR.

Q350 After completing this survey, do you see a use for a detailed word list that is discipline specific in your classrooms as a student?

Q351 In your opinion, what are your questions or concerns about using this word list in your classroom as a student?

Q352 If this word list were published with each word and its following example, what would you, based on your perspective and experience, do to improve the word list in terms of its accessibility and user friendliness?

Q353 Did you find the hyphenated words different in difficulty to the single word examples? Why or why not?

Q354 Would the use of a discipline specific word list such as this provide you with incentive to collaborate and/or explore across disciplines if you had access to it?

End of Block: post-questions