SPECIAL COLLECTION: NEW DIRECTIONS IN DIGITAL MODERN LANGUAGES RESEARCH

ARTICLES – DIGITAL MODERN LANGUAGES

Extended Reality Language Research: Data Sources, Taxonomy and the Documentation of Embodied Corpora

ABDULRAHMAN A. A. ALSAYED

ABSTRACT

This article introduces extended reality linguistics, an emerging interdisciplinary field that explores the use of extended reality technologies in the scientific study of language. It delves into a variety of data sources for language study accessible through extended reality and showcases the advancements in data collection, visualization and analysis that this technology and its unique properties offer linguistic researchers.

A comprehensive taxonomy of extended reality data is also presented, along with a survey of its applications in recent linguistic studies, offering further insights into the potential implications for future work. The article culminates in a pilot report on CEERR, an ongoing extended reality corpus linguistic project. This report highlights the innovative concepts of virtual linguistic fieldwork, communicative event reconstruction and embodied linguistic corpora. The CEERR pilot corpus demonstrates the powerful potential of extended reality for language documentation, prompting consideration of new grounds that traverse previous boundaries of language research data.

CORRESPONDING AUTHOR:
Abdulrahman A. A. Alsayed
University of St Andrews, United Kingdom; The Alan Turing Institute for AI and Data Science, United Kingdom; King Faisal University, Saudi Arabia
alsayed@oxfordalumni.org

TO CITE THIS ARTICLE:
Alsayed, Abdulrahman A. A. 2023 Extended Reality Language Research: Data Sources, Taxonomy and the Documentation of Embodied Corpora. *Modern Languages Open*, 2023(1): 46 pp. 1–20. DOI: https://doi.org/10.3828/mlo.v0i0.441

1 TECHNICAL BACKGROUND

1.1 INTRODUCTION

The use of extended reality (XR) in language studies comes with exciting potential for innovative developments that can enhance existing research methods and give rise to novel approaches for investigating linguistic phenomena. As a uniquely naturalistic way of interacting with digital content, XR has gained market momentum in recent years due to breakthrough developments in its consumer hardware such as wireless virtual reality headsets and augmented reality smart glasses. These developments have contributed to the increase in the adoption rates of the technology by consumers.

Everyday life has been vastly transformed as a direct result of consumer technologies becoming mainstream, as is the case with personal computers and smartphones. This transformation, in turn, has had a resonating effect on different aspects of research where these technologies have become part of how data is collected, stored and used. XR market trends show promise of a similar impact as the technology continues to evolve and become part of the mainstream consumer market.

If these market trends continue, XR will rapidly become integrated into consumers' daily digital experiences, including the use of interactive media and access to the world wide web. Research in various linguistic fields stands to benefit from access to a wealth of data as a result of this increasing growth of XR use. With its unique properties, what this data can be used for in research is unlike any other type of data currently used in language studies, as will be demonstrated in this article.

Since XR encompasses various technologies, this article provides a technical overview of XR and explores the possibilities it offers for gathering data relevant to linguistic work. Despite the technical nature of parts of this discussion, the article does not assume any prior knowledge of the subject from the reader and attempts to define both basic and advanced terms as needed. There are also cases where using XR components may require advanced technical expertise to be available to the researcher. These cases will be made clear throughout the discussion.

Following the technical overview, primary focus is given to XR data. With its use in language research in mind, the article identifies the different sources and properties of this data. Additionally, it presents a taxonomy intended to draw clear boundaries between the different degrees to which XR is involved in producing language-related data. This is an initial step towards establishing clear descriptors for the types of XR data utilized in research, what this data can reveal about linguistic behaviour in virtual spaces and what it cannot.

Linked to the XR data taxonomy is a survey of research work from various linguistic fields. This work is highlighted for its innovative integration of XR into existing methodologies, or for pioneering innovative methodologies built specifically around XR. This survey attempts to showcase the variety of approaches that currently exist in all areas of XR linguistic research. While the focus here is on XR linguistic research, much of the methodological discussion is equally applicable to other areas in the social sciences and the humanities.

The article concludes with a report on a project that aims to establish a methodology for the creation of language corpora extended with virtual reality. This begins by introducing the concepts of virtual linguistic fieldwork, communicative event reconstruction and embodied corpus study. The report details how a pilot corpus for the project was successfully completed by integrating a layer of the language speaker's embodied visual experience in XR into the corpus design. This is provided to showcase how XR can offer advantageous extensions to linguistic fields that have yet to explore its integration in their methodologies.

The different topics discussed in the article ultimately aim to engage the research interests of linguists from all fields and encourage future applications of cutting-edge XR technology to tackle current issues in linguistic enquiry. The possibilities of investigating these issues in virtual environments that surpass the limitations of physical environments is a powerful tool that can redefine the limits of the study of language.

Modern Languages Open
DOI: 10.3828/mlo.v0i0.441

1.2 OVERVIEW OF EXTENDED REALITY (XR)

Extended reality is a cover term for technologies that create a virtual extension of the user's real physical environment by adding virtual elements to it or replacing it completely with virtual environments. They are unique in offering the user a human–computer interactive experience designed around three-dimensional visual perception, gesture and natural body movement. This involves the sophisticated use of hardware and software that is capable of inducing a perception that is convincingly real enough to allow the user to experience virtual elements in the same way they experience real physical objects and environments. The quality of any piece of XR technology is primarily measured by how successful it is at delivering this experience.

In their current form, XR technologies are generally divided into three basic types: augmented reality (AR), mixed reality (MR) and virtual reality (VR). XR devices belonging to these types vary greatly, each occupying a different position on a reality–virtuality continuum (Milgram and Kishino). The degree of virtuality along this continuum is determined by the extent to which virtuality is integrated with real environments or replaces them. For example, a mixed reality experience would involve experiencing a virtual object displayed as part of the physical elements of a real room, and a mixed virtuality experience would mean experiencing a real object keyed in to become part of the virtual elements of a virtual room. The lines dividing these types of XR experiences are continuously thinning with the development of devices capable of a wider range of experiences that can seamlessly and instantly move between the elements of reality and virtuality.

The definitions of XR-related terms are still not consistent in the literature. In the case of MR, for example, Speicher et al. conclude based on expert interviews and a review of sixty-eight papers in the field that there are multiple competing definitions for this term. In specific contexts, it is synonymous with XR as a cover term that includes both AR and VR. In others, it represents the midpoint of the reality–virtuality continuum; the point where real and virtual elements are equally displayed for the user. A more distinct use of MR is to describe technology that makes virtual elements and real elements stand side by side and supplement one another. With this distinction, AR refers to technology that displays imagery that only interacts with the user but not with their real-world environment, while MR refers to technology with virtual elements that interact with both the user and their real-world environment. Table 1 provides an overview of XR technologies.

Table 1 Extended Reality Technologies.

XR TECHNOLOGY	REALITY EXTENSION	VIRTUALITY DEGREE	FORM FACTOR	ADVANTAGES	LIMITATIONS
Virtual Reality (VR)	offers an immersive and full virtual extension to reality that replaces the physical environment	visually isolates the user from their physical surroundings making them fully immersed in the three-dimensional space of the virtual environment	head mounted display (HMD), Cave Automatic Virtual Environment (CAVE)	highly immersive experience, can simulate environments that are dangerous or difficult to replicate in real life	HMDs: inability to see physical surroundings, may cause motion sickness, may feel isolating CAVE: requires a dedicated room with high cost equipment to function
Augmented Reality (AR)	offers a non-immersive and partial virtual extension to reality that overlays the physical environment	displays virtual elements as overlays onto the real elements in the user's physical environment with no interaction between the two	smart/AR glasses, smartphone, head mounted display (HMD)	allows for interaction with real-world environment, does not require isolation from physical surroundings	virtual elements do not offer immersive experiences, lacks spatial awareness of physical elements
Mixed Reality (MR)	offers an immersive and partial virtual extension to reality that blends with the physical environment	displays virtual elements that mix and interact with the real elements in the user's physical environment	smart MR glasses, smartphone, head mounted display (HMD)	offers spatially-aware immersive experiences integrated into the real world	limited in current availability and functionality, less bulky devices have limited processing power and sensor quality

These XR devices come in various shapes and forms depending on their design features. These features are determined by the components of the device and the intended purposes for which it is designed. Table 2 lists the major components of XR technologies. Familiarity with these components will inform the discussion on the sources of XR data and its unique properties provided in the next section.

COMPONENT	DESCRIPTION
Software	The software components of extended reality technology include the operating system, applications and media such as computer-generated and 360-degree video. XR software is responsible for tasks such as interfacing with hardware sensors, managing user input and output and displaying content.
Sensors	Sensors collect data about the user and their surroundings and feed it into XR devices. Commonly used sensors include video cameras, microphones, eye trackers, inertia measurement units (e.g., accelerometers and gyroscopes), time-of-flight sensors for depth measurement, and infrared cameras for tracking equipment that has motion capture markers.
Markers	Extended reality devices use either marker-based or markerless technology to display virtual elements. For example, marker-based AR uses markers with special patterns printed on them to display virtual imagery in reference to the location where the markers are placed. CAVE systems use markers to track the position of eyewear and input devices with infrared cameras.
Input	XR devices register user input through a variety of methods such as controllers, smartphones and hand tracking technology that uses an HMD's video cameras. Experimental input methods such as brain-computer interfaces are also being tested.
Output	The primary output of XR devices is the display of imagery, which can be complemented with audio output through speakers or headphones. Some devices also have actuator components that can output haptic feedback, such as vibrations.
Add-On	Certain XR devices are compatible with additional components that can feed them more input data and serve as extra sensory output sources, such as haptic vests, gloves, or suits. These components offer additional levels of immersion and interaction beyond what can be achieved with standard XR devices alone.

Table 2 Extended Reality Components.

2 SOURCES OF XR DATA FOR LINGUISTIC RESEARCH

The overview of XR technologies and components in the previous section shows that a blend of hardware, software and user input is used to deliver an XR experience. The language researcher has the flexibility to use any or all parts of this unique blend to obtain valuable data for their research. The current section highlights the sources of data obtainable through XR which can be utilized in linguistic studies. These sources and the unique properties of the data associated with them provide the researcher with methods for collecting, visualizing, exploring and analysing data unmatched by any other means currently utilized for language research. Four of these properties are highlighted in this section, each introducing the data sources most closely associated with it. This does not mean that the property is exclusively tied to the sources it introduces, as other sources might contribute to its realization as well, albeit to a lesser extent.

It is important to note that XR data sources can vary based on the device used. High-cost devices are typically equipped with advanced capabilities such as eye tracking and haptic feedback. Additional data can be obtained through these capabilities. Fortunately, the consumer market has recently seen the introduction of affordable devices equipped with comparable features that can match those of high-cost devices. The technical skills needed to acquire and process this data range from basic computer literacy to more complicated programming skills. Programming expertise is required when there is a need to use the software development kits (SDKs) of an XR device. SDKs are development tools designed to give software developers advanced access to the device's software and hardware beyond what is available to the basic user. To aid in identifying which data type is suitable for the researcher, each type discussed below is classified from basic to advanced depending on the technical skill needed to utilize it.

Property 1. *Retrieval of virtual communicative event environments*. Himmelmann defines a communicative event as a linguistic behaviour that includes speech and sign in addition to other aspects of the setting within which communication takes place. The space and time within which communication takes place are a big part of this setting. The setting might also include people and things in the communication environment. Immersive VR offers the unique opportunity for the setting of a communicative event that occurred within VR experiences to be documented in a three-dimensional space that is fully explorable from start to finish after the conclusion of the event. Through the use of VR display capture and by retaining copies of the virtual environment software, this spatio-temporal setting of communication can be retrieved at any point for access by the researcher during the study. It can also be made available to the research community as part of published data.

2.1 VIDEO DISPLAYS, SPEAKERS AND MICROPHONES

The visual and auditory output of XR devices can be used as a source of data that details the user's virtual experience. Capturing the video display feed of an XR device in use provides a recording of the user's complete field of view while they experience the virtual environment. This includes what they see in their binocular vision as well as their monocular peripheral vision. Such recordings offer researchers a first-person look at the user's visual perspective, their range of vision and their head movement reactions within the virtual environment.

Devices that come with built-in microphones can be set to record the user's speech as they engage with the content and experiences delivered by the device. It is now common for devices to have internal video and audio capture capabilities. Using internal capture, it is possible to record the user's display screen, speaker output and microphone input. Once recorded, a time-synchronized copy of the captured video and audio is saved to the device's internal memory. It can then be transferred by connecting the device to a PC.

Video and audio capture can also be performed externally. External capture connects the device to a PC either via a wire or wirelessly and streams both video and audio from the XR device to the PC. Basic technical skills are required to perform both internal and external capture, which, in the case of consumer hardware, is usually explained in the device's user guide. It should be noted that certain software might disable internal video capture for copyright purposes.

2.2 VIRTUAL ENVIRONMENTS

Although capturing the video feed of an XR display enables the researcher to have access to the user's visual field, parts of the environment within which the capture takes place might not be visible or they might only be seen momentarily. A unique feature specific to VR devices is that they offer the researcher the ability to revisit the three-dimensional environment previously experienced by users in cases where there is a need to explore it further. This feature allows for the addition of a complete copy of the communication environment to the linguist's repository of data for a more thorough analysis. This copy of the virtual environment used as a setting for communication can be revisited in its identical form at any time and from any location. Revisiting the environment within which linguistic behaviour took place can be a source of data readily available to the researcher at all stages of analysis without the limitation of time or location.

Access to this environment is dependent on the software it is tied to. Certain software may require the purchase of a licence and can only be downloaded from specific platforms and on specific VR devices. Others can be in the form of 360-degree video files that are compatible with most devices. The technical skills required to utilize this feature range from basic to advanced depending on the software and the specific parts of it the researcher wishes to retain for later access.

Property 2. *Simulation of naturalistic environments for experimental purposes*. The level of experimental control in immersive VR can go further than anything that is possible in a physical experimental setting. Nölle and Spranger discuss language and spatial cognition and how VR can aid in the investigation of frame of reference strategies. These are language strategies that speakers use to refer to objects in 3D environments. VR allows the researcher to investigate this spatial language by systematically controlling various aspects of the virtual environment such as the size of objects and the orientation of participants in relation to them. The virtual environment under control can be a replication of physical locations that the researcher would normally not be able to make systematic alterations to, or that might not be accessible to the researcher and the participants at the time of the study. In addition to simulating locations and objects, they can also include virtual representations of people, as illustrated in section 2.4.

Experimental control and ecological validity are typically viewed as opposing properties of methodological design. A researcher would need to choose between collecting data from within the confines of a controlled experiment or from an uncontrolled natural environment. Peeters ("Virtual") offers an alternative paradigm to this dichotomy by showing how high experimental control can go hand in hand with ecological validity when immersive VR is used. VR offers researchers a simulated naturalistic environment within which they can freely control the necessary elements of this environment for experimental purposes. Through this controllable simulation of naturalistic environments, both experimental control and ecological validity can be maintained in a single study design. Through simulation, linguistic work which

DOI: 10.3828/mlo.v0i0.441

was limited to the confines of the language laboratory environment can be easily moved out of these confines. Participants can be transported to different environments while still physically located in the same room. As mentioned in Peeters ("Virtual"), this allows for the study of brain activity in natural communication settings while maintaining the lab's experimental control required for psycholinguistic investigation.

2.3 VIRTUAL SIMULATIONS

Not all contexts are possible to create in the real world and replicate for the purpose of a single study. This stands as a challenge to studies that rely on the existence of a particular context. XR offers a solution through contextual simulation. For example, Chaffey et al. simulated a wildfire emergency in VR to train human operators who controlled a swarm of rescue robots. The operators gave verbal commands to the swarm, which in turn responded to the operator using computer-generated text messages. The natural language capabilities of this swarm were also used to communicate with the emergency survivors when they were located. The benefit of conducting such a simulation is to analyse and develop more effective human–robot communication systems that can deal with the simulated situations. A real emergency such as the one in the scenario used for this study is unlikely to be readily available to the research team when they are developing the system. This makes VR a powerful tool to use for simulation purposes when developing natural language models that are capable of effectively working in emergency situations.

With virtual simulations proving to be a valuable source for collecting language data, interest in simulations based on real physical locations and events is increasing. Digital twin projects that seek to accurately replicate real physical objects and locations in virtual environments are gaining popularity for cultural heritage preservation (Hutson et al.). Once such projects become widely accessible, they will present an opportunity to collect language data driven by cultural contexts from simulations run in preserved digital twins.

2.4 VIRTUAL CONFEDERATES

A study might make use of a confederate who is instructed by the researcher to behave in a specific way in order to observe the natural reaction of the study's participants to this behaviour. Pan and Hamilton note that despite their training, it is hard for human confederates to maintain the same behaviour when interacting with different participants, especially when different contexts are involved. This poses an issue for the reproducibility of studies that involve human confederates. Using virtual confederates is proposed as a solution to this issue. A virtual confederate can be programmed with artificial intelligence that uses a specific set of behavioural patterns that are fixed throughout the experiment, regardless of changes to participants and contexts. This makes it suitable for designing reproducible studies that can be shared between multiple research groups where identical virtual confederates are used as a common source for collecting data. Figure 1 shows the virtual confederate portraying a hospital patient used in the research conducted by Ochs et al., as discussed in section 3.2 below.

Figure 1 A virtual confederate playing the role of a hospital patient used in Ochs et al.

DOI: 10.3828/mlo.v0i0.441

Property 3. *Access to a portable inventory of 3D objects and facility for their standardization*. The use of physical objects as prompts to elicit language samples has been adopted in different linguistic fields such as language documentation and psycholinguistics. The benefit of using virtual objects for this purpose is that they enable the researcher to have a portable inventory with unlimited 3D models of different shapes and sizes to display to their participants on a device that is usable in both the lab and during fieldwork trips. It can be logistically impossible for the researcher to take a comparable number of physical objects into the field because of weights and sizes that cause issues with transportation. The use of virtual objects completely removes a practical limitation on the use of visual prompts for speech elicitation in language fieldwork.

As noted by Peeters ("Virtual"), virtual 3D objects can be standardized and shared globally between researchers. This is an additional benefit associated with the use of virtual 3D objects for research purposes. This benefit further facilitates the experimental control of these prompts between multiple researchers. To support standardization efforts, Peeters ("Standardized") offers a database of 3D objects that are VR-compatible, with the goal of encouraging more standardization of well-controlled prompts between experimental linguistic labs. He asserts that achieving this goal is important to guarantee experimental control between different labs, which is essential for reaching valid findings.

Additionally, it can be argued that access to a standardized set of 3D objects from any geographical location by the research community as a whole is only possible through the use of digitized objects. XR technology offers a highly practical and cost-effective solution to viewing these objects by research participants in a natural, real-like environment. Other means, such as printing objects with 3D printers, are less practical and come with greater limitations. XR allows the researcher to carry thousands of prompt objects within a small compact device that can fit in something as small as a backpack.

2.5 VIRTUAL OBJECTS

Displaying virtual objects comes as a basic functionality for the majority of XR devices. Virtual objects can either be part of a single virtual environment or they can exist separately and have compatibility with different environments, both virtual and physical. The user can select objects from their inventory to display in different environmental locations and conditions. In current linguistic research, such objects are used as stimuli that prompt linguistic behaviour from the participants.

The objects themselves and the information obtained from analysing their properties can also serve as a source of data. The data acquired through the analysis of objects used as linguistic stimuli can help in establishing patterns between the visual input of the stimuli and the linguistic output produced by the research participants. This approach can provide valuable insights into the relationship between object properties and the resulting linguistic responses.

Inventories of objects are increasingly being made available through online object repositories. These repositories contain various libraries of virtual objects created and freely shared under Creative Commons licences by 3D artists. One such example is the collection of free 3D models available on Clara.io, as illustrated in Figure 2. Another source of virtual objects is databases, such as those described in Tromp et al. and Popic et al., which have been created for research purposes and made openly accessible to the research community.

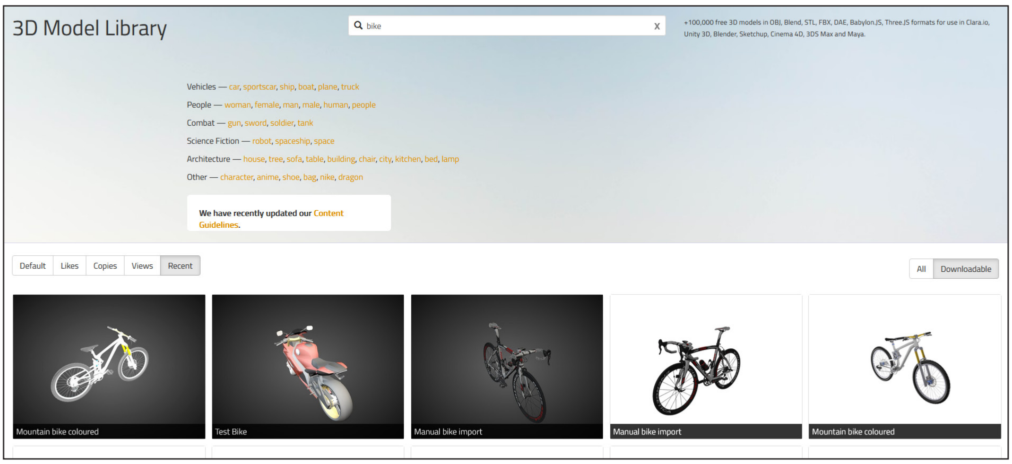

Figure 2 A repository of free 3D models on Clara.io.

2.6 USER INTERACTIONS

Virtual environments vary in their interactivity with the user. Some environments are completely unaffected by any of the user's interactions, while others can change uniquely depending on the actions performed by the users within them, providing a fully naturalistic experience. In environments reshaped by the user's input, a rich source of data comes from the unique interactions of the user with the objects and the environment simulated by the XR software, as well as the effects of these interactions on the environment. This includes data as detailed as changes in virtual physics and haptic generation patterns. Since this data is entirely software-based, it requires highly advanced programming abilities and is only accessible through developer tools.

Property 4. *Access to speaker kinesics within virtual environments*. Verbal and non-verbal communication are interconnected and of interest to work in linguistics and semiotics. The induction of virtual presence and embodiment discussed in section 4 promotes the user to display natural body reflexes and use non-verbal gestures as a response to XR stimuli. Software might also use gestures as a way to control and navigate it. Time-stamped XR-generated data logs can be saved and analysed for a detailed record of the user's body movement in virtual environments.

2.7 SENSOR READINGS

Video cameras, infrared cameras, inertia sensors and eye sensors generate data that tracks the user's movement while they use the device. The data generated by cameras and inertia sensors offers accurate information on the user's head movement position, velocity and orientation at any moment in time.

Eye tracking has recently started to become a standard part of consumer XR devices. XR devices equipped with eye-tracking sensors allow access to a range of data that can be used to determine eye gaze and eye movement. The SDKs for eye tracking enable access to data as detailed as the pupil diameter and blinking rate for each eye. The complex nature of acquiring data from these types of sensors requires highly advanced technical knowledge related to the sensor functionality and the use of SDKs.

2.8 CONTROLLER INPUT

XR attempts to utilize the user's hand movements as a natural way to interact with the device. Controllers held or worn by one or both hands are commonly used to allow the user to control the device's various functionalities. As mentioned previously, there are devices that have the capability to track the user's hands without the need for a controller. This hand-tracking capability is used as a method for users to input commands, interact with the environment and see a mirrored virtual representation of their physical hands, as illustrated in Figure 3. Similar to the previously mentioned sensor data, this type of data can also be accessed through SDKs and requires high technical knowledge to utilize.

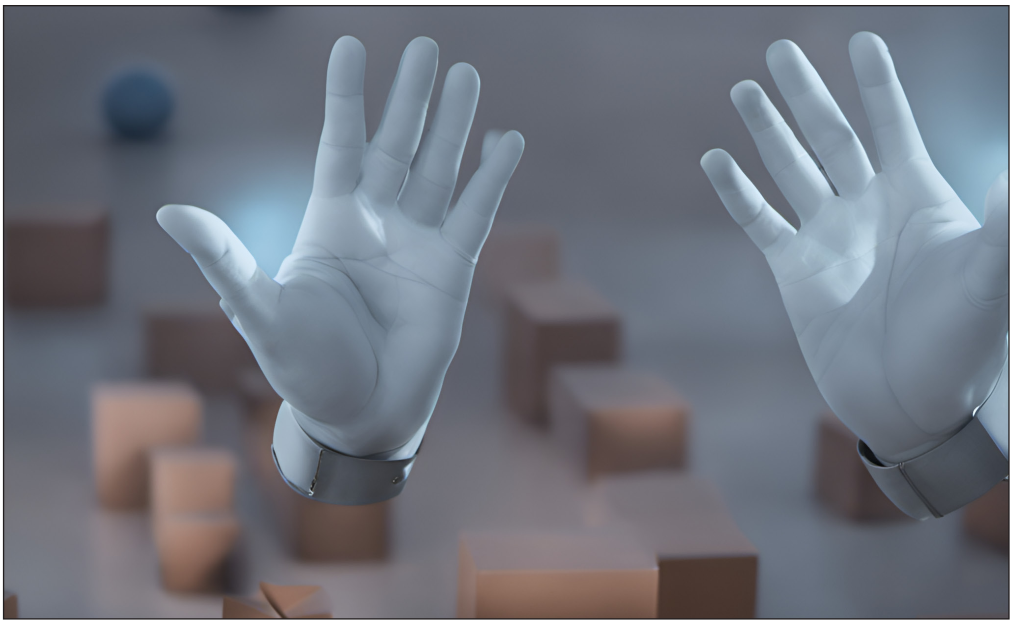

Figure 3 Virtual visualization of data collected from the XR user's physical hands represented in a virtual environment.

Modern Languages Open
DOI: 10.3828/mlo.v0i0.441

3 TAXONOMY OF XR DATA

The previous section illustrated the breadth of data types produced from multiple sources linked to XR technology. It showed how XR hardware, its user, compatible media and add-on hardware can all be useful sources of data for research. This gives language studies the option to choose from different combinations of this data depending on their research needs. The various data choices available here make it beneficial to develop taxonomies that accurately describe each configuration of data type and data source. This, in turn, informs how we can distinguish studies based on the data they use and the extent of the role XR plays in obtaining this data.

The data classification presented in this section is an attempt to develop a taxonomy based on the degree to which XR technology is involved in the production of the data under study. Five different degrees of involvement are proposed, each of which corresponds to a distinct class in the taxonomy. These five classes are XR generation, mediation, affixation, impact and compatibility.

Discussion of each class in the taxonomy examines a variety of linguistic research works in which XR forms a key part of the study. This surveyed work is selected to highlight what distinguishes each class and, additionally, to demonstrate the breadth of linguistic research prospects in this field, with special attention given to the unique properties XR holds for linguistic research. As such, the surveyed work is not meant to be exhaustive of all XR linguistic research to date. Rather, its purpose is to highlight how XR can be successfully integrated into studies from different linguistic fields and how the various types and sources of XR data contribute to linguistic inquiry in these fields. It should also be noted that some studies might make use of a combination of classes in a single methodology. In these cases, the study is discussed under the taxonomy class most central to its methodological design.

3.1 XR-GENERATED DATA

This class of data is produced by an XR device's operating system through readings from the XR hardware. XR devices generate data necessary for their basic functionality and for the different features of the software they run. This data is typically constantly generated by the device during use, independent of any involvement by the researcher. It is common practice for this data to be automatically stored in files referred to as logs as a means for developers to troubleshoot and debug the device and its software if needed.

XR-generated data types include data that is generated by the device's inertia sensors, eye-tracking sensors and haptic actuators. Not surprisingly, these are the types of data that require the highest levels of technical knowledge to utilize, as mentioned above. Additionally, software can be developed to give researchers more control over the generation of this data, allowing for a wider range to be generated as required by the specific needs of the study. Aiming to do so further elevates the level of technical skills required.

Pfeiffer showcases the use of data generated by XR for the creation of an Interactive Augmented Data Explorer (IADE) that solved a problem faced by linguistic researchers investigating the hand gesture of pointing. The linguists were assessing the precision of pointing by measuring pointing trajectory on video recordings of participants. However, the recordings they used did not allow for measurements as accurate as the researchers were anticipating. The IADE was designed as an XR solution for a highly accurate exploration of the multimodal data needed by this study. Instead of using 2D video recordings of participants in a physical environment, data related to pointing was generated by XR input gloves worn by the participants as they engaged in pointing at different objects in a VR environment setting. The IADE allowed the researchers to interactively view both the data and the setting. This interactivity gives the researcher the freedom to switch between the perspectives of the listener and the speaker at different points of the recorded data. It also allows them to rewind the simulation time, reverse it and view the manual annotations associated with specific moments in it. This system offered highly accurate measurements of non-verbal gestures which aided the researchers in achieving the goals of their study.

3.2 XR-MEDIATED DATA

This class encompasses data that comes from the user's observable behaviour resulting from XR's use as a medium for delivering content. It is the data that records the user's visible and audible interactions with this content while subjected to the mediation of XR. Such data allows for studying the effects of mediation on the user and for comparing it with unmediated behaviour.

XR involves the interaction between a human user and a computer device on which content is displayed. While viewing content on the device, the user produces speech and signs to communicate with other users sharing the virtual environment with them. These might also be produced as a way to describe what the user is experiencing in XR to people having a conversation with them while they are in a different physical or virtual environment.

Examples of XR-mediated data that can be obtained from such user experiences include recorded speech taken while speakers are actively using an XR device, video footage that records the speakers' gestures, and notes written by the researcher as they observe the speakers' linguistic behaviour. The uses of this data range from providing a view of the speaker's proxemics to virtual avatars and objects within the virtual experience, to showing the speaker's visible and audible behaviour from the perspective of an external observer. Interpreting mediated data related to speaker kinesics is intricately linked to the virtual environment in which it was produced under mediation.

It is important to note that there might be cases where frequent interruption in the mediation of XR occurs. This might happen in cases such as using AR applications on a smartphone where the speaker is not always looking at their AR-mediated screen. This should be carefully considered by the researcher, as language data collected in such situations will have gaps in the mediation that might have an impact on the results.

Kelly used aerial drone video and three-dimensional 360-degree panoramas to document the language of a Native American tribe. Two participating elders from the Northern Arapaho tribe were immersed in panoramic images of physical locations of cultural significance to the tribe. The researchers state that the locations the elders were able to visit in the VR images would have been impossible for them to visit in person. This enabled the researchers to elicit language samples from the two participants through culturally significant stimuli that were VR-mediated. The researchers note the rich narrative nature of the collected sample compared to traditional interview-style language elicitation.

XR offers a way to keep an accurate record of the different contextual elements that form virtual environments. This can be used to assist in the development of language models for natural language processing and generation that can benefit from direct access to environmental context. XR applications that make use of language models typically use models that are trained on language data collected from outside VR environments. Complementing these models with language data produced under the mediation of VR settings has the potential to enhance key aspects of their functionality. In Ochs et al., XR-mediation was involved in developing a VR application to train new doctors to interact with patients. The application uses a CAVE system. In CAVE, the user is placed in a room within which projectors display VR imagery on the walls of the room and, for some CAVE systems such as the one used in this study, on the ceiling and floor as well. A CAVE environment is used here to simulate the doctor trainees walking into a room which has a virtual patient lying on a hospital bed.

A corpus of human–human interactions between doctors and patients documents the doctors' communication with the patients as they break bad news to them. Ochs et al. began by using this corpus to design a prototype of the VR training platform for new doctors. The prototype was not fully autonomous and required a human operator. The operator was needed to simulate some of the platform's modules. A second corpus of human–machine interactions, now collected through VR mediation using the training prototype, was compiled with the goal of developing a fully autonomous platform. The platform became fully autonomous through substituting for the modules that were previously simulated by the human operator.

Mobile AR has received prominent attention in AR language education research. Parmaxi and Demetriou provide a systematic review of fifty-four research publications that investigated the use of AR in language learning. The review shows that mobile AR was the most used AR device in these studies. Location-based mobile AR applications, in particular, have seen a surge in popularity in the past few years.

Sydorenko et al. investigated the utilization of one of these applications in a language learning setting. Groups of students were asked to use location-based AR software and engage in collaborative dialogue to solve tasks. The students constructed narratives related to environmental sustainability by following the software's trail of green technology projects on and near the university campus. Their AR-mediated interaction was recorded with multiple video cameras. Through analysing these interactions, the study concluded that mobile AR facilitated opportunities for students to engage in open-ended dialogue and learn situation-driven vocabulary focused around a specific subject area.

3.3 XR-AFFIXED DATA

There might be a need to use a device that collects data from the user directly and not from the XR hardware. Such devices are used for additional monitoring of the user in cases such as when tracking their internal physical signs is needed. This includes brain activity data collected with electroencephalograms (EEGs) and heart activity data collected with electrocardiograms (ECGs). Affixed devices are not required for the functionality of XR hardware and are only needed for the purpose of collecting additional data from the user for research purposes. They concurrently collect data from the user while they are actively using the XR device. Studies that benefit from affixed data might also include data collected before and after the active use of XR for comparison purposes.

XR-affixed data of brain activities during immersive VR use has shown that the brain works in a comparable manner when engaged in activities in both physical environments and immersive VR applications (Petukhov et al.). According to this study, this is in contrast to participants using a desktop application where the brain activity data is not comparable with either physical or virtual environments. It is no surprise then that psycholinguistic studies investigating language and the brain have sought to explore the experimental applications of VR technology in their research. The psycholinguistic lab's brain monitoring equipment is used in combination with VR to collect VR-affixed data on brain function and related activity during language production and comprehension.

3.4 XR-IMPACT DATA

This class includes self-reported data that is collected from a research participant before or after their use of XR. It is related to studying the impact of XR's use as perceived by the user. This includes data obtained through user questionnaires, interviews and other experience data related to the use of XR. Acquiring and utilizing data impacted by XR's use does not require technical knowledge outside the limits of the area being investigated by the researcher and their data collection tool of choice.

One of the primary areas that has seen wide utilization of XR-impact data in linguistic studies focuses on the application of XR as a tool to enhance education. This is much like the case with other disciplines in the social sciences and the humanities where the use of XR in education has drawn wide research interest. Language education studies have been quick to investigate the utility of XR technologies for current and future applications in language teaching and learning methods.

A study conducted by Legault et al. presents a systematic investigation of the effectiveness of using immersive VR for learning Mandarin Chinese as a second language. Sixty-four participants were given the task of learning thirty Mandarin words in immersive VR environments and thirty other words in a traditional classroom. The words used in this study belonged to two vocabulary groups: kitchen items and zoo animals. When learning words in VR, the participants were able to see the objects and animals corresponding to the words they were learning displayed in a virtual kitchen and a virtual zoo. The study discusses the role of VR-embodied experiences and their implication for enhancing second language learning effectiveness, particularly for less successful language learners. This study is part of growing research efforts that seek to determine the effectiveness of XR in language education. A review of forty research papers on the topic of using VR for foreign language learning and teaching shows that the majority of research papers in this area focus on the effectiveness of using the technology (Solak and Erdem). Effectiveness continues to be a major topic in XR language education research.

Visiting a country where the target language is spoken gives learners the benefit of immersing themselves in the cultural setting of the language they wish to learn. Travel costs and other difficulties can prevent learners from benefiting from this opportunity. As a solution, Ebert et al. designed a VR system that allows students to immerse themselves virtually in a foreign location and engage in language learning. Study participants were divided into a traditional group, which used traditional language learning methods, and a VR group, which used the immersive VR system. Each group's ability to learn new vocabulary was then compared. The comparison looked at students' retention of the learned material, their opinion regarding the method's effectiveness, and their evaluation of enjoyment when using the system. The collected VR-impact data indicate that the VR group had significantly higher retention, effectiveness and enjoyment. This is taken as an indication that the VR system served a positive role in immersive language learning.

Other fields of linguistics have also utilized this class of data to measure extralinguistic factors affecting language production. In the field of pragmatics, role-play is used as a method of collecting data by giving participants a scenario and observing the pragmatic strategies they use as they each play a role in it. Enhancing methodology such as role-play with XR's powerful capabilities for contextual immersion is showing good promise, as investigated in Taguchi. This study looks at the use of role-play in immersive VR to evaluate participants' competence in producing the speech act of request. In the VR role-play, participants were placed together in a context-rich virtual space to play roles assigned to them. The study compares this VR method with the traditional role-play method by giving the participants a written scenario to read on a computer screen in order to elicit request responses from them. XR-impact data indicates that the participants responded to the VR method with a deeper emotional investment in the role-play compared to the traditional method.

3.5 XR-COMPATIBLE DATA

This is data produced outside the XR ecosystem that has the capability to be viewed on an XR device. While compatible data is not necessarily designed for use natively or exclusively on XR hardware, it can benefit from some of the unique features offered by XR such as immersive visualization.

For example, VR devices have the capabilities to display compatible video recorded with external 360-degree cameras or pre-rendered computer-generated imagery (CGI) created in 360-degree video format. For these types of videos, XR compatibility is superior to 2D video because it allows the user to view the content in a virtual space and feel a sense of immersion within its imagery. This is limited, however, by the position that the 360-degree camera was placed in when the video was made. Unlike native VR content, the displayed environment in these videos cannot be explored any further than the confines of the surroundings of the camera. This prevents the user from moving within the video environment and changing their distance and angle in relation to the elements in the environment.

The field of language documentation and description uses different tools to create records of the languages of the world for research purposes and for language preservation efforts. Recording video footage of a speech community to document its members as they engage in language use is one of these tools. These videos are traditionally recorded in 2D format. Pentangelo investigated the use of the 360-degree video format for this purpose. The endangered language documented in this study is Mohawk Kanien'kéha, an indigenous language spoken in North America. Pentangelo used a 360-degree camera to record a ten-hour video corpus of this language. The study concludes that the use of this video format in combination with VR resolved issues of camera framing associated with the use of unidirectional video formats. Instead of limiting the documentation to what is within the frame of the camera, the 360-degree format allows for documenting all the surroundings of the camera so that all the speakers and their conversation setting are being recorded throughout the documentation. The study also comments on the wide range of devices that the footage can be viewed on, which include PC media players as well as any VR-capable device, and the benefits of recording documentary linguistic videos in this format.

Modern Languages Open
DOI: 10.3828/mlo.v0i0.441

XR-compatible visualization can also be beneficial for exploring complex quantitative data in fields such as sociolinguistics. A pilot study conducted by Alissandrakis et al. tested the use of XR-compatible visualization of the geographical metadata of a corpus of Nordic tweets scraped from Twitter users who live in five Nordic countries. The metadata was used to visualize the languages spoken in these countries on a desktop application. The application has a mode to make the data compatible with a VR device. The visualization allows the user to see graphs of the metadata on a detailed global map in VR. The user can walk on the map and interact with data points on the graph for an immersive experience of the visualization. The study notes the users' ease of use and interpretation of the data using the VR visualization as encouraging signs for further utilization of this method for similar data-rich explorations in linguistics and in other fields of the social sciences and the humanities in general.

In addition to assisting in the visualization of language-related regional data, sociolinguistic research might find XR an ideal method of placing speakers in unique situations that can only occur in the confines of virtual simulations. An innovative demonstration of this use of XR is presented in Staum Casasanto et al., who used a novel method for studying speaker accommodation and its links to social motivations. In this experiment, participants were divided into two groups and asked to have a conversation with a virtual confederate in VR. The confederate could only use pre-recorded responses in the conversation. The participants were made aware that the confederate was incapable of understanding their replies. The confederate's responses were sped up by 12 per cent for one group and slowed down by 12 per cent for the other group. The results showed that the participants accommodated to the speed of the virtual confederate in their replies, even though they realized that it did not understand them. The study concludes that accommodation is not solely motivated by factors of social relationship and social needs. Automatic factors such as matching the speed of the other party involved in the conversation also seem to play a role, according to the results of this experiment.

The utilization of XR in other fields of linguistics continues to gain traction, with further exploration of how the technology can help solve challenges facing linguistic methodologies. Increased familiarity with the various types of XR data and their current uses in language studies has the potential to encourage wider utilization in more fields. Each field stands to gain distinct benefits relevant to the language phenomena it focuses on.

Given these diverse types of data collected by XR devices, careful consideration of the ethical implications associated with their use for research is undeniably necessary. However, a comprehensive exploration of these concerns is beyond the scope of this article.

4 DESIGNING AN EMBODIED LINGUISTIC CORPUS WITH VR-MEDIATED DATA

4.1 INTRODUCTION

This section describes the pilot stage of a project undertaken by the author at the University of St Andrews, United Kingdom, in November 2021. This pilot served as a proof-of-concept, establishing a methodology for creating VR-mediated linguistic corpora that document the embodiment of language speakers. As of the publication of this article, the methodology outlined here has been employed to gather VR-mediated data from fifty-one participants living in the city of Dundee. This data is currently used in a corpus-based study of syntax and information structure.

Embodied corpora collected through VR-mediation are designed to give researchers access to virtual embodiment features retained at all stages of analysis. This is achieved by using a combination of open-source corpus linguistic tools and consumer-grade VR technology, both of which are readily available for researchers who wish to adopt the methodology. The project demonstrates how the VR technology currently available to consumers can readily integrate with existing corpus linguistic tools and highlights the benefits resulting from this integration.

As overviewed in section 3, language documentation work has recently started to utilize new approaches that take advantage of features offered by XR technology, such as video language documentation using 360-degree video format, a format that is VR-compatible (Pentangelo), and the approach to language elicitation using 360-degree panoramic images mediated in immersive VR (Kelly). These approaches highlight some of the possibilities available to language documentation work that uses XR as a data collection and visualization tool.

Language documentation and empirical linguistic work, in general, stand to benefit significantly from XR technology properties that have yet to be utilized, some of which have been detailed in this article. The corpus methodology tested in the pilot presented in this section showcases the new possibilities available when using the unique property of immersive VR to archive and visualize the embodied presence of users. This property has been used to pioneer a methodology for the compilation of a virtual reality corpus of spoken speech, tokenized, transcribed and annotated with linguistic tags and visualized using existing corpus linguistic software integrated with a layer of XR data obtained from consumer-grade VR hardware.

The Communicative Event Extended Reality Repository (CEERR) project started out of a research need to study the linguistic phenomenon of information structure in a low-resource vernacular by using authentic data derived from spoken corpora. Existing spoken corpora that study this phenomenon cross-linguistically, examples of which are overviewed in Lüdeling et al., and spoken corpora that study context-sensitive phenomena in general are limited in how they can document the environment within which the language sample was produced.

What the language speaker sees and experiences by virtue of their presence in a communicative event is generally not included in the documentation, even though it is needed to fully understand the context of speech and the interlocutors' information updates. This is due to the difficulty associated with trying to keep a comprehensive record of the entirety of the communicative event. Attempts to document the communicative scene through field notes, photographs and videos all come with limitations, the most pertinent of which, to the discussion here, is the inability to capture the speaker's embodied presence in the communicative environment and their visual perspective as they engage in the communicative event. This section will demonstrate that this is now possible with the use of VR technology and the combination of virtual linguistic fieldwork, communicative event reconstruction and the compilation of embodied corpora.

4.2 PILOT CONCEPT

CEERR started with the aim of developing a methodology for linguistic fieldwork that uses VR to archive spoken corpora collected from speech communities of low-resource language varieties. The corpus archival process is designed to include a virtual reconstruction of the spatio-temporal elements comprising communicative events for the purpose of comprehensive documentation of the entirety of these events. To achieve this, each corpus is created within an embodied speaker experience in a VR-mediated setting, resulting in the collection of speech samples that can be traced back to a time-synchronized naturalistic setting where the speaker was virtually present. The creation of a CEERR corpus is conceptually built around three basic components: mediation, reconstruction and embodiment.

Conceptual Component 1. *Mediation*. When the spoken language sample is documented, there is extended reality mediation between the speaker and the spatio-temporal setting within which the communicative event is taking place.

Conceptual Component 2. *Reconstruction*. When the spoken language sample is documented, it is time-synchronized with a captured record of the environment within which the speaker was virtually present to allow for the reconstruction of the communicative event.

Conceptual Component 3. *Embodiment*. When the spoken language sample is documented, the speaker is experiencing a naturalistic virtual environment within which a feeling of embodied presence is induced.

These components are based on the three theoretical underpinnings of virtual linguistic fieldwork, communicative event reconstruction and embodied corpora, each of which is introduced below.

4.3 VIRTUAL LINGUISTIC FIELDWORK

Himmelmann distinguishes between three types of communicative events. First, natural communicative events occur without an external factor affecting them, including the presence of a fieldworker. Second, observed communicative events are naturally occurring events in which a fieldworker is present to observe and document them. Third, staged communicative events are orchestrated by the fieldworker and exist for the sole purpose of language documentation. The virtual linguistic fieldworker can collect language data through observation without influencing the virtual environment where the communicative events take place. This is done through observing language use in existing online virtual social hubs and platforms that are frequented by millions of users around the world. In virtual linguistic fieldwork, a fieldworker can be virtually present in communicative events that occur naturally on these platforms without being visible to the participants. This blurs the line between natural and observed communicative events and allows for unique methods of observation for language documentation purposes.

XR also facilitates the process of creating staged communicative events for virtual linguistic fieldwork. With the rapid increase in availability and portability of XR devices, fieldwork opportunities to document languages through virtual experiences have increased exponentially. It is now possible to conduct linguistic fieldwork in virtual settings through portable XR hardware that instantly transports a participant to an environment of the researcher's choice. Through this method multiple participants in different physical locations can be placed in identical environmental settings. This can be used so as to have the participants respond to stimuli in the environment individually or have them interact and communicate with each other, either synchronously or asynchronously. This provides the facility for creating virtually staged communicative events and utilizing them for language documentation.

4.4 COMMUNICATIVE EVENT RECONSTRUCTION

Collecting XR data through virtual linguistic fieldwork gives the researcher the ability to reconstruct the entirety of the communicative event with all its components, from the environment to the representation of the speaker in it, and from the linguistic behaviour to the events that prompted it. The reconstructed communicative event can be retained for all stages of linguistic analysis, virtually allowing the researcher to go back in time and observe the event unfolding in virtual space.

The reconstruction process benefits from the various data sources outlined in section 2. Combinations of data from audiovisual captures, virtual environmental retentions and sensor readings all contribute to different degrees of reconstruction as required by the study. By leveraging the ability to reconstruct virtual communicative events, researchers gain a valuable resource for more in-depth analysis of spoken language in its contextual setting.

4.5 EMBODIED CORPORA

Presence is a property unique to VR which allows the user to feel as if they are physically in the environment on display to them. This property gives the researcher a tool to instantly place their participants in environments that are vastly different from where they are physically located. This can be especially beneficial to fieldwork-based linguistic studies where, for example, maintaining a common environment for all participants is necessary but not easily achievable through physical travel to a specific location.

When immersed in a virtual experience designed to create a feeling of presence, a VR user typically perceives a sense of embodiment of their virtual representation. This sense of embodiment is real enough for participants to react to stimuli directed to their virtual representation as if they were directed to their actual physical bodies. A participant, for example, may step back to maintain distance as a natural reaction when approached by a virtual confederate who they perceive as invading the personal space of their virtual body (Wilcox et al.). The cognitive nature of embodiment in virtual worlds leads to interesting questions regarding other cognitive aspects, especially those related to language, that function in nearly identical ways in both physical and virtual environments.

An embodied corpus is a virtual reality corpus documentation of linguistic behaviour that combines the speaker's linguistic behaviour and their presence in an immersive communicative event that gives them a sense of embodiment. The documentation is achieved through placing language users in embodied experiences and then reconstructing the communicative events that occurred in them, as discussed in section 4.5.

4.6 COMPILING AN EMBODIED CORPUS

Following the design concepts and theoretical underpinnings outlined above, a pilot corpus was created to evaluate how each of the three conceptual components works together to achieve the methodological aims sought by the CEERR project. The first step in creating the corpus was to select a suitable virtual environment that takes the form of a simulated room within which the communicative event can take place. A computer-generated environment was used for this purpose. The environment contains different objects that interact with one another in a temporal linear sequence. Each step in the sequence leads to another until the sequence is complete. For example, an agent nudges a pool ball which starts rolling into a set of dominoes, making them fall. This sequence unfolds in one circular direction around the room to make it easier for the participant to follow the changes happening around them in the environment.

While the participant is present in the virtual environment, what they are viewing is captured using the VR device's display recording capability. Their voice is recorded using the device's internal microphone. This provides a record of the visual perspective of the participant as they look around the environment to follow its sequence of action. It also provides a record of the participant's speech as they talk about the virtual environment and the changes that occur in it. A copy of the full environment is also saved to be explored later in VR, independent of the participant's own visual perspective.

The pilot corpus documented a conversation between two native speakers of Eastern Peninsular Arabic who experienced this environment with the use of a portable HMD with inside-out tracking and video capture capabilities. Each participant was instructed to talk about what they saw to the other participant as an event sequence unfolded around them. They were asked to remain seated and informed that they could move their body freely to look around in any direction as they followed the sequence. While it is possible that the seated position might have limited the participants' movement and use of gestures, it was chosen for this specific pilot to protect the participants from discomfort in case they were prone to motion sickness and to eliminate the possibility of accidental falls. The participants were informed that their speech and visual perspective were being recorded.

After the recording was completed, the captured video was separated into two tracks, one for audio and one for video. Both tracks were kept time-synchronized with one another. Full transcription of the audio was then performed using speech-to-text AI, the output of which was then manually corrected to ensure accuracy. Following the transcription, the text was tokenized at the word level. Each token was time-synchronized with the audio and video tracks. The tokenized text was then annotated with reference to cues from the participant's speech and their visual perspective in the spatial setting of the virtual communicative event. Linguistic annotations included information structure, syntax and morphology. The linguistic annotations were adapted from Dipper et al. The choice of these annotations was determined to evaluate how effectively the VR-enabled immediate access to audiovisual context facilitated the annotation of a context-sensitive phenomenon such as information structure.

Finally, a multi-layer approach (Zeldes) was followed for visualizing the corpus. The included layers contained the tokenized text and the morphological, syntactic and information structural annotations. CEERR-specific annotation layers contained video and audio captured from the participant's embodied experience in VR, a time-stamp for each transition in the action sequence and virtual objects involved in the action. All layers were visualized using the ANNIS multi-layer corpus environment (Krause and Zeldes).

Figure 4 displays a corpus query for all elements annotated with the information status "giv" (given), which include phrases annotated as "NP" (noun phrase) within their parameters using the ANNIS query language syntax: infstat="giv" & NP="NP" & #1 _=_ #2. The search results show all instances of tokens within the parameter of given NPs found in the corpus, along with a video layer showing the captured perspective of the participant in the VR environment and an audio recording of their speech synchronized with the moment they produced the utterance in each instance.

Modern Languages Open
DOI: 10.3828/mlo.v0i0.441

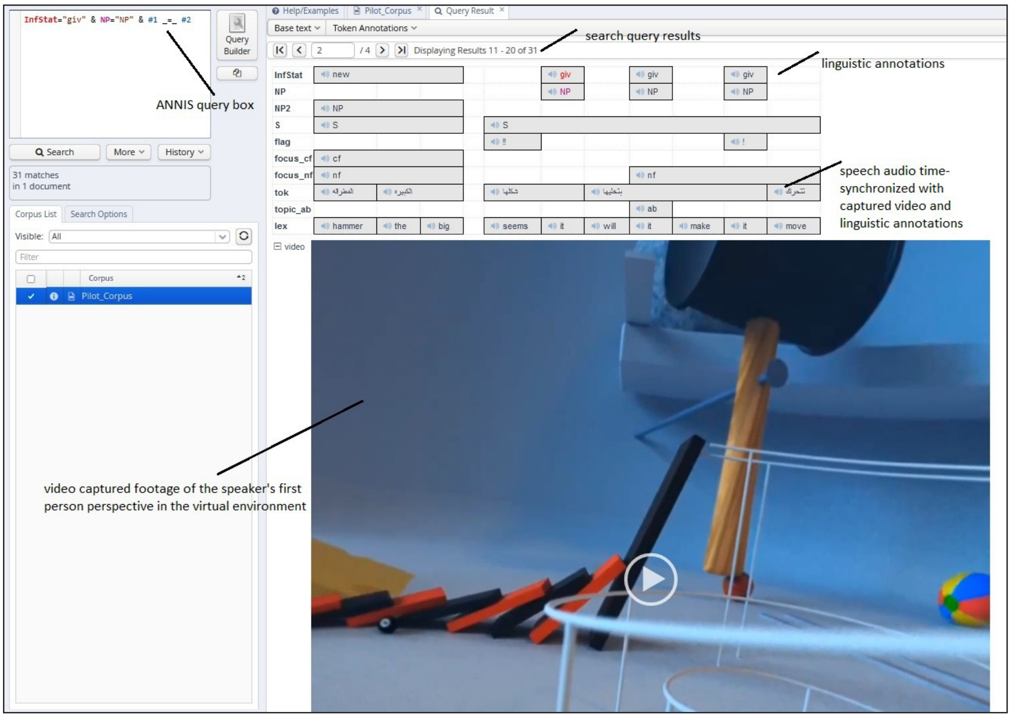

Figure 4 Query results from the CEERR pilot displayed in the ANNIS multi-layer corpus visualization environment.

4.7 INITIAL EVALUATION

With the use of a portable VR HMD, this pilot stage successfully created the first annotated virtual reality linguistic corpus as a proof-of-concept that showcases the capabilities of the project's methodology. A report of this pilot follows. The CEERR project pioneers the creation of an annotated linguistic corpus that combines the following XR features:

- VR-mediated language elicitation with three-dimensional computer-generated stimuli
- direct internal video capture recorded from the first-person perspective of the speaker
- embodied language documentation through virtual linguistic fieldwork
- annotation of linguistic information time-synchronized to visual elements of a virtual communicative event
- embodied corpus visualization through communicative event reconstruction

In its pilot stage, the resulting embodied corpus sufficiently demonstrated the benefits of integrating XR into corpus creation. Data related to the speaker's VR embodied experience could be explored in the corpus at every stage of corpus creation, analysis and use. This includes stages of annotating, visualizing and querying the corpus.

During annotation, the researcher could view the speaker's embodied visual perspective in the virtual environment from a first-person perspective, as if looking through the speaker's eyes. This makes it possible for the annotation process to be more contextually directed by the setting within which the communicative event takes place. Viewing the speaker's visual experience is possible on both a VR device or a desktop during annotation.

For the corpus visualization, the layer containing the VR captured video offers data on the virtual environment and the speaker's experience at any moment in time, fully synchronized with the linguistic annotation layers used in the study. The corpus user has the option to experience the virtual event themselves using a retained virtual environment copy included as part of the corpus metadata. The copy can be explored on a desktop application or immersively on a VR device.

Finally, as illustrated in Figure 4 above, when querying the corpus, the search results display the linguistic data relevant to the query time-aligned with audio of the spoken language and video of the visual perspective of the speaker. This allows for the immediate use of the VR-mediated data whenever needed while viewing the corpus query results. For example, querying specific tags for information structure such as the tag "topic" or "focus" allows for viewing the tokens annotated with these tags and a moment-to-token synchronized video of the visual experience as seen from the speaker's perspective.

The current form of the pilot corpus serves as a model to inform further planning for additional embodied corpus layers that can be used as part of corpora created for future CEERR projects. For instance, adding layers with data derived from eye tracking, hand tracking and haptic sensor readings can further document more aspects of the participant's experience during a virtual communicative event.

An enhanced version of the methodology has been employed in a CEERR documentation project that gathered embodied corpus data from thirty participants in the autumn of 2022 and a total of fifty-one participants in the first half of 2023. The methodology and the documentation project will be further explored in forthcoming work.

5 CONCLUSION

The pilot of the Communicative Event Extended Reality Repository (CEERR) language documentation project, and the new virtual reality corpus methodology it is based on, show the prospects for the combination of virtual linguistic fieldwork, communicative event reconstruction and embodied corpus compilation. They showcase the novel advantages that XR integration can offer to empirical language research in general, and more specifically to corpus-based linguistic inquiry into context-sensitive linguistic phenomena.

The embodiment features employed in this research serve as foundational steps in the nascent field of documenting embodied communication with XR, which can capture not only the speech and signs of speakers but also eye movements, facial expressions, hand gestures and body motion. By integrating these aspects of communication with a replica of the communication environment and the visual perspective of the interlocutors, this methodology enables the complete reconstruction and retention of communicative events, making it one of the most comprehensive documentation approaches available.

As XR technologies continue to evolve rapidly and become better suited for the mainstream consumer market, their user base in turn is growing exponentially. The increasing number of people who own XR devices and use them on a regular basis presents new opportunities for language research. XR is becoming a familiar form of technology, seen and utilized in everyday life for education, business and healthcare. This allows research projects that integrate XR with language studies to have a wider interest in society and consequently more impact.

One implication of this increase in XR usage is that the number of potential participants who might be reluctant to take part in linguistic studies that use XR because of concerns about unfamiliar technology will be reduced. Novelty effects interference is reduced as well. Additionally, the increased communication taking place within virtual spaces under extended reality mediation raises interesting questions related to the nature of language use in this immersive medium. This comes with a significant increase in the production of naturally occurring language under extended reality-mediation which can facilitate access to communicative events and offer a valuable source of language data.

The topics discussed in this article offer a look at how XR has been incorporated into existing methodological design, as well as starting novel approaches for the study of language. XR allows the language researcher to retain copies of virtual communication environments, access portable inventories of standardized 3D objects in unlimited quantities, simulate naturalistic environments for experimental purposes and gain access to highly precise records of speaker kinesics within virtual environments. These are some of the unique properties that XR data offers to language studies, and they come with immense prospects for future research.

There are still unexplored opportunities for investigating linguistic behaviour in virtual environments. Aspects of this behaviour might be shared with physical environments or unique to virtual ones. Future XR linguistic investigation invites exploration of such opportunities using innovative approaches that can further our understanding of XR, language and the research landscape that studies their interaction. Surveying current work that utilizes XR and its unique properties in linguistic research shows that there is still a lot of potential for pioneering research that can redefine the current limits of linguistic data collection and analysis. The different ways of obtaining data through XR by means of mediation, affixation, impact and compatibility further demonstrate the breadth of possibilities available.

Future visions for XR environments aim to shape them into viable spaces in which to study, work and engage in social activities, spaces where a naturalistic social experience would allow the user to embody a digital representation and enter a virtual university campus, attend a business meeting or share immersive social activities with friends and family. Considerable challenges still stand in the way of achieving these visions. Among them are challenges that are linguistic in nature and can benefit from what linguistic research can offer to solve them. The future of extended reality linguistics promises fast-paced advancements moving in tandem with the current rapid speed at which XR technologies, along with other emerging technologies such as artificial intelligence and robotics, are evolving. What we learn from XR linguistics can, in turn, inform the development of XR features related to language and virtual communication.

ACKNOWLEDGEMENT

The author thanks the two anonymous reviewers for their constructive comments and suggestions on the version of this manuscript submitted for review in July 2022.

AUTHOR AFFILIATIONS

Abdulrahman A. A. Alsayed orcid.org/0009-0003-7233-4982
University of St Andrews, United Kingdom; The Alan Turing Institute for AI and Data Science, United Kingdom; King Faisal University, Saudi Arabia

REFERENCES

Alissandrakis, Aris, Nico Reski, Mikko Laitinen, Jukka Tyrkkö, Jonas Lundberg and Magnus Levin. "Visualizing Rich Corpus Data Using Virtual Reality." *Studies in Variation, Contacts and Change in English*, vol. 20, 2019, https://varieng.helsinki.fi/series/volumes/20/alissandrakis_et_al/.

Chaffey, Patricia, Ron Artstein, Kallirroi Georgila, Kimberly Pollard, Setareh Nasihati Gilani, David Krum, David Nelson, Kevin Huynh, Alesia Gainer, Seyed Hossein Alavi, Rhys Yahata, Anton Leuski, Volodymyr Yanov and David Traum. "Human Swarm Interaction Using Plays, Audibles, and a Virtual Spokesperson." *SPIE Proceedings Volume 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, 2020. DOI: https://doi.org/10.1117/12.2557573

Dipper, Stefanie, Michael Götze and Stavros Skopeteas. "Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure." *Interdisciplinary Studies on Information Structure*, vol. 7, 2007, pp. 204–10.

Ebert, Dylan, Sanika Gupta and Fillia Makedon. "Ogma – A Virtual Reality Language Acquisition System." *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments*. New York, Association for Computing Machinery, 2016, pp. 1–5. DOI: https://doi.org/10.1145/2910674.2910681

Himmelmann, Nikolaus P. "Documentary and Descriptive Linguistics." *Linguistics*, vol. 36, 1998, pp. 61–195. DOI: https://doi.org/10.1515/ling.1998.36.1.161

Hutson, James, Joseph Weber and Angela Russo. "Digital Twins and Cultural Heritage Preservation: A Case Study of Best Practices and Reproducibility in Chiesa dei SS Apostoli e Biagio." *Art and Design Review*, vol. 11, 2023, pp. 15–41. DOI: https://doi.org/10.4236/adr.2023.111003

Kelly, Phineas. *Northern Arapaho and Rapa Nui Revitalization and Decolonization*. Unpublished dissertation, University of Wyoming, 2022.

Krause, Thomas, and Amir Zeldes. "ANNIS3: A New Architecture for Generic Corpus Query and Visualization." *Digital Scholarship in the Humanities*, vol. 31, 2014, pp. 118–39. DOI: https://doi.org/10.1093/llc/fqu057

Legault, Jennifer, Jiayan Zhao, Ying-An Chi, Weitao Chen, Alexander Klippel and Ping Li. "Immersive Virtual Reality as an Effective Tool for Second Language Vocabulary Learning." *Languages*, vol. 4, 2019. DOI: https://doi.org/10.3390/languages4010013

Lüdeling, Anke, Julia Ritz, Manfred Stede and Amir Zeldes. "Corpus Linguistics and Information Structure Research." *The Oxford Handbook of Information Structure*, edited by Caroline Féry and Shinichiro Ishihara. Oxford, Oxford University Press, 2016, pp. 599–620. DOI: https://doi.org/10.1093/oxfordhb/9780199642670.013.013

Milgram, Paul, and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays." *IEICE Transactions on Information Systems*, vol. 12, 1994, pp. 1321–9.

Nölle, Jonas, and Michael Spranger. "From the Field into the Lab: Causal Approaches to the Evolution of Spatial Language." *Linguistics Vanguard*, vol. 8, 2022, pp. 191–203. DOI: https://doi.org/10.1515/lingvan-2020-0007

Ochs, Magalie, Philippe Blache, Grégoire Moreau de Montcheuil, Jean-Marie Pergandi, Jorane Saubesty, Daniel Francon and Daniel Mestre. "A Semi-Autonomous System for Creating a Human–Machine Interaction Corpus in Virtual Reality: Application to the ACORFORMed System for Training Doctors to Break Bad News." *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, edited by Nicoletta Calzolari et al. Miyazaki, European Language Resources Association, 2018, https://aclanthology.org/L18-1467. DOI: https://doi.org/10.1111/bjop.12290

Pan, Xueni, and Antonia F. de C. Hamilton. "Why and How to Use Virtual Reality to Study Human Social Interaction: The Challenges of Exploring a New Research Landscape." *British Journal of Psychology*, vol. 109, 2018, pp. 395–417.

Parmaxi, Antigoni, and Alan A. Demetriou. "Augmented Reality in Language Learning: A State-of-the-Art Review of 2014–2019." *Journal of Computer Assisted Learning*, vol. 36, 2020, pp. 861–75. DOI: https://doi.org/10.1111/jcal.12486

Peeters, David. "A Standardized Set of 3-D Objects for Virtual Reality Research and Applications." *Behavior Research Methods*, vol. 50, 2018, pp. 1047–54. DOI: https://doi.org/10.3758/s13428-017-0925-3

Peeters, David. "Virtual Reality: A Game-Changing Method for the Language Sciences." *Psychonomic Bulletin and Review*, vol. 26, 2019, pp. 894–900. DOI: https://doi.org/10.3758/s13423-019-01571-3

Pentangelo, Joseph. *360° Video and Language Documentation: Towards a Corpus of Kanien'kéha (Mohawk)*. Unpublished dissertation, The City University of New York, 2020.

Petukhov, Igor V., Andrey E. Glazyrin, Andrey V. Gorokhov, Luydmila A. Steshina and Ilya O. Tanryverdiev. "Being Present in a Real or Virtual World: A EEG Study." *International Journal of Medical Informatics*, vol. 136, 2020. DOI: https://doi.org/10.1016/j.ijmedinf.2019.103977

Pfeiffer, Thies. "Using Virtual Reality Technology in Linguistic Research." *IEEE Virtual Reality Workshops*, 2012, pp. 83–4, https://ieeexplore.ieee.org/document/6180893. DOI: https://doi.org/10.1109/VR.2012.6180893

Popic, Deian, Simona G. Pacozzi and Corinna S. Martarelli. "Database of Virtual Objects to be Used in Psychological Research." *Plos One*, vol. 15, 2020. DOI: https://doi.org/10.1371/journal.pone.0238041

Solak, Ekrem, and Gamze Erdem. "A Content Analysis of Virtual Reality Studies in Foreign Language Education." *Participatory Educational Research*, Special Issue, 2015–II, pp. 21–6. DOI: https://doi.org/10.17275/per.15.spi.2.3

Speicher, Maximilian, Brian D. Hall and Michael Nebeling. "What is Mixed Reality?" *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, edited by Stephen Brewster et al. New York, Association for Computing Machinery, 2019, pp. 1–15. DOI: https://doi.org/10.1145/3290605.3300767

Staum Casasanto, Laura, Kyle Jasmin and Daniel Casasanto. "Virtually Accommodating: Speech Rate Accommodation to a Virtual Interlocutor." *32nd Annual Meeting of the Cognitive Science Society*, edited by Stellan Ohlsson and Richard Catrambone. Austin, TX, 2010, pp. 127–32.

Sydorenko, Tetyana, John Hellermann, Steven L. Thorne and Vanessa Howe. "Mobile Augmented Reality and Language-Related Episodes." *TESOL Quarterly*, vol. 53, 2019, pp. 712–40. DOI: https://doi.org/10.1002/tesq.507

Taguchi, Naoko. "Application of Immersive Virtual Reality to Pragmatics Data Collection Methods: Insights from Interviews." *Calico Journal*, vol. 38, 2021, pp. 181–201. DOI: https://doi.org/10.1558/cj.41136

Tromp, Johanne, Mert Akbal, Leonardo Pohl, Stephan Krohn, Eva M. Quinque, Felix Klotzsche, Arno Villringer and Michael Gaebler. "OpenVirtualObjects: An Open Set of Standardized and Validated 3D Household Objects for Virtual Reality-Based Research, Diagnostics, and Therapy." *Frontiers in Virtual Reality*, vol. 1, 2020, https://www.frontiersin.org/articles/10.3389/frvir.2020.611091. DOI: https://doi.org/10.3389/frvir.2020.611091

Wilcox, Laurie M., Robert S. Allison, Samuel Elfassy and Cynthia Grelik. "Personal Space in Virtual Reality." *ACM Transactions on Applied Perception*, vol. 3, 2006, pp. 412–28. DOI: https://doi.org/10.1145/1190036.1190041

Zeldes, Amir. *Multilayer Corpus Studies*. Abingdon, Routledge, 2018. DOI: https://doi.org/10.4324/9781315112473

TO CITE THIS ARTICLE:
Alsayed, Abdulrahman A. A. 2023 Extended Reality Language Research: Data Sources, Taxonomy and the Documentation of Embodied Corpora. *Modern Languages Open*, 2023(1): 46 pp. 1–20. DOI: https://doi.org/10.3828/mlo.v0i0.441

Published: 18 December 2023