ORIGINAL RESEARCH

Linguistic modalities and the sources of linguistic utterances

Anna Drożdżowicz[1]

Abstract

As an object of philosophical study, language is typically considered as an abstract object rather than a lived phenomenon that comes with rich and varied phenomenology. And yet our modes of engaging with language are complex and many. The first goal of this paper is to illustrate this variety by looking at some of the linguistic modalities and forms of communication. The second goal is to suggest that at least in some specific philosophical debates, language and communication should be investigated in the context of the various linguistic modalities and forms of communication. This will be done by considering how attention to some of the linguistic modalities and forms of communication may affect philosophical debates concerning: the nature of words, language and linguistic understanding, as well as the relation between linguistic utterances and their sources.

Keywords Linguistic modalities · Linguistic communication · Words · Sources

1 Introduction

Our ways of engaging with language are rich and varied. I can exclaim 'Brilliant!' to express my enthusiasm while reading my friend's chapter. Or I can write the word down on a sticky note and attach it to the paper copy. If they are around, and if they are hearing, they can hear my enthusiasm as it is expressed, or they can discover it later when I give them my copy with the note. In the first case they will be in a position to immediately grasp my reaction as well as its source, in virtue of hearing me exclaim the word 'brilliant'. In the latter case, they will be in a position to grasp my reaction and its source, in virtue of reading the word written on a note and inferring that it comes

✉ Anna Drożdżowicz
anna.drozdzowicz@inn.no

1 Department of Law, Philosophy and International Studies, Inland Norway University of Applied Sciences, P.O. Box 400, 2418 Elverum, Norway

from me.[1] These examples, however, do not exhaust the various forms of linguistic communication that humans have at their disposal. It is not a surprising observation that words, and linguistic utterances more broadly, exist and affect us in different ways.

Words are a much-discussed topic in philosophy. Some of the philosophical debates on words proceed at a high level of abstraction. Among the key questions in the metaphysics of words are the questions of whether and how words exist,[2] as well as how they are individuated, in general and across the board (e.g. Gasparri, 2021; Irmak, 2019; Kaplan, 2011; Miller, 2020, 2021; Nefdt, 2019; Rey, 2008). Other philosophical discussions of words have a much more limited scope. For example, specific groups of words, such as indexicals, proper names, definite descriptions, generics, slurs and pejoratives, have received dedicated, extensive philosophical attention (e.g. Braun, 2017; Cappelen & Dever, 2019; Leslie, 2015; Sosa, 2018). Whatever words are in general and whatever they convey in particular, word tokens, i.e., words produced on a particular occasion, can be produced and recognised by means of different *linguistic modalities* and via different *forms of communication*. So understood, word tokens are expressions that can form more complex phrases and can be thus seen as building blocks of linguistic utterances. The word 'utterance' is often used to designate an uninterrupted chain of spoken language only. An inscription, i.e. a written string of words can be used as a representation of a given spoken utterance. For simplicity, I will here use the term 'linguistic utterance' more broadly to designate an uninterrupted chain (or stream) of e.g.: spoken, written, or signed language. My use of 'linguistic utterance' is meant to apply across linguistic modalities and forms of communication. When a specific case is described, I will use terms such as: 'spoken utterance', 'written utterance' (or 'inscription'). I will use the term 'signed linguistic utterance' for an uninterrupted chain of signed language.[3]

In philosophy, the issue of modality and/or the specific form of communication in which words and linguistic utterances are produced and comprehended is rarely considered.[4] As an object of philosophical study, language is typically considered as an abstract object rather than a lived phenomenon that comes with rich and varied phenomenology. I will use the term '*linguistic modality'* here in a specific, narrow sense

---

[1] The latter is possible because the spoken language English has a written form and because my friend is literate. Not all languages have a written form/tradition of writing and not all users of languages who do have written form are literate.

[2] What is usually understood as 'word' is typically the product of the Greek and Latin tradition of grammar and its application to many languages and language families is problematic. Linguists—outside of informal talk—tend to use more specific terms including 'lexeme', 'linguistic item', and 'phonological word'. I thank Yael Peled and Nick Allott for suggesting that I make this qualification explicit.

[3] I use the term "signed linguistic utterance" following the usage of "utterance" in signed linguistics (e.g. Nicodemus, 2009).

[4] A few exceptions exist and should be mentioned here: in recent discussions on the nature and epistemology of linguistic understanding the questions of whether understanding is perceptual or whether we hear meanings are debated (e.g. O'Callaghan, 2011; Brogaard, 2018). In philosophy of perception that focuses on sounds, there is some, albeit limited, work on speech sound perception (O'Callaghan, 2015). See also Goodman (2020) on reading and Van Woudenberg (2021) on the epistemology of reading and Begby (2017) on homesign communication. Despite these examples, philosophical debates about language and meaning tend to proceed largely in abstraction from linguistic modalities and forms of communication.

to mean spoken and signed (visual or tactile) linguistic communication.[5] By *form of communication* I will mean various ways in which spoken and signed languages can be deployed and communicated with the use of various senses (or *sensory* modalities) and aids, such as for example: writing and haptic speech for spoken English.[6] Linguistic utterances, including words, exist in different ways depending on the linguistic modality and form of communication in which they are produced and comprehended: they are differently materially sustained and result in different experience profiles in language users. This is a complex landscape, for example, an utterance can be: produced and heard in spoken English language (oral-aural linguistic modality) or written and read (a form of spoken English language), signed and seen as visible hand and face gestures, as in British or American Sign Languages (spatial-visual linguistic modality), or signed and felt as touch and haptic sensations, as in tactile Auslan (tactile-spatial linguistic modality).[7] Independently of one's preferred way of charting the landscape and preferred terminology, there is a rich variety of how linguistic utterances exist and are experienced that seems to be missing from many current philosophical debates on the nature of language and communication. At least in some of these areas, as I will try to show, this omission has implications for the philosophical questions under investigation.

The first goal of this paper is to fill that gap by bringing some of the key differences and similarities in linguistic modalities and forms of linguistic communication in which linguistic utterances are produced and comprehended to the fore. Despite the resulting complexity, this paper extends the scope of observations to those that concern not only spoken language (and its written format), but also signed languages, both visual and tactile. It also briefly mentions some of the other forms of cross-modal and aided communication that require further investigation. I will focus on explaining how spoken, written and signed linguistic utterances serve as carriers of linguistic meaning. When relevant, I will also mention prosody across linguistic modalities.

It is important to note that abstraction across different modalities and forms of communication may often be warranted to investigate some questions in philosophy (and linguistics). In some cases it may be beneficial to abstract away from complications that might get in the way of providing systematic generalisations. It is, however, an open and, to my knowledge little discussed, question which philosophical debates concerning language and communication are those where we can and should abstract away from linguistic modalities and forms of communication and which require that we attend to them. In this paper I propose that, at least in *some* philosophical debates, language should be investigated in the context of various linguistic modalities and forms of communication. The second goal of this paper is to illustrate why and how

---

[5] For a similar use of the term 'modality' see literature on signed languages, e.g. Emmorey (2011), Wilbur et al. (2003, Chap. 21), Brentari (2019, Chap. 2).

[6] It is also important to mention that writing systems of spoken languages may provide different affordances to Deaf-and-Hard-of-Hearing interlocutors. For a recent discussion see Snoddon (2022). I thank Yael Peled for suggesting this important qualification.

[7] The above suggested way of drawing the distinction between modality and form is meant to capture a very complex picture where different types of languages (spoken, signed) do not simply match with one of the sensory modalities involved in production and comprehension. However, there may be alternative ways of charting this territory.

this approach can be fruitful for philosophical investigation of language in selected specific cases. I will do so by briefly considering how investigating differences in linguistic modalities and forms of communication may affect philosophical debates concerning: the nature of words, language and linguistic understanding, as well as the relation between linguistic utterances and their sources.

Regarding the nature of words, I will propose that closer attention to linguistic modalities and forms of communication can inform the investigation of how words are individuated, and which properties are crucial for that to happen. Drawing on two examples from recent debates concerning the nature of language and linguistic understanding, I will illustrate how those can benefit from a more inclusive approach to language (Begby, 2017) and to various forms in which languages are produced and comprehended. Regarding the third area, spoken utterances are normally delivered in the medium of a voice of a particular speaker and often reveal important information about the speaker. I will discuss a difference in how salient a source of a spoken linguistic utterance in spoken communication typically is, as compared to the source of a written utterance. I will also propose that there are certain key similarities and differences between how the source and medium are experienced in spoken and signed communication, including visual and tactile signed communication.

This is by no means intended as an exhaustive list of debates where this approach may be fruitful. My choice of the three areas discussed in the second part of the paper is dictated by some principled considerations. First, focusing on theories of words is meant to illustrate that even some of the most abstract debates in the metaphysics of words can be informed by observations of rich and varied forms of human linguistic communication. Second, philosophical discussions of language and linguistic understanding are another area where much of this variety has been neglected. Given that language and linguistic understanding are real-world phenomena, the approach taken in this paper is to succinctly illustrate both how this omission can be successfully rectified (e.g. Begby, 2017), and where more work is still needed (debates concerning linguistic understanding). The third area concerns how words and linguistic utterances, more broadly, relate to their sources and the medium in which they are produced. This is a much less investigated domain, as philosophers tend to abstract away from these relations. However, linguistic utterances almost never come in an author-free format. Information about the author is often unintentionally transmitted in their linguistic production and received by the comprehender and will frequently matter for a communicative encounter between them.

To keep the investigation close to the realities of spoken, written and signed communication, my focus will be on linguistic utterances in a broad sense, including word tokens (spoken, written, signed) that form more complex phrases and can be combined at a sentence level. When relevant, I will indicate that discussion concerns linguistic units at a specific level, e.g. speech sounds, hand configurations, words, or utterances of complete sentences. The proposed investigation is both exploratory and multi-disciplinary—it draws on phenomenological observations about how languages are comprehended, and empirical research on spoken, written and signed linguistic communication, as well as theoretical debates in the philosophy of language. I start by presenting observations of how linguistic utterances exist and are experienced in spoken language and its written form (Sect. 2). Next, I consider signed languages (visual

and tactile) and briefly sketch a map of other forms of linguistic communication (Sect. 3). In Sect. 4 I outline how observations concerning similarities and differences between various linguistic modalities and forms of communication could and should influence at least some specific philosophical debates on language, meaning and communication.

2 Hearing vs reading (and hearing in reading)

Linguistic utterances can be produced and grasped by means of different forms of communication. On the comprehender's side, there is a clear intuitive difference between hearing a given linguistic utterance, e.g. 'Brilliant!', and reading it. In order to capture where the difference comes from, I will here consider how spoken and written linguistic utterances are produced, processed and experienced by language users.[8]

I start with a brief description of the production and processing of speech. Among hearing people, *spoken* linguistic communication is among the most basic forms of linguistic communication available to humans. Hearing children will typically acquire the ability to produce and understand linguistic utterances thanks to average exposure to spoken language (Cutler, 2012). This would normally happen whether or not they can later, thanks to instruction and schooling, acquire the ability to write and read. After all, historically most speakers of most languages have not been literate. Even in literate cultures, children can have very good knowledge of their own language before they can read or write. Presumably, speaking starts with initiating the production of sounds that match a specific thought or idea and the preparation of a specific linguistic form to express those. This leads to articulation (Levelt, 1999). Human speech is produced as a series of events that span in time and involve the use of vocal tract. Acoustic waves are produced in speaking with the air that comes from lungs that is transformed into sounds at the larynx and shaped by articulators in the mouth (Tatham & Morton, 2006). Speech is particularly complex, when compared to other auditory stimuli (Nygaard & Pisoni, 1995). The comprehension of spoken language is possible in part thanks to our ability to perceive speech sounds. Individual words and more complex phrases that make up full utterances can be recognised thanks to the capacity for speech sound perception. The perception of speech sounds takes place at the earliest stages of speech processing and enables the mapping of the time-varying acoustic signal produced by a speaker into a set of discrete linguistic representations. These linguistic representations are typically construed in terms of sequences of phonetic segments, i.e. consonants and vowels. Words of a given language, as produced and recognised in spoken linguistic comprehension, are commonly assumed to be formed by such phonetic segments (Tatham & Morton, 2006). For example, we can describe the word *cat* as being composed of three phonetic segments: an initial consonant (in phonetic notation, symbolized as/k/), a medial vowel (/æ/) and a final consonant (/t/). As simple as this may seem, the mapping between the acoustic signal produced by speakers and the phonemic structure of an utterance that underlies and enables

---

[8] Readers interested in a more detailed and comprehensive comparison between speech and reading based on current psychological studies will benefit from De Gelder and Morais (2017).

the recognition of an individual word is highly complex (for a useful overview see O'Callaghan 2015). This makes the perception of speech sounds both a ubiquitous and unique auditory achievement. Spoken linguistic utterances are processed thanks to their phonetic properties that correspond to and track complex sets of acoustic properties of the speech sounds uttered. Language users typically become experts in the exercise of approximate matching of the articulated acoustic properties with a phonological structure that underlies, on average, seamless and automatic recognition of strings of words and linguistic utterances.[9]

What about reading of a written form of a given spoken language? In this section I focus only on one type of reading, i.e. reading in individuals who can see. People with competency in that language who are visually impaired, blind, or deafblind can read by using tactile writing systems such as Braille (e.g. Bertelson, 2017) (see Sect. 3). Reading can be described as a learned behaviour: it involves looking at a written text and moving one's eyes along the lines of written words (Traxler, 2012). Reading is an optional behaviour, but competent reading is a mandatory process: when you see a series of letters that make up a word in a familiar writing system, you cannot help but read it (Harley, 2013).[10] In this manner, reading is quite similar to speech sound perception, where the segmentation process of the acoustic speech signal into specific phonetic units is mandatory and happens as soon as speech is heard, at least for listeners who are competent in a given language (O'Callaghan, 2015). Reading relies on trained but mandatory visual processing of letters and, consecutively, on *word recognition* (Traxler, 2012). This is possible thanks to saccadic eye movements, where eyes are travelling through the lines of text in jumps of about 20–60 ms and with still period intervals of about 200–250 ms. During these still periods eyes are fixating predominantly on the material that enters the most sensitive part of the visual field, as well as on some of the material that is at the close periphery (Harley, 2013, p. 168). Approximately 10% of eye movements in reading are regressive eye movements that perform a corrective function in cases of decreased comprehension. A widely accepted view on how reading is controlled, supported by many experimental results, is that processing of linguistic properties of *words* (e.g. their frequency, familiarity) regularly influences when and where the eyes should be moving (Harley, 2013, p. 171). Cognitive control theories of reading postulate that the aspects of linguistic processing of specific words and the planning of eye movements take place simultaneously.[11] There is abundant research on what makes word recognition easier or harder, including word frequency and familiarity, as well as different forms of priming (Harley, 2013). While many of these effects are beyond the reader's conscious control, word recognition involves both automatic and mandatory processes that one cannot control, as well as attentional processes that depend on one's expectations and allow one to consciously choose to go back and reread a word.

---

[9] This passage draws on material from Sect. 3 in Drożdżowicz 2021.

[10] The mandatory nature of the reading process is supported by results from the Stroop task (Stroop, 1935).

[11] They often disagree whether these processes operate *sequentially* word after word (e.g. Pollatsek et al., 2006) or *in parallel* for more than one word at a time (e.g. Kliegl et al., 2006). The latter type of views often rely on evidence suggesting that linguistic properties of words in parafovea can affect the control of eye movements.

What is then the difference between hearing and reading a language? One clear difference is that unlike written linguistic utterances, spoken utterances do not come in the pre-segmented format of words and sentences. However, this should be qualified by mentioning prosody[12]—a feature of spoken language that covers various suprasegmental phonetic phenomena, i.e., properties that belong to larger units than phonemes, including syllables, phones, words, various intonation phrases and utterances (Speer & Blodgett, 2006). Prosodic contributions to linguistic communication range from the intentionally produced, properly linguistic, and often language-specific ones (e.g., lexical tone, stress or pitch accent) to spontaneous, involuntary, or 'natural' ones (e.g. an angry, agitated or enthusiastic tone of voice) (Wharton, 2009). A written analog of prosody is quite simple (e.g. punctuation, emojis), whereas speech accompanied by prosody is a rich source of information about the speaker's attitude towards the message and the intended contribution it is supposed to make.

Another difference between hearing and reading language concerns the availability of linguistic input in reading and hearing for the comprehender, resulting in different control mechanisms. In reading, both automatic regressive saccades and consciously controlled attentional processes can steer and repair the reading process as it unfolds. They are the hallmarks of reading and constitute a clear difference with speech sound perception. While written inscriptions are normally available for as long as one needs it, typical spoken linguistic interactions rely on a speech signal available only for a short period of time and do not allow for its rewinding (Harley, 2013, pp. 167–168). Consequently, speech sound perception and the recognition of spoken linguistic utterances involve different mechanisms that can compensate for interference and mistakes, such as phoneme restoration (Shahin et al., 2009; Warren et al., 1972). Hearing and reading places different perceptual requirements on language users and relies on different control and repair mechanisms.

Finally, a clear difference between reading and hearing to consider is this. Reading in syllabic/alphabetic systems involves visual perception of strings of letters that make up individual words (e.g. Traxler, 2012), while comprehension of spoken language necessarily involves auditory perception of the acoustic speech signal that corresponds to phonetic segments which make up words (O'Callaghan, 2015). This intuitive observation suggests a *clear divide*: perceptual processes and experiences involved in spoken language comprehension are auditory, whereas perceptual processes involved in comprehension of written language are visual, resulting in correspondingly different experiences of spoken and written linguistic utterances. But the suggested divide between the auditory and the visual is actually less clear than it might seem, in particular in the case of reading, leading to the following *asymmetry* between the comprehension of spoken and written linguistic utterances.[13] Many

---

[12] I thank Yael Peled for useful comments on this topic.

[13] The divide between auditory and visual is less clear also in the case of spoken language, given that speech sound processing often relies on multimodal perception. Many experiences that accompany spoken linguistic comprehension are multisensory and involve both auditory and visual components. Nevertheless, hearing interlocutors can often perceive speech sounds and comprehend spoken linguistic utterances without seeing the face of the speaker, as cases of speaking on the phone, listening to radio or audiobooks convincingly show.

forms of reading are often accompanied by phenomena that seem to belong to the auditory domain. Reading *aloud* relies on retrieving the representations of speech sounds that make up individual words (Harley, 2013). This form of reading might be quite special, because it is crucially connected to the planning of speech production and speech production relies on auditory representations of speech sounds. *Silent* reading, on the other hand, is geared primarily towards meaning comprehension and does not require speech production. Empirical research suggests that sound (or phonological) representations are important to silent reading as well, at least to some extent (Harley, 2013; Stanovich et al., 1997; Ashby & Martin, 2008; Rayner et al., 2012; Fodor et al., 2017). The source of this phenomenon is most likely the nature of scripts used in written language. In *alphabetic* writing systems, such as the alphabetic writing system of English, individual letters or their small groups correspond to individual speech sounds or phonemes (Traxler, 2012, p. 385). In *logographic* writing systems, such as Chinese, symbols correspond to the units of meaning (morphemes or words) but also typically encode information about the way the character should be pronounced (Lee et al., 2007). Despite these apparent differences in the correspondence between written scripts and phonology of different languages, competent language users of both types of writing systems routinely rely on phonological representations when reading (Hsu et al., 2009; Perfetti et al., 2005).

Sound representations of phonetic properties of words are involved in the process of silent reading. According to the so called dual-route and dual-route cascaded models of reading there are two separate ways in which one can use visual input (strings of letters, words) to access the lexical meanings of words in the mental lexicon (Coltheart, 2005; Coltheart et al., 2013). A reader can access a word's lexical entry by 'sounding out' the word. A reader can also access lexicon entries for many words directly on the basis of the visual input and without first activating phonological codes where letters correspond to phonemes (Traxler, 2012). Notably, the exact nature of the phonological mediation involved in reading is a controversial matter. A careful way of phrasing this claim could be that the recognition of words is influenced by their phonology.[14] In silent reading, the phonological representations of words can be activated. Whether phonological representations take part in *mediating* access to lexical information while reading, they are typically activated as one of the *attributes* of that word. The activation of sound representations influences the typical phenomenology that accompanies silent reading. It results in the experience of a particular type of inner speech that readers routinely have during silent reading (Rayner et al., 2012; Geva et al., 2011; Harley, 2013; see also Magrassi et al., 2015).[15] It has been argued that inner speech accompanies comprehension of written texts and may be beneficial for understanding. The standard explanation for this effect is that inner speech

---

[14] This is compatible with the competing single-route models of reading, according to which every word is represented in the mental lexicon as a distributed pattern of activation across both orthographic and phonological units (McClelland & Rumelhart, 1981). Reading involves activation of both types of information, albeit to a varying degree.

[15] I restrict myself here to observations concerning inner speech in silent reading. Other forms of inner speech (e.g. in thinking, problem solving) may be related but fall outside the scope of this paper. See e.g. Perrone-Bertolotti et al., (2014), Alderson-Day and Fernyhough (2015), Alderson-Day et al., 2020; Geva (2018).

helps to organise sequences of read material and maintain them in working memory (Rayner et al., 2012). Another plausible role for inner speech in silent reading is to provide missing information about the plausible prosodic structure of the written text (Slowiaczek & Clifton, 1980). Inner speech is sometimes also argued to be an epiphe-nomenon of the way readers are taught to read, i.e. by sounding out aloud phonemes and words (Rayner et al., 2012). The exact scope and nature of inner speech involved in silent reading, as many aspects of reading, are debated.

To sum up, spoken and written linguistic utterances exist in two very different ways: spoken utterances can be characterised as events of vocal production, while written utterances are various types of inscriptions. Beyond the clear difference in how linguistic utterances are produced and materially sustained and the choice point of whether to pronounce or write them, there are some clear differences between how language users engage with spoken and written forms of communication. First, spoken utterances are fast fleeting phenomena available for comprehension only for a limited time, whereas inscriptions are typically materially sustained for much longer and can be revisited, which leads to different control and repair mechanisms. Second, processing of spoken and written forms of a language is to a large extent enabled by input from two very different sensory modalities (audition or vision) and depends on different psychological capacities. A closer look at how linguistic utterances are comprehended by language users reveals an *asymmetry* between the word properties that are stored in the mental lexicon and utilised in reading and hearing. The role of sound/phonological and orthographic properties of words and of their respective representations formed by language users is different. The recognition of *both* spoken and written words relies to a varying degree on their sound and phonological properties. Because of that, the phenomenology of reading occasionally involves a particular type of auditory impressions, such as inner speech. These observations pave the way for a more nuanced picture on the exact differences between hearing and reading a language.

3 Signed languages

Among various linguistic modalities, language by ear has traditionally received a lot of attention. Spoken linguistic communication with its written form is by far the dominant form, but it is not the only one. When considering language in abstract, philosophers tend to forget about other linguistic modalities, such as, for example, language articulated and processed by hand and eye. Scientists investigating language and communication have already realised this omission (e.g. Bauman & Murray, 2017; Emmorey, 2011; Sandler & Lillo-Martin, 2001). In this section I briefly introduce both visual and tactile signed languages. I start with a brief historical overview based on a useful chapter by Bauman and Murray (2017). The philosophical payoffs of showing interest in various linguistic modalities will be explained in Sect. 4.

There are more than 300 signed languages around the world used by the overall population of estimated 70 million Deaf people, with an overwhelming majority of them living in developing countries (UN). For a long time, it was wrongly believed that speech is the only modality of natural human language. Although signed languages were even once considered to be a universal form of linguistic communication during

the French enlightenment (Bauman & Murray, 2017; Rosenfeld, 2001), their official status changed dramatically in the nineteenth century (Baynton, 1996). At that time, deafness became a medicalized category and gestural forms of communication became associated with what was considered savage and primitive forms of communication among native people in colonized areas around the world (Baynton, 1996). Moreover, the rise of nation states made a common language part of national identity and led to a gradual suppression of linguistic minorities more broadly, signed languages included. In the twentieth century, the publication of Stokoe's *Sign Languages Structure* (1978), as well as Padden's article "The Deaf Community and the Culture of Deaf People" (1980), mark the advent of establishing signed languages as natural languages on par with spoken ones and of research into various Deaf cultures.

There are some obvious differences between spoken and visual signed languages: the former make use of the auditory modality, the latter make use of visual channels, resulting in very different sensory experiences that accompany the two types of language. More on this soon. There are, however, also substantial similarities between them (Emmorey, 2011): Unlike written and read forms of communication, communication in spoken and visual signed languages relies on dynamic and time-varying signals rather than static symbols (703). Unlike written and read forms of communication, they do not come in the pre-segmented format of words and units (703) but do rely on prosody (Traxler, 2012, p. 455). Finally, unlike writing, spoken and signed languages are acquired in infancy: they do require, however, substantial exposure (456). Neuro-psychological research provides evidence that the human brain can support the use of signed and spoken languages with a similar ease (Hauser & Kartheiser, 2014; Petitto et al., 2000). There is also evidence suggesting that infants show preference for linguistic signals, independently of whether it comes in auditory or visual modality (Krentz & Corina, 2008). Current research in linguistics and psychology of language provides evidence that signed languages involve all fundamental properties of natural languages that spoken languages also have: phonology, morphology, grammar and syntax (Traxler, 2012, p. 447).

How are signed languages produced and comprehended by competent users? Signers quickly and typically effortlessly extract complex messages from the signs produced in the incoming visual signal, in a manner similar to how competent users of spoken languages extract meanings from a time-varying acoustic signal produced by speakers. In both cases this is possible thanks to stored internal representations (Emmorey, 2011, pp. 703–706): phonemic units in speech and sublexical linguistic units that make parsing of visual signed input possible. In signed languages manual and facial gestures are meaningless units that are combined to make distinct signs that carry meaning. Signs can be broken down into four components (Stokoe, 2005; Traxler, 2012): hand shape (or hand configuration); location (place in space where the sign is articulated); movement and orientation of the hand/arm (often represented as part of hand configuration). Linguists agree that signed languages have phonology (without sounds) given that meaningful elements like signs have a structure created from the above listed meaningless units combined in rule-governed ways. Hand shape, location and movement are phonological features of signed languages because they give rise to minimal pairs (Traxler, 2012), i.e. the meanings of two different signs can be differentiated on the basis of these features (just as meanings of two spoken words

can be differentiated on the basis of difference in phonemes that are themselves devoid in meaning, e.g. /p/ and /b/ as in "pat" and "bat"). Several studies show that comprehension of signed languages relies on categorical perception (in a manner similar to comprehension of spoken languages): i.e. small variations in form of hand shape, location and movement do not lead to differences in how a sign is categorized (Baker et al., 2005; Emmorey et al., 2003). Although categorical perception effects are weaker in signed languages than in speech, their presence suggests that categorical perception is a feature of language processing that is independent of linguistic modality.

In signed languages, linguistic expressions analogous to spoken words are produced and materially sustained as manual and facial articulations that are combined into rule-governed structures. Signs in signed languages also have morphology, just as spoken words do, with hand shape and movement as the main types of components used to mark morphological features (Traxler, 2012). Those features are used to produce and comprehend more complex gestures. Signed languages also have grammar and syntax which govern how signs are combined into full sentences, as well as an analogue of prosody. Both manual gestures and facial expressions are used to convey grammatical information about how an action was carried out, to signify linguistic prosody: whether an utterance is a statement or question, as well as emotional prosody (e.g. Emmorey et al., 2009). All this information is produced and comprehended via the visual modality, and results in a specific type of linguistic visual experience. Interestingly, there are reasons to think that the resulting linguistic experiences of hearing a spoken utterance and seeing a signed utterance differ not only in terms of the modality in which the linguistic signal is experienced but also in terms of their internal organisation. While spoken linguistic utterances unfold sequentially and are experienced word by word, signed linguistic production relies on visual phonological information that is made available early and simultaneously for a perceiver, shares less initial phonological shapes across signs (i.e. sequences of produced signs are initially less ambiguous) and initial shapes are more constrained by the phonotactic structure (Emmorey, 2011, p. 708). As a result, the number of word candidates to be activated in the lexicon is quickly narrowed down in signed languages—signs can be recognized faster than spoken words (708).

So far, I have considered only visual signed languages, but the list of modalities used for linguistic communication is longer than that. The term 'tactile signed language' is commonly used to refer to the form of signing used by deafblind people. Specific tactile signed languages often rely on adaptations of visual sign languages for perception through touch and haptic sensations, as e.g. in tactile Auslan (Willoughby et al., 2020).[16] Tactile signed languages are used primarily by deafblind people and their families to enable communication via tactile form of sign languages and haptic sensations (Iwasaki et al., 2019). Deafblind signers have no access to traditional non-manual features, such as eye-gaze, eyebrow and facial expressions. Perceiving distinctions between phonologically similar signs may also be difficult. As a result,

---

[16] As observed by Willoughby et al. (2018) depending on the individual deafblind signer and the signing community in which they are embedded, such adaptions can take various forms and result in minor changes in the existing visual sign language, lead to the emergence of a new language specifically designed for tactile production and reception, or something in between. Consequently, the term 'tactile signed languages' covers a wide variety of cases.

tactile signers must rely on haptic resources to construct new conventions for encoding or inferring information (Iwasaki et al., 2019). Deafblind signers can for example rely on motion, tenseness, and repetition to express adverbial information (Collins 2004). Tactile signed communication has been argued to involve a possibly unique and complex structure that is based on direction, speed and acceleration of movements, pressure, and body position (Dammeyer et al., 2015). Tactile signed communication differs substantially from visual signed communication in several respects: the former more often relies on two-hand symmetrical signing that allows for involving two comprehenders; the sites of articulation in tactile sign communication are different body parts of the addressee, rather than of the signer—both bodies of the signer and addressee are thus used to materially sustain the coproduction of phonological components (Dirksen, Bauman & Murray, 2017). I return to this observation in Sect. 4.3.

The material presented in Sects. 2 and 3 by no means provides an exhaustive list of different forms of linguistic communication available to humans. For example, speech and written language can be perceived via different sensory modalities and with the use of various aids. In the context of reading, we should mention Braille, a tactile writing system used by people who are visually impaired, blind, and deafblind (Bertelson, 2017; Daniels & Bright, 1996). Another cross-modal form of linguistic communication to be mentioned here is haptic speech perception as, for example in Tadoma. Tadoma is a form of communication in which a deafblind person receives speech by placing a hand on the talker's face and monitoring actions associated with speech production. In this case communication is transmitted via a tactile modality through vibrations, motions of the jaw, and facial expressions of the speaker (Reed et al., 1985). A different form of haptic speech makes use of tactile vocoders that filter an acoustic waveform and transduce it into vibratory patterns that are felt on the skin. Finally, assistive communication technologies encompass various devices used in order to enable a person with hearing loss or with a voice, speech, or language disorder to access communication with other interlocutors. Among these, assistive listening devices are used to help amplify the sounds, especially in noisy environments and are often used together with a hearing aid or cochlear implant. Augmentative and alternative communication devices help people with communication disorders to communicate and can take the form of a simple picture board or a computer program that synthesizes speech from text. A detailed discussion of these various forms of communication, many of which rely on processing across different sensory modalities, is called for, as it may improve our understanding of the nature of language and communication. Given limited space, in what follows I focus on linguistic modalities in spoken and signed languages and leave discussion of these other forms for another occasion.

4 Taking stock—the import of linguistic modalities

Can attending to linguistic modalities and forms of communication in which linguistic utterances are produced and comprehended, as described above, bear any fruits for philosophical investigations of language, meaning and communication? I believe it

can. In this section I briefly sketch three areas, where this approach can be useful for investigating specific philosophical questions.

4.1 How words are individuated

Recent debates concerning the metaphysics of words are the first domain that I will look into. Among the key questions debated in this area are: whether and how words exist as well as how they are individuated (e.g. Kaplan, 2011; Miller, 2020; Nefdt, 2019; Rey, 2008). According to the bundle theory of words, recently developed and defended by Miller (2021), word tokens are best conceived as bundles of various properties, such as semantic, phonetic, orthographic and grammatical, whereas word types are bundles (or sets, or collections) of word tokens so conceived. The question of word individuation can thus be rephrased as a question of whether there are any particular properties (or their set) that are necessary for type-membership. According to Miller (2021), a promising answer is that type-membership depends on various linguistic and sociological factors that are relevant for speakers within a particular community.

It is an interesting question to consider which, if any, properties may be crucial for word individuation and how factors pertaining to the nature of linguistic communities determine those. This is where attending to different linguistic modalities comes into the picture. In spoken languages, as explained, words can be produced and comprehended in large part thanks to interlocutors' capacity to articulate and perceive speech sounds. For spoken languages, phonetic properties of words retrieved from acoustic signals produced by the speaker are important, if not necessary, for words' type-membership. However, as explained in Sect. 2, sound (or phonological) properties are also important for word type-membership in written communication, in addition to orthographic properties, given that information about both types of word properties is stored in the mental lexicon and utilized in spoken and written communication (Harley, 2013). Although sound properties can be shown to have a particular importance in a word bundle for users of spoken and written languages, the scope of this conclusion is limited. Words exist beyond spoken and written communication forms. In visual signed languages words are expressed by means of manual and facial articulations and comprehended by language users thanks to their ability to parse visual signals produced by signers into a stream of meaningful linguistic expressions. Deaf users of sign languages do not use phonetic or sound-related properties to individuate words, but they do rely on meaningless phonological properties determined by sublexical linguistic units produced in the form of manual and facial gestures. Furthermore, deafblind signers utilize signs that are composed of gestures produced on the body of the comprehender. Manual, facial and tactical properties are what determines word type-membership in visual and tactile sign languages. The list of properties that are crucial for word-type individuation across different types of language is thus substantially broader than what was initially suggested.

To the best of my knowledge, the proponents of the bundle theory and other theories of words tend to keep their investigation in abstraction from linguistic modalities, which is why they do not discuss signed languages. Attending to linguistic modalities can however inform their accounts, as illustrated with the bundle theory, where one

can argue that a more inclusive list of properties is called for.[17] More generally: one might expect that philosophical views on the nature and individuation of words should be compatible with basic facts about what it is for language users to be competent users of words in their language. Some of these facts, as illustrated here, cannot be established in abstraction from the realities of different linguistic modalities and forms of communication. But this leaves us with the following dilemma: On the one hand, not allowing for a more inclusive approach that goes beyond spoken languages would leave the bundle theory rather limited. On the other hand, expanding the list of properties to track the realities of linguistic communication across different modalities and forms may lead to a disjunctive characterization,[18] i.e., the bundled properties are either of this sort … (in the case of spoken language) or that sort … (in the case of signed language). This may make us question whether a more abstract characterization of these properties can be provided.[19] It is not clear in which direction this debate should proceed, given what we can learn about different properties involved in word individuation across modalities and forms of communicating.

4.2 Language and linguistic understanding

The second domain of philosophical research where attention to different linguistic modalities and forms of communication can be fruitful are debates concerning the nature of language and linguistic understanding. To illustrate that, I consider two examples from the philosophical literature. One recent example comes from philosophical discussions concerning the nature of language and meaning. Endre Begby (2017) has argued that the study of homesign communication provides evidence that puts pressure on several philosophical views on language where emphasis is placed on public languages that are static, structurally well-defined objects handed over by generations. According to Begby, homesign communication, where idiosyncratic gesture systems are devised spontaneously by deaf children who communicate with their hearing parents, shows that semantic properties are not governed by public norms

---

[17] Other theories according to which words are individuated by their causal-historical ancestry propose that how words are coined, i.e. their origination properties, are of key importance for individuation (Gasparri, 2021). Looking at how meaningful linguistic units are coined in spoken and signed languages can also be revealing for those accounts. For example, some signed languages start as homesign languages, a classic example being Nicaraguan Sign Language (Emmorey, 2001), which gave researchers a unique opportunity to observe the spontaneous formation of a new language developed by deaf children brought together to a school, from its early to later stages where complex grammatical features were introduced. Origination properties of expressions and their coinage can thus be closely studied in such spontaneously formed signed languages.

[18] I thank Endre Begby for drawing my attention to this problem and suggestions presented in this passage.

[19] An important observation that sets some other limitations on this debate/approach is that the dominant notion of word was historically developed on the basis of synthetic European languages, making its applicability a contested issue in the case of some agglutinative polysynthetic languages, e.g. Kanien'kéha (Mohawk), which do not straightforwardly fit the neat distinction between "words" and "sentences". This may put even more pressure on the project of arriving at a theory/account of what words are (across the board). I thank Yael Peled for drawing my attention to this fact. It should be also noted that when differences and details matter, linguists tend to use technical terms such as e.g. morpheme or lexeme. None of these seem to match perfectly with the folk notion of a word that some philosophers want to characterise. I thank Nick Allott for comments on this issue.

that determine the use of words in a linguistic community (pace the Wittgensteinian tradition), given that in such languages meanings are spontaneously devised from the ground up (p. 697). Moreover, Begby argues that homesign languages challenge the Gricean divide between semantics and pragmatics and call for a more nuanced picture, given that pre-established conventional codes are not necessary for homesign communication.[20] Rather, they draw our attention to the primacy of speaker meaning and pragmatics (pp. 698–702). Finally, homesign seems to put pressure on the Peircean tradition where capacity for symbolic representation is dependent on the acquisition of a public language. As Begby points out, homesign languages exhibit both arbitrariness and iconicity and there are no reasons to suppose that homesign users are incapable of symbolic thinking (705–707). Begby's case study of homesign languages shows that taking a close look at different forms of linguistic communication can help us correct some theoretical misconceptions about the nature of language and meaning.

Another example to consider here are recent philosophical discussions concerning linguistic understanding. One strand of research in this area concerns the nature of states of linguistic understanding that language users enter in communication, such that they could serve as a basis or justification for beliefs (or knowledge) about what was communicated. The dominant focus in these discussions has been on speech that is used as a toy example for investigation of such states. However, it seems natural to assume that states of linguistic understanding arise in different linguistic modalities: it seems thus warranted to aim for a general account. The discussion of whether and in what sense such states are perceptual-like focuses on arguments from phenomenology and psychology of spoken and occasionally written communication (e.g. Bayne, 2009; O'Callaghan, 2011; Nes, 2016; Brogaard, 2018).[21] Supposing that such arguments can tell us something about states of linguistic understanding more generally, minimally, one can argue that philosophical views on the nature and functional significance of such states, when appealing to such phenomenological and psychological arguments and considerations, should be able to encompass linguistic plurality across modalities and forms of communication, which is why expanding the usual domain of cases may be required. There are no principled reasons to assume that observations concerning psychology and phenomenology of spoken and written communication should have priority over other linguistic modalities to inform our philosophical accounts of linguistic understanding.

An inclusive approach that sees language as a varied phenomenon naturally arising within (and across) different linguistic modalities and forms of communication is becoming part and parcel of modern language research in psychology and linguistics (e.g. Emmorey, 2011), but is still exceptionally rare in the philosophical study of language and communication (cf. Begby, 2017). In some cases, preconceptions concerning linguistic modalities and forms of communication may direct the focus and influence what counts as the object of philosophical study.

---

[20] One may resist some of Begby's points here by observing that at some point homesign users do establish some conventional codes and thus homesign communication can be seen as compatible with the views on language that emphasize the importance of conventional codes.

[21] On the other hand, it seems that some of the recent philosophical discussions on states of linguistic understanding have proceeded largely in abstraction from issues of the sensory modality to which linguistic utterances are presented (e.g. Fricker, 2003; Reiland, 2015; Longworth, 2018).

4.3 The sources and medium of linguistic utterances

Except for some occasions, linguistic utterances are typically intentionally produced by a person.[22] Information about the author of a linguistic utterance is often unintentionally transmitted or revealed in their linguistic production to the comprehender and will frequently influence the communicative encounter between them. Who is the author of a linguistic utterance (e.g. a known individual, a stranger giving an impression of being (or not being) trustworthy, competent, foreign) often matters for how the message carried by the utterance will be received. Which information about the source is conveyed, how linguistic utterances are related to their sources and how this relation is experienced by comprehenders are thus important questions to consider. The relation between a linguistic utterance, the medium in which it is produced and who produces it is another area where a closer look at different linguistic modalities can be instructive. I propose that there are some key differences in how spoken, written and signed linguistic utterances are related to their sources and the medium in which they are produced. Arguably, those are reflected in the information transmitted, as well as in experiences on the comprehender's side. One difference concerns how salient a source of a spoken and signed linguistic utterance is, as compared to the source of a written utterance. Another difference concerns experiencing spoken and signed linguistic utterances as intrinsically intentionally produced by the source. Finally, it will be suggested that the source and medium in tactile signed communication relies on the body of the signer and addressee allowing for a much tighter relation between those three than in other linguistic modalities.

First, comparing hearing and reading of spoken and written linguistic utterances and their accompanying phenomenologies reveals an interesting difference in how *salient* a source of a spoken linguistic utterance typically is, as compared to the source of a written utterance. Consider the following example. Among hearing people, when a speaker says "Don't forget to buy some tea" as a hearer are about to go to the shop, by hearing their voice the hearer can recognize that it is the friend speaking, whether they can see them or not. Under such circumstances, when one hears a person speaking one perceives the speech sound they produce, recognise individual words and comprehend entire utterances, *as well as* their voice. The capacity for voice perception is an ability to recognize and differentiate between human voices that is typically developed and possessed by users of spoken languages (Belin et al., 2004; Schweinberger et al., 2014). In spoken linguistic interactions, voice perception allows human communicators to reliably track the source of spoken linguistic utterances (e.g. Schweinberger et al., 2014). This is so whether or not a hearer is familiar with the speaker's identity, e.g. speech is normally identified as coming from *a* particular speaker.

Consider now the second case: When one reads a note written by their friend that says "Don't forget to buy some tea" as one is about to go to the shop, one has very different resources at one's disposal to assess who the source of this message is. One may be able to recognize one's friend's handwriting. Even without being able to do so, one may infer from the context that it must have been her who has left the note. In

---

[22] Exceptions include e.g.: automated speech, chat bots, speech and documents produced on behalf of a group, collective or institution.

other cases, one might know that the note was written by another person but conveys a message from the friend who has a broken arm and cannot write herself. Other forms of written texts can indicate their sources in various idiosyncratic ways. In case one's friend sent a text message on the phone, one would immediately see her name displayed with the message. When one reads a book, the author's name usually stands on the cover. There are various ways in which one can communicate by writing. In contrast to the case of spoken linguistic utterances that are normally delivered in a voice of a particular speaker, in the case of reading there is no unique way in which the author of the message has to be present or revealed, depending on various factors, such as, handwriting, (un)characteristic style, explicit authorship statement, conventions and shared assumptions.

These observations suggest an important difference in what one experiences when reading and hearing language. The voice of the speaker is *a medium* in which spoken utterances are delivered. An auditory experience of a speaker's voice is thus part and parcel of one's experience in the case of spoken linguistic communication (Smith, 2009; Drożdżowicz, 2021). In typical cases of hearing spoken language, speech sounds produced by the speaker and their voice appear audibly related. Try as you might, it is impossible to hear speech sounds in an abstract voice-less way. Speech sounds spoken in two different voices result in two markedly different experiences. Thus, a key feature of hearing speech in a voice is that speech sounds produced and the vocal characteristics of a speaker appear united.[23] Spoken linguistic utterances come in a medium that has specific vocal properties. This has important repercussions for how spoken linguistic utterances are experienced: they often *appear* to us as packed with rich information about the speaker (e.g. their age, class, nationality, perceived gender, identity).

Writing, on the other hand, does not involve a clear and unique parallel to the human voice that is a medium in which a linguistic utterance is delivered. This intuitive difference is nicely captured in J.L. Austin's observation that speech acts made with spoken utterances are "tethered" to their origin, a speaker, whereas inscriptions are not and may require appending them with a signature (Austin, 1975, pp. 60–61). Reading an utterance may thus often require some further steps to identify their sources, e.g. reading a signature or an email header. This suggests that, typically at least, the source of a spoken linguistic utterance is particularly salient to the comprehender as a medium in which linguistic utterance is produced. The source of a written utterance can also be made salient to the reader, but in much more varied ways, often requiring further cognitive steps on the comprehender's part. On the comprehender's side, reading seems *prima facie* devoid of auditory vocal impressions that almost inevitably accompany comprehension of spoken language. Intuitive as this observation may seem it too needs to be qualified. Auditory impressions of voices and vocal sounds have been argued to

---

[23] The nature of this particularly salient and tight relation between spoken linguistic utterances and voice that produces them can be further investigated by drawing on a recent debate concerning the relation between sounds and their sources. In another paper (Drożdżowicz, 2021), I have suggested that one promising way to account for the audible relation between speech sounds and their sources is in terms of the mereological view on sounds and sources (O'Callaghan, 2010). On that view, the audible relation between sounds and sources is a relation between parts and a whole.

systematically accompany reading, silent reading in particular. Their exact nature and epistemic significance, however, are much debated.[24]

A closer look at other linguistic modalities and forms of communication reveals further complexity in how linguistic utterances are related to their sources and the medium in which they are produced. This results in different information and experience profiles. Written linguistic utterances, as mentioned, are usually (at least to some extent) severed from their sources and the intentional production at the point when they are comprehended. In visual signed languages, the physical presence, the body and movements of the person who is producing signs, the source, are of key importance. The body, including the handshape configuration, the place of articulation, the face, as well as movement and gestures of the body, and face, are part of the overall visual linguistic experience. Spoken linguistic utterances are constituted by phonation: the movement and modulation of a person's vocal apparatus and experienced by competent comprehenders as intentionally produced speech sounds in the medium of a particular voice. Linguistic production in visual signed languages is constituted by the movement and configuration of specific body parts, including the face, and experienced by competent comprehenders as intentionally produced bodily actions with a specific communicative purpose. Unlike written linguistic utterances, whose relation to the source can take various idiosyncratic forms and will often require additional inference steps on the part of the reader to uncover their source, the sources in spoken and signed visual linguistic utterances are necessarily present to physically sustain linguistic production (in online communication) and are thus immediately salient in the comprehenders auditory and visual experience.

An interesting similarity here is this: The experience of speech and sign production is an experience (as) of a particular intentional action. The experienced intentional nature of speech can be to some extent, at least, explained by the propensity for recognizing spoken linguistic utterances produced in a voice as having ostensive and communicative significance. In producing an utterance an individual speaker tries to convey a message—an utterance is produced with a certain communicative intention (Smith, 2009). This typically happens even in cases when one does not recognise the speaker's identity. This kind of auditory propensity is also present in cases where the voice and speech do not correspond to a particular communicative intention of a specific, individual speaker, e.g. when one is listening to a pre-recorded automated message on a telephone line or to a voice personal assistant. The impression of a linguistic utterance being intentionally produced by *someone specific* remains, despite one's knowledge that the voice in which a message is produced is merely a proxy for, e.g. a service, an institution. A parallel immediate experience accompanies comprehension of signed languages, where signs produced by the source are immediately

---

[24] Evidence suggests that phonological representations activated in silent reading and occasionally experienced by readers as inner speech are much closer to phonological representations of normal external speech sounds produced in voices. Various impressions of prosody can influence processing of individual words, as well as parsing, and interpretive processes during silent reading (e.g. Breen & Clifton, 2011; Chen et al., 2016). Silent reading may be accompanied by a vocal impression of our own voice and influenced by the vocal properties exhibited in the external speech of the reader, e.g. regional accent (Filik & Barber, 2011). Finally, the imagined voice of a literary character has been argued to influence our ability to engage with fictional words represented in written texts (e.g. Petkov & Belin, 2013; Yao et al., 2011).

recognized by competent interlocutors as having ostensive and communicative significance, rather than movements and gestures formed without a communicative intention. This seems compatible with research on categorical perception of phonological units in signed languages (Emmorey, 2011), where categorical perception is found only for distinctive hand configurations used to communicate, and not for other hand configurations. Thus, part of the experience of how both spoken and signed linguistic utterances relate to their sources is experiencing linguistic utterances as intentionally produced by the source, i.e. indicating an ostensive communicative act sustained in a specific vocal or bodily action.

This parallel attunement to the immediate presence of the source in spoken and signed linguistic communication comes with different benefits. Users of spoken languages become expert in differentiating across spoken utterances produced in different voices and attributing different spoken utterances to the same source, even without seeing the speaker and knowing their identity, as well as inferring information about the speaker from their speech production, which may serve as basis for further beliefs about speakers' characteristics (e.g. Schweinberger et al., 2014). Users of signed languages become entrained in linguistic skills that rely on visual modality. It has been argued that they focus a lot on the face during conversations which leads to better face processing skills, such as recognition of facial features and signer's identity and recognition of emotional states expressed in the face of the signer (Hauser & Kartheiser, 2014).

My final point concerning the source and medium of linguistic utterances is that the source and medium in tactile signed communication rely on the body of the signer and addressee allowing for a special, more intimate relation between those four than is common in other linguistic modalities. In tactile signed communication the relation between the source of a linguistic utterance and the medium in which it is delivered is more complex. In spoken and visual signed communication, the medium in which linguistic utterances are produced is an action of the source and to a large extent belongs to the source, be it their vocal production or movement, configuration of handshape, orientation in the body and face gesture. In tactile signed communication the medium and source of linguistic utterances seem to take a very special *participatory* form (Mesch et al., 2015). I draw here on a useful paper from Mesch et al., (2015) that analyzes utterances in tactile signed languages of Swedish and Norwegian signers who are both deaf and blind from a cognitive linguistics perspective. Tactile signs are produced by the source with the use of the body of the addressee and frequently together with the addressee's bodily and tactile action. Linguistic production in tactile signed communication relies on interlocutors who remain in contact with each other's hands while signing or other body parts, e.g. face, neck (Mesch et al., 2015). The bodies of the author of an utterance and of the addressee both provide the medium in which tactile signed languages are produced. Linguistic utterances in tactile signed communication are sustained by the signer producing signs with his or her own hands (or other body parts) in contact with the body of the addressee, some articulations also recruit the other addressee's hands (or another body part) in a more active manner. Mesch et al. (2015) argue that touch, joint movement, and haptic sensations provide key resources for linguistic expression. This very special configuration and relation between the source, the medium of an utterance and the addressee in tactile signed

communication is described by them as involving a *blended space* between the bodies and movements of the signer and the addressee (Mesch et al., 2015). The bodily and participatory form of linguistic communication is what makes the relation between the source, the medium and the addressee special in this case. It seems that, the author of a linguistic utterance and the addressee jointly provide a medium in which that utterance is produced. Moreover, one could argue that in some cases, where the addressee's body is recruited and actively co-forming an articulation, there may be a sense in which both of them may be deemed to be a source of a linguistic utterance.

The observations presented in Sect. 4.3 suggest, in my opinion, that our understanding of the relation between language and its source can be greatly improved by taking a closer look at and across different linguistic modalities and forms of communication. These observations allow for a varied and more nuanced conception of how the author of a linguistic utterance may be revealed and experienced by an addressee. A subject for further investigation is whether and how differences between linguistic modalities may support specific patterns of information about the author of an utterance (and addressee in the case of tactile signed communication) transmitted in linguistic production and if those can feed differently into linguistic interactions. For example, impressions of spoken linguistic utterances produced in different voices and coming from different speakers may vary depending on how we experience their vocal characteristics. Low pitch has been shown to be experienced as an indicator of dominance and leadership skills (Klofstad et al., 2012). There is evidence that voices vary substantially in perceived pleasantness: voices perceived as unpleasant voices are often raspy, grating, husky, or shrill (Gentsch et al., 2020). Both associations between low-pitch and competence and the perceived unpleasantness of a voice triggering aversive response may have an impact on the overall uptake and one's attitudes towards spoken linguistic utterances coming from particular speakers. Depending on who exclaims 'Brilliant!' and the impression their voice makes, the same message of enthusiasm may be differently received and appreciated. A written note or a signed utterance of 'Brilliant!' may have an entirely different effect.

5 Concluding remarks

This exploratory paper advocates for greater inclusion of various linguistic modalities and forms of communication into some, selected philosophical debates concerning language, meaning and communication. I have provided a short glance at some linguistic modalities and forms of communication and described some results that such inclusive approaches can provide. Abstraction can be warranted to work with some philosophical questions concerning language and communication. But whether and how observations concerning various linguistic modalities and forms of communication can enrich our understanding of language and communication is a possibility philosophers should consider. This paper has a modest aim and can be seen as just a first step in such an investigation.

to develop this manuscript. I am also grateful to two anonymous reviewers for this journal for very helpful comments and suggestions. I would also like to thank the editors of this issue: Thomas Hughes and James T. Miller.

Declarations

Conflict of interest　The author declares no conflict of interest.

References

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin, 141*(5), 931.

Alderson-Day, B., Moffatt, J., Bernini, M., Mitrenga, K., Yao, B., & Fernyhough, C. (2020). Processing speech and thoughts during silent reading: Direct reference effects for speech by fictional characters in voice-selective auditory cortex and a theory-of-mind network. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01571

Ashby, J., & Martin, A. E. (2008). Prosodic phonological representations early in visual word recognition. *Journal of Experimental Psychology, 34*, 224–236.

Austin, J. L. (1975). *How to do things with words*. Oxford University Press.

Baker, S. A., Idsardi, W. J., Golinkoff, R. M., & Petitto, L. A. (2005). The perception of handshapes in American sign language. *Memory & Cognition, 33*(5), 887–904.

Bauman, H. D., & Murray, J. J. (2017). Sign languages. In *Handbook of language and society* (pp. 243–260). Springer.

Bayne, T. (2009). Perception and the reach of phenomenal content. *Philosophical Quarterly, 59*(236), 385–404.

Baynton, D. C. (1996). *Forbidden signs: American culture and the campaign against sign language*. University of Chicago Press.

Begby, E. (2017). Language from the ground up: A study of homesign communication. *Erkenntnis, 82*(3), 693–714.

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129–135.

Bertelson, P. (2017). Language by touch: The case of braille reading. In *Speech and reading* (pp. 91–108). Routledge.

Braun, D. (2017). Indexicals. In Edward N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2017 Edition). Retrieved from https://plato.stanford.edu/archives/sum2017/entries/indexicals/.

Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language, 64*, 153–170.

Brentari, D. (2019). *Sign language phonology*. Cambridge University Press.

Brogaard, B. (2018). In defense of hearing meanings. *Synthese, 195*(7), 2967–2983.

Cappelen, H., & Dever, J. (2019). *Bad language*. Oxford University Press.

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition, 154*, 11–21.

Collins, S. (2004). *Adverbial Morphemes in Tactile American Sign Language*, Ph.D. thesis. Graduate College of Union Institute and University.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In *The science of reading: A handbook* (pp. 6–23).

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (2013). *Models of reading aloud: Dual-route and parallel-distributed-processing approaches* (pp. 381–422). Psychology Press.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Mit Press.

Dammeyer, J., Nielsen, A., Strøm, E., Hendar, O., & Eiríksdóttir, V. K. (2015). A case study of tactile language and its possible structure: A tentative outline to study tactile language systems among children with congenital deafblindness. *Journal of Communication Disorders, Deaf Studies & Hearing Aids.* https://doi.org/10.4172/2375-4427.1000133

Daniels, P. T., & Bright, W. (1996). *The world's writing systems*. Oxford University Press on Demand.

De Gelder, B., & Morais, J. (2017). *Speech and reading: a comparative approach* (Vol. 2). Routledge.

Drożdżowicz, A. (2021). Bringing back the voice: On the auditory objects of speech perception. *Synthese, 199*(1–2), 661–687.

Emmorey, K. (2001). *Language, cognition, and the brain: Insights from sign language research*. Psychology Press.

Emmorey, K. (2011). The psycholinguistics of signed and spoken languages: How biology affects processing. In *Oxford handbook of psycholinguistics.* Oxford University Press.

Emmorey, K., Thompson, R., & Colvin, R. (2009). Eye gaze during comprehension of American sign language by native and beginning signers. *Journal of Deaf Studies and Deaf Education, 14*(2), 237–243.

Emmorey, K., McCullough, S., & Brentari, D. (2003). Categorical perception in american sign language. *Language and Cognitive Processes, 18*(1), 21–45.

Filik, R., & Barber, E. (2011). Inner speech during silent reading reflects the reader's regional accent. *PLoS ONE, 6*(10), e25782.

Fodor, J. D., Nickels, S., & Schott, E. (2017). Center-embedded sentences: What's pronounceable is comprehensible. In R. G. De Almeida & L. R. Gleitman (Eds.), *On concepts, modules, and language: Cognitive science at its core* (pp. 139–168). Oxford University Press. https://doi.org/10.1093/oso/9780190464783.001.0001

Fricker, E. (2003). Understanding and knowledge of what is said. In A. Barber (Ed.), *Epistemology of language* (pp. 325–366). Oxford University Press.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*, 71–99.

Gasparri, L. (2021). A pluralistic theory of wordhood. *Mind & Language, 36*(4), 592–609.

Gentsch, K., Beermann, U., Wu, L., Trznadel, S., & Scherer, K. R. (2020). Temporal unfolding of micro-valences in facial expression evoked by visual, auditory, and olfactory stimuli. *Affective Science, 1*(4), 208–224.

Geva, S. (2018). Inner speech and mental imagery. In *Inner speech: New voices* (p. 105).

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J. C., & Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion–symptom mapping. *Brain, 134*(10), 3071–3082.

Goodman, J. (2020). On reading. *Acta Analytica, 35*(1), 51–59.

Hauser, P. C., & Kartheiser, G. (2014). Advantages of learning a signed language. In *Deaf gain: Raising the stakes for human diversity* (pp. 133–145).

Harley, T. A. (2013). *The psychology of language: From data to theory*. Psychology Press.

Hsu, C. H., Tsai, J. L., Lee, C. Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: An event-related potential study. *Brain and Language, 108*(1), 56–66.

Irmak, N. (2019). An ontology of words. *Erkenntnis, 84*(5), 1139–1158.

Iwasaki, S., Bartlett, M., Manns, H., & Willoughby, L. (2019). The challenges of multimodality and multisensoriality: Methodological issues in analyzing tactile signed interaction. *Journal of Pragmatics, 143*, 215–227.

Kaplan, D. (2011). Words on words. *The Journal of Philosophy, 108*(9), 504–529.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135*(1), 12.

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society b: Biological Sciences, 279*(1738), 2698–2704.

Krentz, U. C., & Corina, D. P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science, 11*(1), 1–9.

Lee, C. Y., Tsai, J. L., Chan, W. H., Hsu, C. H., Hung, D. L., & Tzeng, O. J. (2007). Temporal dynamics of the consistency effect in reading Chinese: An event-related potentials study. *NeuroReport, 18*(2), 147–151.

Leslie, S. J. (2015). Generics oversimplified. *Noûs, 49*(1), 28–54.

Levelt, W. J. (1999). Models of word production. *Trends in Cognitive Sciences, 3*(6), 223–232.

Longworth, G. (2018). Understanding what was said. *Synthese, 195*(2), 815–834.

Mesch, J., Raanes, E., & Ferrara, L. (2015). Co-forming real space blends in tactile signed language dialogues. *Cognitive Linguistics, 26*(2), 261–287.

Magrassi, L., Aromataris, G., Cabrini, A., Annovazzi-Lodi, V., & Moro, A. (2015). Sound representation in higher language areas during language generation. *Proceedings of the National Academy of Sciences, 112*(6), 1868–1873.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375.

Miller, J. T. M. (2020). The ontology of words: Realism, nominalism, and eliminativism. *Philosophy Compass, 15*(7), e12691.

Miller, J. T. M. (2021). A bundle theory of words. *Synthese, 198*(6), 5731–5748.

Nefdt, R. M. (2019). The ontology of words: A structural approach. *Inquiry, 62*(8), 877–911.

Nes, A. (2016). On what we experience when we speak. *Phenomenology and Mind, 10*, 58–85.

Nicodemus, B. (2009). What is prosody. Chapter 2 in Prosodic Markers and Utterance Boundaries in American Sign Language Interpretation. Studies in Interpretation Series, Vol. 5. Gallaudet University Press.

Nygaard, L. C., & Pisoni, D. B. (1995). *Speech perception: New directions in research and theory*. Elsevier.

O'Callaghan, C. (2010). Constructing a theory of sounds. *Oxford Studies in Metaphysics, 5*(5), 247.

O'Callaghan, C. (2011). Against hearing meanings. *The Philosophical Quarterly, 61*(245), 783–807.

O'Callaghan, C. (2015). Speech perception. In *Handbook of the philosophy of perception.* OUP.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112*(1), 43.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Loevenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research, 261*, 220–239.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., & Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences, 97*(25), 13961–13966.

Petkov, C. I., & Belin, P. (2013). Silent reading: Does the brain 'hear' both speech and voices? *Current Biology, 23*(4), R155–R156.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the EZ reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1–56.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading*. Psychology Press.

Reed, C. M., Rabinowitz, W. M., Durlach, N. I., Braida, L. D., Conway-Fithian, S., & Schultz, M. C. (1985). Research on the tadoma method of speech communication. *The Journal of the Acoustical Society of America, 77*(1), 247–257.

Reiland, I. (2015). On experiencing meanings. *The Southern Journal of Philosophy, 53*(4), 481–492.

Rey, G. (2008). In defense of folieism: Replies to critics. *Croatian Journal of Philosophy, 8*(23), 177–202.

Rosenfeld, S. A. (2001). *A revolution in language: The problem of signs in late eighteenth-century France*. Stanford University Press.

Sandler, W., & Lillo-Martin, D. (2001). Natural sign languages. In *Handbook of linguistics* (pp. 533–562).

Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., & Zäske, R. (2014). Speaker perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(1), 15–25.

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage, 44*(3), 1133–1143.

Slowiaczek, M. L., & Clifton, C., Jr. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior, 19*(5), 573–582.

Smith, B. C. (2009). Speech sounds and the direct meeting of minds. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception* (pp. 183–210). Oxford University Press.

Snoddon, K. (2022). Writing as being: On the existential primacy of writing for a deaf scholar. *Qualitative Inquiry*. https://doi.org/10.1177/10778004211073071

Sosa, D. (2018). *Bad words: Philosophical perspectives on slurs*. Oxford University Press.

Speer, S., & Blodgett, A. (2006). Prosody. In *Handbook of psycholinguistics* (pp. 505–537). Academic Press.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*(1), 114.

Stokoe, W. C., Jr. (2005). Sign language structure: An outline of the visual communication systems of the American deaf. *Journal of Deaf Studies and Deaf Education, 10*(1), 3–37.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643.

Tatham, M., & Morton, K. (2006). Speech perception: Production for perception. In *Speech production and perception* (pp. 218–234). Palgrave Macmillan.

Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science* (p. 10). Wiley.

Van Woudenberg, R. (2021). Reading as a source of knowledge. *Synthese, 198*(1), 723–742.

Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Science, 176*(4039), 1149–1151.

Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press.

Wilbur, R. B., Marschark, M., & Spencer, P. (2003). *Modality and the structure of language: Sign languages versus signed systems* (pp. 332–346). Oxford University Press.

Willoughby, L., Iwasaki, S., Bartlett, M., & Manns, H. (2018). Tactile sign languages. In *Handbook of pragmatics.* (Vol. 21). https://doi.org/10.1075/hop.21.tac1

Willoughby, L., Manns, H., Iwasaki, S., & Bartlett, M. (2020). From seeing to feeling: How do deafblind people adapt visual sign languages? In K. Allan (Ed.), *Dynamic language changes: Looking within and across languages* (pp. 235–252). Springer.

Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience, 23*(10), 3146–3152.