A tale of two modalities:

How modality shapes language production and visual attention

Francie Manhardt

ISBN: 978-94-92910-24-0

Cover design and illustration by Marlijn ter Bekke and Francie Manhardt

Printed and bound by Ipskamp Drukkers b.v.

The research reported in this dissertation has been supported by NWO VICI Grant awarded to H. A. Özyürek.

The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University - the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond.

A tale of two modalities: How modality shapes language production and visual attention

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 12 maart 2021
om 14:30 uur precies

door

Francie Manhardt
geboren op 11 april 1991
te Erfurt (Duitsland)

Promotor:

Prof. dr. H. A. Özyürek

Copromotor:

Dr. S. M. Brouwer

Manuscriptcommissie:

Prof. dr. F. Huettig

Dr. R. G. Bosworth (Rochester University, Verenigde Staten)

Prof. dr. M. Gullberg (Lunds Universitet, Zweden)

Prof. dr. C. R. Marshall (University College London, Verenigd Koninkrijk)

Prof. dr. P. M. Perniss (Universität zu Köln, Duitsland)

Für Oma Flora und Opa Thilo

Contents

Chapter 1

General Introduction

Language is a powerful tool that enables us to talk about our experiences with the world. However, languages differ fundamentally from one another in how they package and represent different aspects of our experiences. Hence, there is linguistic variation across speakers of different languages as well as among bilinguals that speak different languages (see Evans & Levinson, 2009). This diversity has motivated researchers to explore whether different expressive options across different languages influence how speakers perceive and conceptualise the world differently. In the current thesis, I aim to enhance our understanding of how language diversity can affect cognition for language production by comparing users of a spoken language versus a sign language, focusing on spoken Dutch and Sign Language of the Netherlands (*NGT, Nederlandse gebarentaal*). In doing so, I zoom into the domain of spatial expressions where spoken and sign languages differ radically from each other due to different affordances of the vocal versus visuo-spatial modality. This enables me to go beyond previous investigations on the effect of language on cognition that have focussed on differences between spoken languages only.

Many theories on the relationship between language and cognition suggest that human perception and cognition are universal. Thus, irrespective of which language one speaks, all humans perceive and conceptualise the world in the same way (e.g. Jackendoff, 1996; Pinker, 1995). However, others claim that diversity across languages can influence speakers' cognition differently, although scholars share different beliefs in terms of which aspects of language and the language production process could influence cognition. Some argue that if language influences cognition, these effects are robust and should extend to domains of cognition outside of the contexts of language use (e.g. Sapir, 1912; Whorf, 1997). However, more recent views argue that language can influence cognition only in the moment of or immediately prior to speaking, which is captured in the so-called *thinking for speaking* hypothesis (Slobin, 2003). In particular, during message preparation, speakers of different languages attend to different features of the world linked to the way they plan to speak about them. Thus linguistic categories, which often differ across languages, seem to influence cognition during the process of converting thoughts into language (e.g. Flecken, Carroll, Weimar, & Stutterheim, 2015; Flecken, Von Stutterheim, & Carroll, 2014; Gleitman, January, Nappa, & Trueswell, 2007; Griffin & Bock, 2000; Slobin, 2003).

However, this body of research on the effects of linguistic diversity on cognition have exclusively focussed on spoken languages and until now overlooked sign languages, the natural languages of deaf communities. Consequently, it is less known to what extent differences in modality; that is, using the vocal versus visuo-spatial modality, can shape how conceptualization for language production differs between speaking (i.e. *thinking for speaking*) and signing (i.e. *thinking for signing*). Furthermore, previous studies have mainly looked at the effects of language diversity on cognition across populations that use different languages but often left out linguistic diversity *within* populations that speak multiple languages; that is, bilinguals.

The current thesis attempts to go beyond the previous literature to contribute to the debate on language and cognition universals and diversity. I introduce differences in language modality as a new angle to study the effects of language diversity on cognition and bilingual diversity. I explore this by comparing diversity between a spoken (Dutch) and a sign language (NGT) across hearing speaking and deaf[1] signing populations as well as in hearing bilinguals who can both speak and sign, so-called *bimodal bilinguals*[2]. This

[1] With the term "deaf" I refer simply to the hearing status and do not intend to make any implications about cultural belongings.

[2] With the term "bimodal bilinguals" I follow Emmorey, Borinstein, Thompson, and Gollan (2008) and refer to the hearing individuals born to and raised by deaf parents who can speak and sign. In the literature, they are also referred to as *CODAs* (children of deaf parents), sign-speech bilinguals or also as hearing native signers (e.g. De Quadros, 2018; Emmorey, Borinstein, Thompson, & Gollan, 2008; Pichler, Lillo-Martin, & Palmer, 2018; Quadros & Lillo-Martin, 2018). They form a special group of bilinguals because they have early exposure to both a spoken and a sign language through uninstructed (i.e. no schooling) exposure to language from early on and are thus highly proficient in two languages of different modalities. For more information see section 4.1 Participants.

will provide not only new insights into effects of linguistic diversity on cognition but also reveal novel knowledge about how two languages can influence each other across modalities; that is, between speech and sign, which has not been examined before. To do so, I will focus on the domain of spatial language, specifically, on expressions of spatial relations, such as *the pen is to the right of the glass*, where previous research has identified clear modality-specific differences between spoken and sign languages in their linguistic patterning (e.g. Özyürek, Zwitserlood, & Perniss, 2010; Emmorey, 2002; Perniss, Zwitserlood, & Özyürek, 2015; Zwitserlood, Perniss, & Özyürek, 2012).

Modality differences and spatial language as a domain to study effects of language diversity on cognition

Sign languages are considered as fully developed languages that are expressed with the visible parts of the body (Stokoe, 2005) and operate in the visuo-spatial modality using hand, face and body movements for linguistic expressions. Although sign languages are found to share similar linguistic patterning to those used by spoken languages (see e.g. Meier, 2002; Özyürek & Woll, 2019), they also contain modality-specific linguistic aspects (Perniss, Özyürek, & Morgan, 2015). These are most prominent in the expressions of spatial language, such as *the pen is to the right of the glass* (Figure 1A), which require the identification of the figure object (i.e. small object, e.g. pen) and the ground object (i.e. bigger object, e.g. glass) as well as expressing the spatial relation between them (Talmy, 1985; 2003). In such expressions, modality-specific patterning in sign languages arise, first, due to sign languages' preferences to use iconic form-to-meaning mappings, and secondly, due to the consistent word order they choose for spatial language expressions (e.g. Emmorey, 2002; Kimmelman, 2012; Perniss et al., 2015).

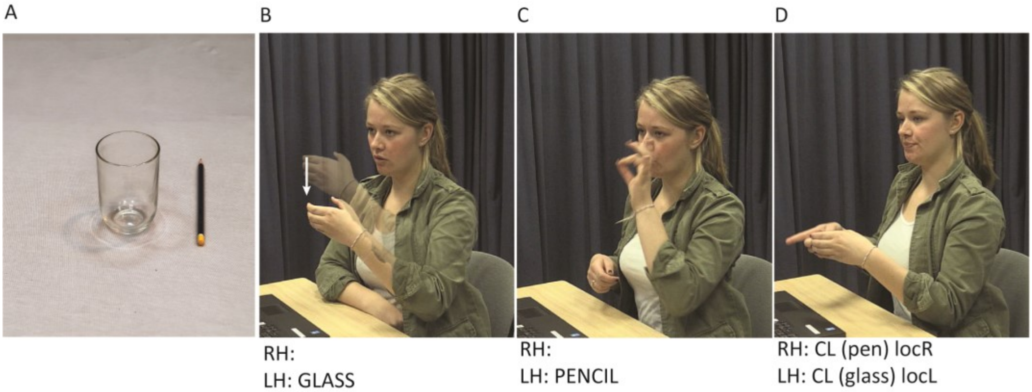
Sign languages, unlike spoken languages[3], use a large proportion of visually motivated forms. This so-called *iconicity* refers to the motivated link between a visual (in the case of sign languages) or auditory (in the case of spoken languages) linguistic form and its meaning. Visual iconicity in sign languages can occur at different levels of language expressions, ranging from the lexical level to the sentence and discourse levels (e.g. Dingemanse, Perlman, & Perniss, 2020; Emmorey, 2014; Perniss, 2007; Perniss, Thompson, & Vigliocco, 2010; Sherman Wilcox, 2004; Slonimska, Özyürek, & Capirci, 2020; Taub, 2000).

For example at the lexical level, the NGT sign for GLASS resembles the round shape of the glass as well as the action of putting a glass onto a surface as illustrated in Figure 1B. Hence, the sign for GLASS resembles visually and perceptually the GLASS in the real world (also known as *imagistic iconicity*; e.g. Russo, 2004; Taub, 2001). In the case of expressing spatial relations, signed sentences can resemble structural relationships between entities in the signing space (also referred to as *diagrammatic iconicity*; e.g. Perniss, 2007; Russo, 2004; Slonimska et al., 2020) by not only mirroring shape, size and orientation features of the glass and the pen but also the spatial relationship between them by using so-called *classifier constructions* (*CLs*). In the example below (Figure 1D) the spatial relationship (pen to the left of glass) between the pen and the glass is depicted through diagrammatic iconicity; that is, where and how two objects are placed in space in relation to each other. Such diagrammatic visual iconicity is not possible in the spoken expressions of spatial language where spatial expressions are expressed by using arbitrary and categorical labels such as LEFT and RIGHT.

[3] Some of these visually motivated iconic mappings can also be optionally used by speakers using co-speech gestures alongside speech and spatial language (e.g. McNeill, 1992; Sauter, Uttal, Alman, Goldin-Meadow, & Levine, 2012). Spoken languages can also involve iconicity. As opposed to resembling visual perceptions as in the case of sign languages, iconicity in spoken languages resembles auditory perception like onomatopoeias, such as *bang* in English or ideophones such as *sinisinisini* (closely woven; e.g. Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015).

Figure 1

Iconic expressions in NGT to describe "the pen is to the right of the glass" by using classifier constructions.



Panel (A) shows a pen that is located to the right of the glass. Panel (B) shows the introduction of the lexical sign for GLASS. Panel (C) shows the introduction of the lexical sign for PEN. Panel (D) shows an iconic expression (glossed CL to refer to so-called classifier constructions) where the hands represent size, shape, and orientation of the two entities located in space (glossed as LocL/R). RH stands for the right hand, LH stands for the left hand.

In addition to iconicity, the use of the visuo-spatial modality in sign languages also drives a specific word order preference when it comes to describing spatial relations. In particular, signers are found to mention grounds (e.g. glass, Figure 1B) before figures (e.g. pen, Figure 1C) which are then followed by CLs (Figure 1D, although CLs can occur in differed positions of a sentence). This order of mentioning first grounds and then figures is similar in all sign languages studied so far (e.g. Emmorey, 1996 for American Sign Language; Morgan, Herman, Barrière, & Woll, 2008 for British Sign Language; Sümer, 2015 for Turkish Sign Language; Perniss, 2007 for German Sign Language; and Kimmelman, 2012 for Russian Sign Language). Interestingly, the same preference of mentioning grounds first has also been observed in silent gestures (e.g. Goldin-Meadow et al., 2008; Laudanna & Volterra, 1991); that is, when hearing people are asked to use their hands only to depict relations between objects without using language. Thus, it appears that using the hands to express the ground first and the figure later is a modality-driven preference shared by many sign languages of the world (e.g. Emmorey, 2002; Kimmelman, 2012; Perniss et al., 2015). Across spoken languages, however, there are substantial differences in word order preferences in the domain of spatial language.

In the current thesis I will be focusing on these two modality-specific differences between spoken and sign languages, namely iconicity and word order preference, and in particular in spoken Dutch and NGT to explore three broader questions that are investigated in four experimental studies in the current thesis:

(1) To what extent do differences in modality influence **production** across different speaking (hearing) and signing (deaf) populations?
(2) Do these modality-driven differences in production patterns found in (1) influence **cognition**, as measured by visual attention to referents during message preparation?
(3) How do patterns observed in (1) and (2) across different speaking (hearing) and signing (deaf) populations manifest themselves in bilinguals (hearing); that is, in the same individuals who can both speak and sign (i.e. bimodal bilinguals)?

By answering these questions and combining knowledge from three research areas, namely, language and cognition, sign languages, and bilingualism, this thesis offers novel contributions for enhancing our understanding of how language diversity can affect cognition across different populations as well as within the same individuals. Alternatively, it is possible that differences in modality do not affect cognition across different language populations or within bilinguals. To explore these possibilities and answers to the questions addressed above, the thesis combines elicited productions with an eye-tracking experiment to investigate possible relations between modality-driven production patterns and cognition.

Previous research

In the remainder of this section, I will present a literature overview to provide a theoretical background for the studies described in the next chapters. In section 1, I will introduce approaches to language diversity and its possible links to cognition. In section 2, I will describe in detail two modality-specific differences explored in this thesis, namely iconicity and word order preference in the domain of spatial language. More specifically, I will focus on left/right relations, the empirical domain of the thesis. In section 3, I will then talk about what is known about diversity in bilinguals of two spoken languages as well as bilinguals of a spoken and a sign language and possible relations between bilingualism and cognition.

1 Language diversity and cognition

A widespread assumption among linguistic scholars has been that all languages are the same concerning their underlying structures and simply consist of different sound systems and vocabularies (e.g. Jackendoff, 1996; Pinker, 1995). However, contrary to this universalist view, a recently more empirically supported opinion is that diversity can be found at almost every level of linguistic organization (Evans & Levinson, 2009). Consequently, languages are assumed to differ fundamentally from one another in the way they offer different expressive options to highlight certain aspects of our experiences of the world and to de-emphasise others.

The variation found across the worlds' languages has raised questions on whether speakers of different languages conceptualise aspects of the visual world differently. Some scholars claim that language does not influence cognition (e.g. Jackendoff, 1996; Pinker, 1989). In this view, language has been traditionally considered as simply the *vehicle* for expressing thought and therefore, language cannot influence cognition (for a review, see e.g. Gleitman & Papafragou, 2005; Gentner & Goldin-Meadow, 2003; Ünal & Papafragou, 2016; Wolff & Holmes, 2011). Others argue that all types of cognition are influenced by language, even outside of the context of language (e.g. Sapir, 1912; Whorf, 1997). Within this view, language is seen as *toolkit* or *augmenter* to highlight or represent certain aspects of the world which can then change the perception of this world. More recently, empirical evidence supports a hybrid view that language can influence cognition but only those aspects that are relevant for speaking, as captured in the *thinking for speaking* hypothesis (Slobin, 2003). To empirically study this hypothesis, scholars have often measured visual attention to static or animated scenes during message preparation or during online language production.

These studies have found that speakers guide their visual attention to those features that they plan to speak about, in ways that are guided by the language-specific grammatical patterns in their language (e.g. Bunger, Skordos, Trueswell, & Papafragou, 2016; Flecken et al., 2015, 2014; Gleitman et al., 2007; Griffin & Bock, 2000; Papafragou et al., 2008; Slobin, 2003; Trueswell & Papafragou, 2010; van de Velde, Meyer, & Konopka, 2014). Thus, language-specific ways in which messages are formulated become automatized and shape visual attention during the preparation of messages even before they are produced (e.g. Bock, 1995; Levelt, 1989). This is in line with Slobin's claim that speakers attend to those aspects of the world that they intend to talk about, given the constraints or possibilities within their language (Slobin, 2003). For example, cross-linguistic differences between English and Greek, in how a motion event is expressed, direct speakers' visual attention to different parts of the event. Greek predominantly expresses the Path (i.e. the direction) of the motion in the main verb (Talmy, 1985; Talmy, 2003), whereas English typically expresses Manner of the motion (i.e. how the motion is performed) in the main verb. Reflecting these grammatical differences, during the planning of event descriptions (e.g. someone is running up the stairs), Greek speakers looked more at the Path (e.g. ascending the stairs), while English speakers looked more at the Manner of the event (e.g. running).

While this influence of language on cognition has been studied across speakers of different languages, these findings, however, are restricted to the study of spoken languages. Sign languages have until now been ignored. In the current thesis, I compare a spoken and a sign language for the first time in this debate and ask whether and how differences in modality can provide new insights into the effects of language diversity on cognition across speaking and signing populations in **Chapter 2**. In addition, I investigate language diversity in (bimodal) bilinguals (**Chapter 3**-**5**) and its possible effects on cognition by comparing this diversity to differences across speaking and signing populations.

2 Modality differences between spoken and sign languages in the domain of spatial language

In this section, I will elaborate on the modality-specific aspects of sign languages that differ from spoken languages. In this thesis, I will highlight two aspects, namely iconicity and word order preference in the domain of spatial language. Specifically, I will focus on those expressions that are used to describe relations between two static objects because previous research has already identified this as an area where spoken and sign languages radically differ from each other. For each aspect, I will first discuss how spoken languages express spatial relations and second how sign languages express them.

Generally, to express three-dimensional aspects of spatial relations between two objects, language users identify the objects involved and also specify the spatial configuration in which these objects are related to each other through some predicative structure. They also need to assign the objects into themes of *figure (*smaller, movable, foregrounded object*)* and *ground* (bigger, less movable, backgrounded object) categories (i.e. to specify what is in relation to what) and order them accordingly in line with conventions of their languages (Talmy, 1985, 2003).

Spatial relations can be classified into two types which differ in terms of perspective taking. The first type contains descriptions of topological relations (e.g. in/on/under) which are not viewpoint-dependent; that is, they do not require perspective-taking to express them. For the second type of spatial relations, language users need to take different viewpoints into account. That is, lateral (left/right) and sagittal relations (front/behind) require language users to assign a viewpoint to describe the figure's relation with respect to the ground object by using forms such as left/right/in/front in English (Levinson, 1996, 2003). In this thesis, I focus on the expressions of viewpoint-dependent relations and specifically on left/right relations where differences in modality have the most impact on the expressions between spoken and sign languages.

2.1 Iconicity

To demonstrate how spoken and sign languages differ in the way they express spatial relations, I will first elaborate on what is known about how spoken languages express the spatial relationship between two objects, and after that explain how iconic ways of expressing spatial relations in sign languages differ from the patterns in spoken languages.

Spoken languages. To talk about space, spoken languages typically rely on categorical linguistic forms that have an arbitrary relationship to their meaning (an exception are co-speech gestures as mentioned earlier in the footnote on page 13, but the emphasis I have here is on speech). To do so, spoken languages predominantly use adpositions, which typically involve prepositions that can indicate support or contact (e.g. in English: the pen is in/on/under the glass) or morphological case markers, which mostly occur as postpositions to mark relations between entities (for example, in Turkish -de/da can be attached to the noun expressing the ground object or to spatial nouns such as left/right (sag/sol) to indicate the pen's position in respect to the glass; e.g. Evans & Levinson, 2009; Levinson & Wilkins, 2006; Sümer, 2015). Importantly for this thesis, expressions of all spoken languages in the world have an arbitrary relation to the actual spatial scene and are thus typically, non-iconic. Furthermore, they also express spatial relations in categorical ways such as in/on/left/right (e.g. Emmorey & Herzig, 2003; Talmy, 2003).

Furthermore, possibly due to using arbitrary form-to-meaning mappings, spoken languages also exhibit an enormous amount of variation in the way they categorise and encode spatial relations in non-iconic ways. For example, in English one single spatial preposition *on* can be used to describe various spatial relations, such as “the pen is *on* the table”, “the shelf is *on* the wall” as well as “the ring is *on* the finger”. However, in Dutch, three different devices must be used to describe these spatial scenes (i.e. *op*, *aan* and *om*, respectively; van Staden, Bowerman, & Verhelst, 2006).

Spoken languages can also vary in the amount of semantic specificity in which they express spatial relations. For example, in Tzeltal, speakers can use positionals that include information about the shape of figure objects in their spatial expressions; that is, the predicate *waxal* can be used for oval-shaped containers (e.g. a tall glass or a bucket) or *lechel* indicates flat-bottomed objects (e.g. a frying pan; e.g. Brown, 1994). Whereas languages such as Dutch do not express any information about physical features of objects in the way they categorise relations between entities. Spoken languages might also differ from each other by specifying the spatial relation concretely or not. Some spatial expressions might be under informative as in saying “the pen is *next* to the glass” unlike saying “the pen is to the *right* of the glass”. Thus, different languages employ fully specific versus ambiguous categories to different degrees (e.g. Levinson, 2003; Levinson & Wilkins, 2006).

Thus, spoken languages classify and express spatial relations in different ways across languages and they do so through arbitrary form-to-meaning mappings by using the vocal modality to express such relations in categorical ways. This is in strong contrast to how sign languages can express spatial relations. In sign languages, the nature of the visuo-spatial modality allows for a large proportion of visually-motivated iconic mappings allowing similarities across the worlds' sign languages studied to date and with relatively little variation among them (e.g. Aronoff, Meir, Padden, & Sandler, 2008; Emmorey, 2014; Perniss et al., 2010).

Sign languages. Sign languages studied to date mostly use iconic strategies to express spatial relations albeit at different levels of iconicity. One common and the most investigated strategy is the use of *classifier constructions* (CLs, see Figure 1D, page 14). In these constructions, the handshapes depict specific information about the size and shape of the objects and the hands' location and orientation in the signing space express information about the location and orientation of the objects in ways that mirror how they appear in real space (e.g. Emmorey, 2002; Engberg-Pedersen, 1993; Perniss, 2007; Schembri, 2003).

Researchers initially have suggested that CLs are the most common and universal strategy across different sign languages for expressing space (see also Cuxac, 1996; Sallandre & Cuxac, 2002; Talmy, 2003). However, more recent comparative work has shown that there might be also slight differences between sign languages and that other strategies also exist to express the spatial relationship between objects using different types of iconic depictions. To provide an overview of how sign languages can express spatial relations and highlight how they differ from the spatial expressions in spoken languages I will describe these strategies and different views in more detail below. I will start with a detailed description of CLs and their function and move on to another iconic linguistic strategy relevant for this thesis, so-called *relation lexemes (RLs)*. This is to show that signers can use different linguistic devices that vary in how relations and objects are expressed iconically, which is important for this thesis.

Figure 1 (page 14) illustrates an example ofusing CLs to describe a pen that is located to the left of the glass in NGT. It shows that signers may first introduce the lexical signs of the two entities, first the ground object *glass* (Figure 1B) and then the figure object *pen* (Figure 1C). After this introduction of the two objects, signers can select a round handshape to classify and represent the round shape of the glass and the index finger to represent the thin, elongated shape of a pen. Both hands are then placed consecutively next to each other to match the signers' view of the relative relations of the two objects to each other (Figure 1D). Overall, the way the hands are formed and placed in space, mirror and resemble both the object properties and the spatial relation. These devices are thus highly iconic (e.g. Emmorey & Herzig, 2003; Perniss, Özyürek, et al., 2015; Perniss et al., 2010), making full use of the visual correspondences between the described objects, the event space in the real world and the signing space. In this sense, CLs also express high semantic specificity about size, shape and orientation of the objects as well as their relative relations to each other in space.

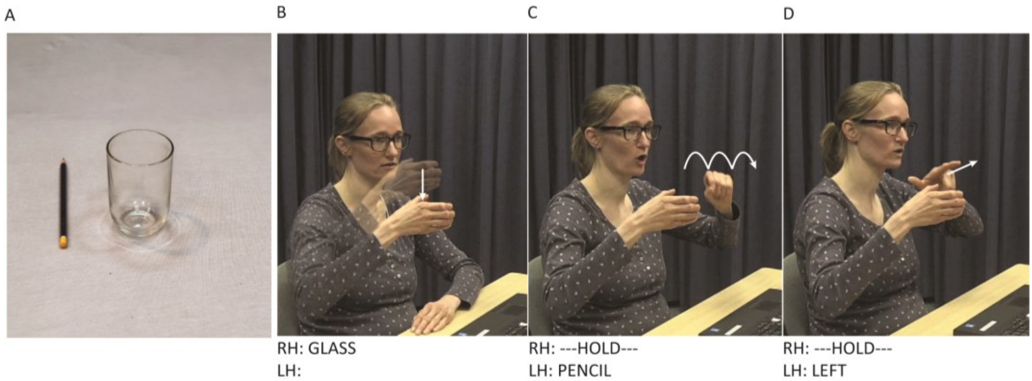
The grammatical nature of these CLs has been the subject of debate in sign language linguistics. For some researchers*,* the handshape in CLs is considered to be morphological. The morphemes' primary function is to express information about entities by "classifying" these entities in respect to specific semantic properties, like shape, size and orientation as well as their locations (Zwitserlood, 2003), rather than using merely analogue mappings. Thus in this view, CLs are assumed to obligatorily mark and classify locations and object properties. Others consider this view as problematic because a strict morphological analysis cannot fully account for the potentially infinite possible locations in which these spatially modifiable signs can be produced (e.g. Liddell, 1995, 1998). Furthermore, the handshapes used for size and shapes that are used to locate objects can be quite variable from context to context and do not seem to be fixed. For this reason, others prefer the term *depicting signs,* which emphasise their optional function to highlight and depict (i.e. to show; for the concept of depicting in language, see Clark, 1996), rather than classify (Ferrara & Halvorsen, 2017; Ferrara & Hodge, 2018; Liddell, 2003; Schembri, Jones, & Burnham, 2005). What is important for this thesis is that they are highly iconic to the size, shape and orientation of the objects as well as to the locations of these entities. Without taking a position in this debate, for the rest of the thesis, I will use the term CLs to refer to these expressions.

While it has been suggested for a long time that CLs are the most prominent way in which sign languages express spatial relations, there is only little empirical evidence and there are only a few quantitative investigations for this claim. Furthermore, recently it has become more evident that different sign languages, in fact, might employ also different iconic linguistic forms and there might also be quantitative differences across sign languages in terms of the different strategies they use (Perniss et al., 2015).

One other such strategy in addition to using CLs is using so-called *relational lexemes*, which seem to be more common in signers' spatial descriptions than previously assumed (e.g. Arik, 2013; Perniss, 2007; Perniss et al., 2015; Sümer, 2015). RLs are lexical signs that express the spatial relation between objects in iconic ways but, unlike CLs, do not convey semantically specific information about the physical properties of the objects involved. While in CLs the round shape of the glass and the long elongated shape of the pen become salient, RLs do not express information about size, shape or orientation features of the glass and the pen (see Figure 2). For example, RLs for LEFT/RIGHT in NGT involve the finger-spelled handshape of the initial letters (*L* vs. *R*). At the same time, the hand moves to the left or right corresponding to the direction in which the figure is located in respect to the ground (Figure 2D). However, there is also cross-linguistic variation across sign languages in how RLs are represented. For example, in Turkish Sign Language RLs for LEFT/RIGHT involve the tapping of the right or left upper arm (Karadölller, Sümer, & Özyürek, in press; Özyürek et al., 2010; Perniss et al., 2015; Sümer, 2015). Overall, this categorical characteristic of RLs makes them more alike to the spoken forms such as LEFT/ RIGHT. Yet, RLs retain visual iconicity concerning the iconic resembling of relations in real space to the signing space. Overall, the types of iconicity in spatial expressions can differ in sign languages by using space and object resemblance through CLs and space resemblance only through RLs.

Figure 2

Iconic NGT expressions to describe "the pen is to the left of the glass" by using relational lexemes (RLs).



Panel (A) shows a pen that is located to the right of the glass. Panel (B) shows the introduction of the lexical sign for the ground object GLASS. Panel (C) shows the introduction of the lexical sign for the figure object pen. Panel (D) shows RLs where the handshape fingerspelling for L moves to the left side in space to indicate the location of the pen in relation to the signers' body. RH stands for the right hand, LH stands for the left hand.

In summary, while the vocal modality in spoken languages does not allow visual resemblance between form and meaning and express spatial relations in categorical ways, in sign languages the visual motivated form-to-meaning mappings are a universal feature, which is especially prominent in the domain of spatial language. These iconic forms albeit can vary in their degree of visual resemblance to the spatial relations described (i.e. degree of iconicity) ranging from space and object resemblance in CLs to only space resemblance in RLs. Through the use of CLs and affordances of the visuo-spatial modality, sign languages can also express more semantically specific information about objects size, shape and orientation than most spoken languages, where semantic specificity and ambiguity can vary. Spoken and sign languages therefore differ radically in the way they express spatial relations.

In addition to iconicity, there seems to be another modality-driven difference in how spoken and sign languages describe spatial relations, namely, in terms of the order in which ground and figure objects are introduced in spatial descriptions.

2.2 Word order preference

Spoken languages. In addition to the above-mentioned diversity found in spatial relations across spoken languages, there are also differences in the order in which ground and figure objects can be expressed. However, there is little comparative research in this area. For example, while languages like Korean and Turkish predominantly mention first the ground (i.e. glass) followed by mentioning the figure (i.e. pen; see among others Ahn, Gollan, & Ferreira, 2019 for Korean; Sümer, 2015 for Turkish), languages like English, however, predominantly prefer to mention the figure first, (i.e. pen), followed by the ground (i.e. glass; e.g. Ahn et al., 2019). Some other languages have more mixed options, such as Dutch, which might not have a pre-set linguistic word order for describing spatial relations. Rather, both figure-first and ground-first can be valid and acceptable word orders (Hartsuiker, Kolk, & Huiskamp, 1999).

Sign languages. Unlike spoken languages, sign languages across the globe and studied to date show, homogeneity when it comes to the order in which figures and grounds are mentioned in spatial descriptions. Signers universally prefer mentioning first the lexical sign for the ground (Figure 1B on page 14 & Figure 2Bon page 24) followed by the lexical sign for the figure, which are then followed by CLs or RLs (Figure 1C on page 14 & Figure 2C on page 24; see among others Emmorey, 1996 for American Sign Language; Morgan, Herman, Barrière, & Woll, 2008 for British Sign Language; Sümer, 2015 for Turkish Sign Language; Perniss, 2007 for German Sign Language; and Kimmelman, 2012 for Russian Sign Language).

This preference for ground-first has been suggested to be driven by the affordances of the visuo-spatial modality (Kimmelman, 2012; Perniss et al., 2015). That is, ground-first seems to reflect that grounds are most conceptually salient in the manual modality (Perniss et al., 2015). This is in line with Gestalt and linguistic conceptual theories that identify grounds as bigger and more stable as well as permanent and backgrounded objects (e.g. Rubin, 1958). Supporting this claim, previous research has demonstrated that when even speakers are prohibited of speaking and are instead asked to silently gesture about spatial relations they also show a clear preference of ground-first order (Goldin-Meadow, So, Özyürek, & Mylander, 2008; Laudanna & Volterra, 1991), even though they might prefer another word order when describing similar pictures through speech in their spoken language. This can be due to a cognitive bias in perceiving relations between two objects and having to map them onto space using the visuo-spatial modality.

In sum, the evidence provided in this section so far shows that sign languages can differ in expressing spatial relations from spoken languages, both in terms of iconicity of the expressions as well as in their order of figure and ground mention in spatial descriptions. These differences arise due to specific affordances of the visuo-spatial modality of sign languages. Only recently, it has been debated whether these modality-specific expressions, in fact, matter and influence cognitive processing differently in signers compared to speakers.

2.3 Effects of iconicity and modality on cognition and processing

Previous studies that investigated consequences of iconicity for language processing have focussed on the lexical level (i.e. imagistic iconicity) and mostly looked at its consequences for comprehension processes and word learning (Anderson & Reilly, 2002; Bosworth & Emmorey, 2010; Orlansky & Bonvillian, 1984; Poizner, Bellugi, & Tweney, 1981). Earlier studies have claimed that iconicity does not matter and that it does not facilitate language processing or learning. For example, it has been shown that highly iconic signs are not remembered better than arbitrary signs (Poizner et al., 1981) and that iconic signs are not learnt earlier than arbitrary signs (e.g. Anderson & Reilly, 2002).

However, other studies show the opposite and claim that iconicity does impact sign language processing and learning (e.g. Campbell, Martin, & White, 1992; Grote & Linz, 2003; Navarrete, Peressotti, Lerose, & Miozzo, 2017; Sümer, 2015; Thompson, Vinson, & Vigliocco, 2010; Thompson, Vinson, Woll, & Vigliocco, 2012). For instance, iconic signs are comprehended faster (e.g. Grote & Linz, 2003; Thompson, Vinson, & Vigliocco, 2010) or learnt earlier by children (Thompson et al., 2012) than arbitrary signs.

Recently, work from both spoken and sign languages view iconicity as a more dynamic and not binary property. In these recent views, iconicity involves semiotic relations and manifests itself in various forms and degrees and involves different types of form-to-meaning mappings (e.g. Bellugi & Klima, 1976; Dingemanse et al., 2020; Ortega, Sümer, & Özyürek, 2017; Perniss et al., 2015). For example, deaf children preferred iconic lexical forms that were action-based (e.g. a writing hand for the sign PEN) over other variants that for example resembled perceptual features of objects (e.g. index finger to resemble the thin, elongated shape of the pen, Ortega et al., 2017). These studies showed that the *type* of iconicity also matters for language processing. Further supporting this claim, there is neuroimaging evidence showing that the planning of signed descriptions with different types of iconicity (space and object resemblance in CLs vs. space resemblance in RLs) recruits different brain areas (parietal regions for during CL planning vs. left hemisphere language regions for RL planning, see Emmorey, McCullough, Mehta, Ponto, & Grabowski, 2013).

In conclusion, there are a handful of studies assessing whether iconicity can affect cognition and language processing. Furthermore, previous research has mostly focussed on language comprehension or learning. This evidence is also limited to iconicity at the lexical level. In this thesis, I aim to go beyond this previous literature by looking at different types of iconicity (as in CLs vs. RLs) as well as at the sentence level (i.e. diagrammatic iconicity) and its possible effects on cognition.

3 Language diversity in bilinguals and cognition

The previous evidence presented so far shows that language diversity across spoken and sign languages can shape cognition across speakers of different languages. However, less is known about how language diversity manifests itself in bilinguals and how this might shape cognition. Even less is known about this topic in bimodal bilingual individuals who can both speak and sign. To assess these gaps, I will test bimodal bilinguals as a new window into bilingual diversity and its possible effects on cognition. In doing so, I will assess whether the diversity found in bilinguals resemble diversity found across different populations that either speak or sign. Bilinguals demonstrate a different type of language diversity because often bilinguals' two languages shape each other (e.g. Dijkstra, 2005; Kroll, Dussias, Bogulski, & Kroff, 2012). In the next section, I will talk about these influences and how they might shape cognition as well as introduce bimodal bilinguals.

From a large body of research on spoken languages, it has been widely recognized that bilinguals are not two monolinguals in one person and that the co-existence of two languages creates a unique language user (see also Grosjean, 1989). Often the two languages influence each other due to the mental activation of these languages during both language production and comprehension (e.g. Blumenfeld & Marian, 2005, 2007, 2013; Costa, 2005; Dijkstra, 2005; Kroll, Bobb, & Wodniecka, 2006; Kroll & Gollan, 2014). Due to this *cross-linguistic influence* (also known as *cross-linguistic transfer)* one language can influence the structure or concepts of the other language or both languages can converge and become more similar to each other than when comparing each language to that of monolinguals (e.g. Brown & Gullberg, 2008; Indefrey, şAhin, & Gullberg, 2017; Pavlenko & Jarvis, 2002; Pavlenko & Malt, 2011). Note that cross-linguistic influence or convergence might not always be present but instead, bilinguals can also maintain the specific patterns of each language (e.g. Ahn et al., 2019; Azar, Özyürek, & Backus, 2019; Daller, Treffers-Daller, & Furman, 2011; Athanasopoulos et al. 2015).

This bilingual diversity also raises questions on whether bilinguals conceptualise aspects of the world differently than language users that only speak one language. Some findings show that bilinguals process non-linguistic spatial tasks flexibly depending on the linguistic environment the bilinguals are in (e.g. different language modes due to different language of instruction; Athanasopoulos et al., 2015; Kersten et al., 2010). For example, when Spanish-English second language learners (*L2*) judged similarities of motion events, bilinguals flexibly shifted their classification preferences based on the language in which they were tested (Lai, Rodriguez, & Narasimhan, 2014). Thus bilinguals' cognition was similar to differences found between monolingual speakers of each language. This reflected the cross-linguistic differences between Spanish and English in terms of which aspect of the motion is encoded in the main verb; that is, the path of motion in Spanish and the manner of motion in English.

Other findings, however, indicate that bilinguals' cognition converges into a new system and differs from that of the monolingual peers (Wolter, Yamashita, & Leung, 2020). For example, Japanese and English differ in how they express narrowness. While in Japanese there is only one broader adjective "semai" that expresses small spaces in general, the English equivalent "narrow" refers exclusively to the width. When judging the narrowness of rooms English monolinguals judged narrowness based on the width and Japanese monolinguals in terms of differences in overall area. English-Japanese late L2 bilinguals, however, judged rooms in an in-between pattern which differed when being instructed in English versus Japanese and also differed from that of their monolingual peers.

The studies described so far looked either at cross-linguistic influence or the processing of non-linguistic tasks. To my knowledge, these two approaches have not been directly combined to assess whether cross-linguistic influences found in linguistic productions can further influence conceptualization during the preparation of these messages in individuals who use different languages. To offer new insights into bilingual diversity and its effect on cognition, the current thesis uses such a combined approach. In addition, it also investigates bimodal bilinguals as a unique population that uses languages from different modalities to study language diversity.

Bimodal bilingualism. Bimodal bilinguals[4] form a special group of hearing bilinguals because they have early exposure to a spoken and sign language and are thus highly proficient in two languages of different modalities. In childhood, bimodal bilinguals typically interact with their deaf parents in a sign language and learn the spoken language from their hearing relatives and the surrounding hearing community. In most of these contexts, sign language is a minority language because it differs from the spoken language used most commonly in the community (i.e. majority language). Similarly to other bilingual contexts involving pairs of spoken languages, the language

[4] Note that deaf signers can also be considered to be bimodal bilinguals as in addition to being proficient in a sign language in some countries (but not all) they are also exposed to a secondary code (print) of the spoken language (see Lillo-Martin, de Quadros, Chen Pichler, & Fieldsteel, 2014). That is, in some countries deaf individuals can learn the sign language from their deaf parents at home and learn the spoken language in school mostly through the medium of written language. In this thesis, however, I distinguish between bilinguals' statuses for hearing and deaf bilinguals and do not equate them. While deaf bimodal bilinguals have learnt spoken/written language through instruction at school, hearing bimodal bilinguals have acquired both languages naturally in the home environment and have more comparable access to both languages. This thesis thus focuses on hearing bilinguals.

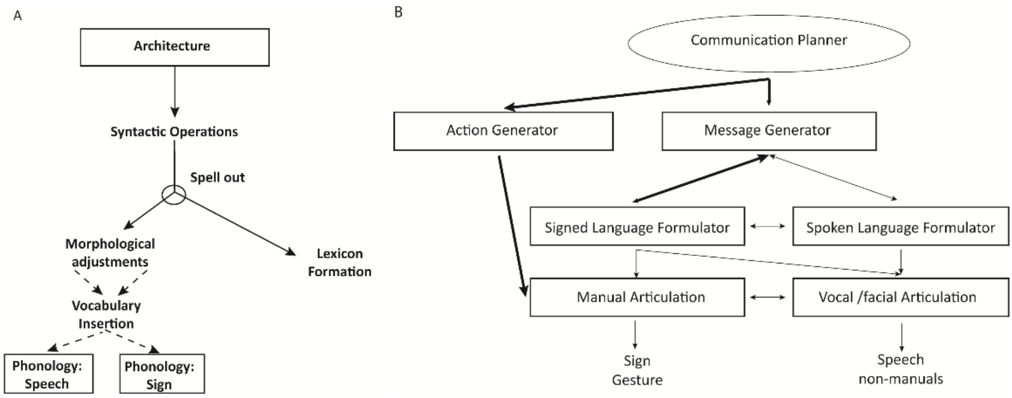
used at home (i.e. the sign language) is referred to as the so-called *heritage* language. Thus, bimodal bilinguals are heritage signers and have two different cultural and linguistic experiences (De Quadros, 2018; Pichler, Lillo-Martin, & Palmer, 2018; Quadros & Lillo-Martin, 2018). Investigating language diversity and cognition in this group enables me to explore cross-linguistic influences between languages of different modalities and their influences on cognition *within* the same individuals comparing them to the variation found *across* different language user groups. Studying the use of a spoken and a sign language within the same individuals, that is bilinguals, raises the issue of whether these two languages influence one another or whether they look similar to those used by different populations such as hearing individuals who do not know a sign language or deaf individuals who have restricted or no access to a spoken language.

Unlike the rigorous amount of research on cross-linguistic influence between two spoken languages; that is, *within* a single modality, it is less known whether such influences can occur between speech and sign; that is, *across* modalities. Theoretically, cross-linguistic influences across modalities could occur because also in bimodal bilinguals the two languages are known to be activated during language production and comprehension (e.g. Emmorey, Borinstein, Thompson, & Gollan, 2008; Giezen, Blumenfeld, Shook, Marian, & Emmorey, 2015; Giezen & Emmorey, 2016; Shook & Marian, 2012). Further, it has been shown that sign can influence the manual modality accompanying speech; that is, the occurrence of co-speech gestures. For example, bimodal bilinguals used more iconic co-speech gestures than their hearing speaking non-signing peers (Casey & Emmorey, 2009; Weisberg, Casey, Sevcikova Sehyr, & Emmorey, 2020). Bimodal bilinguals are also found to often produce signs while speaking, such as signing the sign for GLASS when uttering the word GLASS in speech. This simultaneous production of signs and words is referred to as *code-blends* (e.g. Emmorey, Borinstein, & Thompson, 2005). Code-blending is a behaviour unique to bimodal bilinguals due to the possibility of using two different main articulators through which the two languages can be produced. Yet, these previous findings have shown only influences within the same modality; that is, from speech to speech or from hands (sign) to hands (co-speech gestures). No study so far has explored whether cross-linguistic influences across modalities are possible in bimodal bilinguals or whether expressions in speech and sign in bimodal bilinguals are the same as in those used by different hearing and signing populations.

There have been two theoretical models proposed to explain cross-linguistic influence in bimodal bilinguals' production and the occurrences of code-blends. First, the Language Synthesis model (Lillo-Martin, de Quadros, & Pichler, 2016) claims to account for cross-linguistic influence in any language pair in the same way because it is simply generated by one architecture that allows features from different languages to enter into one single source (i.e. *language synthesis*). Second, the model by Emmorey et al., (2008) (adapted from models by Kita & Özyürek, 2003; Levelt, 1989, see Figure 3B) focuses on the processing aspects of cross-linguistic influences and contains a shared *Message Generator* (i.e. pre-verbal message generation) and separate but interacting production systems, so-called *Formulators*, for sign and speech. Within-modality influences (sign to co-speech gestures) have been suggested to occur at the level of the Message Generator. Specifically, this Message Generator is the place where sign languages frequently encode spatial information through iconic representations and can consequently influence co-speech gestures through the Action Generator (i.e. gesture generation) (see bold lines in Figure 3B; Casey & Emmorey, 2009). However, there is yet no empirical evidence for possible influences between speech and sign, even though this is allowed by the model.

Figure 3

Bimodal bilingual language production models.



Panel (A) illustrates an adapted figure from Lillo-Martin et al.'s (2016) model. Panel (B) illustrates an adapted figure from Emmorey et al.'s (2008) model. Bold lines in Panel (B) illustrate the suggested route for the occurrence of within-modality influences (sign to co-speech gesture; Casey & Emmorey, 2009).

In sum, bilinguals are found to experience cross-linguistic influence between their two spoken languages and this has not been shown between different modalities in bimodal bilinguals. Furthermore, to my knowledge, there is no study so far that has investigated whether message conceptualization for language production differs in bilinguals. The current thesis aims to fill these gaps by looking at bimodal bilinguals to enhance our notion about potentials and constraints of bilingual diversity. While there appears to be cross-linguistic influence between sign and co-speech gestures it has yet to be asked whether such influences can occur between sign and speech. It is further unknown whether being bimodal bilingual has consequences for cognition, specifically on the conceptualization of messages relevant for language production and how they compare to differences found between spoken and sign language users from different hearing and signing populations.

4 The current thesis

In this thesis, I intend to bridge the above-described types of literature together, namely, language and cognition, sign languages, and bilingualism. This interdisciplinary and multimodal approach of all three research areas will enhance our understanding of how language diversity affects cognition, especially during conceptualization during message preparation. To do so, I will present four elicited production and eye-tracking studies, which were all designed to examine how differences in modality can influence production and cognition in different speaking and signing populations as well as how those manifest themselves in the same individuals; that is, in bimodal bilinguals.

4.1 Research questions

With this thesis, I aim to reveal three overarching questions:

(1) To what extent do differences in modality influence **production** across different speaking (hearing) and signing (deaf) populations (in the case of spoken Dutch and NGT)?

(2) Do these modality-driven differences in production patterns found in (1) influence **cognition**, as measured by visual attention to referents during message preparation?

(3) How do patterns observed in (1) and (2) across different speaking (hearing) and signing (deaf) populations manifest themselves in bilinguals (hearing); that is, in the same individuals who can both speak and sign (i.e. bimodal bilinguals)?

To answer these questions, especially regarding the effects of language modality on cognition (2), I focus on the conceptualization of messages that takes place during message preparation. As claimed in Levelt's (1989) language production model, speaking begins with the conceptualization of general details of an event when preparing to speak. This message conceptualization can include information about people, objects, places, time and relations among them and is linked to the linguistic constraints on how this information is packaged into different linguistic structures. Consequently, speakers of different languages attend to different aspects of the world that they intend to talk about (i.e. *thinking for speaking*, Slobin, 2003). I assess possible differences between using a spoken versus a sign language in message conceptualization by measuring visual attention to spatial relations before describing left/right relations.

I chose the domain of expressing spatial relations as this is an area where differences in modality between spoken and sign languages have been demonstrated to be most explicit (e.g. Emmorey, 2002; MacSweeney et al., 2002; Perniss et al., 2015) as described before in section 2. Furthermore, I focus specifically on left/right relations for two reasons. First, due to their requirement to take a perspective they can be mapped onto the body and space and are contrastive to each other, which allows me to study visual attention effects more easily. They then can reveal differences in the conceptualization of messages between speakers and signers more clearly. Secondly, left/right relations do not necessarily involve occlusion of the figure objects in contrast to in/on/behind relations, where the figure object touches or partly hides behind the ground object and thus are more appropriate to measure visual attention. Thus, left/right relations are especially suitable for measuring visual attention prior to language production.

I explore both language production and visual attention within a single experimental paradigm that allows me to assess languages users' eye gaze before they produce their spatial expressions and link them to their different linguistic choices within and across languages. In particular, I use a so-called *Visual World Eye-Tracking Paradigm* which is typically used to measure language comprehension (for a review, see Huettig, Rommers, & Meyer, 2011). I use this paradigm for language production in a novel way. I present participants with four-picture displays during the message preparation phase, but as different from previous research, each picture contains the same two objects but in different spatial relations to each other (e.g. left, right, front, in) because I am interested in relational encoding. After an initial viewing phase of the four pictures, participants see a visual cue in the form of an arrow indicating the picture they have to describe, the so-called *target picture*. This visual cue then disappears and the four pictures remain on the screen until it was time to describe (i.e. speaking/signing) the target picture. I measure visual attention during the conceptualization phase during message preparation which corresponds to the viewing phase after the arrow disappeared and until participants described the target picture.

Assessing differences in visual attention to target pictures and other competitive pictures in the four-picture displays allows me to operationalize differences in message preparation. For example, to use iconic expressions of spatial relations (e.g. having to place a figure object to the left or right in the signing space and to map the orientation of the objects shown in the display), signers might need to allocate more attentional resources to the relative locations and/or physical features of the objects to pictures other than the target picture in relation to each other than when using a spoken language. Consequently, deaf signers might look more at non-target pictures than hearing speakers. Also considering differences in word order preferences across spoken and sign languages, having to prepare messages for ground-first mentions might attract more attention to ground objects during message preparation in deaf signers compared to hearing speakers.

4.1 Participants

To answer the three overarching questions mentioned above, I used data from the same participants that can be divided into three populations: (a) *hearing speakers* of Dutch (N = 20, **Chapters 2-5**), (b) *deaf signers* of NGT (N = 20, **Chapters 2-5**), and (c) NGT-Dutch hearing *bimodal bilinguals* (N = 21, **Chapters 3-5**). Comparing data from (a) and (b) allows me to make comparisons of language production and cognition across populations. Comparing these to (c) allows me to look into bilingual diversity. All participants were born and raised in the Netherlands by Dutch parents but differed in their linguistic background.

I determined sample size based on previous research (e.g. Casey & Emmorey, 2009; Pyers & Emmorey, 2008; Weisberg et al., 2020) that used sample sizes of bimodal bilingual population lower than N = 15 per group. Since the Netherlands is a small country and many bimodal bilinguals do not acquire NGT I aimed to increase this previously used sample size out of convenience to 21 bimodal bilinguals and determined the sample size of the hearing speakers and deaf signers accordingly.

In (a), hearing speakers acquired Dutch as their only native language and learnt additional spoken languages (mostly English, German, and French) only later in instructional settings. Importantly, none of the speakers had any exposure to NGT or any other sign language. In (b), deaf signers were born to and raised by deaf parents in the Netherlands. They therefore acquired NGT from birth at home from their deaf parents and are as such, the most native NGT signing population available in the Netherlands. Yet, they have access to written Dutch when entering school around age four through formal instruction. NGT was the primary language of instruction. In (c), bimodal bilinguals were hearing bilinguals born to at least one deaf parent and thus acquired NGT at home from their parents and acquired Dutch in the surrounding hearing community without formal instruction.

Thus, the first two groups, namely hearing speakers (a) and deaf signers (b), have no or limited access to the other spoken or sign language. In contrast, the other group (c), namely bimodal bilinguals, have full access to both a spoken and a sign language and acquire those without formal instructions. Comparing the use of a spoken versus a sign language also forms the basis to detect differences in language production and cognition across populations. These groups are considered or labelled as “*controls”* in

Chapters 3-5 when I investigate possible cross-linguistic influences in bimodal bilinguals in their language production and/or cognition.

A complete overview about the language background and other specific characteristics of hearing speakers and deaf signers can be found in **Chapter 2** and of bimodal bilinguals can be found in **Chapter 3**.

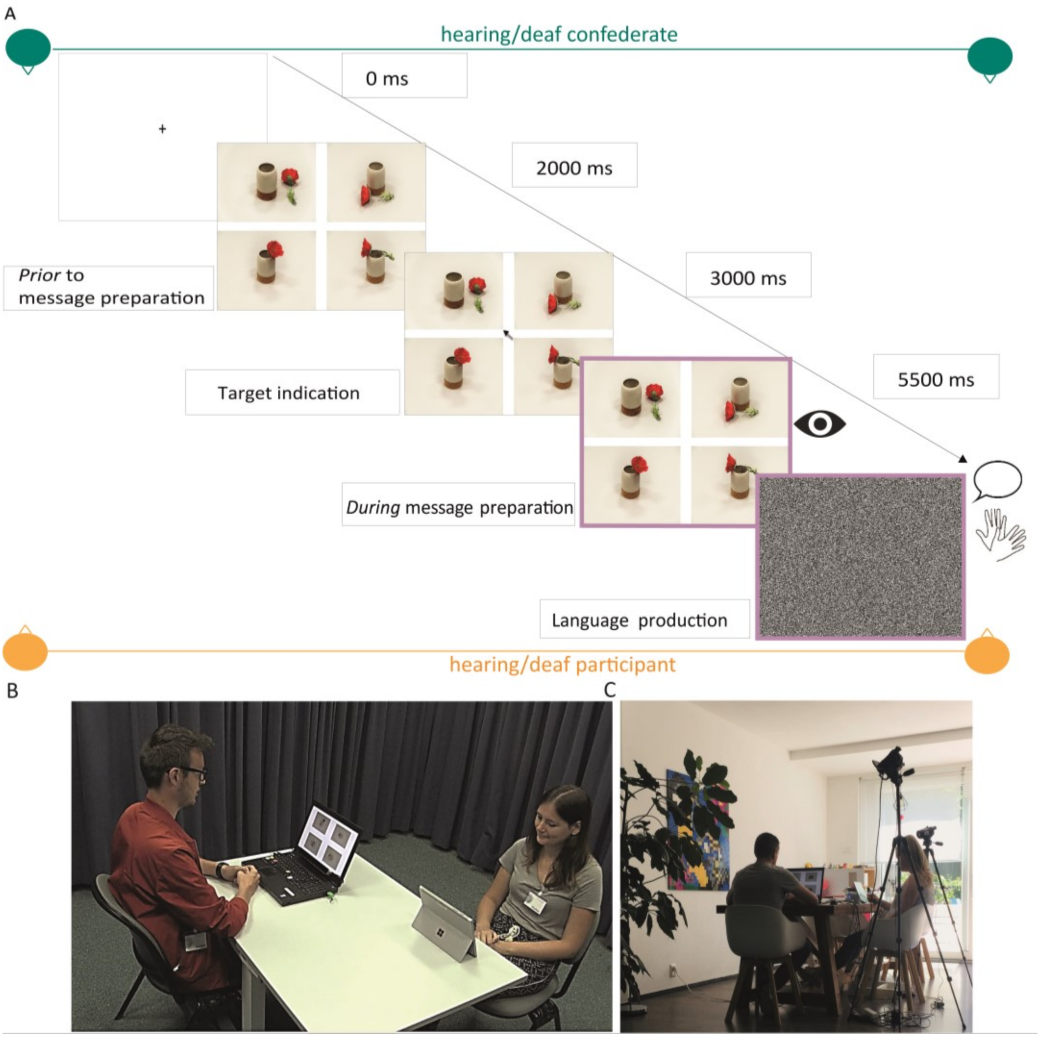
4.2 Method and materials

I examined picture descriptions as a measure for language production. In addition, I assessed visual attention during message preparation of these picture descriptions as a measure for message conceptualization. The two aspects are assessed within the same experimental paradigm.

As briefly mentioned in section 4.1, I adapted the Visual World Eye-Tracking Paradigm to study visual attention for language production. As mentioned above, this paradigm has been designed to measure how language comprehension interacts with attention to visual referents (for a review, see Huettig, Rommers, & Meyer, 2011). This novel paradigm is the first of its kind in sign language production research. It allows exploring both sign language production and visual attention within a single experimental paradigm (see Figure 4). Furthermore, the paradigm can be used for both speaking and signing populations in a communicative, yet controlled, face-to-face setting by involving a confederate (trained addressee).

Figure 4

Visual World Language Production Eye-Tracking Paradigm used in all chapters.



Panel (A) illustrates the timeline of the Visual World Language Production Eye-Tracking Paradigm. Panel (B) shows an experimental set up in the lab in which the participant sits in front of the eye-tracker (eye-tracker is mounted underneath the laptop screen) across the confederate with a tablet. Panel (C) shows an example of the experimental set-up at a bimodal bilinguals' home.

As in a traditional Visual World Paradigm set-up, I used displays containing four pictures. Importantly, I presented two referents; that is, static objects, which were in different spatial relations to each other (e.g. a pen that is in/on/to the left/in front of a glass; see Figure 4A). In the displays of interest, participants always had to describe left/right spatial relations. The remaining three pictures of the display involved other spatial relations, such as front, behind, in, or on. Displays varied in complexity by including contrastive spatial relations, that is, including left *and* right relations within the same display (i.e. more complex displays), or either left *or* right relations within one display (i.e. less complex displays).

Thus this paradigm uniquely includes visual complexity by involving a 2x2 grid in which each picture included two objects in different spatial configurations. As such, it differs from previous studies that used pictures with a single referent or object (e.g. Altmann & Kamide, 2009; Cooper, 1974; Davies & Kreysa, 2017; Ferreira, Foucart, & Engelhardt, 2013; Hintz, Meyer, & Huettig, 2019; Huettig & Altmann, 2005; Lieberman, Borovsky, Hatrak, & Mayberry, 2015; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Thus as such, the paradigm provides many possibilities to expand our knowledge onto the relationship between spatial language production, both spoken as well as signed and visual attention. Because descriptions of these complex visual displays were combined with the eye-tracking technology, the paradigm required rigorous experimental control. In particular, involving spatial configurations required a careful balancing act of distributing the different spatial configurations (left, right, front, behind, in and on) across the different displays. Overall, this balancing was necessary to avoid any biases to a certain type of spatial relations.

All three participants groups, that is, hearing speakers, deaf signers and bimodal bilinguals, were tested using the same experiment and materials. A complete overview of the experimental paradigm and materials will be given in **Chapter 2.**

4.3 Procedure

Participants were seated in front of a laptop at which the eye-tracker was mounted (Figure 4B, for details about the eye-tracker and sampling rate used, see Chapter 2). First, participants had some time to look at the four pictures until a visual cue in form of an arrow pointed at one of the pictures, namely the *target picture*. Using a visual cue instead of the commonly used auditory cue in spoken sentences makes the paradigm suitable for both hearing and deaf participants.

The arrow disappeared after a short amount of time and the four pictures remained on the screen. During that time, eye-gaze was measured to assess visual attention during message preparation (Figure 4A). Once a grey noise screen appeared, the participants' task was to describe the target picture to a confederate. These descriptions served as a measure for language production. At this time participants were free to move to describe the pictures. This approach to assess visual attention before describing pictures or movie clips is inspired from previous research that investigated whether cross-linguistic differences to describe motion events also guides their visual attention to these events differently (e.g. Flecken et al., 2015; Papafragou et al., 2008; Trueswell & Papafragou, 2010).

Once a picture description was given, the confederate pretended to choose the correct picture out of four identical pictures on a separate tablet. After this description-matching routine, participants initiated the next trial by pressing the ENTER button and were asked to sit still again until describing the next picture. The entire session was video-recorded using two cameras (frontal view and top view, Figure 4 C) to later code participants' descriptions. While hearing speakers and deaf signers participated once in the eye-tracking experiment, bimodal bilinguals were tested twice; that is, once in each language. Furthermore, different confederates were present during spoken and signed sessions. That is, during Dutch sessions the confederate was always a hearing

Dutch native sign-naive speaker. During NGT sessions, the confederate was a deaf native NGT signer.

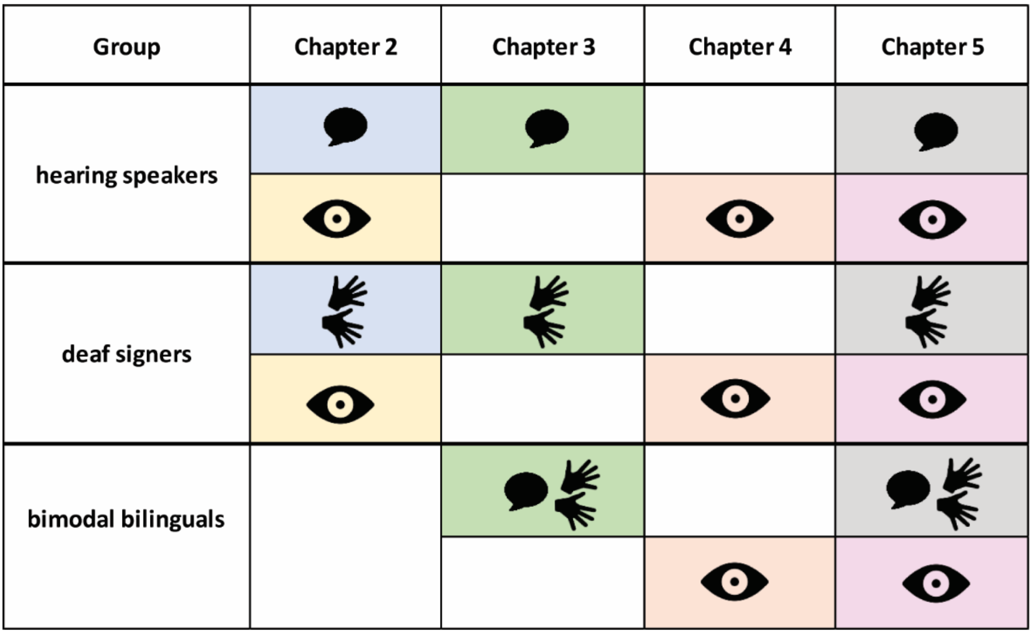
A complete overview of the procedure will be given in **Chapter 2** for testing hearing speakers and deaf signers and **Chapter 3** for testing bimodal bilinguals.

4.4 Outline of the thesis

In all experimental chapters (**Chapters 2-5**), I assessed to what extent differences in modality influence the production of spatial language in spoken Dutch and NGT. In addition to language production, I assessed in **Chapters 2, 4 and 5** to what extent differences in modality influence message conceptualization by measuring visual attention to referents during message preparation. While **Chapters 2-4** focussed on effects of iconicity, **Chapter 5** focussed on word order preference. As previously described, I used the same experimental paradigm to measure both language production and visual attention between populations (see Figure 5 for an overview); that is, hearing speakers and deaf signers (**Chapters 2-3 and 5**) as well as within subjects; that is, bimodal bilinguals who can both speak and sign (**Chapters 3-5**).

Figure 5

Overview of comparisons across groups of language users used throughout this thesis.



The graphic illustrates the different groups; that is, hearing speakers, deaf signers and bimodal bilinguals as well as methods used across the 4 experimental chapters, namely, language production (spoken and sign, indicated by the speech bubble and hand symbol respectively) and message conceptualization by measuring visual attention (indicated by the eye symbol). Cells with the same colour indicate which groups were compared with which measures (speech/sign/visual attention).

The general premise of the thesis was to find out how differences in modality influence cognition for language production and more specifically conceptualization during message preparation as measured by their visual attention to spatial relations before producing messages. In general, this would be against the assumption that language does not influence cognition. However, if language diversity influences cognition then I would expect to find that different types of expressions between spoken and sign languages, as well as subtle differences within a sign language (e.g. different types of iconicity in CL vs. RL), would influence conceptualization of messages. I also expect these effects to manifest themselves differently across (hearing speakers and deaf signers) versus within subjects (i.e. bimodal bilinguals) due to cross-linguistic influences (alternative hypotheses will be covered in more detail in Chapter 6). Furthermore, I assess whether this possible cross-linguistic influence across modalities can also change bimodal bilinguals' cognition compared to that of different speaking and signing populations.

In **Chapter 2,** I first examined the influence of iconicity on spoken and signed productions of left/right relations (research question (1), page 34) by comparing spatial descriptions between hearing speakers and deaf signers; that is, across populations. I also assessed within deaf signers to what extent they employ different types of iconicity, such as those that map both spatial and physical properties of objects (CLs) or only spatial properties (RLs; see also e.g. Emmorey, 2002; Emmorey & Herzig, 2003; Perniss et al., 2010). Secondly, I examined whether and how using different types of linguistic expressions influences message conceptualization (research question (2), page 34) by assessing how hearing speakers and deaf signers allocate their visual attention to spatial pictures during message preparation of the target picture (see section 4.2). If iconicity is relevant for message conceptualization, then deaf signers, but not hearing speakers, would look more at other pictures than the target picture to map orientation and shape of the objects and to place them onto the sign space. I also assessed within signers to what extent different types of iconic expressions within a sign language modulate visual attention. This has the advantage that possible differences between hearing speakers and deaf signers cannot be simply attributed to different hearing statuses (hearing vs. deafness) but is rather specific to using iconic expressions. In sum, with **Chapter 2** I aimed to reveal that *thinking for signing* differs from *thinking for speaking* due to the iconic expressions in spatial language and that linguistic diversity even within a sign language can shape visual attention and thus conceptualization during message preparation.

In **Chapter 3,** I examined whether bimodal bilinguals show similar production patterns as found across different speaking and signing individuals in Chapter 2 (research question (3), page 34) or if they speak and sign differently providing evidence for cross-linguistic influence across modalities. Thus in this chapter, I investigated whether cross-linguistic influence is present across modalities in two directions, namely, between speech and sign, as well as between co-speech gestures and sign as found in previous research (Casey & Emmorey, 2009; Gu, Zheng, & Swerts, 2018). Although plenty is known about how bi-directional influence can occur across spoken languages it is unknown whether and how such influences can occur across modalities. To explore this, I compared language production patterns of bimodal bilinguals in each of their languages to that of their hearing speaking and deaf signing peers (who are reported in Chapter 2). I explored various types of influences: (1) from sign to speech, (2) from sign to the manual modality accompanying speech (co-speech gestures and signs accompanying speech), and (3) from speech to sign. I hypothesised that iconic expressions in sign and arbitrary and categorical expressions in speech might shape each other in a way that bilinguals' productions will differ from that of their hearing speaking and deaf signing peers. I predicted that speech and sign influence each other bi-directionally. More specifically, as expressions in sign are more semantically specific due to iconic forms I predicted that this might influence and increase semantically specific expressions in speech. In return, using arbitrary categorical expressions might decrease iconicity in signed expressions. Nevertheless, rather than experiencing cross-linguistic influence bimodal bilinguals might also maintain their language-specific patterns as previously shown in proficient bilinguals (e.g. Ahn et al., 2019; Athanasopoulos et al., 2015; Azar, Özyürek, & Backus, 2019). Therefore, in this chapter, I aimed to reveal whether and to what extent cross-linguistic influence across modalities can shape bilingual diversity and how this diversity compares to that found across populations.

In **Chapter 4**, I examined whether bimodal bilinguals' visual attention also differs from those of the controls (as found in Chapter 2) when preparing to speak versus to sign. That is, I investigated to what extent message conceptualization can be changed within individuals as a consequence of bi-directional cross-linguistic influence shown in Chapter 3. Although previous research has suggested that bilingual language diversity can influence cognition, studies have separately investigated either cross-linguistic influence or cognitive processes during non-linguistic tasks but not whether conceptualization for messages differs in bilinguals. The thesis is the first to combine these two approaches by assessing how productions, which are shaped by cross-linguistic influence, can affect cognition. To explore this, I assessed bimodal bilinguals' visual attention during spoken versus signed message preparation and then compared those to that of hearing speakers and deaf signers (similar to and reported in Chapter 2). I predicted that if bimodal bilinguals' speak and sign about spatial relations differently in Chapter 3, then this might influence how bimodal bilinguals visually attend to these spatial relations when preparing to speak versus to sign compared to that of the controls. Therefore, this chapter aimed to investigate if cross-linguistic influence across modalities found in Chapter 3, has further consequences for cognition and more specifically influences message conceptualization in bimodal bilinguals.

Finally, in **Chapter 5,** I wanted to investigate if and how another modality-driven difference across spoken and sign languages; that is, word order preference, influences production patterns and conceptualization during message preparation across populations and in bimodal bilinguals. I aimed to show that the effects shown in Chapter 3 and 4 are not restricted to iconicity, but can also occur in a different domain, such as word order. Rather than comparing first across populations and then in bilinguals (as in Chapters 2, 3 and 4), in this chapter, I compared all 3 populations in one study in the domain of word order preference.

For production, I compared bimodal bilinguals' descriptions of the target picture in each of their languages to that of their hearing speaking and deaf signing peers. Descriptions were assessed with respect to whether grounds or figures were mentioned first to examine which word order all three populations preferred. I hypothesised that hearing and deaf populations would differ; that is, deaf signers would consistently use ground-first but hearing Dutch speakers would use both types of word order. If there is cross-linguistic influence from sign, then bimodal bilinguals would mention grounds first more often than figures in their speech compared to their hearing speaking peers. Thus, the modality-specific word order in sign might shape speech, as I have predicted for the influence of iconic expressions on speech in Chapter 3. Regarding a reverse influence, I did not expect a strong effect. In particular, if word order in sign is the modality-driven order and grounded in cognitive perceptual biases, it might be more resistant to be influenced from the spoken language. Thus when signing bimodal bilinguals might not differ from their deaf signing peers, unlike predicted in Chapter 3. Alternatively, Dutch as majority language (i.e. the most commonly used language in The Netherlands) might influence word order preferences in NGT, the minority language (i.e. the less commonly used language in The Netherlands; as evident from spoken languages in various domains, e.g. Backus, 2005; Muysken, 2000; Polinsky, 2008). However, the recruited bimodal bilinguals in this thesis are highly fluent in both languages. Consequently, they might maintain their language-specific patterns in each language (for similar findings in spoken language bilinguals, see e.g. Ahn et al., 2019; Athanasopoulos et al., 2015; Azar, Özyürek, & Backus, 2019).

Next, I examined the influence of word order preference on message conceptualization by comparing visual attention to grounds and figures within the target picture during spoken versus signed message preparation in all the three groups. My basic assumption was that language users allocate more visual attention to those aspects of a visual scene that are mentioned first within an utterance (e.g. Griffin, 2004; Griffin & Bock, 2000; van de Velde, Meyer, & Konopka, 2014). When comparing hearing speakers to deaf signers, I expected deaf signers to attend more to grounds than hearing speakers. Furthermore, if cross-linguistic influence is found from sign to speech, I also expected bimodal bilinguals to visually attend to grounds more than their hearing speaking peers when preparing to speak. However, when preparing to sign I expected that their visual attention would not differ from those of the deaf signers. Thus, I predicted that in bimodal bilinguals, the one-way cross-linguistic influence from sign to speech, if found, will change how bimodal bilinguals visually attend to grounds versus figures when preparing to speak but not when preparing to sign. Therefore, in this chapter, I aimed to reveal (as in Chapter 3) that bilinguals' language production can be shaped by cross-linguistic influence across modalities in the domain of modality-driven word order preference and can at the same time shape conceptualization of messages.

Finally, in **Chapter 6,** I bring together the results of all the experiments in this thesis. I also reflect on the thesis' contribution to our broader understanding of how linguistic diversity can affect cognition based on differences in modality. I also provide a discussion of the implications of this work for (bimodal) language production models. Finally, I discuss the theoretical and methodological contributions to the field of bilingualism and sign language research more generally.

Chapter 2

Iconicity in spatial language guides visual attention: A comparison between deaf signers' and hearing speakers' eye-gaze during signed and spoken message preparation

Chapter adapted from:

Manhardt, F., Ozyurek, A., Sumer, B., Mulder, K., Karadöller, D. Z., & Brouwer, S. (2020). Iconicity guides visual attention: A comparison between signers' and speakers' eye-gaze during message preparation. *Journal for Experimental Psychology: Learning, Memory, and Cognition. 46*(9), 1735–1753.

Abstract

To talk about space, spoken languages rely on arbitrary and categorical forms (e.g. LEFT/RIGHT). In sign languages, however, the visuo-spatial modality allows for iconic expressions (motivated form-meaning mappings) of space in which form and location of the hands bear resemblance to the objects and spatial relations depicted. We assessed whether the iconic expressions in sign languages guide visual attention to spatial relations differently than spatial expressions in spoken languages during message preparation at the sentence level.

Using a visual world production eye-tracking paradigm, we compared hearing speakers' (N = 20) and deaf signers' (N = 20) visual attention to describe left versus right configurations of objects (e.g. pen is to the left/right of glass). Participants viewed four-picture displays in which each picture contained the same two objects but in different spatial relations (lateral (left/right), sagittal (front/behind), topological (in/on)) to each other. They described the target picture (left/right) highlighted by an arrow.

*Durin*g message preparation, deaf signers, but not hearing speakers, experienced increasing eye-gaze competition from other spatial configurations. This effect was absent during picture viewing *prior to* message preparation of relational encoding. Moreover, deaf signers' visual attention to lateral and/or sagittal relations was predicted by the type of iconicity (i.e. object and space resemblance vs. space resemblance only) in their spatial descriptions. Findings are discussed in relation to how *thinking for speaking* differs from *thinking for signing* and how iconicity can mediate the link between language and human experience and guides deaf signers' but not hearing speakers' attention to visual aspects of the world.

Introduction

As humans, we constantly explore the visual world around us with our eyes by selecting and attending to relevant details, while ignoring the irrelevant ones. Over the past decades, researchers have studied and claimed a link between eye-gaze patterns and production and comprehension of spoken language (for language comprehension, see e.g. Cooper, 1974; Tanenhaus & Trueswell, 2006; for language production within languages at the lexical level as well as for complex messages, see e.g. Griffin & Bock, 2000; Konopka & Meyer, 2014; van de Velde, Meyer, & Konopka, 2014; for language production across different languages, see e.g. Flecken, Von Stutterheim, & Carroll, 2014; Papafragou, Hulbert, & Trueswell, 2008; Slobin, 2003). This link has been extensively studied for spoken languages but we know little about whether these effects are similar between spoken languages versus sign languages or whether the modality (vocal vs. visuo-spatial) of language guides visual attention differently. Recent studies have provided evidence that eye-gaze and sign language comprehension are also linked as found in spoken languages (e.g. Lieberman, Borovsky, Hatrak, & Mayberry, 2015; Lieberman, Borovsky, & Mayberry, 2017; Thompson, Vinson, Fox, & Vigliocco, 2013). However, this evidence is limited to comprehension and it is not known whether the modality of expressions guides visual attention differently for linguistic production at the sentence level and during message preparation.

The present study investigates for the first time the link between visual attention and language production in hearing speakers and deaf signers during message preparation at the sentence level. More specifically, it tests whether a modality-specific aspect of sign languages, that is iconicity in the expressions, guides visual attention differently for deaf signers than for hearing speakers during message preparation for spatial language (e.g. the pen is to the right of the glass). Sign languages, unlike spoken languages, encode spatial relations between entities in iconic ways where linguistic forms resemble the forms of the objects and the spatial relations between them. In the current study, hearing speakers and deaf signers are asked to view pictures depicting different spatial relations between items and describe one of them while their eye-gaze patterns are recorded *prior to* and *during* message preparation. Eye movements are analysed to see whether hearing speakers and deaf signers show similar or different eye-gaze patterns, and more specifically if differences in visual attention can be linked to the arbitrary and categorical versus iconic ways of expressing information in spoken language versus sign languages.

The link between language and eye-gaze in spoken languages

Previous research has shown that during language production speakers direct their eye-gaze at the referents they are describing in the order that they mention them, reflecting the incremental characteristics of spoken language planning (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer, Sleiderink, & Levelt, 1998; van de Velde et al., 2014). Based on this evidence, it has been claimed that there is a tight link between incremental speech production and eye-gaze (i.e. speech-gaze link). Furthermore, linguistic variation across different languages in which different elements of a scene are encoded seems to guide speakers' visual attention to different components of these visual scenes during message preparation (i.e. *thinking for speaking*, see Slobin, 2003; e.g. Bunger et al., 2016; Flecken et al., 2015, 2014; Papafragou et al., 2008; Trueswell & Papafragou, 2010). For example, Papafragou and colleagues (2008) used eye-tracking to show that language influences visual attention during message preparation. The study assessed whether cross-linguistic differences between English and Greek, in how a motion event is encoded, direct visual attention to different parts of the event. Greek predominantly encodes the Path (i.e. the direction) of the motion in the main verb (Talmy, 1985; Talmy, 2003), whereas English typically encodes Manner of the motion (i.e. how the motion is performed) in the main verb. During the planning of event descriptions (e.g. someone is running up the stairs), Greek speakers encoded and paid more attention to the Path (e.g. ascending the stairs), while English speakers encoded and paid more attention to the Manner of the event (e.g. running). Therefore, eye-gaze patterns differed between Greek and English speakers during message preparation. Similarly, categorization differences between German and Korean speakers impacted the way these speakers visually attended to spatial relations between two objects (Goller, Lee, Ansorge, & Choi, 2017). Korean speakers based their linguistic categorization of space on the degree of fit between the two objects (e.g. *kkita* for loose fit vs. *netha* for tight-fit) whereas German speakers did not. Instead, German differentiates between support and containment (e.g. *auf* for support vs. *in* for containment). Eye-gaze patterns indicated that Korean speakers looked equally likely at the two objects, while German speakers looked more frequently at the ground object. Overall, previous research suggests that during the planning of describing spatial scenes, speakers of different spoken languages focus on those aspects of scenes that are relevant for linguistic expression. However, this body of evidence for language production influencing visual attention is limited to the study of spoken languages. It yet needs to be investigated whether *thinking for speaking* (Slobin, 2003) might differ from *thinking for signing* because of the iconic ways sign languages are organised for expressing spatial relations, unlike spoken languages.

Iconicity in sign languages and effects in language processing

Some aspects of sign languages differ from spoken languages in terms of iconicity which is defined as the motivated *mapping* between meaning and a visual or auditory linguistic form (e.g. Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Emmorey, 2014; Perniss, Thompson, & Vigliocco, 2010). In spoken languages, arbitrariness mostly governs the relation between form and meaning (but see Assaneo, Nichols, & Trevisan, 2011; Dingemanse, 2012 for an overview of existing iconic (i.e. motivated) forms in various spoken languages). In contrast, the visuo-spatial nature of the modality of sign languages allows a large proportion of motivated one-to-one mappings in their linguistic structures. Within sign languages, iconicity can be found at the lexical level (e.g. in the sign for aeroplane in NGT) the extended thumb and pinkie depict the wings and the movement of the sign depicts the movement up into the sky) and also at the sentence level in which signers can map visual features of spatial relations between entities onto their hands and signing space.

While the investigation on the influence of iconicity on spatial language processing in sign languages has just begun, a number of behavioural studies suggest that iconicity at the lexical level can affect sign language processing during language comprehension (e.g. Grote & Linz, 2003; Thompson, Vinson, & Vigliocco, 2009; Thompson et al., 2010, 2012). For instance, Grote and Linz (2003) found that deaf signers of German Sign Language were faster in judging picture-sign relationships for iconic properties of signs. That is, signers were faster to judge object-property relations when the iconic property of the sign (e.g. tracing the beak of the eagle for the sign for eagle) matched the following picture compared to when it did not match (e.g. a picture of a beak vs. a picture of a wing). Interestingly, this effect was not found in speakers of German, who received a version of the same task containing the written translation equivalents in German. In another study, signers of British Sign Language were slower in deciding whether a sign involved a straight or bent handshape (i.e. phonological parameters of a sign) when signs were iconic compared to when they were arbitrary (Thompson et al., 2010). Moreover, it has been claimed that iconicity is not only relevant for language processing but can actually mediate the link between language and human experience (i.e. action and perception) because sign languages' visuo-spatial mappings of many signs are salient aspects of the mental representations of their corresponding entities in the world (e.g. Perniss & Vigliocco, 2014; Thompson et al., 2012; Vigliocco, Perniss, & Vinson, 2014; Vigliocco, Vinson, Woolfe, Dye, & Woll, 2005; Vinson, Thompson, Skinner, & Vigliocco, 2015). However, the role of iconicity in semantic language processing at the lexical level is actively debated since many studies have also failed to show iconicity effects on language comprehension (e.g. Bellugi & Klima, 1976; Bosworth & Emmorey, 2010). These findings suggest that iconicity effects on language comprehension depend on the investigated sample (e.g. native vs. late signers), the task, and the type of iconicity. For sign language production, however, Navarrete, Peressotti, Lerose, and Miozzo (2017) presented evidence for a facilitating effect of iconicity during a target-distractor picture naming task. That is, picture distractors with iconic signs induced faster responses than when picture distractors were arbitrary, suggesting that iconicity can drive the activation of unproduced lexical signs and can therefore affect linguistic production planning.

Overall, previous studies have shown that iconicity can affect sign language processing during language comprehension and production, although the latter has been less studied. Furthermore, these findings are limited to the lexical level. To our knowledge, there are no studies so far looking at whether and how message preparation at the sentence level guides visual attention. Within the domain of spatial language (e.g. the pen is to the right of the glass) and sign languages, relations between two entities and their similarity to the linguistic forms are important (i.e. iconicity). This allows us to assess how visual attention relates to visual similarity between linguistic forms and visual relations between two objects and shape features of the objects. Therefore, the present study aims to investigate whether and how the iconic forms in sign languages at the sentence level guide visual attention to spatial relations differently than spoken language production during message preparation.

The present study

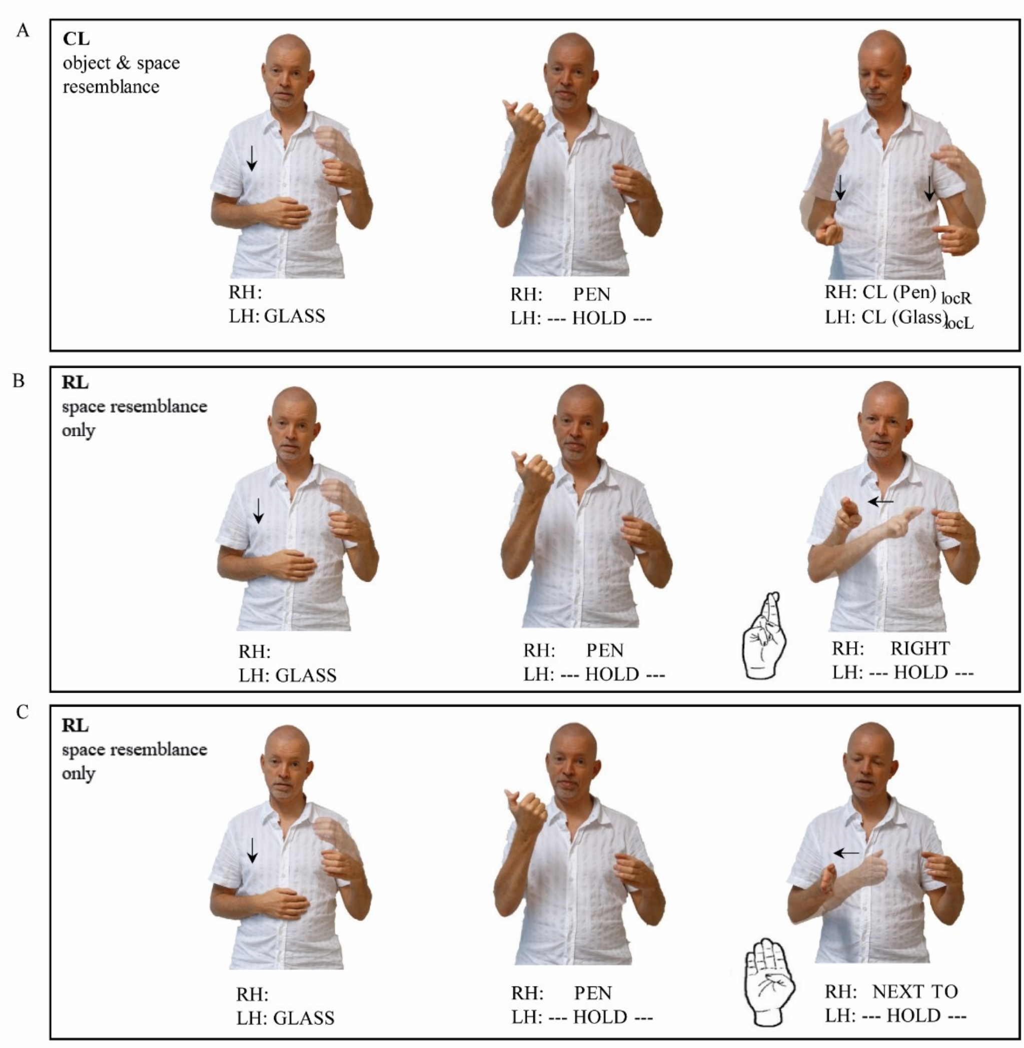
As a consequence of the vocal modality, speech forms in spoken languages rely on expressing three-dimensional events and spatial relations into arbitrary and categorical forms (see also Perniss, Zwitserlood, & Özyürek, 2015). The visuo-spatial nature of the modality of sign languages, however, allows iconic representations of space by directly mapping entities and their spatial relations onto the hands and the signing space (e.g. Emmorey, 1996, 2002; Emmorey et al., 2013; Perniss et al., 2015; Talmy, 2003). In the present study, we investigated whether these iconic spatial expressions in sign languages guide deaf signers' attention to spatial relations differently than that of hearing speakers *during* message preparation. That is, whether *thinking for signing*, thus the way iconic forms depict aspects of referents and relations, differs from *thinking for speaking* (Slobin, 2003). In addition, we aim to explore whether the type of iconicity within deaf signers' spatial expressions influences their visual attention differently. Below, we first give some more detailed information about iconicity and different types of iconicity in sign languages to describe spatial relations.

Iconicity in expressing spatial relations in sign languages

Sign languages use iconic forms to encode spatial relations, unlike spoken languages. In particular, different constructions of spatial languages in sign languages exploit different types of iconic expressions. The common constructions to express spatial relations between entities, called *classifier constructions* (CLs). In using CLs, signers predominantly map shape and orientation properties of objects and relations among them onto their hands and the signing space in front of them by placing both hands in front of the body (e.g. Emmorey, 1996, 2002; Perniss et al., 2015; Zwitserlood, 2012). For example in NGT, signers may first introduce the lexical signs of the two entities (e.g. the pen and the glass), followed by a round handshape to represent the shape of the glass and the index finger to represent the thin, elongated shape of a pen, placing both hands next to each other to match the signers' view of the relative relations of the entities to each other (Figure 6A). These mappings "mirror" or visually resemble both the object properties (i.e. shape) as well as the spatial relations between entities onto the signing space as viewed from the signers' own perspective (e.g. Emmorey & Herzig, 2003; Perniss et al., 2015; Perniss, Thompson, & Vigliocco, 2010).

Figure 6

Types of iconicity in different linguistic expressions for "pen is to the right of the glass" in NGT.



Panel A illustrates a classifier construction (CL) that visually resembles both object properties and space. Panel B and C demonstrate relation lexemes (RL) for right (R-handshape) (B) and next to (to the right) (B-handshape) (C) which resemble space only. RH stands for the right hand, LH for the left hand, and loc for locative placement (to the right or left).

However, there are other iconic forms that signers can use, which encode only the spatial relation *without* expressing specific information about object properties (e.g.

shape; *relational lexemes*: RLs, e.g. Arik, 2013; Perniss et al., 2015; Sümer, Perniss, Zwitserlood, & Özyürek, 2014). RLs thus differ from CLs in their type of iconicity. RLs resemble space (i.e. figure object to the right or left of the ground object), while CLs resemble object properties and space (for neurobiological evidence for different processing of CLs vs. RLs in signers see Emmorey et al., 2013). Specifically for left/right spatial configurations in NGT, there are two forms of RLs available to encode spatial relations on the lateral axis. First, NGT signers can use RLs for LEFT and RIGHT (Figure 6B). These lexemes contain first letters of the spoken words (i.e. *L* handshape for Dutch LEFT: *LINKS; R* handshape for Dutch RIGHT: *RECHTS*) as well as the hand movement to the right or to the left (corresponding to the position of the pen in regard to the glass).

Second, NGT signers can use an RL meaning *next to* (Figure 6C) consisting of a *B*-handshape moving to the left or right (corresponding to the location of the figure object in regard to the ground object, but without changing the *B*-handshape) (see Perniss, 2007; Perniss et al., 2015). Both forms of RLs in NGT express the relative spatial relation between two entities in iconic ways (one-to-one mappings of the pictures) through space resemblance but not through resemblance to object properties, unlike CLs. Therefore, the types of iconicity in spatial expressions can differ and raise novel questions whether these different types of iconic forms selected guide visual attention to spatial relations differently.

Visual world production eye-tracking paradigm

In order to assess whether the above-mentioned types of iconic forms in NGT influence eye-gaze to relations between two objects differently than spoken Dutch during message preparation, we used a visual world eye-tracking paradigm. Instead of assessing language comprehension, we used this paradigm for language production in a novel way (for a similar approach, see Davies & Kreysa, 2017, 2018). We presented hearing speakers and deaf signers with the traditional four picture-displays, but as different from previous research, each picture contained the same two objects but in different spatial relations to each other (i.e. left, right, front, behind, in or on) as we were interested in relational encoding. After an initial viewing phase of the four pictures, participants saw a visual cue in the form of an arrow indicating the target picture. This visual cue then disappeared and the four pictures remained on the screen until a grey screen appeared signalling that it was time to describe (i.e. speaking/signing) the target picture. We recorded hearing speakers' and deaf signers' eye-gaze during two crucial display phases. First, we assessed the initial viewing phase before the target was indicated (i.e. *pre-arrow window*, Figure 8) serving as a baseline to explore whether hearing speakers and deaf signers observe the visual displays differently *prior to* message preparation for encoding of spatial relations (e.g. due to deafness, enhanced peripheral vision or sign language experience) and also to check whether our displays had intrinsic features that guided visual attention. Second, we assessed the viewing phase after the arrow disappeared until the production screen appeared (i.e. *post-arrow window*, Figure 8) which was our main interest to investigate visual attention *during* message preparation.

In the four picture displays, we manipulated the ways in which spatial relations in non-target pictures (i.e. competitors and distractors) competed with the target picture (i.e. right/left) either due to being *visually similar* and/or *semantically relevant* to the target picture (see Figure 7). In this manipulation, when non-target pictures depicted figure objects placed to the sides of the ground object (i.e. left/right/front/behind; lateral and sagittal relations) they were considered more visually similar to the target than when they were in in/on relations (i.e. topological relations). The reason to consider lateral and sagittal relations as more similar to the target compared to topological relations is based on the view that the former are viewpoint dependent and therefore require perspective taking from the language user (Levinson, 1996, 2003). Thus sagittal and lateral relations share more similarities to each other and to the target object in terms of viewpoint dependency and perspective taking and differ in those terms from topological relations, which are not viewpoint dependent and therefore do not require perspective taking. In addition to visual similarity, non-target pictures were also defined as semantically relevant if they shared the same lateral symmetrical axis as the target object (i.e. left/right more semantically related than left/ front or left/in).

Following this approach, displays included a) a s*ymmetrical competitor* (i.e. left/right) which was visually similar to the target picture and also semantically relevant, b) a *sagittal competitor* (i.e. front/behind) which was only visually similar to the target picture (i.e. figure object is on the sagittal side of the ground object) but not semantically relevant, and c) *a topological distractor* (i.e. in/on) which was neither visually nor semantically relevant. We reasoned that if iconicity of relational expressions matters, then the competitors would attract deaf signers' attention more than hearing speakers' and also where deaf signers allocate their visual attention would depend on the type of iconicity used in their spatial expressions.

We further manipulated the complexity of displays in two ways: we either included the symmetrical competitor, creating more complex displays, or excluded it, creating less complex displays (for a similar approach using a contrast manipulation, see Davies & Kreysa, 2017, 2018; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). We therefore decreased display complexity when a symmetrical competitor was absent due to its high degree of visual similarity to the left/right target picture. Thus, we assumed that the presence of a symmetrical competitor is more perceptually complex (i.e. more complex displays) compared to when it is absent (i.e. less complex displays). In these so-called less complex displays, we presented the target, one sagittal competitor and two topological distractors. This manipulation served as a baseline for exploring whether hearing speakers' and deaf signers' language productions and eye-gaze are further influenced by the complexity of displays.

Predictions

For linguistic productions, we predicted that hearing speakers will use the most informative strategy available, independent of the complexity of displays (Davies & Kreysa, 2017). That is, hearing speakers will prefer to use specific spatial relation adverbs (e.g. LEFT/RIGHT) over general spatial relation adverbs (e.g. *next to*) despite the presence or absence of a symmetrical competitor. For deaf signers, we predicted a preference of using CLs over RLs as demonstrated previously for American Sign Language (Emmorey, 2002), German Sign Language (Perniss, 2007) and Turkish Sign Language (Sümer, 2015). It is less known how these strategies might differ in different complexity contexts. Due to the high level of informativeness of CLs (resembling objects and space) we predicted, that deaf signers prefer the most informative strategy (i.e. CLs), independent of the presence or absence of a symmetrical competitor.

For eye-gaze *prior to* message preparation (*pre-arrow window*), we predicted that hearing speakers and deaf signers do not differ in the way they observe the visual displays. Previous research suggests that only *during* the planning of describing spatial scenes, speakers of different spoken languages look at those aspects of scenes differently and in ways that are relevant for linguistic expression. During only viewing; that is, without the goal of language production, these cross-linguistic effects were absent (e.g. Papafragou et al., 2008; Slobin, 2003). In our study during the pre-arrow window, subjects do not know which spatial relation to encode. We therefore consider that eye-gaze patterns in the pre-arrow window do not reflect differences in encoding for spatial relations (i.e. the focus of the current study). We cannot rule out that eye-gaze might be guided by preparation for object labels, similar in all of the four pictures, or other general visual attention differences between hearing speakers and deaf signers. However and importantly, we did not expect differences between hearing speakers and deaf signers to arise *prior to* message preparation neither for more complex displays nor for less complex displays that would reflect differences in visual attention related to relational encodings.

For eye-gaze *during* message preparation (*post-arrow window*), we expected to find differences between hearing speakers and deaf signers related to relational encodings. More specifically, we predicted that in order to plan to describe the relation depicted in the target picture deaf signers experience more eye-gaze competition from the symmetrical competitor or the sagittal competitor than hearing speakers. Moreover, this enhanced competition from deaf signers might increase as time unfolds and gets close to the actual message production (unlike in *pre-arrow window*). Due to the visual similarity of those competitors, deaf signers need to identify the relative locations and/or the orientation, size and shape of the objects in relation to each other to map information onto the signing space iconically (e.g. orientation/shape of figure hand placed in relation to hand resembling the ground object). Consequently, disambiguating where to move or place the hands in space might require more effort for deaf signers during production planning, resulting in an increase in visual attention to the symmetrical and/or sagittal competitors in order to describe the target picture. In contrast, for hearing speakers, this kind of iconic information is not relevant for their categorical and arbitrary spatial expressions. Thus, we expected that hearing speakers experience less competition than deaf signers from the visually similar competitors (i.e. symmetrical and sagittal competitor). For the topological distractor, we did not expect differences in eye-gaze competition for both hearing speakers and deaf signers, since it is neither semantically relevant nor visually similar to the target picture. Following the same line of reasoning, for less complex displays we predict that deaf signers experience no enhanced eye-gaze competition from the topological competitor compared to hearing speakers. However, we do predict more eye-gaze competition from the sagittal competitor for deaf signers compared to hearing speakers. We do, however, take into account the possibility that competitors might be viewed differently depending on the complexity of the displays.

Additionally, we also expected that within deaf signers, the type of iconicity (CLs or RLs) preferred in their linguistic expressions would modulate visual attention differently, providing more evidence for linking gaze-allocation of deaf signers to the iconicity of their expressions. As explained above deaf signers can either use CLs which resemble both space and object properties in a one-to-one mapping (Figure 6A) or they could use RLs in which iconicity focusses on spatial properties only (see Figure 6B, Figure 6C). Depending on which strategy deaf signers are planning to use we expect different eye-gaze patterns to arise. On the one hand, preparing to use CLs might lead to more eye-gaze competition from visually relevant pictures (i.e. symmetrical and/or sagittal competitor) because they require more effort to resemble the object properties and where to place the hands in space (e.g. placing a long-elongated handshape representing the pen to the left vs. right vs. front vs. behind of the round handshape representing the glass). On the other hand, planning to use RLs might lead to more competition from only the semantically relevant symmetrical competitor but not from the other visually similar sagittal competitor. RLs' iconicity focuses on the relative location in space in a more abstract and categorical fashion than CLs' and thus only semantically relevant competitors (i.e. symmetrical competitor) might elicit competition with the target picture. As above we expected these types of competitions to change over time and increase getting close to the message production. We do not expect more eye-gaze competition from the topological distractor as it is neither visually nor semantically relevant for describing the left/right target configuration.

Methods

The methods reported in this experiment were approved by the Humanities Ethics Assessment Committee of the Radboud University.

Participants

The participants in this study were 20 hearing speakers of Dutch (11 female), and 20 deaf signers of NGT (16 females; see Table 2 for participants' descriptive statistics). To assess similarities in age and language proficiency across hearing speakers, and deaf signers we conducted Bayesian t-tests in which we assessed the probability of the mean difference greater than zero and less than zero, using the R package *BayesianFirstAid* (version 0.1; Bååth, 2014).The groups were similar in age (Bayesian two sample t-test: M_{DIFF} (-5) > 0: p = 0.129, M_{DIFF} (5) < 0: p = 0.871). All deaf signers were born to deaf parents and acquired NGT from birth (i.e. from their deaf signing parents), except one who acquired sign supported Dutch (i.e. manually coded form of Dutch, thus spoken Dutch is being supported by NGT signs) from birth. Two additional deaf signers and one hearing speaker were tested but were excluded from the study due to another native language than NGT (N = 1), or high errors in describing the pictures (hearing speakers: N = 1; deaf signers: N = 1). Additionally, one deaf signer was discarded due to high eye-tracking loss (larger than 45%).

Table 1

Descriptive statistics for hearing speakers and deaf signers.

	Hearing speakers	**Deaf signers**
Age (years)	32 (11)	37 (12)
Self-rated proficiency (comprehension)	4.87 (0.4)	4.80 (0.4)
Self-rated proficiency (production)	4.80 (0.4)	4.60 (0.8)
Visuo-spatial memory span	6.65 (1.1)	6.40 (1.1)

Self-rated proficiency scores contain evaluations of Dutch from hearing speakers and of NGT from deaf signers from 0 (no knowledge) to 5 (native like). Visuo-spatial memory span contains the average span per group (out of 9) assessed by the Corsi Block Tapping Task. SDs are reported between parentheses. SDs are reported between parentheses.

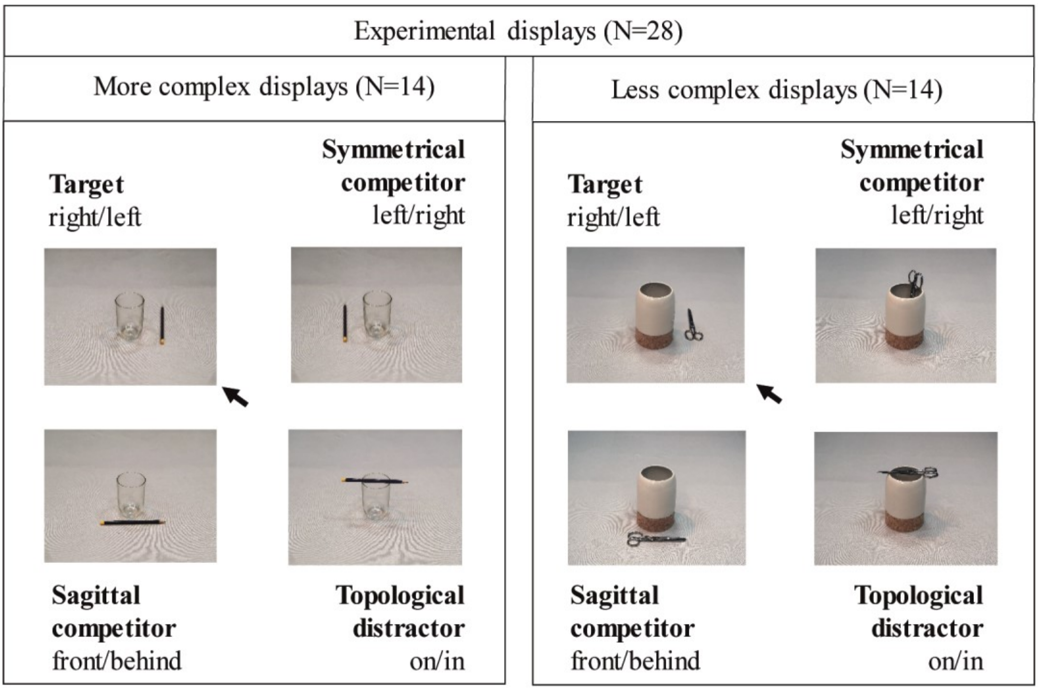
The self-rated language proficiency scores ranging from 0 (no knowledge) to 5 (native like) were collected from hearing speakers for Dutch and deaf signers for NGT separately for comprehension and production (for validity of self-ratings to measure proficiency, see Wilson, 1999). Comprehension scores of Dutch included scores for *reading* and *listening*, while the scores for NGT included *understanding*. Production scores of Dutch included *speaking* and *writing*, while the scores for NGT included *signing*. Bayesian two-sample t-tests indicated similar comprehension scores (M_{DIFF} (-0.07) > 0: p = 0.501, M_{DIFF} (0.07) < 0: p = 0.499) and production scores (M_{DIFF} (-0.2) > 0: p = 0.502, M_{DIFF} (0.2) < 0: p = 0.498) between hearing speakers and deaf signers. Paired Bayesian t-tests showed no difference between self-rated language production scores and comprehension scores within hearing speakers (M_{DIFF} (-0.2) > 0: p = 0.506, M_{DIFF} (0.2) < 0: p = 0.494) and within deaf signers (M_{DIFF} (-0.07) > 0: p = 0.503, M_{DIFF} (0.07) < 0: p = 0.497). We used a Corsi Block Tapping Task (Corsi, 1972) to control for differences in visuo-spatial memory between hearing speakers and deaf signers, which yielded similarities between hearing speakers' and deaf signers' working memory span (Bayesian two sample t-test: M_{DIFF} (-0.25) > 0: p = 0.726, M_{DIFF} (0.25) < 0: p = 0.274). All hearing speakers and the majority of the deaf signers were tested at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. Three deaf signers were tested at their homes. Participation was voluntary and all participants received a financial compensation.

Materials

Participants were presented with 84 visual displays containing four pictures (see Figure 7). Each picture displayed the same two objects in a spatial relation to each other (i.e. left, right, front, behind, on, in). A visual cue in the form of an arrow appeared in the middle of the screen and pointed at one of the four pictures, indicating the target picture. 28 visual displays were used as experimental trials, half of them containing more complex displays and the other half containing less complex visual displays. The target in both conditions was always a left or right configuration. More complex displays included a symmetrical competitor (left or right, depending on whether the target was left or right), a sagittal competitor (front or behind) and a topological distractor (in or on). Less complex displays, importantly, did not contain a symmetrical competitor but included instead a sagittal competitor (front or behind) and two topological distractors (i.e. both in and on).

Figure 7

Examples of experimental displays.



Displays are specified for more complex and less complex displays.

Next to the 28 experimental displays, we included 56 additional displays as filler trials to distract participants from the critical left/right configurations during the experiment. The filler displays consisted of four different patterns to offer visual variation and avoid biases to the left/right symmetry. Importantly, all filler displays contained front/behind/in/on targets instead of left/right, which were equally distributed across all 56 displays (i.e. targets: 14 x front, 14 x behind, 14 x in, 14 x on). Thus, in filler displays, the arrow pointed at any other spatial relation but left/right. First, 14 filler displays contained the same combinations of spatial relations as the 14 experimental more complex displays, thus contained left, right, front/behind, in/on. Second, 14 filler displays resembled the 14 experimental less complex displays (i.e. containing left/right, front/behind, in, on). Third, 14 visual displays contained both a front and behind configuration to distract participants from our critical lateral contrast (i.e. front vs. behind, left/right, in/on). Finally, 14 visual displays also contained both front and behind relations, however, did not include any left/right spatial relations (i.e. front vs. behind, in, on).

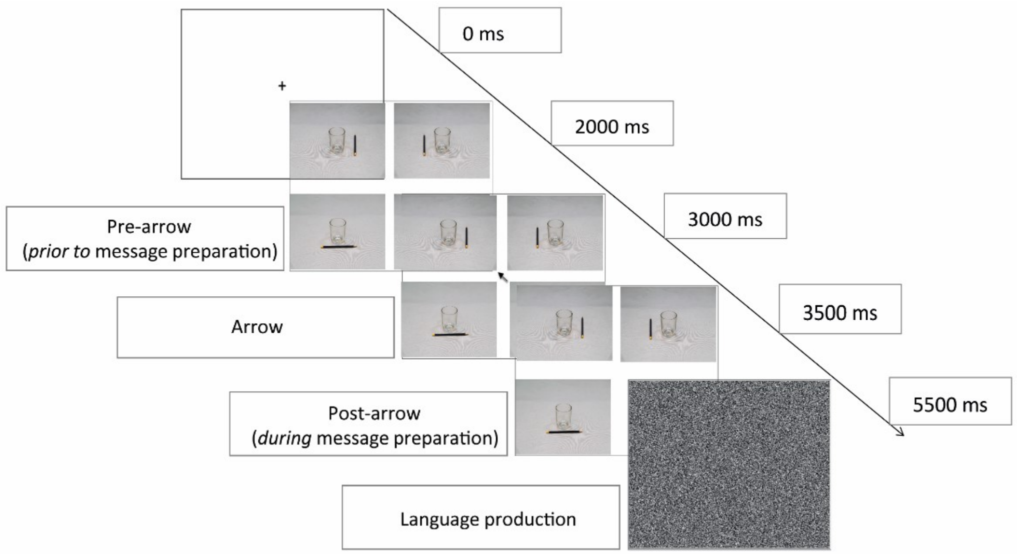
Figure objects (e.g. the pen) were presented once, while Ground objects (e.g. the jar) occurred four times, but always in combination with another figure (e.g. jar-lemon, jar-watch, jar-screwdriver, and jar-jar lid). One Ground object was never presented more than twice in a row. Ground objects in experimental displays were always non-intrinsic objects to allow left/right interpretations rather than intrinsic interpretations. Additionally, target spatial relations (i.e. left/right) were presented not more than twice in a row to avoid biases to one type of spatial relations. The location of each configuration was counterbalanced across participants and trials. All visual displays have been piloted to assure that participants can name the presented objects.

Procedure

Before the actual experiment, participants performed a familiarization task using displays similar to those in the actual eye-tracking experiment to introduce participants to the general complexity of our displays (i.e. 2x2 grid with two objects in different spatial relations to each other). After answering some questions about the displays, we proceeded with the actual eye-tracking description task. The experiment was preceded by three practice trials, which were repeated if necessary, and a five-point calibration and validation procedure. Trials started with a fixation cross shown for 2000 ms, followed by a pre-arrow display introducing the four pictures for 1000 ms (see Figure 7). After that, an arrow indicated the target picture by pointing at one of the four pictures and disappeared after 500 ms. The four pictures remained on the screen for 2000 ms until a grey, visual noise screen was presented as a cue for linguistic production (see Figure 8). Participants' task was to describe (i.e. speak/sign) the picture at which the arrow was pointing (i.e. the target picture), to a confederate when the grey screen appeared. After each description, the confederate pretended to choose the described picture out of four pictures on a tablet. We instructed participants in such a general manner because the confederates' cover task provided the participants with enough indications of what is expected from them. Furthermore, with this general instruction, we aimed to allow unbiased linguistic descriptions, as well as individual scanning of the display that is guided by the way participants, will describe the pictures. The confederate was present to elicit a natural conversation and the confederates' hearing status was always matched to that of the participant (i.e. a deaf confederate for deaf participants, a hearing confederate for hearing participants). Importantly, participants were told that confederates were naïve participants and the confederate was always another person than the experimenter to assure that the picture descriptions to the addressee are as informative as in natural conversations. Participants never questioned the status of the addressee during or after the experiment. All hearing speakers were tested with the same hearing confederate and all deaf signers were paired with the same deaf confederate. Participants initiated the next trial by pressing a button after they had described the target picture. Importantly, participants did not receive feedback regarding their spatial descriptions.

Figure 8

Timeline of trial structure.



After this eye-tracking task, participants took part in a digital Corsi Block Tapping task to control individual differences in spatial memory. The eye-tracking task and the Corsi Bock Tapping task were presented on an SMI RED-250 mobile laptop. The software package *Presentation NBS 16.4* (Neurobehavioural Systems, Albany, CA) was used to control the eye-tracker, present the stimuli, and register button presses during each trial. Triggers were sent from this software to the eye-tracker. Eye-gaze was recorded binocular at a rate of 250Hz (every 4 ms). Instructions were always given orally/visually in form of a video. A language background questionnaire was given at the end of the session to assess language use, language proficiency, deafness in family, etc. In total, the experimental session lasted approximately 45 minutes.

Data analysis

In this section, we will first describe how we analysed hearing speakers' and deaf signers' spatial descriptions of the target pictures during the eye-tracking task. Next, we describe the analysis of the eye-gaze data first *prior to* message preparation followed by *during* message preparation. Finally, we will combine the two types of data (i.e. linguistic productions and eye-gaze) by linking deaf signer's linguistic productions to their eye-gaze competition *during* message preparation.

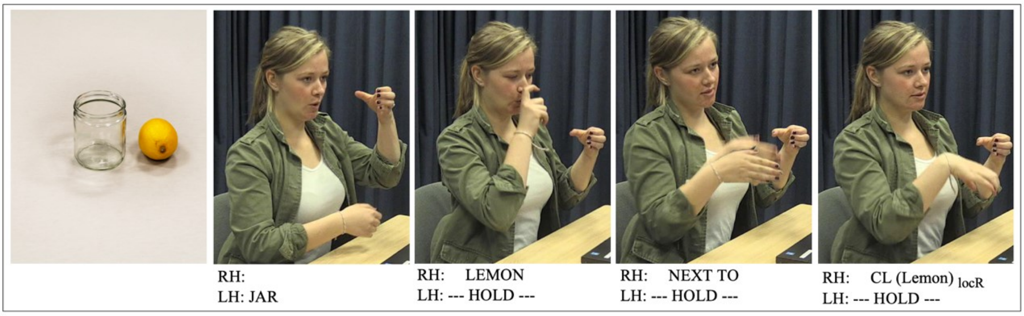
Linguistic production. To assess the frequency of certain types of spatial expressions used by hearing speakers and deaf signers, we coded descriptions of the experimental items to describe left/right spatial configurations across more complex and less complex displays. In order to do this, we used ELAN, a free annotation tool (http://tla.mpi.nl/tools/tla-tools/elan/) for multimedia resources developed by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Trained, native annotators (i.e. native Dutch annotator for Dutch data, deaf native NGT annotator for NGT data) performed annotations and codings of the data. All codings were checked by an additional coder to find consensus. If no consensus could be found, the trial was excluded from further analyses (3.61% of the data).

For hearing speakers' picture descriptions in Dutch, we distinguished between two types of spatial expressions: categorical forms (i.e. specific spatial relation adverb LEFT *LINKS* or RIGHT *RECHTS*) and other alternative non-categorical forms (i.e. NEXT TO *NAAST*). In addition to speech, we also coded hearing speakers' co-speech gestures, albeit very few, during their picture descriptions and coded per description whether any additional information was encoded iconically with the hands (i.e. about the object properties or spatial relation) or not.

For deaf signers' spatial descriptions in NGT, we categorised linguistic strategies into two main groups depending on their type of iconicity. First, CLs (Figure 6A), which resemble visually space and object properties. Second, RLs (Figure 6B and Figure 6C), which resemble space only. In addition, we coded for double strategies, thus descriptions that contained both RLs followed by CLs within one response, resembling space and additionally space and object properties (see Figure 9; CLs followed by RLs can also occur, although less frequently; see also Sümer, 2015 for Turkish Sign Language).

Figure 9

Example of a double strategy for "lemon is to the right of the jar" in NGT using both RL (next to) and CL within one description.



This coded data were analysed in R (version 3.3.1) (R Core Team, 2013). We used a logistic regression analysis with binomial link separately for hearing speakers and deaf signers. For hearing speakers, we compared categorical (LEFT/RIGHT) versus non-categorical (NEXT TO) strategies (Spatial Expression Type) across more complex and less complex displays (Display Complexity) by using numeric contrasts (Helmert contrast). That is, for Spatial Expression Type we coded categorical as +1/2 and non-categorical as -1/2. For Display Complexity, we coded more complex displays as +1/2 and less complex displays as -1/2. Similarly, for deaf signers we compared linguistic strategies such as CLs and RLs (Spatial Expression Type) across more complex and less complex displays (Display Complexity) by using numeric contrasts (Helmert contrast). That is, for Spatial Expression Type we coded one the first level double strategies as -2/3, CLs as +1/3, and RLs as +1/3. On the second level, we coded CLs as +1/2 and RLs as -1/2. For Display Complexity, we coded more complex displays as +1/2 and less complex displays as -1/2.

Visual attention. For each trial, eye movements were recorded from pre-arrow onset (0 ms) until the four-picture display disappeared (3500 ms). We analysed fixation proportions (right eye only) across 50 ms continuous time bins for two specific time windows: the pre-arrow window and post-arrow window. The pre-arrow window initiated immediately after presenting the fixation cross (0 ms) until arrow onset (1000 ms) (see Figure 8), thus participants did not know yet which spatial relation to describe. The pre-arrow window served as a baseline *prior to* message preparation (related to spatial relation encoding) to ensure that different aspects of our visual displays did not differ in their tendencies to attract gaze. The post-arrow window initiated directly after arrow offset (1500 ms) until production onset (3500 ms) (see Figure 8) capturing participants' linguistic planning phase linked to relational encoding. Thus, in the post-arrow window, we were able to assess hearing speakers' and deaf signers' viewing *during* message preparation (for relational encoding) over time.

We defined five different square-shaped Areas of Interest (AoI): one for each picture and one for the arrow region. We kept the size and positions of the five AoIs equal across all trials. Fixation data were preprocessed in R (version 3.3.1) (R Core Team, 2013). First, it was determined for each participant whether a fixation fell into a particular AoI in each of 40 consecutive bins of 50 ms. Participants with more than 45% track loss across all trials were excluded from the analysis (N = 1). Additionally, we excluded trials in which track loss was higher than 50% (4.55%).

To examine the eye-gaze patterns, we analysed the fixation data by using linear mixed-effects regression models (Baayen, Davidson, & Bates, 2008) in R (version 3.3.1) (R Core Team, 2013) using the packages *lme4* (version 1.1-19) (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (version 3.0-1) (Kuznetsova, Brockhoff, & Christensen, 2017) to retrieve p-values. We used the package *emmeans* (Length, 2019; Searle, Speed, & Milliken, 1980) to interpret significant interactions in these models.

Fixation proportions were corrected in both time windows for 200 ms to plan a first saccade (Matin, Shao, & Boff, 1993). We logarithmically transformed mean fixation proportions and time bins, which mathematically required shifting zeros and ones by +1.01.

Fixation proportions were corrected in both time windows for 200 ms to plan a first saccade (Matin, Shao, & Boff, 1993). We logarithmically transformed mean fixation proportions and time bins, which mathematically required shifting zeros and ones by +1.01.

Eye-gaze competition prior to message preparation in hearing speakers versus deaf signers (pre-arrow window). First, we wanted to ensure that hearing speakers and deaf signers did not differ from each other in their eye-gaze patterns before they knew which relation they would describe and that our displays did not have intrinsic features that guided attention. For more complex displays, Group (categorical predictor: hearing speakers vs. deaf signers), Bin (continuous predictor: for each 50 ms bin) and Spatial Relation Type (categorical predictor: *lateral* vs. *sagittal* vs. *topological*), were entered as fixed effects (predictors). Group was coded as a numeric contrast (Helmert contrast); that is, hearing speakers as -1/2 and deaf signers as +1/2. Spatial Relation Type was also coded as a numeric contrast (Helmert contrast). Due to the presence of two lateral configurations (left and right), fixations across these two pictures were averaged. On the first level, we compared fixations to topological relations versus the averaged lateral and sagittal relations by coding *topological* as -2/3, *lateral* as +1/3, and *sagittal* as +1/3. On the second level, we compared fixations to lateral versus sagittal relations by coding *lateral* as +1/2 and *sagittal* as -1/2.

For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays and we applied the same Helmert contrast coding. Due to the absence of a second lateral relation but inclusion of a second topological relation the fixations for both topological relations (*in* and *on*) were averaged.

Eye-gaze competition during message preparation in hearing speakers versus deaf signers (post-arrow window). To analyse visual attention to left/right spatial relations *during* message preparation (i.e. post-arrow window) in more complex displays, we created three measures that assessed different types of eye-gaze competition: 1) mean of looks to target minus symmetrical competitor (*symmetrical competition*), 2) mean of looks to target minus sagittal competitor (*sagittal competition*), 3) mean of looks to target minus topological distractor (*topological competition)*. In the model, Group (categorical predictor: hearing speakers vs. deaf signers), Bin (continuous predictor: for each 50 ms bin) and Competition Type (categorical predictor: *symmetrical competition* vs. *sagittal competition* vs. *topological competition*) were entered as fixed effects (predictors). As for the pre-arrow window, Group was coded as a numeric contrast; that is, hearing speakers as -1/2 and deaf signers as +1/2. Competition Type was also coded as a numeric contrast. On the first level, we compared competition from the topological distractor versus competition from the other pictures by coding *topological competition* as - 2/3, *symmetrical competition* as +1/3, and *sagittal competition* as + 1/3. On the second level, we compared competition from the symmetrical competitor (coded as +1/2) versus competition from the sagittal competitor (coded as -1/2).

For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays, except for the *symmetrical competition* measurement, due to the absence of the symmetrical competitor. As in the above models, categorical predictors were coded as numeric contrast, however, for Competition Type, *topological competition* was coded as -1/2 and *sagittal competition* as +1/2.

The link between type of iconicity in linguistic production and eye-gaze competition during message preparation in deaf signers only (post-arrow window). To further assess whether deaf signers' eye-gaze competition can be predicted by the type of iconicity in their spatial expressions, we linked deaf signers' linguistic productions to their eye-gaze data *during* message preparation in more complex and less complex displays. For this, we excluded descriptions that could not be easily divided between types of iconicity (i.e. double strategies). Thus, we excluded double strategies (44.44% in complex displays and 41.97% in less complex displays) and consequently included only descriptions that contained only CLs or RLs. In the more complex display model, Type of Iconicity (categorical predictor: object + space resemblance vs. space resemblance), Competition Type (categorical predictor: symmetrical competition vs. sagittal competition vs. topological competition) and Bin (continuous predictor: for each 50 ms bin) were entered as fixed effects (predictors). Type of Iconicity was coded as numeric contrast; that is, visual resemblance as +1/2 and semantic relevance as -1/2. Competition type was coded as numeric contrast (see above for eye-gaze competition model). For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays, except for the *symmetrical competition* measurement, due to the absence of the symmetrical competitor.

For all models, we conducted a backward selection procedure in which insignificant predictors were removed to obtain the most parsimonious model. The maximal random effects structure that converged in the model was implemented, which included random intercepts for participants and items, as well as random slopes for Group by items.

Results

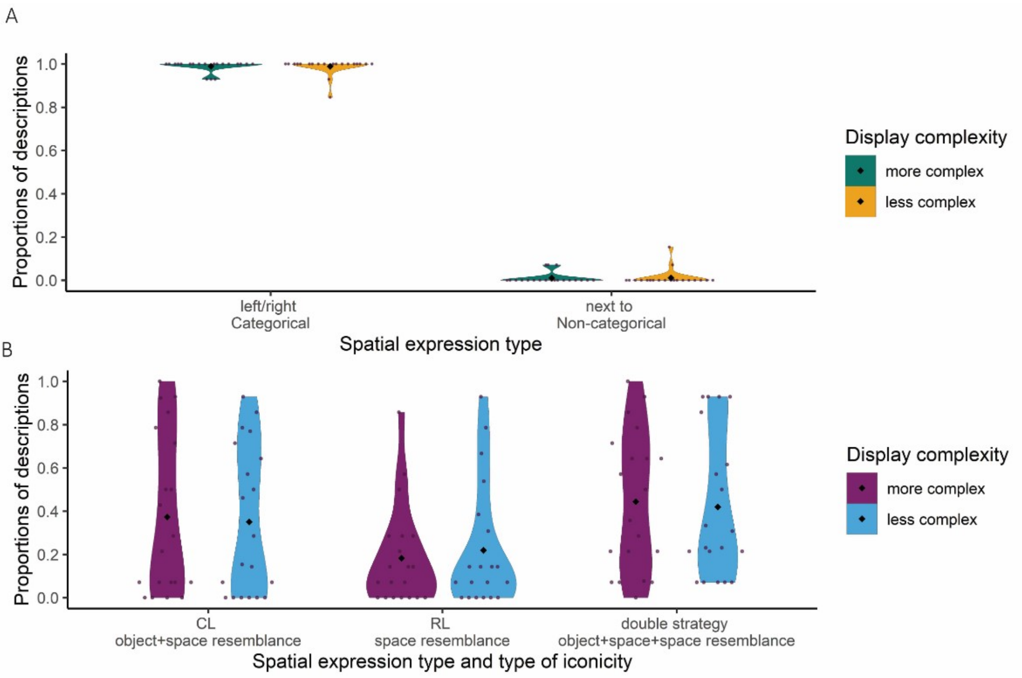
In this section, we will first report the linguistic production data to assess the most frequently used linguistic forms by hearing speakers and deaf signers across more complex and less complex displays, including the type of iconicity in deaf signers' expressions. After this, we will report the eye-gaze data beginning with the results *prior to* message preparation separately for more complex and less complex displays. Next, we will report results from eye-gaze competition *during* message preparation with the distinct measures separately for more complex and less complex displays. Finally, we will combine the two types of data (i.e. type of iconicity in linguistic productions and eye-gaze) and link deaf signers' linguistic productions to their eye-gaze data across more complex and less complex displays.

Linguistic production

Figure 10 shows the most frequently used linguistic forms by hearing speakers and deaf signers during their picture descriptions (see Figure 10A for hearing speaker data; see Figure 10B for deaf signer data).

Figure 10

Types of spatial expressions used by hearing speakers of Dutch deaf signers of NGT.



Panel (A) illustrates the types of spatial expressions used by Dutch hearing speakers. Panel (B) shows the type of spatial expressions and iconicity in NGT deaf signers. Rectangles represent the mean. Dots represent each data point (participant). The width of the violins represents the data distribution's density, the length of the violins depicts the range of data points.

For hearing speakers, we investigated the frequency of categorical and non-categorical picture descriptions as a binary factor across the more complex and less complex displays. To investigate whether there are differences in hearing speakers' descriptions, a logistic regression analysis yielded a main effect of Spatial Expression Type *(β* = 0.094, *SE* = -0.581, *z* = 15.573, *p* < 0.001). This main effect suggests a preference of categorical spatial expressions (98.92% across trials) over non-categorical spatial expressions (1.08% across trials) (see Figure 10A). There was no main effect of Display Complexity (*β* = 0.001, *SE* = 0.581, *z* = 0, *p* > .1) and no significant interaction between Display Complexity and Spatial Expression Type (*β* = -0.022, *SE* = 0.016, *z* = -0.019, *p* = 0.985).

In addition to speech, we coded hearing speakers' co-speech gestures during their picture descriptions to assess whether any additional information was encoded with the hands (i.e. about the object properties or spatial relation). Results indicate a small amount of co-speech gestures (13% in total). In all those instances, the co-speech gestures did not add any extra information about the location nor the properties of the figure and ground object. Rather, they indicated roughly the direction of the figure independent of the ground, which was already encoded in speech (e.g. small head or hand movement to the left/right of the speakers' body).

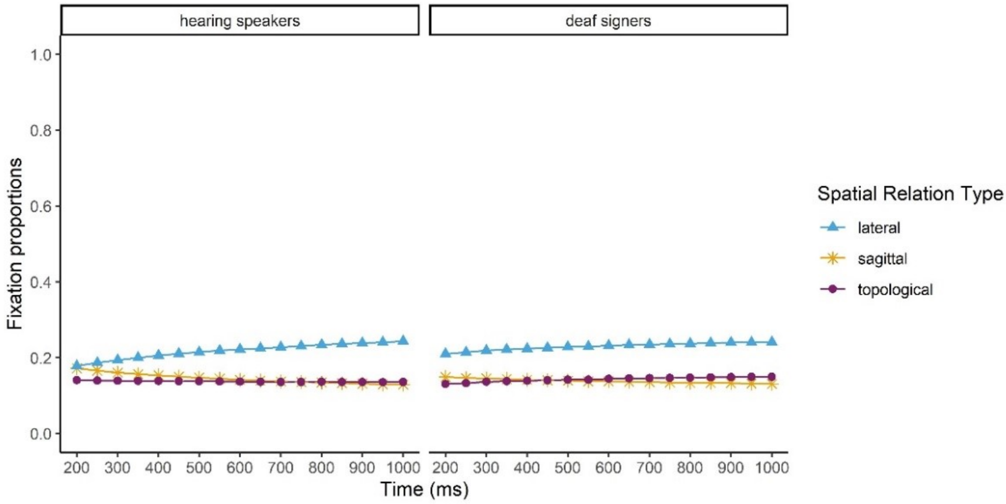
Furthermore, we assessed the frequency of linguistic strategies varying in their type of iconicity in deaf signers. To investigate whether there are differences in deaf signers' descriptions, a logistic regression analysis yielded a main effect of Spatial Expression Type (comparison $1_{[double\ vs.\ CL+RL]}$: $\beta = -0.742$, $SE = 0.054$, $z = -13.668$, $p < 0.001$, comparison $2_{[CL\ vs.\ RL]}$: ($\beta = 0.816$, $SE = 0.138$, $z = 5.890$, $p < 0.001$). The first comparison suggests that double strategies (43.22% across trials) are used more compared to using the average of RLs only and CLs only (see Figure 10B). In addition, the second comparison reveals that using CLs only (36.12% across trials) is the more preferred strategy than using RLs only (20.17% across trials). There was no main effect of Display Complexity ($\beta = 0.009$, $SE = 0.109$, $z = -0.086$, $p = 0.931$) and no interaction between Display Complexity and Spatial Expression (comparison $1_{[double\ vs.\ CL+RL]}$: $\beta = -0.165$, $SE = 0.221$, $z = -0.749$, $p = 0.454$, comparison $2_{[CL\ vs.\ RL]}$: $\beta = 0.323$, $SE = 0.277$, $z = 1.166$, $p = 0.244$).

Eye-gaze patterns *prior to* message preparation in hearing speakers versus deaf signers (pre-arrow window)

More complex displays. First, we assessed whether hearing speakers and deaf signers did not differ from each other in their eye-gaze patterns before they knew which relation they would encode and that our displays did not have intrinsic features that guided attention in complex displays. The most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for Group. The model contained one main effect (Spatial Relation Type$_{[lateral\ vs.\ sagittal]}$), two two-way interactions (Spatial Relation Type$_{[lateral\ vs.\ sagittal]}$*Group; Spatial Relation Type$_{[lateral\ vs.\ sagittal]}$*Bin) but no significant three-way interaction (Group*Bin*AoI; Figure 11). The first interaction with Group suggests that while both hearing speakers and deaf signers look at lateral relations more the difference between the two spatial relation types is bigger for hearing speakers than for deaf signers, independent of the time course of display viewing. The second interaction with Bin suggests increasing looks over time to lateral relations but not to sagittal relations independent of group. However, the lack of an interaction between Group and Bin suggests that there is no difference between hearing speakers and deaf signers over time in the way they view the visual displays *prior to* message preparation (for relevant statistics and corresponding coefficients, see Table 2).

Figure 11

Eye-gaze competition prior message preparation for more complex displays.



The left panel shows hearing speakers. The right panel shows deaf signers during the pre-arrow window. Smaller y-axis values indicate more eye-gaze competition.

Table 2

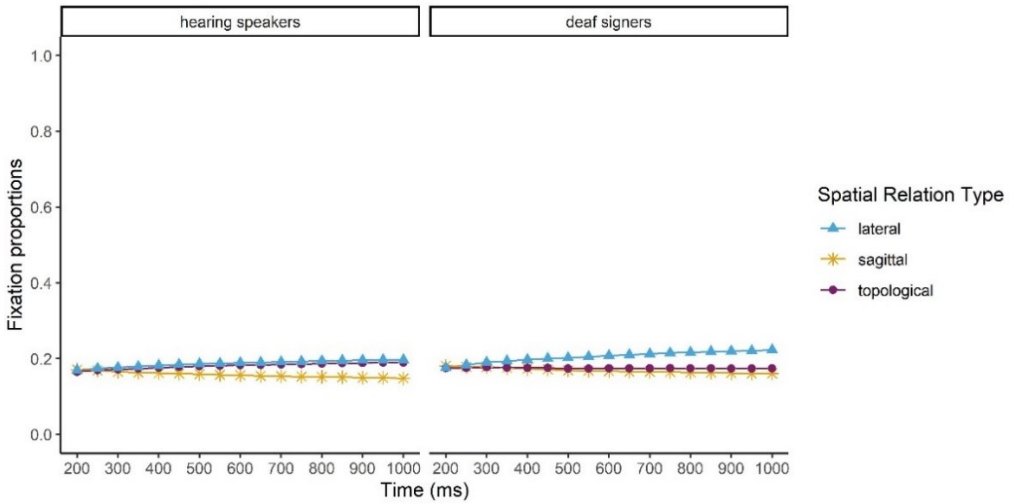
Estimates, standard error, t-values, and p-values of the main effects of the eye-gaze baseline model for more complex displays for the pre-arrow window.

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.061	0.009	6.430	0.001
Relation Type$_{[\text{other vs. topological}]}$	0.015	0.019	0.757	0.449
Relation Type$_{[\text{lateral vs. sagittal}]}$	-0.082	0.022	-3.678	0.001
Group	-0.006	0.019	-0.295	0.768
Bin	0.004	0.003	1.116	0.264
Relation Type$_{[\text{other vs. topological}]}$:Group	0.040	0.039	1.044	0.296
Relation Type$_{[\text{lateral vs. sagittal}]}$:Group	0.092	0.045	2.057	0.040
Relation Type$_{[\text{other vs. topological}]}$:Bin	0.001	0.007	0.049	0.961
Relation Type$_{[\text{lateral vs. sagittal}]}$:Bin	0.041	0.008	5.039	0.001
Group:Bin	0.002	0.007	0.353	0.724
Relation Type$_{[\text{other vs. topological}]}$:Group:Bin	-0.015	0.014	-1.078	0.281
Relation Type$_{[\text{lateral vs. sagittal}]}$:Group:Bin	-0.031	0.016	-1.915	0.055

Less complex displays. For less complex displays, the most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for Group. The model contained one main effect (Spatial Relation Type$_{[\text{lateral vs. sagittal}]}$) and a two-way interaction (Spatial Relation Type$_{[\text{lateral vs. sagittal}]}$*Bin) but no significant three-way interaction (Group*Bin*AoI; Figure 12). The two-way interaction suggests, as in complex displays, increasing looks over time to lateral relations but not to sagittal relations, independent of group. The lack of an interaction with Group and Bin, as in complex displays, suggests that over time hearing speakers and deaf signers do not view the visual displays differently prior *to* message preparation (for relevant statistics and corresponding coefficients, see Table 3).

Figure 12

Eye-gaze competition prior message preparation for less complex displays.



The left panel shows hearing speakers. The right panel shows deaf signers during the pre-arrow window. Smaller y-axis values indicate more eye-gaze competition.

Table 3

Estimates, standard error, t-values, and p-values of the main effects of the eye-gaze baseline model for less complex displays for the pre-arrow window.

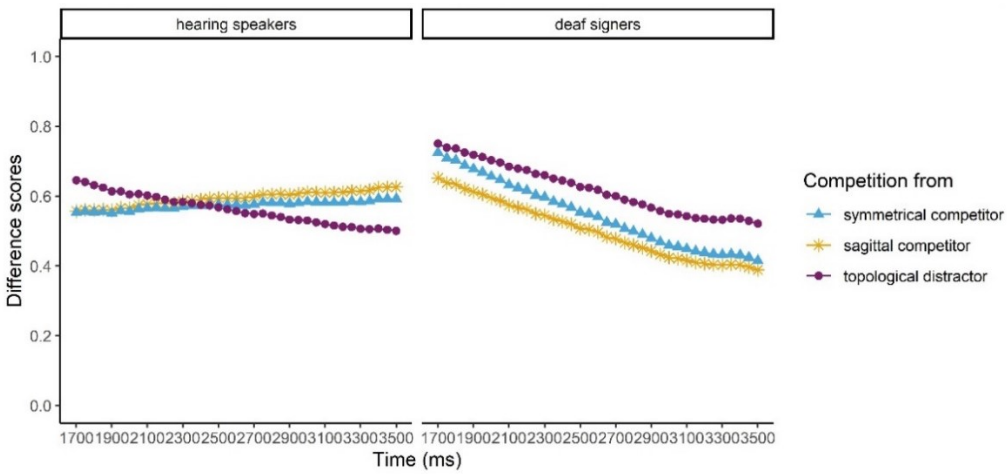
Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.064	0.010	6.358	0.001
Relation Type$_{[\text{other vs. topological}]}$	0.007	0.020	0.352	0.725
Relation Type$_{[\text{lateral vs. sagittal}]}$	-0.071	0.024	-3.024	0.002
Group	0.008	0.020	0.407	0.684
Bin	0.004	0.004	1.164	0.244
Relation Type$_{[\text{other vs. topological}]}$:Group	-0.044	0.041	-1.077	0.282
Relation Type[lateral vs. sagittal]:Group	-0.019	0.047	-0.403	0.687
Relation Type$_{[\text{other vs. topological}]}$:Bin	-0.002	0.007	-0.324	0.746
Relation Type$_{[\text{lateral vs. sagittal}]}$:Bin	0.031	0.009	3.547	0.001
Group:Bin	-0.002	0.007	-0.275	0.784
Relation Type$_{[\text{other vs. topological}]}$:Group:Bin	0.019	0.015	1.266	0.206
Relation Type$_{[\text{lateral vs. sagittal}]}$:Group:Bin	0.008	0.017	0.454	0.650

Eye-gaze competition *during* message preparation in hearing speakers versus deaf signers (post-arrow window)

More complex displays. We assessed visual attention to spatial relations during message preparation in complex displays. The most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for Group. The model contained a main effect of Group, Bin and a two-way interaction (Group*Bin). Thus, both groups experience competition equally from symmetrical competitors, sagittal competitors and topological distractors (see Figure 13; note that lower difference scores indicate more competition). The two-way interaction suggests, however, that deaf signers but not hearing speakers experience increased eye-gaze competition over time independent of the type of competition. The relevant statistics and corresponding coefficients are reported in Table 4.

Figure 13

Eye-gaze competition during message preparation for more complex displays.



The left panel shows hearing speakers. The right panel shows deaf signers during the post-arrow window. Smaller y-axis values indicate more eye-gaze competition.

Table 4

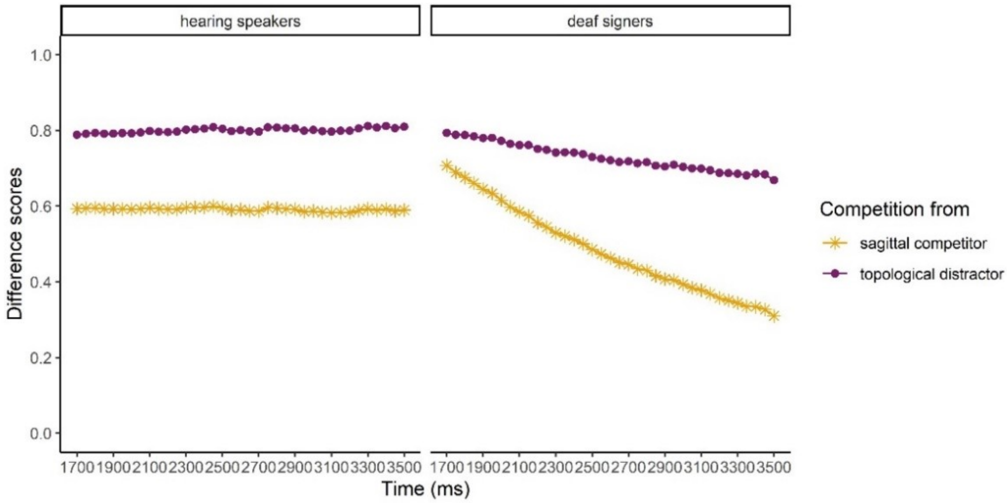
Estimates, standard error, t-values, and p-values of the main and interaction effects of the eye-gaze competition model for more complex displays for the post-arrow window.

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.628	0.086	7.274	0.001
Competition Type$_{[\text{other vs. topological}]}$	0.033	0.218	0.153	0.878
Competition Type$_{[\text{symmetrical vs. sagittal}]}$	0.230	0.222	1.037	0.300
Group	0.609	0.172	3.549	0.001
Bin	-0.131	0.025	-5.302	0.001
Competition Type$_{[\text{other vs. topological}]}$:Group	-0.632	0.437	-1.448	0.148
Competition Type$_{[\text{symmetrical vs. sagittal}]}$:Group	-0.328	0.445	-0.737	0.461
Competition Type$_{[\text{other vs. topological}]}$:Bin	-0.009	0.064	-0.137	0.891
Competition Type$_{[\text{symmetrical vs. sagittal}]}$:Bin	-0.065	0.065	-0.993	0.321
Group:Bin	-0.182	0.049	-3.697	0.001
Competition Type$_{[\text{other vs. topological}]}$:Group:Bin	0.187	0.128	1.457	0.145
Competition Type$_{[\text{symmetrical vs. sagittal}]}$:Group:Bin	0.108	0.131	0.825	0.409

Less complex displays. For less complex displays, the most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for Group. The model contained three main effects (Competition Type, Group; Bin), three two-way interactions (Competition Type*Group; Competition Type*Bin, Group*Bin) and a three-way interaction (Competition Type*Group*Bin). This three-way interaction suggests that, both groups experience more competition from the sagittal competitor compared to the topological distractor (averaged over groups: β = -0.060, *SE* = 0.003, z = -17.520, $p < 0.001$) (see Figure 14; note that lower difference scores indicate more competition). However, as in more complex displays, deaf signers but not hearing speakers experience increased eye-gaze competition over time. The relevant statistics and corresponding coefficients are reported in Table 5.

Figure 14

Eye-gaze competition during message preparation for less complex displays.



The left panel shows hearing speakers. The right panel shows deaf signers during the post-arrow window. Smaller y-axis values indicate more eye-gaze competition.

Table 5

Estimates, standard error, t-values, and p-values of the main and interaction effects of the eye-gaze competition model for less complex displays for the post-arrow window.

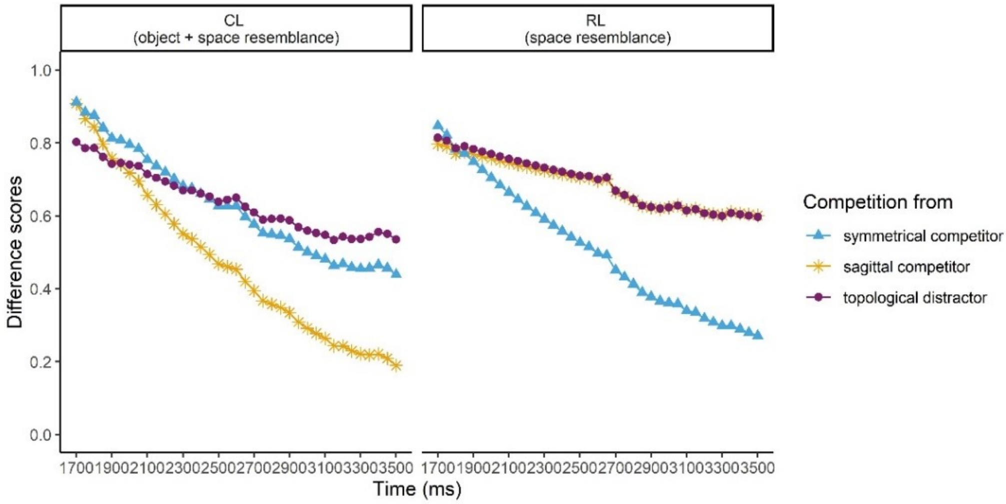
Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.611	0.066	9.300	0.001
Competition Type	0.423	0.127	3.328	0.001
Group	0.727	0.131	5.536	0.001
Bin	-0.118	0.019	-6.320	0.001
Competition Type:Group	0.823	0.254	3.235	0.001
Competition Type:Bin	-0.142	0.037	-3.804	0.001
Group:Bin	-0.221	0.037	-5.914	0.001
Competition Type:Group:Bin	-0.246	0.075	-3.292	0.001

The link between language production and eye-gaze competition during message preparation in deaf signers only (post-arrow window)

Finally, to assess whether deaf signers' eye-gaze competition can be predicted by the type of iconicity in their spatial expressions, we linked the type of iconicity for each picture description (i.e. object + spatial resemblance vs. spatial resemblance) to deaf signers' eye-gaze competition *during* the preparation of that description. For more complex displays, the most parsimonious model incorporated random intercepts for participants and for the by-items random slope for Type of Iconicity. The model contained three main effects (Competition Type$_{[\textit{symmetrical vs. sagittal}]}$, Type of Iconicity, Bin), four two-way interactions (Competition Type$_{[\textit{other vs. topological}]}$*Type of Iconicity; Competition Type $_{[\textit{symmetrical vs.}\ \text{sagittal}]}$*Type of Iconicity; Competition Type $_{[\textit{symmetrical vs.}\ \text{sagittal}]}$*Bin; Type of Iconicity*Bin), and two three-way interactions (Competition Type$_{[\textit{other vs. topological}]}$ *Type of Iconicity*Bin; Competition Type$_{[\textit{symmetrical vs.}\ \text{sagittal}]}$ *Type of Iconicity*Bin). The first comparison of this three-way interaction suggests that when planning CLs, deaf signers experience more competition from sagittal and symmetrical competitors than from topological distractors compared to planning RLs. In addition, the second comparison reveals that when deaf signers plan RLs, they experience more competition over time from the symmetrical competitor than from the sagittal competitor compared to planning CLs (see Figure 15). The relevant statistics and corresponding coefficients are reported in Table 6. Relating this to the results shown in Figure 13, deaf signers in fact *do* divide their attention to the different competitors presented depending on the linguistic strategies they are preparing for. Thus, the current analysis on eye-gaze linked to types of iconic forms reveals that differences in attention to the different competitors is attributed by iconicity in the relational expressions.

Figure 15

Type of iconicity predicts eye-gaze competition during message preparation for more complex displays.



Left panel shows eye-gaze competition when planning CLs. Right panel shows eye-gaze competition when planning RLs. Smaller y-axis values indicate more contrast competition.

Table 6

Estimates, standard error, t-values, and p-values of the main and interaction effects of the model assessing whether the type of iconicity in NGT spatial expressions predict deaf signers' eye-gaze competition in more complex displays during message preparation (post-arrow window).

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	1.412	0.143	9.855	0.001
Competition Type$_{[other\ vs.\ topological]}$	-0.505	0.291	-1.734	0.083
Competition Type$_{[symmetrical\ vs.\ sagittal]}$	-0.680	0.336	-2.024	0.043
Type of Iconicity	0.584	0.277	2.109	0.035
Bin	-0.357	0.040	-8.822	0.001
Competition Type$_{[other\ vs.\ topological]}$:Type of Iconicity	1.272	0.582	2.184	0.029
Competition Type$_{[symmetrical\ vs.\ sagittal]}$:Type of Iconicity	-1.478	0.672	-2.199	0.028
Competition Type$_{[other\ vs.\ topological]}$:Bin	0.153	0.086	1.785	0.074
Competition Type$_{[symmetrical\ vs.\ sagittal]}$:Bin	0.207	0.099	2.101	0.036
Type of Iconicity:Bin	-0.190	0.081	-2.352	0.019
Competition Type$_{[other\ vs.\ topological]}$:Type of Iconicity:Bin	-0.395	0.171	-2.307	0.021
Competition Type$_{[symmetrical\ vs.\ sagittal]}$:Type of Iconicity:Bin	0.449	0.198	2.272	0.023

For less complex displays, the model did not contain a three-way interaction (Competition Type*Type of Iconicity*Bin) (β = -0.244, *SE* = 0.157, t = -1.56, p = 0.119), suggesting that independent of the type of iconicity, deaf signers experience more eye-gaze competition over time from the sagittal competitor compared to the topological distractor in less visually complex displays (for relevant statistics and corresponding coefficients, and graphs see Table S1 in the supplementary materials).

Discussion

To talk about space, spoken languages rely on arbitrary and categorical forms. In sign languages, however, the visuo-spatial modality allows for visually motivated, that is iconic, form-to-meaning mappings of space onto the signing space. The aim of this study was to assess whether the iconic expressions about spatial relations in sign languages guide visual attention to spatial relations differently than the arbitrary and categorical spatial expressions in spoken languages during message preparation at the sentence level. We introduced an adapted version of the visual world eye-tracking experiment to investigate the link between message preparation (spoken vs. signed) and visual attention for spatial relations during message preparation. We expected differences in how visual attention is deployed *during* but not *prior to* message preparation, due to differences in the way hearing speakers and deaf signers linguistically encode spatial relations. Previous studies linking eye-gaze to language production at the sentence level have focussed only on spoken languages and have also found differences in eye-gaze in line with cross-linguistic patterning. This is the first study to investigate whether modality differences such as iconicity across sign and spoken languages also guide visual attention differently when planning messages at the sentence level.

Our study has several key findings. First, hearing speakers and deaf signers described spatial configurations in the most informative manner available (according to the expected modality preferences) regardless of the complexity of displays (i.e. presence vs. absence of a symmetrical competitor). That is, hearing speakers used more categorical than non-categorical expressions while deaf signers preferred iconic forms, CLs over RLs as well as double strategies (i.e. using both CLs and RLs within one description). Second, deaf signers, but not hearing speakers, experienced more eye-gaze competition over time from non-target spatial configurations *during* message preparation (post-arrow window). Importantly, *prior to* message preparation (pre-arrow window), eye-gaze patterns between hearing speakers and deaf signers were similar and did not indicate differences in eye-gaze competition over time as we have observed during message preparation for relational encoding (i.e. post-arrow window). Furthermore, deaf signers' eye-gaze competition was predicted by the type of iconicity in deaf signers' spatial expressions. These results together indicate that the way sign languages are organised (i.e. iconic patterning) modulates deaf signers' visual attention to spatial relations differently than that of hearing speakers *during* message preparation. Below we will discuss the implications of our results in more detail.

Hearing speakers and deaf signers prefer modality-specific ways of expressing spatial relations but both use the most informative strategy to describe left/right spatial relations

We determined to what extent hearing speakers of Dutch and deaf signers of NGT used different linguistic strategies to encode spatial relations across more or less complex displays. Hearing speakers used more categorical than non-categorical expressions. Deaf signers used iconic expressions and within those iconic preferences, they favoured CLs over using RLs, confirming previous findings on such a CL preference (e.g. Emmorey, 2002; Perniss, 2007; Sümer, 2015). Interestingly, NGT deaf signers overall preferred to use double strategies (i.e. CLs and RLs within one description) compared to using one strategy alone. We are the first to demonstrate a preference in NGT for using both CLs and RLs within one description for expressing left/right relations.

Crucially, however, the complexity of displays influenced neither hearing speakers' nor deaf signers' linguistic strategies. That is, hearing speakers did not produce more specific spatial relation adverbs (e.g. LEFT/RIGHT*)* to describe more complex displays and more general spatial relation adverbs (e.g. NEXT TO) to describe less complex displays. Nor did deaf signers use more double strategies or CLs in more complex displays compared to less complex displays. Instead, hearing speakers and deaf signers described spatial configurations in the most informative way available, regardless of the presence or absence of a symmetrical competitor. This is in line with previous research in which speakers were as informative about contrast objects (e.g. big vs. small sausage) during complex displays (eight picture grid) as well as simple displays (four picture grid; Davies & Kreysa, 2017). The current study demonstrated that these findings can be extended to even more complex displays (i.e. 2x2 grid with spatial relations) and to the visuo-spatial modality.

Deaf signers, but not hearing speakers, experienced more eye-gaze competition from spatial configurations *during* message preparation (post-arrow window) but not prior to message preparation (pre-arrow window)

Post-arrow window (during message preparation). For the post-arrow window, that is where we expected message planning to occur specifically for relational encoding, hearing speakers and deaf signers showed competition equally from both the competitors and distractor. However, deaf signers, but not hearing speakers, experienced increasing eye-gaze competition from competitors and distractor over time. For hearing speakers, planning categorical linguistic forms to describe spatial relations did not increase their eye-gaze competition. Deaf signers, unlike hearing speakers, have to map iconic information on to the signing space and thus need to identify the orientation, size, shape, and the relative locations of the objects in visually similar pictures. This seems to result in enhanced visual attention to spatial configurations, independent of whether it is visually or semantically related or not.

During the viewing of less complex displays, similar differences arose in how hearing speakers and deaf signers allocate their visual attention. Thus as for more complex displays, deaf signers, but not hearing speakers, experienced an increase in competition from spatial configurations *during* message preparation for less complex displays. Additionally, we found more competition from the sagittal competitor than from the topological distractor in less complex displays for both groups (which we did not observe in more complex-displays). This could be indeed due to fact that sagittal relations (require viewpoint) are perceived differently than topological relations (Levinson, 2003; Levinson, 1996).

The reason that this difference between sagittal competitor and topological distractor only arises in less complex displays and not in more complex displays could be due to the visual complexity of displays (see Sorensen & Bailey, 2007, for evidence on display complexity overshadowing the eye movement-language processing link). Furthermore, it has been proposed that visual context such as shape similarity between presented stimuli within a visual world paradigm can constrain language-mediated anticipatory eye-movements (Hintz et al., 2019). That is, the presence of too many competitors (and additionally the visual similarity between sagittal and lateral relations) and a distractor might have masked this difference in the complex displays. Furthermore, the use of different types of iconicity by deaf signers might have also masked differences between types of spatial relations in complex displays in this analysis. As we will see below in the following section, eye-gaze patterns linked to the type of iconicity in deaf signers, in fact, reveal differences between sagittal and lateral competitors versus topological distractors in the complex displays depending on which linguistic strategy they are preparing for (CLs vs. RLs).

Pre-arrow window (prior to message preparation). We had expected that during the pre-arrow window, hearing speakers' and deaf signers' viewing would not differ in ways linked to linguistic message preparation of spatial relations. Even though our results unexpectedly indicate a difference between hearing speakers and deaf signers, we argue that this difference is unrelated to the way hearing speakers and deaf signers prepare their linguistic message. In particular, hearing speakers were more likely to prefer looking at lateral relations compared to deaf signers, which was, however, independent of time (i.e. no interaction with Bin, unlike what we see in the post-arrow window). It appears that our visual displays create an overall preference to lateral relations for both groups, which seems to occur more strongly in more complex than in less complex displays. This preference might be rather due to the visual similarity between the two lateral relations and thus might be more driven by characteristics in the visual input (more strongly for hearing speakers) but not related to message preparation (as there is no interaction with Bin). Note that at this stage participants do not know which spatial relation they will encode and thus these looks are unlikely to be related to the relational encoding we are interested in. However and most importantly, when hearing speakers and deaf signers concretely prepare their linguistic message (post-arrow window), different patterns arise which progressed differently in time between hearing speakers and deaf signers while message preparation was unfolding (i.e. increasing competition for deaf signers but not for hearing speakers). Thus, once hearing speakers and deaf signers start to prepare the linguistic message after the appearance of the arrow (i.e. knowing which relation to describe), differences in visual attention arise that relate to the way they prepare their messages and these patterns appear to be different than what we can observe in the pre-arrow window (i.e. when participants do not know which spatial relation to describe). Nevertheless, we do not fully exclude the possibility that during the pre-arrow window participants engage in some sort of linguistic preparation for the lexical items, but importantly, we believe this does not occur for the relations between items. We argue that possible linguistic preparation ongoing in the pre-arrow window does not reflect the active pre-planning stage of the concrete linguistic message regarding the spatial relations. Rather, the pre-arrow window might reflect attentional preferences based on visual similarity or dissimilarity of the pictures (as we indeed find a slight preference to look at lateral relations than at sagittal relations), but also by other factors, for instance, covert object naming.

Note that such differences in the pre-arrow window might influence the interpretation of eye-gaze patterns in the post-arrow window (see Barr, 2008). However, our results provide evidence against such a baseline influence. In particular, the differences between hearing speakers' and deaf signers' eye-gaze pattern in the pre-arrow window arise specifically for lateral relations over sagittal relations (i.e. more looks at lateral relations than sagittal relations for hearing speakers compared to deaf signers).

However, in the post-arrow window, the difference in visual attention between hearing speakers and deaf signers occurs *across all* competitor types and seems thus not to be a result of the group difference observed in the pre-arrow window.

Deaf signers' visual attention was predicted by the type of iconicity in their spatial descriptions

Our data also revealed that in the post-arrow window not only the modality differences of spatial language (arbitrary and categorical speech vs. iconic sign) influenced visual attention to spatial relations *during* message preparation, but also the type of iconicity *within* deaf signers modulated their visual attention. That is, for more complex displays preparing to use CLs to describe left/right configurations led to eye-gaze competition from other spatial competitors compared to the planning of producing RLs. When planning to use CLs, deaf signers experienced more eye-gaze competition from pictures that visually resembled the target picture, thus from the symmetrical and sagittal competitors. Planning CLs requires where to place the hands in space, thus, placing the figure handshape (e.g. long-elongated handshape for pen) to the sides of the ground handshape (e.g. left vs. right vs. front vs. behind the round handshape for glass). Consequently, this mapping of object shapes and spatial relation onto the signing space requires more visual attention to visually resembling spatial configurations. On the other hand, planning RLs resulted in enhanced competition only from pictures that were also semantically relevant, thus from the symmetrical competitor only. Planning RLs require the identification of the relative location in space, thus only semantically relevant pictures (i.e. symmetrical competitor) competed for visual attention. Competition from the topological distractor was not predicted by the type of iconicity as it is neither visually nor semantically relevant for describing left/right target configurations. This confirms our initial expectations about differences between viewpoint dependent spatial relations (sagittal and lateral) versus topological relations in complex displays when specific expressions are taken into account and justifies the choice of spatial relations used in our design.

For less complex displays, deaf signer's visual attention was not predicted by the type of iconicity in their spatial descriptions. Due to the absence of a symmetrical competitor, only one visually similar picture was present in the display (sagittal competitor) and no picture was semantically relevant to the left/right target picture. Thus, deaf signers experienced more competition from sagittal competitors than from topological distractors, independent of whether they planned CLs or RLs.

Taken together, the way deaf signers planned to describe the target picture (i.e. CLs vs. RLs) shifted their attention towards different aspects of the visual display (visual resemblance vs. semantic relevance) that were relevant for message preparation. However, when a symmetrical competitor was absent (i.e. in less complex displays) the type of iconicity did not predict eye-gaze competition from the sagittal competitor or topological distractor. This finding again corroborates the above-mentioned finding that the complexity of the visual display itself influences how visual attention is allocated for language planning.

Conclusion

Overall, our results provide novel knowledge that message preparation for language production affects how hearing speakers and deaf signers allocate visual attention to spatial relations and thus suggests that *thinking for speaking* (Slobin, 2003) differs from *thinking for signing*. We further show that sign production and eye-gaze seem to be linked as speech production and eye-gaze are linked at the sentence level (e.g. Griffin, 2004; Griffin & Bock, 2000; Koneke & Meyer, 2014; van de Velde et al., 2014), although further sign-eye-gaze links during actual signing need to be investigated.

Previous research has demonstrated that iconicity can facilitate (e.g. Grote & Linz, 2003; Thompson et al., 2009) or hinder (e.g. Thompson et al., 2010) sign language comprehension. Our findings reveal that iconicity can affect not only language comprehension but also visual attention during the planning of language production. Our findings show that iconicity effects can surpass the lexical level and can be extended to the sentence level. Furthermore, we extend previous evidence from neuroimaging studies showing that planning CLs not only activate different brain regions involved in expressing spatial language (e.g. parietal regions) than expressing with RLs (e.g. left hemisphere language regions; Emmorey et al., 2013), but can also affect visual attention during message preparation differently.

Overall, the current study provides evidence for the first time that the way languages encode information across modalities (arbitrary speech vs. iconic sign) influences visual attention to entities and relations between them during message preparation at the sentence level. It further suggests that the influence of language on visual attention goes beyond cross-linguistic differences found across spoken languages and that *thinking for speaking* might indeed differ from *thinking for signing*. This study opens up new avenues for looking at whether iconicity involved in different aspects of sign language (e.g. syntactic, discourse level) also influences visual attention. Finally, this study offers novel insights into the relationship between sign production and eye-gaze by showing first-ever that the type of form-meaning mapping in a language can guide visual attention and can be taken further to the view that iconicity can bridge the gap between language and human experience.

Acknowledgements

This work has been supported by NWO VICI Grant awarded to the second author. We thank our deaf and hearing assistants Tom Uittenbogert and Renske Schilte for their help collecting, annotating, and coding data as well as, our deaf and hearing addressees Marcel Koopmans and Julia Merkus. The authors are very grateful to Nick Wood† and Jeroen Geerts for helping us with processing the video data.

Supplementary materials

Table S1

Estimates, standard error, t-values, and p-values of the main and interaction effects of the model assessing whether the type of iconicity in NGT spatial expressions predict deaf signers' eye-gaze competition during message preparation (post-arrow window) in less complex displays.

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.840	0.137	6.115	0.001
Competition Type	0.954	0.263	3.619	0.001
Type of Iconicity	1.099	0.266	4.129	0.001
Bin	-0.186	0.039	-4.813	0.001
Competition Type:Type of Iconicity	0.824	0.527	1.563	0.118
Type of Iconicity:Bin	-0.300	0.077	-3.884	0.001
Type of Iconicity:Bin	-0.343	0.077	-4.433	0.001
Competition Type:Type of Iconicity:Bin	-0.244	0.155	-1.579	0.114

Chapter 3

A tale of two modalities: Iconic signs and speech influence each other in bimodal bilinguals

Chapter adapted from:

Manhardt, F., Brouwer, S., & Özyürek, A. (in press). A tale of two modalities: Signs and speech influence each other in bimodal bilinguals. *Psychological Science.*

Abstract

Bimodal bilinguals are hearing individuals fluent in a spoken and a sign language. Can the two languages influence each other in such individuals despite differences in the vocal (speech) and visuo-spatial (sign) modalities of expression?

We investigated cross-linguistic influences on bimodal bilinguals' expression of spatial expressions. Unlike spoken languages, sign uses iconic linguistic forms that resemble physical features of objects in a spatial relation and thus expresses specific semantic information. Bimodal bilinguals (N = 21) fluent in Dutch and Sign Language of the Netherlands and their hearing speaking and deaf signing peers (N = 20 each, reported in Chapter 2) described left/right relations between two objects.

Bimodal bilinguals expressed more specific information about physical features of objects in speech than hearing speakers, showing influence from sign. They also used fewer iconic signs with specific semantic information than deaf signers, demonstrating influence from speech. Bimodal bilinguals' speech and sign are shaped by two languages from different modalities.

Introduction

In spoken language, bilinguals activate their two languages simultaneously during language use, allowing the languages to influence each other (e.g. Costa, 2005; Indefrey, Sahin, & Gullberg, 2017; Kroll & Gollan, 2014). It is not known, however, if cross-linguistic influence between speech and sign also occurs in bimodal bilinguals, hearing individuals fluent in a spoken (vocal) and a sign (visuo-spatial) language. Here we investigated cross-linguistic influence in bimodal bilinguals who are fluent in Dutch and NGT. We focussed on the domain of spatial language in which differences in modality are mostly visible with respect to how spatial information is encoded (e.g. Emmorey et al., 2013).

Sign languages use the hands and body for linguistic expression. This allows for the use of visually iconic forms, where there are varying degrees of visual resemblance between the form of the linguistic expression and its meaning. This is especially prominent in the domain of spatial language (Figure 16). Most expressions of spatial language are highly iconic and allow signers to express specific semantic information about the physical features of objects in a spatial relation, such as their shape or orientation. For example, in order to describe a picture of a pen located to the right of a glass, signers first introduce the lexical signs for the objects involved (e.g. glass and pen) and later choose iconic handshapes that visually resemble features of each object, such as a round handshape depicting the round shape of the glass and an index finger depicting the elongated shape of the pen (Figure 16A). These handshapes are then placed into the signing space, corresponding in an iconic way to the relative relation and orientation of the objects in the picture (Figure 16A). The iconic forms, called *classifier constructions* (*CLs*, e.g. Emmorey & Herzig, 2003) allow signers but not hearing speakers to express specific information about the physical features of the objects in addition to the location of objects relative to each other.

Signers may also use a less iconic linguistic strategy. They can use *relational lexemes* (*RLs*), such as LEFT and RIGHT (Figure 16B; see also Chapter 2). These are performed towards the left or the right of the body to represent the spatial relations between objects. RLs do not express any specific information about the physical features of the objects. As such, they are more akin to categorical forms such as LEFT and RIGHT in spoken languages.

Figure 16

Two types of iconic linguistic expressions to describe "the pen is to the right of the glass" in NGT.

Panel (A) shows CLs where the hands represent specific information about the shape and orientation of the entities located (glossed as LocL/R) in space (right-most picture). Panel (B) shows the RL for RIGHT, where the handshape for fingerspelling the letter R moves to the right side in space to indicate the location of the pen in relation to the signer's body (right-most picture).

Unlike sign languages, the vocal modality in spoken languages does not allow visual resemblance between form and meaning. Spoken languages also show more variation among each other in the domain of spatial expressions than sign languages which are more similar to each other due to the presence of iconicity. Even though some spoken languages can express more specific semantic information about objects in expression of spatial relations (e.g. Aikhenvald, 2006; Brown, 1994), most spoken languages use arbitrary and categorical speech forms, such as LEFT/RIGHT. These forms lack specific information about physical features of objects (see Chapter 2, for differences in Dutch and NGT in depicting left/right relations).

The present study

Focusing on such differences between spoken and sign languages, we investigated whether and how cross-linguistic influences across modalities occur between Dutch and NGT in bimodal bilinguals' descriptions of left/right relations. NGT-Dutch bimodal bilinguals described spatial relations in each language. These were compared to their hearing speaking and deaf signing peers. Previous studies showed that bimodal bilinguals used more iconic co-speech gestures than their non-signing peers, indicating an influence from sign language (e.g. Casey & Emmorey, 2009; Weisberg et al., 2020). However, this influence is shown only within the same modality but not across modalities; that is, from sign to gesture, but not between sign and speech.

Predictions

In our study, we first explored cross-modal influences from sign to speech and asked whether bimodal bilinguals' use of iconic forms can influence expressions in their speech by enhancing semantically specific information about object shape and orientation in their spoken expressions. Alternatively, such influence might not be present due to differences in modality and iconic versus categorical formats of representation in the two languages. Secondly, we examined influences within one modality. We asked whether bimodal bilinguals produce more iconic gestures as well as CLs along with their spoken utterances (so-called code-blends to refer to signs used along with speech in bimodal bilinguals, Emmorey et al., 2005). Thirdly, in cases where we find cross-modal as well as within-modality influences, we checked whether these influences occur independently of each other. Finally, we explored whether cross-linguistic influences can be bi-directional, in that speech can also influence sign. Bimodal bilinguals might prefer using more lexical signs for LEFT/RIGHT than their deaf signing peers. For example, their forms may be reduced in iconicity and thus be less semantically specific due to the influence of using categorical and arbitrary speech forms such as LEFT/RIGHT in Dutch. If we find bi-directional influences, this will constitute novel evidence that bimodal bilinguals do not resemble two monolinguals in one, as shown for spoken language bilinguals (see Grosjean, 1989); instead, they can be shaped by two languages from different modalities.

Methods

The methods reported in this experiment were approved by the Humanities Ethics Assessment Committee of the Radboud University.

Participants

The participants were 21 hearing native bimodal bilinguals of Dutch and NGT (11 female, M_{age} = 34.33, SD_{age} = 16.62). Additionally, 20 hearing sign-naive speakers of Dutch (11 female, M_{age} = 33.25, SD_{age} = 10.95) and 20 deaf NGT signers (16 females, M_{age} = 34, SD_{age} = 2.5) were recruited as control groups. We determined sample size based on previous research (e.g. Casey & Emmorey, 2009; Pyers & Emmorey, 2008; Weisberg et al., 2020) that used bimodal bilingual population sample sizes lower than

N = 15 per group. However, no power analyses were conducted in this previous research.

All deaf signers were born deaf and raised by deaf parents. They all acquired NGT at an early age from their deaf signing parents and also acquired Dutch in its written form when entering school (M_{age} = 3.5, SD_{age} = 2.8). NGT was the primary language of instruction (for self-rated literacy skills in Dutch, see Table S2 in Supplemental Materials). Four deaf signers received a cochlear implant later in their lives (at ages 12, 30, 37, 48). Hence, none of the deaf signers had access to auditory Dutch from birth. Overall, our deaf signing participants had considerable knowledge of written Dutch. Consequently, they might have experienced some sort of influence from Dutch. However, we expected these possible influences to be less than for bimodal bilinguals, who have more exposure and access to both spoken and written Dutch in production as well as comprehension.

All hearing speakers acquired Dutch as their native language and learnt additional languages (mostly English, German, and French) later in their lives through instructional settings (for more information on hearing speakers' language background, see Table S3 in Supplemental Materials). We selected hearing speakers and deaf signers as control groups because they acquired Dutch and NGT as their first languages without formal instruction.

All bimodal bilinguals were born to at least one deaf parent, except one who grew up with a deaf sister (younger by year; when she was born the family began to learn NGT). Despite this delay in the NGT input, we included this bimodal bilingual participant in our study due to their very early and naturalistic exposure to NGT. Consequently, all bimodal bilinguals had access to NGT naturally at home and acquired Dutch from the surrounding community from early on (i.e. similar to language populations studied in De Quadros, 2018; Pichler, Lillo-Martin, & Palmer, 2018; Quadros & Lillo-Martin, 2018).

Eight of the 21 bimodal bilinguals were trained sign language interpreters. On a five-point Likert scale, bimodal bilinguals rated their language use (1 = never; 2 = rarely; 3 = sometimes; 4 = most of the time; 5 = all the time) and their proficiency (1 = beginner, 2 = intermediate, 3 = advanced, 4 = native-like, 5 = native) in Dutch and NGT separately, for both comprehension and production. Comprehension scores for Dutch included scores for reading and listening, while the scores for NGT included understanding. Production scores for Dutch included speaking and writing, while the scores for NGT included signing.

Bimodal bilinguals indicated that they use Dutch ($M = 4.80$, $SD = 0.41$) more often than NGT ($M = 3.65$, $SD = 0.93$; paired samples t-test: $t(20) = -5.21$, $p<0.001$, *Cohen's d* = -1.59; using package *effsize* (version 0.7.6; Torchiano, 2019) to calculate effect sizes). All bimodal bilinguals rated their proficiency for production in NGT and Dutch to be somewhere between advanced and native-like, although scores were significantly higher for Dutch ($M = 4.55$, $SD = 0.51$) than for NGT ($M = 3.85$, $SD = 0.93$; paired samples t-test: $t(20) = -2.77$, $p = 0.012$, *Cohen's d* = -0.89; for similar rating asymmetry in bimodal bilinguals see, De Quadros, 2018; Quadros & Lillo-Martin, 2018). For comprehension, bimodal bilinguals rated themselves as almost native in both Dutch ($M = 4.75$, $SD = 0.44$) and NGT ($M = 4.50$, $SD = 0.69$) (paired samples t-test: $t(20) = -1.56$, $p = 0.14$, *Cohen's d* = -0.43). All bimodal bilinguals reported Dutch as the language they speak best and feel most comfortable with. We additionally assessed language proficiency using objective assessment tools for NGT and Dutch. To do so, we used fluency measures such as grammatical and narrative skills for NGT and speech rate in Dutch.

For NGT, we used an assessment tool for narrative production created for British Sign Language development (Herman, Grove, Holmes, Morgan, Sutherland, & Woll, 2004). Retellings of a 3.41 min video were scored by a deaf native NGT signer on narrative structure, grammar, and, in particular, the use of CLs. Scores indicated no differences between bimodal bilinguals and deaf signers in narrative structure (independent samples t-test: $t(38) = 0.22$, $p = 0.83$, *Cohen's d* = 0.07) or in the use of CLs and other grammatical aspects (independent samples t-test: $t(38) = 1.52$, $p = 0.14$, *Cohen's d* = 0.05), showing high NGT proficiency and appropriate use of CLs in our bimodal bilingual sample.

For Dutch, we assessed speech fluency by measuring participants' speech rate (number of syllables/time) using the speech analysis software Praat (Boersma & Weenink, 2001). Previous research suggests that speech rate and fluency are tightly linked. In particular, higher speech rates are associated with higher language proficiency (e.g. Daller, Yıldız, de Jong, Kan, & Başbaĝi, 2011; Polinsky, 2008). In order to measure speech fluency, we selected a 30-second speech sample from participants' retelling the same narrative that was used for the NGT assessment and calculated participants' speech rate (for the script, see de Jong & Wempe, 2009). Bimodal bilinguals' speech rate ($M = 3.56$, $SD = 0.46$) did not significantly differ from that of Dutch hearing speakers ($M = 3.39$, $SD = 0.43$) (independent samples t-test: $t(39) = -1.21$, $p = 0.23$, *Cohen's d* = -0.39), showing evidence for high Dutch proficiency in the bimodal bilingual sample.

Materials

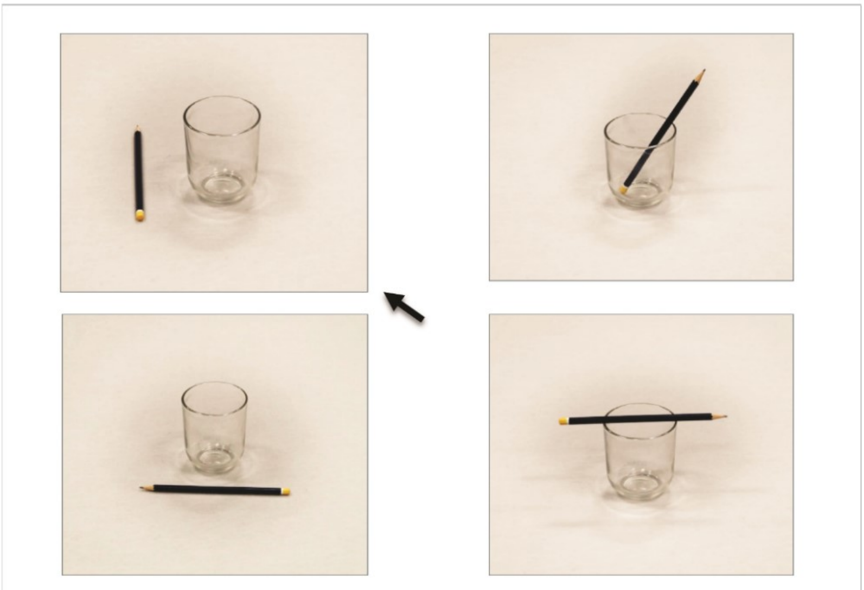
We used the same stimuli as in Chapter 2. These stimuli consisted of 84 displays of four pictures. Each display contained the same two objects but in different spatial relations to each other (i.e. left, right, front, behind, on, in; see Figure 17). An arrow appeared in the centre of the screen and indicated the target picture. We used 28[5] displays as experimental trials in which the target picture always displayed a left or right configuration, while the remaining three pictures depicted other spatial configurations (i.e. front/behind, in/on). In our analysis, we focussed on left/right relations. Chapter 2

[5] In this chapter we did not distinguish between less and more complex displays, because Chapter 2 indicated that the type of displays does not influence language production in hearing speakers and deaf signers. We therefore assumed that also for bimodal bilinguals this would not influence their language productions and thus collapsed both less and more complex displays.

had shown that deaf signers of NGT use a variety of CLs and RLs in describing left/right relations. However, for other spatial relations signers preferred mostly one type of sign; that is, RLs for pictures displaying topological relations such as in/on and CLs for pictures showing sagittal relations such as front/behind. Thus, we decided to focus on left/right relations because they offer variation in the type of signs that signers can choose to use. This allowed us to detect possible cross-linguistic influences in bimodal bilinguals. We included 56 fillers in which target pictures displayed other spatial relations such as front/behind and in/on in addition to left/right.

Figure 17

Example of an experimental display with ground and figure objects in different spatial relations.



The arrow shows the target picture that needs to be described to the addressee.

We pre-tested all visual displays to ensure that participants could name the presented objects. We used two sets of lists which were counterbalanced within and across participants and also across sessions. Consequently, bimodal bilinguals did not describe the same pictures across their two sessions.

All experimental displays contained a ground object in the middle and figure objects located around the ground object. Ground objects (e.g. glass in Figure 17) did not have intrinsic directionality or orientation such as top or bottom. In this way, we made sure participants used left/right expressions rather than front/behind. The majority of figure objects (e.g. pencil in Figure 17) did have intrinsic directionality (N = 48 out of 56), allowing a differentiation in object orientation. All ground and figure objects were everyday objects that were familiar to the participants and allowed the use of basic handshapes for CLs (for a complete list of objects, see Table S4 in Supplemental Materials). We counterbalanced the location of each configuration across participants and trials. Furthermore, we kept the orientation of intrinsic figure objects equal across different types of displays. That is, in left/right configurations, we placed elongated objects vertically in relation to the ground object, while the top of the object was pointing upwards on the screen. In front/behind/on configurations, we placed the figure object horizontally in relation to the ground object. For in configurations, we placed the figure object vertically inside the ground object. We kept the distance between the ground and figure objects equal across displays for the different spatial relations (i.e. for left/right, front/behind, etc.).

Procedure

We tested bimodal bilinguals twice, once in NGT and once in Dutch. The two sessions were three to five weeks apart and the order of sessions was counterbalanced across participants. Hearing speakers and deaf signers were tested only once in their native language. In addition to the speaker/signer participant, each session also included an addressee to elicit a natural communicative situation and informative descriptions. The addressee was not a naïve subject but was employed by the experimenter to create a communicative setting for the actual participant.

We tested participants individually on a laptop. Three practice trials preceded the experiment. Trials started with a fixation cross shown for 2000 ms, followed by a display introducing the four pictures for 1000 ms to familiarize participants with the objects. After that, an arrow pointed at one of the four pictures and disappeared after 500 ms. The four pictures remained on the screen for 2000 ms so that participants had time to plan their spatial descriptions. At the end of 2000 ms, a grey visual noise screen was presented. Participants had to describe the picture indicated by the arrow to the addressee after the grey screen had disappeared. Once a description was completed, the trained addressee pretended to choose the correct picture from the same four-picture set on a separate tablet. In each trial participants' and interlocutors' picture displays were arranged differently from each other (e.g. the participant's target picture was in the upper-left corner, while the same picture was located in the lower-right corner on the interlocutor's tablet). This four-picture set up ensured that the participant had to give a full and informative description of the target picture so that it could be identified by the addressee. Participants initiated the next trial by pressing the ENTER button. We controlled the timing of each trial element (e.g. fixation cross, introduction of four-picture display) to ensure that all participants had equal viewing times of the visual displays before describing them.

Importantly, bimodal bilinguals were not tested in a bimodal bilingual context since that could have enhanced cross-linguistic influence. Thus, during the Dutch sessions, the addressee was always a hearing Dutch hearing speaker. During NGT sessions, both the interlocutor and the experimenter were deaf native NGT signers. Addressees gave no feedback during the sessions in order to not influence or bias the descriptions. Participants were free to choose how much information they wanted to convey in their picture descriptions to help the addressee identify the correct spatial relation.

Data coding

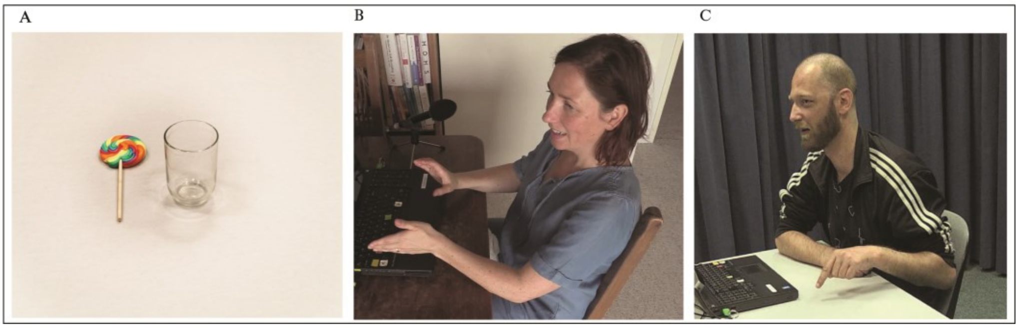
We coded the linguistic descriptions using ELAN, a free annotation tool (http://tla.mpi.nl/tools/tla-tools/elan/). Trained hearing native Dutch and deaf native signers of NGT performed annotations and coded the data. We coded Dutch data for speech and manual productions (i.e. iconic co-speech gestures/CLs) that occurred accompanying speech across hearing speakers and bimodal bilinguals. We coded NGT data for signed productions. An additional coder checked all coding to find consensus. If no consensus could be found, we excluded the descriptions from further analyses (5.81% of all descriptions/trials).

Speech productions (Dutch spoken sessions). For speech, we assessed for each picture description whether it contained semantically specific information or not. If descriptions contained only an introduction of the objects and type of the spatial relation (e.g. “the lollypop is to the left of the glass”), we considered them as expressions that were not semantically specific. If descriptions contained additional information about objects’ shape or orientation (e.g. “the lollypop is to the left of the glass and the sugar part of the lollypop is pointing upwards”), we considered them as semantically specific expressions.

Manual productions accompanying speech (Dutch spoken sessions). For manual productions accompanying speech, the deaf native NGT signer distinguished CLs from co-speech gestures to the best of his abilities. We then assessed for each picture description whether or not it contained manual productions with semantically specific information (either through iconic co-speech gestures or CLs, see Figure 18). We considered manual productions as *semantically specific* if they depicted object shape or object orientation (Figure 18B). We considered manual productions as *not semantically specific* if they only indicated the location of the figure object in relation to the ground object without depicting shape or orientation information (e.g. with small beat-like hand movements, pointing to the left or right of the participant's body, or using RLs for LEFT/ RIGHT; see Figure 18C).

Figure 18

Examples of manual productions accompanying bimodal bilinguals' speech in Dutch.



Panel (A) shows an example of a target picture. Panels (B) and (C) show examples of different types of manual productions by bimodal bilinguals used to describe the target picture in Dutch. Panel (B) illustrates an example of a manual production that expresses semantically specific information. While uttering, "the lollipop is lying left of the glass" the left hand shows the shape and orientation of the lollypop's stick and the right hand the shape of the glass. Panel (C) illustrates a manual production, which conveys non-specific semantic information. A pointing gesture to the left indicates the location of the lollipop accompanying the spoken expression "a glass with a lollipop on the left side".

Sign productions (NGT signed sessions). For each signed picture description in NGT, we first distinguished between two main types of linguistic strategies: a) CLs, which are highly iconic and semantically specific (Figure 16A) and b) RLs for LEFT/RIGHT (Figure 16B), which are more categorical than CLs and not semantically specific. Second, we coded for double strategies; that is, descriptions that contained RLs for LEFT/RIGHT followed by CLs within one response (CLs followed by RLs for LEFT/RIGHT can also occur, although less frequently, see also Chapter 2 for NGT; Sümer, 2015 for Turkish Sign Language).

Data analysis

For Dutch (during Dutch spoken sessions), we first analysed whether bimodal bilinguals produced more semantically specific picture descriptions than their non-signing peers. Next, we analysed whether bimodal bilinguals were more likely to produce spoken descriptions together with manual productions that express semantically specific information than hearing speakers. Finally, we combined the two types of data and analysed occurrences of semantically specific descriptions per modality (i.e. speech only, manual only, speech + manual) in bimodal bilinguals only. The aim of this analysis was to assess whether sign can influence speech and the expressions in manual modality accompanying speech in an independent manner. For example, if we find changes in speech, this could be due to either the adaptation of speech to the increased use of iconic co-speech gestures/CLs in the manual modality or to direct cross-linguistic influences from sign to speech. This analysis aimed to distinguish between these two types of influences.

For NGT (during signed descriptions), in order to assess whether speech can influence sign, we analysed whether bimodal bilinguals produced more descriptions with RLs for LEFT and RIGHT than their deaf signing peers, reflecting the influence of using categorical expressions in Dutch.

For all analyses, we used binominal data as the dependent variable. Data were averaged over picture descriptions/trials and participants. All analyses were conducted using R (version 3.3.1; R Core Team, 2013). We used the package *lme4* (version 1.1-19; Bates, Mächler, Bolker, & Walker, 2015) to analyse the data in a logistic regression or in a generalized linear mixed-effects regression (Baayen et al., 2008). Categorical predictors were coded as numeric contrasts. We used a backward selection procedure in which insignificant predictors were removed.

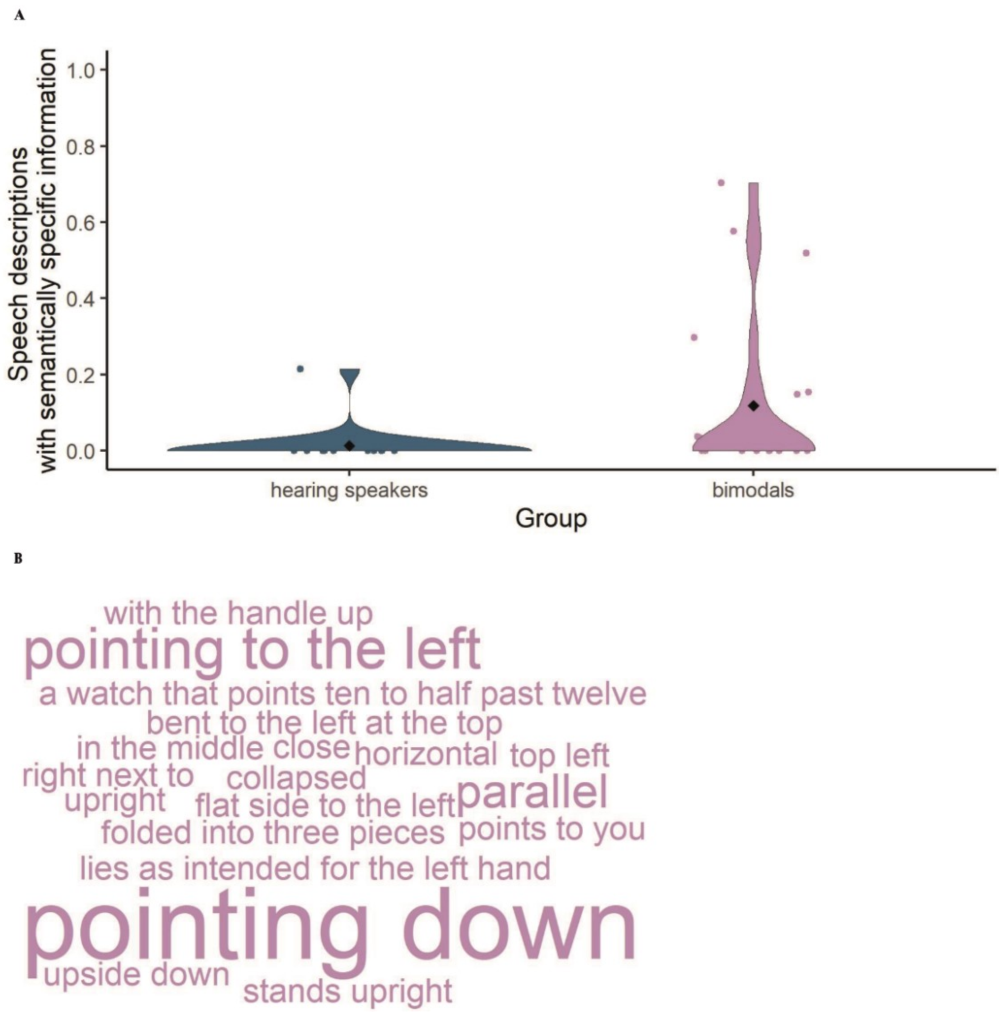
Results

Sign influences speech

The results of a logistic regression comparing whether bimodal bilinguals produced more picture descriptions with semantically specific information than hearing speakers demonstrated a main effect of group ($\beta = 2.75$, $SE = 0.31$, $z = 9.01$, $p<0.001$; see Figure 19A). Bimodal bilinguals produced more speech descriptions that were semantically specific ($M = 0.12$, $SD = 0.32$) than their non-signing peers ($M = 0.03$, $SD = 0.11$). Figure 19B illustrates the variety of semantically specific descriptions used by bimodal bilinguals (for the original Dutch version, see Figure S1 in Supplemental Materials). Object orientation was the most commonly emphasised semantically specific feature. There was no effect of session order for bimodal bilinguals ($\beta = 3.52$, $SE = 2.11$, $z = 1.67$, $p = 0.10$), indicating that bimodal bilinguals' semantically specific descriptions are not due to or primed by describing similar pictures in an earlier session in NGT (for information on the order analysis, see Supplemental Materials).

Figure 19

Speech descriptions with semantically specific information in hearing speakers and bimodal bilinguals and the variety of such descriptions in bimodal bilinguals.



Panel (a) shows proportions of speech descriptions that contained semantically specific information in hearing speakers and in bimodal bilinguals. Dots represent each data point (participant). Squares represent the mean. The width of the violin plots indicates the density and the length of the violins show the range of data points. Panel (b) illustrates the variety of semantically specific descriptions used by bimodal bilinguals (English translations given). Larger font size indicates a higher frequency of usage.

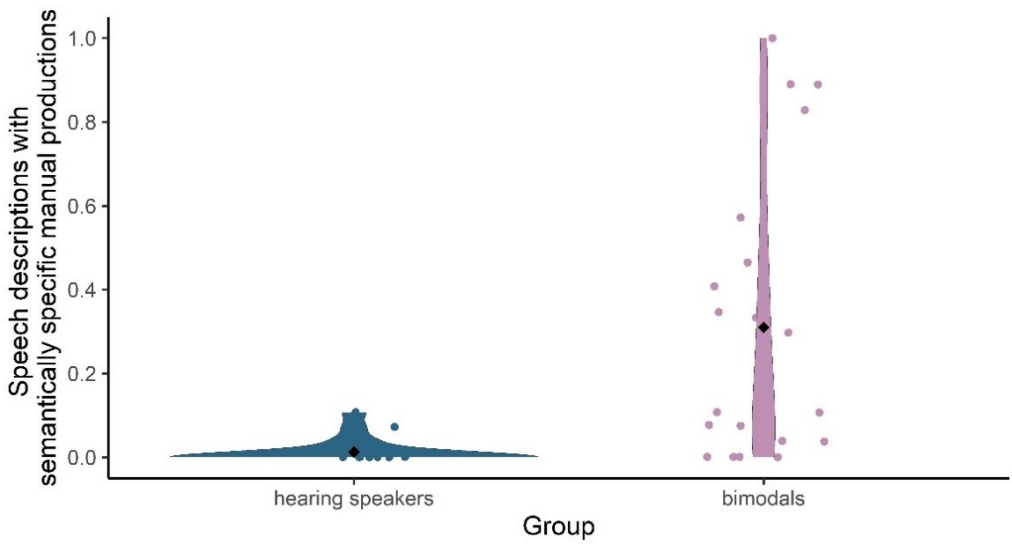
Sign influences manual productions accompanying speech

Bimodal bilinguals' overall manual productions were more likely to consist of CLs accompanying speech (*M* = 0.69, *SD* = 1.13) compared to co-speech gestures (*M* = 0.11, *SD* = 0.36), while hearing speakers rarely gestured (*M* = 0.05, *SD* = 0.26). We assessed whether picture descriptions were accompanied by manual productions that expressed semantically specific information across hearing speakers and bimodal bilinguals. Note that we did not distinguish between CLs accompanying spoken expressions and iconic co-speech gestures because we aimed to assess the amount of semantically specific information that each manual production conveyed in general.

A logistic regression analysis showed that bimodal bilinguals produced more speech descriptions that were accompanied by semantically specific manual productions (*M* = 0.31, *SD* = 0.46) than their non-signing peers (*M* = 0.01, *SD* = 0.11; *β* = 3.77, *SE* = 0.39, *z* = 9.70, *p*<0.001; see Figure 20). There was no effect of session for bimodal bilinguals (*β* = -0.50, *SE* = 1.52, *z* = -0.33, *p* = 0.74), indicating that the increase in bimodal bilinguals' descriptions containing semantically specific manual productions is not due to, or primed by describing similar pictures in an earlier session in NGT (for information on the order analysis, see Supplemental Materials).

Figure 20

Speech descriptions produced with semantically specific manual productions in hearing speakers and bimodal bilinguals.



Dots represent each data point (participant). Squares represent the mean proportions. The width of the violin plots indicates the density and the length of the violins shows the range of data points.

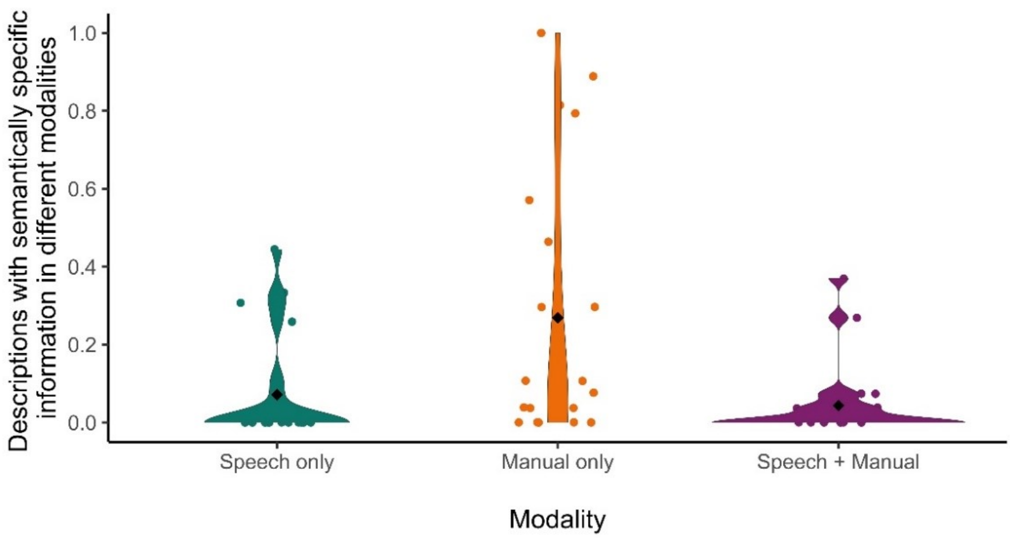
Semantically specific information expressed in different modalities

After showing that sign influences both speech and the manual modality accompanying speech, we investigated next whether either type of influence occurs independently or whether expressions in one modality are simply adaptations to the information expressed in the other. To do so, we assessed whether semantically specific expressions occurred in speech or in the manual modality only, or in both modalities simultaneously. The former would indicate independent influences from sign and the latter would indicate adaptation of changes in one modality to the other.

The most parsimonious generalized linear mixed-effects regression model included random intercepts for participants and items, as well as uncorrelated random slopes for modality by items. We found that bimodal bilinguals produced more descriptions that are semantically specific in speech only (*M* = 0.07, *SD* = 0.26) or in the manual modality only (*M* = 0.27, *SD* = 0.44) than descriptions in both modalities simultaneously (*M* = 4.36, *SD* = 0.20; *β* = 1.50, *SE* = 0.37, *z* = 3.89, *p*<0.001; see Figure 21). Thus, an increase in semantic specificity in speech descriptions or in manual productions can occur independent of each other and show evidence of influence from sign directly. In a second analysis we compared speech only to manual only productions. This analysis revealed that within-modality influence from sign to manual productions is more common than cross-modal influence from sign to speech (*β* = -2.17, *SE* = 0.34, *z* = -6.23, *p*<0.001).

Figure 21

Semantically specific information expressed in different modalities in bimodal bilinguals.



Dots represent each data point (participant). Squares represent the mean proportions. The width of the violin plots indicates the density and the length of the violins show the range of data points.

Speech influences sign

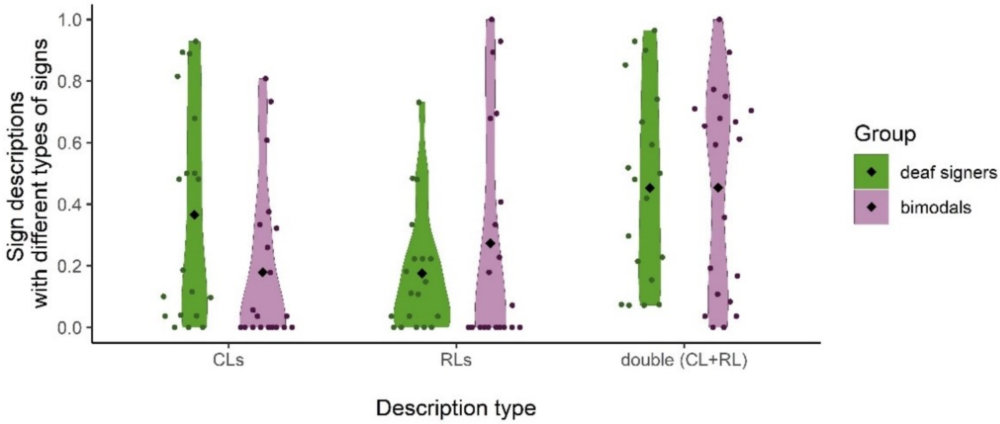
To investigate whether speech influences signed expressions, we assessed whether bimodal bilinguals expressed fewer semantically specific expressions; that is, more RLs of LEFT/RIGHT and fewer CLs in NGT due to the influence from using categorical forms in Dutch (e.g. LEFT/RIGHT). We therefore analysed the different description types such as RLs only versus CLs only versus double strategies in deaf signers and bimodal bilinguals (group).

The most parsimonious generalized linear mixed-effects regression model included random intercepts for participants and items, as well as random slopes for description type by items. The analysis yielded main effects of description type and of group. Most importantly, the interaction between description type and group was significant when comparing bimodal bilinguals' descriptions with RLs (M = 0.27, SD = 0.45) and CLs (M = 0.18, SD = 0.38) to their deaf native signing peers (RLs: M = 0.18, SD = 0.38; CLs: M = 0.37, SD = 0.48; β = 1.76, SE = 0.27, z = 6.52, p<0.001; see Figure 22). The bimodal bilinguals used more descriptions with RLs and fewer with CLs compared to their signing peers. No significant interaction was found when comparing bimodal bilinguals and deaf signers on their descriptions containing double strategies (β = -0.25, SE = 0.25, z = -0.97, p = 0.33).

There was a main effect of session order for bimodal bilinguals (β = 1.75, SE = 0.26, z = 6.66, p<0.001; see Supplemental Materials for more details). This suggests that when NGT sessions followed Dutch sessions, bimodal bilinguals used fewer mixed descriptions with double strategies (M = 0.41, SD = 0.49) than descriptions with CLs or RLs alone, compared to when NGT sessions preceded Dutch sessions (M = 0.50, SD = 0.49). Crucially, the increased use of RLs in bimodal bilinguals' signed descriptions was not related to the order of sessions. Rather, it was related to cross-linguistic influences across modalities from Dutch to NGT.

Figure 22

Sign descriptions with different types of signs in deaf signers and bimodal bilingual.



Black dots represent each data point (participant). Squares represent the mean proportions. The width of the violin plots indicates the density and the length of the violins shows the range of data points.

Discussion

Bimodal bilinguals expressed more specific information about the physical features of objects, such as their shape and orientation, in speech than did hearing speakers, showing an influence from sign to speech. This extends findings regarding cross-linguistic influences within a single modality, specifically in the domain of spatial language in bilinguals (e.g. Indefrey et al., 2017), to cross-linguistic influences across modalities. It further shows that language distance is not an obstacle for cross-linguistic influences in bilinguals (contrary to previous suggestions, e.g. Kellerman, 1979) and can be found even between languages expressed in completely different formats.

Furthermore, in line with previous research (e.g. Casey & Emmorey, 2009; Gu, Zheng, & Swerts, 2018), bimodal bilinguals produced more semantically specific manual productions such as iconic gestures/CLs accompanying their speech than their non-signing peers. We also showed that both cross-modality and within-modality influences from sign to speech or to manual productions could occur independently. Interestingly, the influence from signs to manual productions accompanying speech was stronger than from CLs to speech. Furthermore, enhanced semantic specificity in speech was mostly limited and evident through orientation information about the objects. Semantic specificity in manual productions, however, was found through a variety of specific features, namely, size, shape and orientation. Therefore, even though sign can influence speech, this influence is more constrained than influence from sign to the manual modality accompanying speech. This suggests that cross-linguistic influences are stronger within the same modality; that is, from sign to manual productions, than across modalities; that is, from sign to speech. This can be due to the lack of corresponding forms between sign and speech.

Lastly, we showed that cross-modal influences can occur in bimodal influences not only from sign to speech but also from speech to sign. Bimodal bilinguals produced fewer descriptions with CLs and more descriptions with RLs for LEFT/RIGHT than their deaf signing peers. This demonstrates that bimodal bilinguals used fewer semantically specific expressions in their signs due to influence from speech.

Overall, both the bi-directional influences between sign and speech and the influence from sign to the manual modality suggest that bimodal bilinguals activate NGT while speaking Dutch, and vice-versa. Moreover, these co-activations can occur across modalities. This supports previous claims about the co-activation of sign languages while using spoken languages (e.g. Emmorey, Borinstein, Thompson, & Gollan, 2008; Giezen & Emmorey, 2016), extending evidence from the lexical to the sentence level of spatial descriptions.

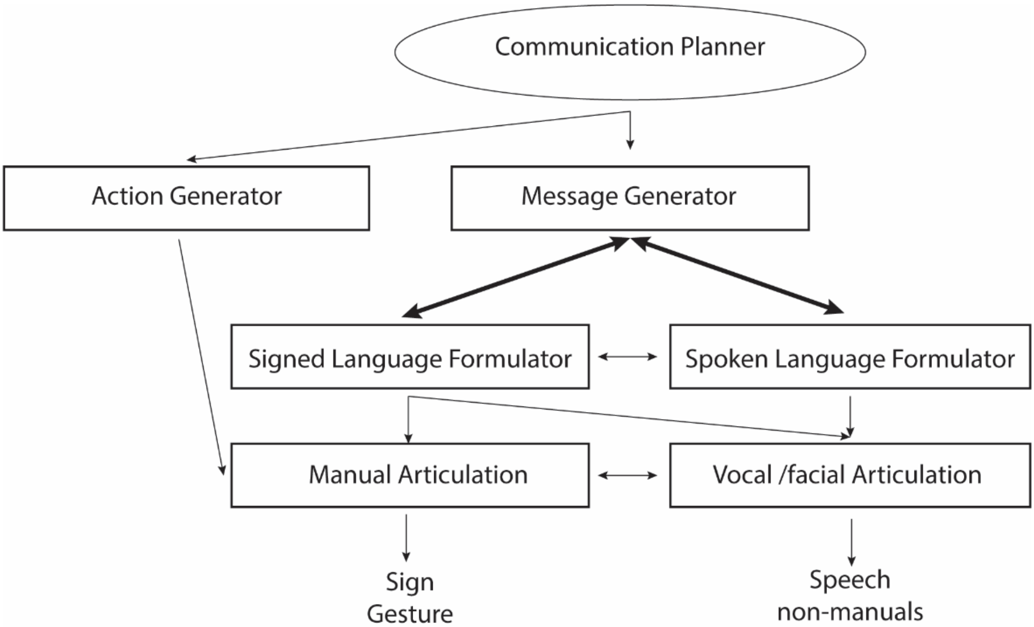
Taken together, these findings provide further evidence for an existing bimodal bilingual language production model (see Emmorey et al., 2008; based on Kita & Özyürek, 2003, Levelt, 1989; additionally see Lillo-Martin, de Quadros, & Pichler, 2016).

The model proposes a shared Message Generator (preverbal message) but separate and interfacing production systems (i.e. Formulators) for sign and speech (Figure 23). Additionally, it involves an Action Generator (a general mechanism for creating action plans) responsible for the production of co-speech gestures and which interacts with the Message Generator.

Accordingly, we propose that semantically specific information in the Message Generator used for iconic NGT expressions such as CLs, could influence bimodal bilinguals to produce more semantically specific expressions in speech through the Spoken Language Formulator. This then results in cross-linguistic influence from NGT to Dutch. At the same time, manual productions accompanying Dutch could be generated through the route from the Message Generator influenced by NGT to the Action Generator, as has been proposed before for American Sign Language (ASL) (Casey & Emmorey, 2009). Speech can influence sign through the Message Generator, which is affected by categorical forms in Dutch. The Message Generator influenced by Dutch might then influence the Signed Language Formulator, resulting in more lexical, less semantically specific signs in NGT. Our study provides first evidence for the routes between the Message Generator and both Signed and Spoken Language Formulators (bold lines in Figure 23 below). This is also in line with assumptions that commonly used expressions in a specific language would guide attention to certain aspects of states and events (Chapter 2; Slobin, 2003) at the level of the Message Generator (Kita & Özyürek, 2003), extending it to the mind of bimodal bilinguals.

Figure 23

A model of bilingual language production.



Bold arrows indicate where our findings provide additional evidence for the model. Figure adapted from Emmorey et al. (2008).

Finally, our findings contribute to the debate on using sign language as well as spoken language with deaf and hard of hearing children, such as for those receiving cochlear implants. Often speech and hearing professionals advise against the use of sign language after receiving cochlear implants, arguing for a negative interference of sign on the development of speech (e.g. Humphries et al., 2012). Our findings show that being exposed to sign from birth together with spoken language does not hinder speech and communication but might make it even richer, as bimodal bilinguals' speech and signs are shaped by two dynamically intertwined languages from different modalities.

Acknowledgements

This work has been supported by NWO VICI Grant awarded to the last author. We thank our bimodal bilingual participants and our deaf research assistant Tom Uittenbogert as well as Dilay Z. Karadöller, Bezya Sümer and Ercenur Ünal for their feedback and help with the study design. We also thank our interlocutors Eveline van Wijk, Julia Merkus, Marlijn Metzlar, Maud van der Wouw, and Renske Schilte, Els Gielen, Madeleine Pairan, Linda Lamers, Nick Wood† and Jeroen Geerts for their help in handling, processing, and coding data.

Supplementary materials

Table S2

Self-rated scores (means, SD in brackets) from deaf native NGT signers for Dutch reading and writing proficiency on a five-point Likert scale (1 = beginner, 2 = intermediate, 3 = advanced, 4 = native-like, 5 = native) and for Dutch and NGT use (1 = never, 2 = rarely, 3 = sometimes, 4 = most of the time-like, 5 = all the time).

deaf NGT signers			
language proficiency		language use	
Dutch-reading	**Dutch-writing**	**Dutch-frequency**	**NGT-frequency**
3.90 (0.85)	3.95 (0.69)	4.10 (0.97)	4.80 (0.52)

Table S3

Language background from hearing Dutch speakers about their second language, its age of acquisition and frequency of use on a five-point Likert scale (1 = never, 2 = rarely, 3 = sometimes, 4 = most of the time-like, 5 = all the time).

hearing Dutch speakers		
Most frequent other languages	**AoA L2**	**Frequency of L2 use**
English, German	13.56 (7.38)	3.10 (0.85)

Table S4

List of objects presented across experimental displays and filler displays. Ground objects were repeated and paired with different figure objects. Objects were distributed across two different lists (i.e. each ground object appeared four times within each list, always paired with another figure).

	Ground	Figure	Ground	Figure	Ground	Figure	Ground	Figure
Experimental displays	BASKET	NEWSPAPER	BOWL	TOMATO	MONEY BOX	LOCK	VASE	GLOVE
		HAMMER		RULER WOODEN		CAR KEY		CORN
		NAIL POLISH		TREE BRANCH		LEAF		WALLET
		LADLE PLASTIC		COOKING SPOON METAL		RIBBON		COMB
		KNIFE		WHISKER METAL		MONEY		CANDLE LONG
		FLUTE		TONGS		STRAW		BANANA
		TOILET PAPER ROLE		CUTTING BOARD		STONE		FLOWER POPPY
		REMOTE		BREAD		BALLON ORANGE		SCISSORS BIG
	CUP	TAPE	JAR	WHISTLE	PLANT POT	TORCH		
		PASTRY		JAR LID		HAND FAN		
		EGG		SOAP BAR		SMART PHONE		
		CHOCOLATE		RAZOR		DISH SPONGE		
		PENCIL		LEMON		ORANGE		
		FORK		WRIST WATCH		SUN GLASSES		
		TOOTHBRUSH		MANDARIN		WOOL		
		LOLLIPOP		SCREWDRIVER		CARROT		
Filler displays	BAG	PEAR	ENVELOPE	CLOTHES PIN	PAN	COOKING BRUSH	SHOE	LEAF
		CHAIN		THEORMOMETER		SMALL ROLLING PIN		BALLOON GREEN
		TISSUE		BATTERY		TOWEL		SHOE SPONGE
		CHARGER		NAIL BRUSH		STRAINER		LIGHT BULB
		CARD		PILLS		EGGPLANT		SOCK
		HEADPHONES		BRACELET		SPATULA WOODEN		CAMERA
		PARFUME2		NAIL CLIPPER		BELLPEPPER GREEN		LIGHTER
		GIFTBOX		KEY		COOKING SPOON PLASTI		BOARDMARKER SHORT
	BOOK	MAKEUP BRUSH	FOODBOX	LID	PAPER	CAP	SLIPPER	CREAM TUBE
		CRAYON		BANNER		HAIR BAND		PURSE
		TEABAG		FLY FLAP ROUND		STAMP		INJECTION
		CROWN		NAPKIN TEXTILE		KIWI		CD
		PACIFIER		AGENDA		PEN		GLASSES
		BOARDMARKER LONG		PENCIL CASE LONG		ERASER		BOX METAL
		DUSTER		FLOWER ROSE		ONION		PHONE
		GARLIC		PAINT BRUSH PLASTIC		COKE		PARFUME1
	BOX	ALARMCLOCK	HAT	TIE1	PLATE	LEMONJUICER	TRAY	SALT
		VIDEO CAMERA		TSHIRT		GREEN FLAG		FUNNEL
		CALCULATOR		SCARF		GLASSES CASE		STRING
		MICROPHONE		IPAD		TIE2		BOTTLE OPENER
		HAIRDRYER		UMBRELLA LONG		NAPKIN PAPER		SMALL BOTTLE
		PENCIL CASE ROUND		WATER BOTTLE		CHEESE		BALL
		TREE BRANCH WITH LEAVE		RULER PLASTIC		CUTTING KNIFE		CANDLE SMALL
		MAGNIFIER		NOTEBOOK		HAIR BRUSH		APPLE
	COOKING POT	UMBRELLA SHORT	MUG	COMPUTER MOUSE				
		FLY FLAP SQUARE		DISH SPONGE				
		WHISKER PLASTIC		SIEVE				
		GRATER		MIRROR				
		LADLE METAL		SPOON				
		POINTED BELLPEPPER RED		SCISSORSS MALL				
		KITCHEN BRUSH		CREDIT CARD				
		SPATULA PLASTIC		PAINT BRUSH WOOD				

Information about session order analysis

For Dutch, we assessed possible effects of session order on the frequency of semantic specificity in speech and its manual modality in bimodal bilinguals. We used two logistic regression analysis, of which one assessed semantic specificity in speech and the other assessed semantic specificity in the manual modality. For NGT, we assessed possible effects of session order on description type in bimodal bilinguals. The most parsimonious generalized linear mixed effects regression model included random intercepts for participants and items, as well as uncorrelated random slopes for description type by items.

Results

Figure S1

Variety of semantically specific descriptions used by bimodal bilinguals (original Dutch version)

The larger the font size, the higher the frequency.

Chapter 4

Iconicity influences visual attention during signed and spoken message preparation in bimodal bilinguals

Abstract

Speech and sign influence each other bi-directionally in bimodal bilinguals when describing spatial relations. Chapter 3 showed that bimodal bilinguals increased the use of semantically specific forms in speech as well as the number of iconic co-speech gestures, and at the same time, decreased the use of semantically specific forms during signing compared to hearing speaking and deaf signing controls. This chapter asks whether this cross-linguistic influence changes the way bimodal bilinguals conceptualise these spatial relations during message preparation before speaking and signing differently than hearing speaking and deaf signing controls using similar materials and procedure.

We compared visual attention from the same bimodal bilinguals (N = 21, reported in Chapter 3) when preparing to sign in NGT versus when to speak in Dutch to that of hearing speaking and deaf signing controls (N = 20 each, reported in Chapter 2). All participants viewed four-picture displays in which each picture contained the same two objects but in different spatial relations to each other, that is, lateral (left/right), sagittal (front/behind) and topological (in/on)) relations. They described the target picture (left/right) highlighted by an arrow. Previous comparisons in Chapter 2 across hearing speaking and deaf signing populations have demonstrated that deaf signers, but not hearing speakers, experienced increased eye-gaze competition during message preparation from non-target pictures due to using iconic forms in sign.

The current results demonstrate that, during spoken message preparation, bimodal bilinguals, but not hearing speakers, experienced increased eye-gaze competition from non-target pictures. During signed message preparation, both bimodal bilinguals and deaf signers experienced increased eye-gaze competition. However, bimodal bilinguals' increase in eye-gaze competition was less steep than that of deaf signers. Overall, these findings reveal that cross-linguistic influence across modalities between sign and speech can change message conceptualization in bimodal bilinguals.

Introduction

Bimodal bilinguals are not two monolinguals in one person. Instead, they speak and sign differently than their hearing speaking and deaf signing peers, as shown in Chapter 3. Specifically, bimodal bilinguals' two languages can influence each other when describing spatial relations, that is, from speech to sign as well as from sign to speech, even if they use non-corresponding formats such as arbitrary and categorical forms in speech and iconic forms in sign. Analyses of bimodal bilinguals' speech, as well as co-speech gestures, show that this influence from sign increased semantic specificity in speech as well as the use of iconic co-speech gestures that convey detailed information about the physical features of objects. In reverse, influence from the spoken language decreased the use of semantically specific forms in sign. In the current chapter, we investigated whether cross-linguistic influence across modalities, as found in Chapter 3, can have cognitive consequences that go beyond language production and guide visual attention to spatial relations during spoken versus signed message preparation. To do so, we measured bimodal bilinguals' visual attention while they prepared to speak versus sign (for eye-tracking methods, see Chapter 2, page 65-69). These measurements were compared to that of hearing speakers and deaf signers. In this chapter, we therefore bring together the results of Chapter 2 on differences in visual attention between hearing speakers and deaf signers and on cross-linguistic influence in bimodal bilinguals' language production from Chapter 3.

Previous research has not yet asked whether cross-linguistic influence across modalities can occur (as investigated for the first time in Chapter 3) and whether this influence can also affect bimodal bilinguals' message conceptualization. However, in the field of spoken languages it has become evident that bilinguals have language-specific conceptualizations during non-linguistic spatial tasks based on the linguistic environment the bilinguals are in (e.g. Athanasopoulos et al., 2015; Kersten et al., 2010; Kousta, Vinson, & Vigliocco, 2008; Lai, Rodriguez, & Narasimhan, 2014). For example, Spanish differs from English in terms of which aspect of motion is expressed in the main verb. That is, Spanish expresses the path of motion in the main verb, while English expressed the manner of motion in the main verb. When Spanish-English bilinguals, as well as Spanish and English monolingual controls, had to judge similarities of motion events, bilinguals flexibly shifted their classification preferences based on the language in which they were tested (Lai et al., 2014). Specifically, when the language of instruction was Spanish, bilinguals and Spanish controls were more likely to base their similarity judgments on the path of motion compared to bilinguals that were instructed in English and English controls. This type of evidence demonstrates that bilinguals can switch between two conceptual systems, suggesting that the human conceptualization system is adaptive and flexible (Barsalou, 2009) as well as dependent on the language of operation.

However, instead of switching between two systems, bilinguals' conceptualization can also converge into a new system which differs from that of monolingual peers (Wolter et al., 2020). For example, Japanese and English express narrowness of rooms differently. While the English "narrow" refers to the width in relation to its height or length, the Japanese translation equivalent "semai" can be used similarly to the English "narrow" but can also be used to describe spaces that are small in general (like cramped or confined in English). When English-Japanese late bilinguals and their monolingual peers had to judge the narrowness of rooms, English controls judged narrowness based on the width and Japanese controls in terms of differences in overall area. English-Japanese bilinguals, however, judged rooms in an in-between pattern which differed when being instructed in English versus Japanese as well as differed in their choices from their monolingual peers.

Taken together, based on this sparse evidence it is until now unclear under which circumstances bilinguals switch between two conceptual systems or when the two systems converge into one. Furthermore, this body of evidence has exclusively focussed on spoken languages and not on bilinguals who use two languages from different modalities that have different formats of expressions. The current chapter will fill these two gaps by investigating message conceptualization during message preparation in highly proficient bimodal bilinguals who can both speak and sign.

The present study

Cross-linguistic influence is present across modalities between speech and sign in bimodal bilinguals, where the production of spatial language differs in terms of arbitrary and categorical versus iconic expressions. In the present study, we examined whether this cross-linguistic influence across modalities can influence bimodal bilinguals' message conceptualization. Although previous research has suggested that bilinguals process non-linguistic tasks differently than their monolingual peers, studies have not looked at the conceptualization of messages during message preparation in bilinguals. This chapter is the first to assess how production, which is shaped by cross-linguistic influence, can affect conceptualization during message preparation. To do so, we measured bimodal bilinguals' visual attention while they prepared to speak versus sign and compared those patterns to that of hearing speakers and deaf signers (for eye-tracking methods and design, see Chapter 2, page 65-69).

The chapter is based on prior findings of Chapter 2 and 3, that is, a) hearing speakers and deaf signers (i.e. controls) who do not have equal access to both languages show differences in visual attention during message preparation based on using arbitrary and categorical speech versus iconic sign, and b) there is cross-linguistic influence across modalities in bimodal bilinguals between arbitrary/categorical speech and iconic sign.

To do so, we will use the same eye-tracking paradigm as in Chapter 2 to assess whether bimodal bilinguals show similar eye-gaze patterns than hearing speakers when preparing to speak and compared to deaf signers when preparing to sign (i.e. compared to controls). In particular, Chapter 2 demonstrated that iconicity was relevant for message conceptualization and resulted in increased eye-gaze competition to non-target pictures for deaf signers but not for hearing speakers. Thus, mapping physical features of the figure object on to one hand and to place it in relation to the other hand resembling the ground object resulted in increased visual attention to non-target pictures for signers during message preparation. Moreover, within deaf signers, the planning of expressions that contain a high degree of semantic specificity (CLs) in NGT resulted in eye-gaze competition from different non-target pictures (i.e. those that are visually and semantically relevant, such as left/right/front/behind) than when preparing those that are less semantically specific (RLs, i.e. those that are only semantically relevant such as left/right). Based on the findings in Chapter 3 that bimodal bilinguals speak and sign differently than the hearing speaking and deaf signing controls, the question thus arises whether these differences in language production might also influence their visual attention differently compared to the controls, which will be assessed in this chapter.

Predictions

Visual attention during spoken message preparation. As shown in Chapter 3, bimodal bilinguals produced increased semantically specific information about physical features of objects in their speech than hearing speakers. Furthermore, they produced more iconic co-speech gestures and CLs that accompanied speech. This increase in semantic specificity in bimodal bilinguals' speech and the manual modality was due to cross-linguistic influence from the iconic and thus semantically specific expressions in sign.

If this cross-linguistic influence also changes message conceptualization, we expected to find differences in visual attention between hearing speakers and bimodal bilinguals. Chapter 2 found no increased eye-gaze competition while hearing speakers prepared to produce categorical expressions compared to deaf signers. For bimodal bilinguals, however, we predicted that they would experience increased eye-gaze competition from non-target pictures over the time course of Dutch message preparation. This might be due to the increased semantically specific information in their spoken expressions and manual modality accompanying speech. Thus, in order to identify and produce this semantically specific information in both modalities regarding the physical features of the objects, bimodal bilinguals might need to allocate increased visual attention to non-target pictures. This is because in non-target pictures the figure object is located and orientated differently in respect to the ground object (e.g. the pen is to the right and is orientated vertically versus the pen is to the front and is oriented horizontally). This effect of semantic specificity on visual attention would be in line with the findings in Chapter 2 in which the planning of expressions that contain a high degree of semantic specificity (CLs) in NGT also lead to more eye-gaze competition from all non-target pictures for deaf signers and thus showing influence from sign even when preparing to speak.

Visual attention during signed message preparation. As shown in Chapter 3, bimodal bilinguals produced more descriptions with RLs, which do not express semantically specific information about physical features of objects, and less with CLs than deaf signers. Thus, cross-linguistic influence from speech decreased semantically specific information in sign.

If cross-linguistic influence changes message conceptualization, we also expected to find differences in visual attention between deaf signers and bimodal bilinguals. Deaf signers experienced increased eye-gaze competition over time in Chapter 2. For bimodal bilinguals, we predicted that they would also experience increasing eye-gaze competition from non-target pictures over the time course of NGT message preparation. This is because of the visual form-to-meaning mappings in deaf signers' and bimodal bilinguals' spatial expressions. That is, both deaf signers and bimodal bilinguals need to disambiguate the objects' shape, size and orientation as well as the location of placing the hands in space. However, since bimodal bilinguals produced less semantically specific expressions than deaf signers due to cross-linguistic influence from speech, bimodal bilinguals' increase in eye-gaze competition might be less steep compared to their deaf signing peers.

If we find such parallelism between cross-linguistic influence and visual attention during message preparation, it would reveal that due to cross-linguistic influence bimodal bilinguals do not only speak and sign differently about spatial relations but also conceptualise these spatial relations differently compared to hearing speakers and deaf signers when preparing to speak versus sign. Furthermore, it would suggest that bimodal bilinguals do not switch between two conceptual systems that are similar to those of hearing speakers and deaf signers respectively, but instead, their languages might converge two systems into one which is shaped by bi-directional influence.

If we do not find differences in visual attention between bimodal bilinguals and their two respective control groups, this would indicate that despite cross-linguistic influence in bimodal bilinguals' language production, they might maintain their modality-specific conceptualization patterns. Consequently, bimodal bilinguals' patterns in visual attention would not differ from that of hearing speakers when preparing spoken messages and from that of deaf signers when preparing signed messages. In other words, cross-linguistic influence might not have cognitive consequences that can go beyond language production in bimodal bilinguals and they rather switch between two different conceptual systems.

Method

Participants

We used the same sample as recruited in Chapter 2 and 3. We will briefly summarise the most important information below but for details about hearing speakers and deaf signers we would like to refer the reader to Chapter 2 (page 64-65) and about bimodal bilinguals to Chapter 3 (page 104-107).

The current sample consisted of 19 hearing bimodal bilinguals of Dutch and NGT (11 female, M_{age} = 34.77, SD_{age} = 16.62) as well as two control groups. Namely, 20 Dutch hearing speakers (10 female, M_{age} = 33.25, SD_{age} = 10.95) and 20 deaf native NGT signers (16 females, M_{age} = 34, SD_{age} = 2.5). Originally, this sample consisted of two additional bimodal bilinguals and two signers, but they were excluded due to high eye-tracking loss (larger than 45%).

All deaf signers were born deaf and raised by deaf parents and had no access to auditory Dutch from birth. Thus, all deaf signers acquired NGT from early on (i.e. from their deaf signing parents) and also acquired Dutch in its written form when entering school (M_{age} = 3.5, SD_{age} = 2.8). NGT was the primary language of instruction. All hearing speakers acquired Dutch as their native language and learnt additional languages (mostly English, German, and French) later through instructional settings.

All bimodal bilinguals were born to at least one deaf parent, except one who grew up with a deaf sister. Consequently, all bimodal bilinguals had access to NGT naturally at home from early on and acquired Dutch from the surrounding community. Eight of the 21 bimodal bilinguals were trained sign language interpreters. As in Chapter 3, bimodal bilinguals participated twice, once in Dutch and once in NGT, three to five weeks apart.

Materials and procedure

We used the same stimuli set as in in the previous three chapters. Namely, 84 pre-tested four-picture displays. 28 out of these 84 were experimental displays which contained left/right target pictures. Half of these experimental displays contained more complex displays (N = 14) and the other half containing less complex visual displays (N = 14). More complex displays included a symmetrical competitor (left or right, depending on whether the target was left or right), a sagittal competitor (front or behind) and a topological distractor (in or on). Less complex displays, importantly, did not contain a symmetrical competitor but included instead a sagittal competitor (front or behind) and two topological distractors (i.e. both in and on). In this chapter, we only included more complex displays because in Chapter 2 less complex displays simply served as baseline and most of the key findings were observed in the more complex displays. Therefore, in this chapter, we omitted less complex displays for simplicity of the analysis.

The remaining 56 displays were fillers in which target pictures did not include left/right relations (i.e. front, behind, in, on; for details, see Chapter 2, page 65-67). Participants were asked to describe the target picture. The procedure was identical to that described in Chapter 2 (see page 67-69).

Data analysis

In order to compare bimodal bilinguals' visual attention to those of hearing speakers and deaf signers of Chapter 2, the eye-tracking data have been preprocessed and analysed identically to that chapter (page 69-75). However, unlike in Chapter 2 we diverged from the analysis in two ways. First, we did not assess visual attention *prior to* message preparation (corresponding to the "pre-arrow window" in Chapter 2) but only assessed visual attention *during* message preparation (corresponding to the "post-arrow window" in Chapter 2). Second, we did not include displays in which the symmetrical competitor was absent. Both of these measures served as baseline to verify the paradigm in Chapter 2. Hence, we did not include the pre-arrow window nor the less complex displays in this current chapter for the simplicity of the analysis.

Importantly to note is that during the first inspection of the data we found effects for bimodal bilingual's eye-gaze fixations to non-target pictures of participating twice in two similar testing sessions. While this was not the case in Chapter 3 for bimodal bilinguals' productions, it significantly influenced bimodal bilinguals' eye-gaze behaviour despite counterbalancing the order of sessions (β = 0.63, *SE* = 0.22, *t* = 2.90, *p*<0.01). Particularly, during the second session bimodal bilinguals were less attentive, resulting in only a few looks to non-target pictures (for more information about this analysis, see the supplementary materials). Consequently, we excluded data from the second sessions and included only those from the first sessions (Dutch in first sessions: N = 10, NGT in first sessions: N = 9). The hearing speakers and deaf signers were unaffected by this as they were tested only once in Dutch and the deaf signers only once in NGT. Consequently, this chapter includes fewer numbers of bimodal bilinguals in each language.

Following Chapter 2, for each trial, we recorded eye movements from pre-arrow onset (0 ms) until the four-picture display disappeared (3500 ms). We analysed fixation proportions (right eye only) across 50 ms continuous time bins. Since we were interested in examining eye-gaze tied to message preparation, our analyses focussed on a subset of the time course. Namely, we selected the post-arrow window initiating immediately after target indication (1500 ms) until linguistic production onset (3500 ms, see Chapter 2, Figure 8, page 69). This time window captures best participants' message preparation phase linked to relational encoding as shown in Chapter 2.

As in Chapter 2, we defined five different square-shaped AoIs, one for each picture and one for the arrow region. The size and positions of the five AoIs were equal across all trials. Fixation data were preprocessed in R (version 3.6.2) by determining for each participant whether a fixation fell into a particular AoI in each of 40 consecutive bins of 50 ms. Participants with more than 45% track loss across all trials were excluded from the analysis (N = 1). Additionally, we excluded trials in which track loss was higher than 50% (4.55%). Fixation proportions were corrected for 200 ms to plan a first saccade (Matin et al., 1993). We logarithmically transformed mean fixation proportions and time bins, which mathematically required shifting zeros and ones by +1.01.

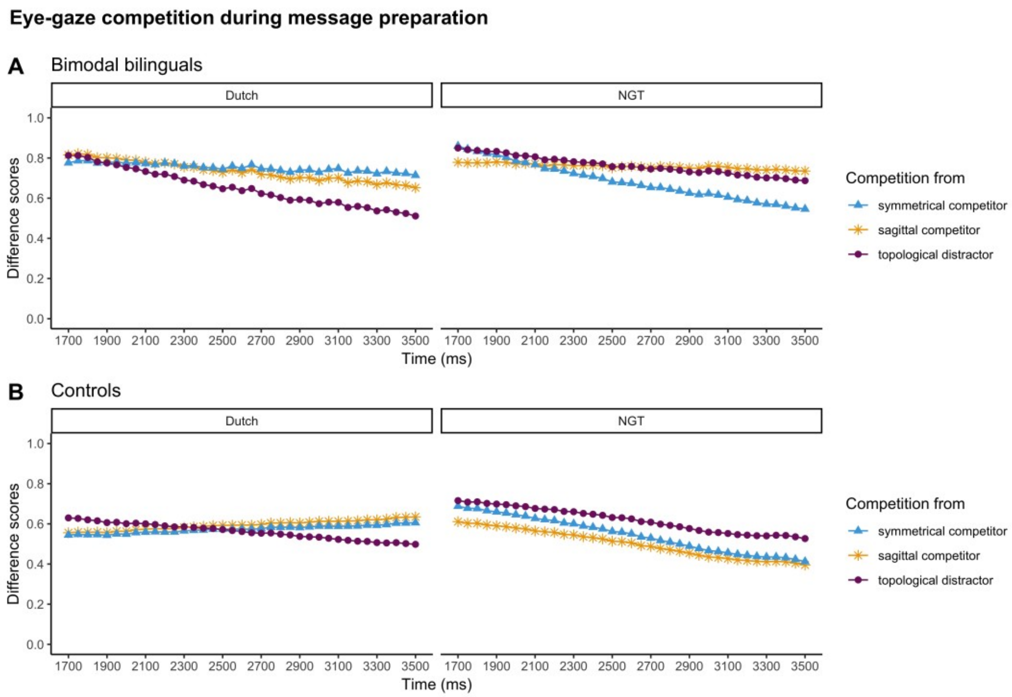
To analyse competition in visual attention to left/right spatial relations *during* message preparation (i.e. corresponding to the “post-arrow window” in Chapter 2) in the same way as in Chapter 2, we created three measures that assessed different types of eye-gaze competition: 1) mean of looks to target minus symmetrical competitor (i.e. left/right, *symmetrical competition*), 2) mean of looks to target minus sagittal competitor (i.e. front/behind, *sagittal competition*), 3) mean of looks to target minus topological distractor (i.e. in/on, *topological competition)*. Due to multiple comparisons of the two control groups, we conducted a Bonferroni correction on the p-values ($p < .025$). Fixation proportions were corrected in both time windows for 200 ms to plan a first saccade (Matin, Shao, & Boff, 1993).

Results

Figure 24 illustrates eye-gaze competition during message preparation in bimodal bilinguals (panel A) in Dutch (left) and NGT (right) compared to their hearing speaking (panel B, left) and deaf signing peers (panel B, right), respectively. Plots show difference scores averaged over picture descriptions (i.e. trials) across groups (i.e. bimodal bilinguals, deaf signers, and hearing speakers) in successive 50 ms time bins, initiating immediately after arrow offset (1500 ms plus 200 ms saccade correction) until linguistic production onset (3500 ms). Hence, smaller y-axis values indicate more eye-gaze competition.

Figure 24

Eye-gaze competition during message preparation in bimodal bilinguals.



Panel (A) shows eye-gaze competition in bimodal bilinguals in Dutch (left) and NGT (right). Panel (B) shows hearing Dutch speakers (left) and deaf native NGT signers (right) reported in Chapter 2. Smaller y-axis values indicate more eye-gaze competition. The x-axis represents the time course of message preparation initiating from arrow offset (1500 ms plus 200 ms saccade correction) until linguistic production onset (3500 ms).

To assess visual attention during message preparation we used a linear mixed-effects regression model (Baayen et al., 2008) with Group (hearing speakers/ deaf signers, numerically contrast coded as -1/2 vs. bimodal bilinguals, numerically contrast coded as +1/2), Competition type (categorical, *symmetrical competition* vs. *sagittal competition* vs. *topological competition*) and Bin (continuous, centred and scaled) as fixed effects. For Competition Type we compared first competition from the topological distractor versus competition from the other pictures by coding *topological competition* as - 2/3, *symmetrical competition* as +1/3, and *sagittal competition* as +

1/3. Second, we compared competition from the symmetrical competitor (coded as +1/2) to that of the sagittal competitor (coded as -1/2). The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group.

For visual attention during spoken message preparation, the model yielded a significant main effect of Group and Bin and a significant interaction between Group by Bin. The two-way interaction suggests that both groups experience competition equally from symmetrical competitors, sagittal competitors and topological distractors. However, bimodal bilinguals, but not speakers experience increasing eye-gaze competition over time during Dutch spoken message preparation (Figure 24, left panels of A and B). The relevant statistics and corresponding coefficients are reported in Table 7.

Table 7

Estimates, standard error, t- values, and p-values of the main and interaction effects of the model for Dutch message preparation. Note that we conducted a Bonferroni correction on the p-values (p < .025) due to multiple comparisons of the two control groups.

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.597	0.097	6.159	0.001
Competition Type[other vs. topological]	0.523	0.245	2.132	0.033
Competition Type[symmetrical vs. sagittal]	0.160	0.250	0.639	0.523
Group	0.545	0.193	2.819	0.005
Bin	-0.118	0.028	-4.244	0.001
Competition Type[other vs. topological]:Group	0.347	0.491	0.708	0.479
Competition Type[symmetrical vs. sagittal]:Group	-0.469	0.500	-0.939	0.348
Competition Type[other vs. topological]:Bin	-0.157	0.072	-2.178	0.029
Competition Type[symmetrical vs. sagittal]:Bin	-0.049	0.073	-0.671	0.502
Group:Bin	-0.156	0.055	-2.812	0.005
Competition Type[other vs. topological]:Group:Bin	-0.109	0.144	-0.759	0.448
Competition Type[symmetrical vs. sagittal]:Group:Bin	0.139	0.147	0.947	0.344

For visual attention during signed message preparation, the model yielded a significant main effect of Bin and significant interactions between Competition $\text{Type}_{\text{[other vs. topological]}}$

by Bin as well as Group by Bin. The first interaction between Competition Type$_{[\text{other vs. topological}]}$ by Bin suggests that both groups experience less increasing competition over time from the topological distractor compared to the other two competitors. The second interaction between Group by Bin indicates that both groups experience increasing competition equally from symmetrical competitors, sagittal competitors and topological distractors. However, deaf signers experienced increased eye-gaze competition over time than bimodal bilinguals during NGT message preparation (Figure 24, right panels of A and B). The relevant statistics and corresponding coefficients are reported in Table 8.

Table 8

Estimates, standard error, t- values, and p-values of the main and interaction effects of the model for NGT message preparation.

Fixed effects	**Estimate**	**Std. error**	***t* value**	***p* value**
(Intercept)	0.597	0.097	6.159	0.001
Competition Type[other vs. topological]	0.523	0.245	2.132	0.033
Competition Type[symmetrical vs. sagittal]	0.160	0.250	0.639	0.523
Group	0.545	0.193	2.819	0.005
Bin	-0.118	0.028	-4.244	0.001
Competition Type[other vs. topological]:Group	0.347	0.491	0.708	0.479
Competition Type[symmetrical vs. sagittal]:Group	-0.469	0.500	-0.939	0.348
Competition Type[other vs. topological]:Bin	-0.157	0.072	-2.178	0.029
Competition Type[symmetrical vs. sagittal]:Bin	-0.049	0.073	-0.671	0.502
Group:Bin	-0.156	0.055	-2.812	0.005
Competition Type[other vs. topological]:Group:Bin	-0.109	0.144	-0.759	0.448
Competition Type[symmetrical vs. sagittal]:Group:Bin	0.139	0.147	0.947	0.344

Discussion

Based on the results that bimodal bilinguals speak and sign differently compared to their hearing speaking and deaf signing peers in Chapter 3, the aim of this chapter was to assess whether bimodal bilinguals' visual attention also differs to that of hearing speakers and deaf signers when preparing to speak versus sign. Results revealed that message conceptualization, measured through visual attention, can be changed within individuals as a consequence of bi-directional cross-linguistic influence in bimodal bilinguals' language productions shown in Chapter 3.

In particular, during the preparation of spoken messages, bimodal bilinguals, but not speakers, experienced increasing eye-gaze competition from non-target pictures over the time course of message preparation. This reflects bimodal bilinguals' language productions that were shaped by cross-linguistic influence from NGT. That is, speech descriptions contained more semantically specific information and were frequently accompanied by iconic co-speech gestures or CLs. Thus, in order to produce such semantically specific expressions in their speech and manual modality, bimodal bilinguals needed to identify physical features of the objects, especially orientation and shape. Consequently, while preparing to describe the target picture, bimodal bilinguals experienced increased eye-gaze competition during spoken message preparation. This is in line with results from Chapter 2 showing that those iconic expressions that resemble physical features of objects in NGT led to increased eye-gaze competition to non-target pictures for deaf signers.

When preparing NGT descriptions, both bimodal bilinguals and deaf signers experienced increased eye-gaze competition over time from non-target pictures. This overall increase of eye-gaze competition in both signing groups reflects the influence of semantic specificity in the NGT expressions. More specifically, the act of depicting spatial relations and object properties onto the signing space in a one-to-one fashion requires more effort for signers and bimodal bilinguals during message preparation. This then results in an increase in visual attention to non-target spatial relations in order to describe the target picture. Crucially, bimodal bilinguals' increase in eye-gaze competition was less steep compared to their deaf signing peers. This reflects bimodal bilinguals' language productions from Chapter 3 as well, namely, bimodal bilinguals produced more descriptions with RLs, which are less semantically specific, than CLs compared to deaf signers.

Nevertheless, we want to mention a few limitations due to the few numbers of bimodal bilinguals in each language (Dutch: N = 10, NGT: N = 9). As mentioned earlier, there was an effect of the second session on bimodal bilinguals' visual attention, which allowed us to only include data of the first session. Consequently, eye-gaze data from Dutch message preparation comes from different bimodal bilinguals that from NGT message preparation (because half of the bimodal bilinguals were tested first and Dutch and later in NGT, while the other half were tested first in NGT and late in Dutch). Ideally, this study should have been carried out by using NGT and Dutch data of the same bimodal bilingual individuals. Thus, we cannot fully exclude the possibility that the observed effects are driven by other factors than cross-linguistic influences (e.g. individual differences, amount of language exposure). Nevertheless, bimodal bilinguals' language productions (reported in Chapter 3, page 114-119) did not depend on the session order. Thus, we would expect that with more participants and reliable eye-tracking data of the second session, we would find similar links between bimodal bilinguals' linguistic behaviour and their eye-gaze patterns.

Overall, these differences in visual attention are in contrast to Wolter et al. (2020) who found no bilingual effects in eye-gaze during the processing of a non-linguistic task. Rather, our results suggest that during message preparation bimodal bilinguals do not switch between two conceptual systems that are similar to those of hearing speakers and deaf signers respectively but instead seem to converge into a system, which is shaped by bi-directional influence and differs from those of the controls. Furthermore, the parallelism between language production and visual attention patterns reveal that cross-linguistic influence across modalities can change the conceptualization of messages in bimodal bilinguals when preparing to speak versus sign. This is in line with previous claims that language production and eye-gaze are tightly linked (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014). Even more, our results show that such links between language production and eye gaze do not only arise during language production but already during message preparation and also in bilingual language users.

Lastly, these results are in line with and strengthen our proposed contribution to the bimodal bilingual language production model by Emmorey et al. (2008) in Chapter 3 (page 123). In particular, we have claimed that the observed cross-linguistic influence occurs not simply at the linguistic level but might also have cognitive consequences. In particular, we argued that the NGT-influenced Message Generator used for iconic NGT expressions (CLs), influenced bimodal bilinguals' speech by producing more semantically specific expressions through the Spoken Language Formulator (bold lines in Figure 23, Chapter 3, page 123). Reversely, the Message Generator shaped by preparing messages for Dutch influenced the Signed Language Formulator by producing more RLs, which are less semantically specific. The current eye-tracking results confirm these hypotheses that such cross-linguistic influences are not only occurring at the linguistic level (i.e. route between the Formulators) but rather that those expressions that are commonly used in a language would influence attention to certain aspects of states and events (i.e. *thinking for speaking/signing*, Chapter 2; Slobin, 2003) at the level of the Message Generator (Kita & Özyürek, 2003). To conclude, cross-linguistic influence across modalities between sign and speech can have cognitive consequences that go beyond language production and can change message conceptualization in bimodal bilinguals.

Supplementary materials

Information about session order analysis

To assess possible effects of session order on bimodal bilinguals' visual attention to non-target pictures we used a linear mixed-effects regression model with difference scores as dependent variable and entered Session (first/ second, numerically contrast coded as -1/2 and +1/2) and Bin (continuous, centred and scaled) as fixed effects. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model revealed two main effects (Session and Bin) a significant interaction between Session and Bin ($\beta = 0.03$, $SE = 0.03$, $t = 6.124$, $p<0.001$).

Chapter 5

Word order preference in sign influences speech in bimodal bilinguals but not vice versa: Evidence from language production and visual attention

Chapter adapted from:

Manhardt, F., Brouwer, S., van Wijk, E., & Özyürek, A. (under review). Word order preference in sign influences speech in bimodal bilinguals but not vice versa: Evidence from language production and eye-gaze.

Abstract

The use of the visuo-spatial modality in sign languages drives a specific word order preference when it comes to describing spatial relations (e.g. the pen is to the right of the glass) by mentioning grounds (big objects) prior to figures (smaller objects). We investigated whether cross-linguistic influence of word order preferences occurs between speech and sign in bimodal bilinguals, proficient in a spoken and sign language. Furthermore, we assessed whether this influence has cognitive consequences that go beyond language production and also modulates message conceptualization. To do so, we measured visual attention during spoken versus signed message preparation. Dutch-NGT bimodal bilinguals (N = 21, reported in Chapter 3) and hearing speaking and deaf signing controls (N = 20, each, reported in Chapter 2) viewed and described spatial relations between two objects.

For language production, results indicated that deaf signers consistently mentioned grounds first and figures later, however, hearing speakers would use both types of word order. Bimodal bilinguals mentioned grounds first more than figures in speech compared to hearing speaking controls, showing influence from sign. However, when signing bimodal bilinguals mentioned grounds first equally often then deaf signing controls, showing no influence from speech.

For visual attention, deaf signers attended more to grounds than hearing speakers during message preparation. For bimodal bilinguals, during spoken message preparation they visually attended more to grounds than hearing speaking controls. During signed message preparation, bimodal bilinguals attended to grounds equally often than deaf signing controls. This one-way influence across modalities from sign to speech, but not vice-versa, provides new insights into mechanisms of cross-linguistic influences in bilinguals and reveals that this one-way influence can modulate message conceptualization during message preparation.

Introduction

Bilinguals activate the languages they know during language use, enabling cross-linguistic influence between these languages (Costa, 2005; Giezen & Emmorey, 2016; Kroll & Gollan, 2014; Loebell & Bock, 2003; Shook & Marian, 2012). It is less known, however, whether and how cross-linguistic influence between speech and sign also occurs in bimodal bilinguals, hearing individuals fluent in a spoken (vocal) and a sign (visuo-spatial) language. While there is first evidence for cross-linguistic influence between categorical and arbitrary expressions in speech and iconic expressions in sign (Chapter 3), here we focussed on the domain of word order in spatial expressions. In this domain, differences in modality are mostly visible with respect to the order in which objects are mentioned (e.g. Emmorey, 2002; Kimmelman, 2012; Perniss et al., 2015). That is, sign languages universally prefer to mention grounds (bigger objects) before figures (smaller objects) to describe spatial relations between two objects (e.g. *a glass with on the left a pen* instead of *a pen is to the right of the glass*). This universal preference seems to be motivated by perceptional biases (rather than linguistic biases) based on principles of Gestalt perception. Spoken languages, however, vary in the order in which objects are mentioned to describe spatial relations (Levinson, 2003; Levinson, 1996). In addition, we also explored whether cross-linguistic influence of word order can have cognitive consequences and also guides visual attention, as found in the domain of iconicity (Chapter 4). In particular, the question is whether language users allocate more visual attention to those aspects of a visual scene that are mentioned first within an utterance (e.g. Griffin & Bock, 2000). This would reveal that cross-linguistic influences can go beyond language production and can also shape message conceptualization.

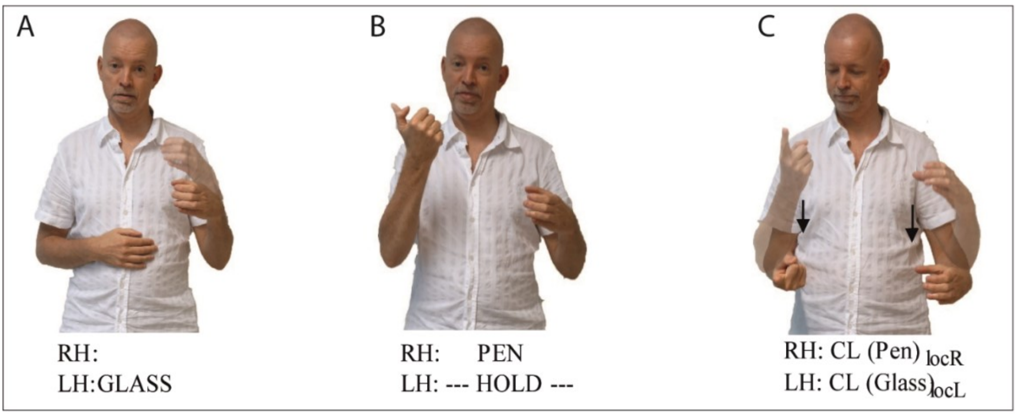
Modality-driven word order preference in sign languages

One modality-specific aspect of sign languages can be found in the domain of spatial language (e.g. the pen is to the right of the glass; e.g. Kimmelman, 2012; Perniss, Zwitserlood, & Özyürek, 2015) and relates to the order in which two objects are mentioned. Following imaged-based Gestalt principles, the two objects (i.e. glass and pen) differ perceptually. Particularly, one is visually perceived as smaller and more foregrounded (i.e. *figures,* e.g. the pen) and the other as bigger and more backgrounded (i.e. *grounds*, e.g. the glass*;* e.g. Rubin, 1958). In the field of linguistics, figures are assumed to be smaller and more movable entities and their location is characterized with respect to the ground (Talmy, 1978, 2003). Grounds are typically assumed to be the reference entity since they are bigger and more permanent compared to figures. These two objects can be mentioned in different orders in linguistic utterances.

While word order preferences differ across sign languages in other domains (e.g. mental verbs), when describing such spatial relations, sign languages universally prefer establishing first the lexical sign for the ground (Figure 25A) followed by introducing the lexical sign for the figure (Figure 25B; see among others Emmorey, 1996 for American Sign Language; Morgan, Herman, Barrière, & Woll, 2008 for British Sign Language; Sümer, 2015 for Turkish Sign Language; Perniss, 2007 for German Sign Language; and Kimmelman, 2012 for Russian Sign Language). Next to this universal preference, there are multiple pieces of evidence pointing out that this ground-first preference is more dominant in the manual modality compared to the spoken modality where word order preferences widely vary in spatial expressions (Levinson, 2003; Levinson, 1996). For one, during the most common linguistic strategy to express spatial relations in sign languages, known as *classifier constructions* (*CLs*), a similar ground-first preference is observed. More specifically, after introducing the objects using lexical signs, signers dominantly map object properties (e.g. size and shape) and relations between them onto the signing space by placing both hands in front of the body, resembling the actual spatial configuration (e.g. Emmorey, 1996, 2002; Perniss, Zwitserlood, & Özyürek, 2015; Figure 25C). Similarly to the order of mentioning lexical signs for grounds and figures, in CLs, the hand representing the ground is typically mapped onto signing space first followed by the hand representing the figure (e.g. Emmorey, 2002; Perniss, 2007; Perniss, Zwitserlood, & Özyürek, 2015; Sümer, 2015). Finally, when hearing speakers are asked to silently gesture about spatial relations they also show a clear preference of ground-first order (Goldin-Meadow et al., 2008; Laudanna & Volterra, 1991), even though they prefer another word order when describing similar pictures through speech. Overall, this provides evidence that ground-first is not simply a linguistically preferred word order but rather driven by the visuo-spatial modality, to describe spatial relations (e.g. Kimmelman, 2012; Perniss et al., 2015).

Figure 25

An example of describing "the pen is to the right of the glass" in NGT.



Panel (A) shows the introduction of the lexical sign for GLASS first, the ground (ground-first). Panel (B) shows the introduction of the lexical sign for PEN second, the figure. Panel (C) illustrates the simultaneous placement of the object properties onto the signing space respectively to the spatial relation between them.

This prominent ordering of the ground first also reflects that grounds are most conceptually salient and anchored in the visuo-spatial modality (Perniss et al., 2015) which is in line with Gestalt and linguistic conceptual theories that identify grounds as the bigger and more stable and permanent object (e.g. Rubin, 1958). Based on this, it is worthwhile to investigate whether this robust word order preference modulates cross-linguistic influence between spoken and sign language production as well as whether it increases preferences in visual attention to look at grounds over figures in bimodal bilinguals depending on in which language they prepare their messages.

Cross-linguistic influence of word order preferences in spoken languages

A wide range of studies has shown that word order preferences can cross-linguistically influence the two spoken languages in bilinguals but has mostly focussed on active/passive alternations or datives (e.g. Hatzidaki, Branigan, & Pickering, 2011; Kootstra, van Hell, & Dijkstra, 2012; Loebell & Bock, 2003; Pickering & Ferreira, 2008; Torres Cacoullos & Travis, 2016; Weber & Indefrey, 2009). A popular way to test this is using priming paradigms in which word order of one sentence is reflected in the word order of a second sentence that is otherwise unrelated to the first. In these studies, bilinguals are typically exposed to both of their languages within the same testing session. For example, when English-German bilinguals were asked to repeat English sentences in prepositional dative structure in English (e.g. *the lawyer sent the contract to his client),* they were more likely to describe unrelated pictures in German using the same prepositional dative structure (e.g. *die Großmutter nähte ein Kleid für ihre Enkeltochter [the grandmother sewed a dress for her granddaughter],* Loebell & Bock, 2003). When repeating English sentences in a double-object dative structure in English (e.g. *the lawyer sent his client the contract*) bilinguals were more likely to describe unrelated pictures in German using the same double-object dative structure (e.g. *die Großmutter nähte ihrer Enkeltochter ein Kleid [the grandmother sewed her granddaughter a dress]*). Overall, such effects of cross-linguistic influence have been evident between different languages and in both directions (i.e. from first to second language and vice versa; e.g. Schoonbaert, Hartsuiker, & Pickering, 2007), suggesting that this kind of priming is mostly unaffected by a specific word order or type of language.

However, most of this evidence stems from spoken languages and from bilinguals acquiring a second language later in their lives. Moreover, there seem to be limitations to finding effects of influence. For instance, it is typically only observed when structures were similar across bilinguals' two languages (Cleland & Pickering, 2003; Loebell & Bock, 2003; Salamoura & Williams, 2007). Furthermore, in absence of these typical priming paradigms, influence of word order preferences has not always been observed. For instance, in a task in which bilinguals are not directly primed, Korean-English bilinguals did not show any effect of influence (Ahn et al., 2019) but rather maintained the specific patterns of each language. Particularly, Korean differs from English in respect to the order in which objects are mentioned to describe spatial relations between objects (e.g. a cat that is above the piano). Korean prefers mentioning the location first ("[piano] [above] [cat]") while English prefers mentioning figures first (the cat above the piano). When describing spatial relations in each of their languages, bilinguals' descriptions did not differ from those of their monolingual peers. These differences shown across studies raise questions about whether word-order influence is specific to the languages tested and/or the experimental paradigms used. Even though it has been found that word order can be primed within American Sign Language (i.e. pre-nominal vs. post-nominal sentence structure) in deaf signers (Hall, Ferreira, & Mayberry, 2015), whether and how cross-linguistic influence of word order preference occur in bimodal bilinguals across speech and sign has not yet been investigated. Furthermore, it is unknown whether such influences - if found - have cognitive consequences beyond language production such as guiding preferences of visual attention to objects during message preparation.

The present study aims to assess cross-linguistic influence of word order in a special group of language users, namely, bimodal bilinguals. This population is special because they are hearing and highly proficient in both a spoken and a sign language from birth. Bimodal bilinguals are typically exposed naturally to sign language from early on as a home language, as they are hearing children born to deaf parents. The sign language (minority language) they acquire at home differs from the spoken language used most commonly in the community (i.e. majority language). Thus, they are *heritage* signers (De Quadros, 2018; Pichler et al., 2018; Quadros & Lillo-Martin, 2018) who have early exposure to a spoken and a sign language enabling us to explore cross-linguistic influences and possible effects on visual attention between languages of different modalities.

The link between language production and visual attention

Previous research has shown that during message preparation, cross-linguistic variation between languages guides speakers' visual attention to different components of these visual scenes in respect to which elements of a scene are encoded (i.e. *thinking for speaking,* Slobin, 2003; e.g. Bunger, Skordos, Trueswell, & Papafragou, 2016; Flecken, Carroll, Weimar, & Stutterheim, 2015; Flecken, Von Stutterheim, & Carroll, 2014; Goller, Lee, Ansorge, & Choi, 2017; Papafragou, Hulbert, & Trueswell, 2008; Trueswell & Papafragou, 2010). Furthermore, during actual language production, speakers are found to look at the referents they are describing in the order that they mention them (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer, Sleiderink, & Levelt, 1998; van de Velde, Meyer, & Konopka, 2014). For example, when English speakers had to describe picture scenes in which *the turtle is squirting the mouse with water* versus *the mouse is squirting the turtle with water,* they directed their gaze first to the turtle and then to the mouse when uttering the former, while they fixated first on the mouse and then on the turtle while uttering the latter (Griffin & Bock, 2000). Based on this evidence, researchers have argued for a tight link between the way speakers linguistically encode visual scenes and how they visually attend to such scenes (i.e. speech-gaze link).

Recently, this evidence has been extended to the visuo-spatial modality as well by providing evidence for a sign-gaze link motivated by the iconicity of spatial expressions (Chapter 2). In this work, another modality-specific aspect to describe spatial relations has been assessed, namely, iconic expressions in sign versus categorical and arbitrary expressions in speech (Figure 25C). This modality-specific difference has been found to guide deaf signers' visual attention differently to those spatial relations than that of hearing speakers during message preparation showing evidence for *thinking for signing.* Furthermore, previous work has shown that cross-linguistic influence between categorical and arbitrary expressions in speech and iconic expressions in sign in bimodal bilinguals also modulated visual attention during spoken and signed message preparation (Chapter 4). However, whether the influence of word order on eye-gaze, found for spoken languages, extends to sign productions has not been explored yet, let alone in bilinguals.

Taken together, previous studies have presented robust evidence for cross-linguistic influence of word order preferences in bilinguals as well as for links between language production and eye-gaze. The present study aimed to combine these two lines of research by investigating whether cross-linguistic influence of word order preferences can occur across modalities and whether such influences can go beyond language production and influence also message conceptualization during message preparation in bimodal bilinguals.

The present study

In the domain of spatial language, sign languages prefer ordering grounds first[6] and figures second to describe spatial relations between two objects. This ground-first preference appears to be driven by the nature of the visuo-spatial modality (e.g. Kimmelman, 2012; Perniss et al., 2015). Spoken languages, however, vary in word order preferences to describe spatial relations (Levinson, 2003; Levinson, 1996) and hence lack a modality-driven word order related to spatial expressions. In the present study, we investigated whether this word order preference, driven by the visuo-spatial modality, can influence the word order preference in a spoken language and/or in turn be influenced by the spoken language preference in a special group of language users. That is, bimodal bilinguals who are hearing and acquire both a spoken language (Dutch) and a sign language (NGT) from birth. In particular, we assessed cross-linguistic influence at the production level, specifically between spoken and signed descriptions, and also at the level of visual attention, specifically whether there are eye-gaze preferences to look more at grounds versus figures depending on the language in which messages are prepared, reflecting cross-linguistic influence in language production.

To do so, we used a visual world language production eye-tracking paradigm (Chapter 2) in which we presented four-picture displays of which each picture contained two objects that were arranged in different spatial configurations (i.e. left, right, front, behind, in and on). After briefly introducing the four pictures, we indicated the target picture by presenting an arrow in the middle of the screen that pointed to one of the pictures. We recorded eye-gaze once the arrow disappeared until participants had to describe the target picture to a confederate. This paradigm allowed us to assess linguistic word order preferences during a naturalistic but controlled picture description task (i.e. refraining from using priming paradigms), and at the same time,

[6] In order to facilitate comparisons between word order in sign and speech we established ground-first order based on the order of mentioning lexical signs not CLs.

preferences in allocating visual attention to grounds versus figures in the target picture during message preparation.

Overall, we first assessed language-typical word order preferences in Dutch and NGT by comparing hearing Dutch speakers' and deaf NGT signers' picture descriptions in respect to which word order they had produced. We then assessed cross-linguistic influence by comparing word order preferences of bimodal bilinguals in both of their languages to that of their control groups, respectively. Concerning visual attention, we investigated language-typical preferences in Dutch and NGT to look at grounds versus figures by comparing hearing Dutch speakers' and deaf NGT signers' eye-gazes during message preparation. Finally, we examined cross-linguistic influence of visual attention, by assessing bimodal bilinguals' and controls' eye-gaze preferences during message preparation to look more at grounds or figures related to which object is mentioned first.

Predictions

Language production. As for our control groups, we predicted that deaf NGT signers produce more ground-first descriptions than hearing Dutch speakers as sign languages typically prefer ground-first order and often allow less variability, while in Dutch, both figure-first and ground-first are valid and acceptable word orders to describe spatial relations (Hartsuiker, Kolk, & Huiskamp, 1999). However, frequency counts on preferences in NGT or Dutch are unavailable.

For bimodal bilinguals, if there is cross-linguistic influence from the robust modality-driven word order of NGT to Dutch, we predicted that, while describing spatial relations in Dutch bimodal bilinguals prefer ground-first over figure-first order and will thus produce less figure-first descriptions compared to their hearing speaking peers. However, since NGT is a minority language this effect might not be there due to sociolinguistic factors such as language status, prestige, and group identity (e.g.

Michael, 2014). Concerning influence in the opposite direction, Dutch as majority language might influence word order preferences in NGT, the minority language (as evident from spoken languages in various domains, e.g. Backus, 2005; Muysken, 2000; Polinsky, 2008). However, if ground-first is the modality-driven word order in sign and is grounded in cognitive perceptual biases, then this word order might be invariant and more resistant to change than the flexible word order in Dutch speech. Consequently, speech might not influence sign. Thus, we should find influence in only one direction, namely, from sign to speech due to the robust modality-driven word order in sign and variability in speech, but not reversely, from speech to sign due to variability in speech and no variability and robust modality-driven word order in sign.

Finally, instead of experiencing cross-linguistic influence across modalities between sign and speech, bimodal bilinguals might maintain their language-specific patterns as previously observed in such highly proficient bilinguals (e.g. Ahn et al., 2019; Azar et al., 2019). That is, when speaking Dutch, bimodal bilinguals are like hearing speakers, and when signing in NGT, they are like deaf signers in respect to their word order preferences. This would also indicate that outside of priming paradigms these one- or bi-directional influence effects might not occur (as observed in e.g. Ahn et al., 2019).

Visual attention. Generally, we expected eye-gaze effects to arise from early on when message preparation is unfolding, assuming that eye-gaze preferences are related to the order in which ground and figures are mentioned. Moreover, our predictions are based on the assumption that the aspect that is mentioned first in a sentence is most salient and foregrounded in the language users' mind (Gundel, 1985; Macwhinney, 1977). Thus, mentioning grounds first might lead to visually attending more to grounds, while mentioning figures first might lead to visually attending more to figures during message preparation.

As for our control groups, we predicted that deaf NGT signers prefer looking at grounds more than hearing Dutch speakers over the time course of message preparation. This might reflect that deaf-signers' modality-driven preference to produce predominantly ground-first descriptions also guides more attention to grounds compared to hearing speakers. This would indicate, that *thinking for speaking* extends to *thinking for signing* (Chapter 2) in the domain of word order.

For bimodal bilinguals, we predicted that cross-linguistic influence can go beyond language production and also influence message conceptualization based on the findings of Chapter 4 for iconicity. Thus, following the predictions on language production mentioned above, we should find influence on visual attention in only one direction, namely, from sign to speech, but not reversely from speech to sign. In particular, we predicted that if cross-linguistic influence from sign to speech – if found- changes bimodal bilinguals' message conceptualization during spoken message preparation, bimodal bilinguals would not only mention grounds first more often but would also allocate more attention over time to grounds than to figures compared to hearing speakers. In the reverse direction, if there is no cross-linguistic influence from sign to speech, then bimodal bilinguals' would not differ from the deaf signing controls and allocate more visual attention to grounds than figures over time during signed message preparation. Thus overall, we expected that effects of language production on message conceptualization in bimodal bilinguals can be found only when there is cross-linguistic influence, thus from sign to speech but not from speech to sign.

If we find such parallelism between language production and visual attention during message preparation, it would indicate that such eye-gaze-language production links do not only arise during language production (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014) but already during message preparation. Finally, if we do not find eye-gaze effects during message preparation, this might indicate that modality-specific influence – if found – has no cognitive consequences that go beyond the level of language production.

Method

Participants

The participants were the same as reported in Chapter 2-4. This sample consisted of 21 bimodal bilinguals of Dutch and NGT (11 female, M_{age} = 34.77, SD_{age} = 16.62) as well as two control groups consisting of 20 Dutch hearing speakers (10 female, M_{age} = 33.25, SD_{age} = 10.95) and 19 deaf native NGT signers (16 females, M_{age} = 34, SD_{age} = 2.5). Originally, this sample consisted of two additional bimodal bilinguals and two signers, but they were excluded due to high eye-tracking loss (larger than 45%). More details about the language background of the controls can be found in Chapter 2 (page 64-65) and of bimodal bilinguals in Chapter 3 (page 104-107).

Materials and procedure

We used the same stimuli set as in in the previous chapters, consisting of 84 four pre-tested picture displays containing the same two objects but in different spatial configurations to each other (i.e. left, right, front, behind, in and on; see Figure 17 in Chapter 3 on page 109). An arrow pointing at one of the pictures indicated the target picture participants had to describe. 28[7] experimental displays consisted of left/right target relations, while the remaining three pictures included other spatial relations (i.e. front, behind, in and on). We included 56 filler displays in which targets were any other spatial relation (i.e. front, behind, in and on) to avoid emphasis on left/right relations during the whole experiment. We focussed on left/right relations because these allowed us to assess eye-gaze preferences to grounds and figures without overlapping locations or occlusions of the two objects (as this the case with in/on/behind relations). The distance between the ground and figure object was always kept equal across

[7] In this chapter we did not distinguish between less and more complex displays, because we expected word order preference to be independent of the type of displays. We therefore collapsed both less and more complex displays.

displays for the spatial relations respectively (i.e. for left/right, front/behind etc.). Ground objects were always located in the centre of the pictures and figures were always placed to the left, right, front, behind, inside, or on top of the ground (Figure 17 in Chapter 3 on page 109). Thus grounds were distinguished from figures based on their size and mobility. Namely, grounds as bigger and permanent objects and figures as smaller and more mobile objects (Talmy, 2003, 1978). The procedure was identical to that described in Chapter 2 (see page 67-69).

Data analysis

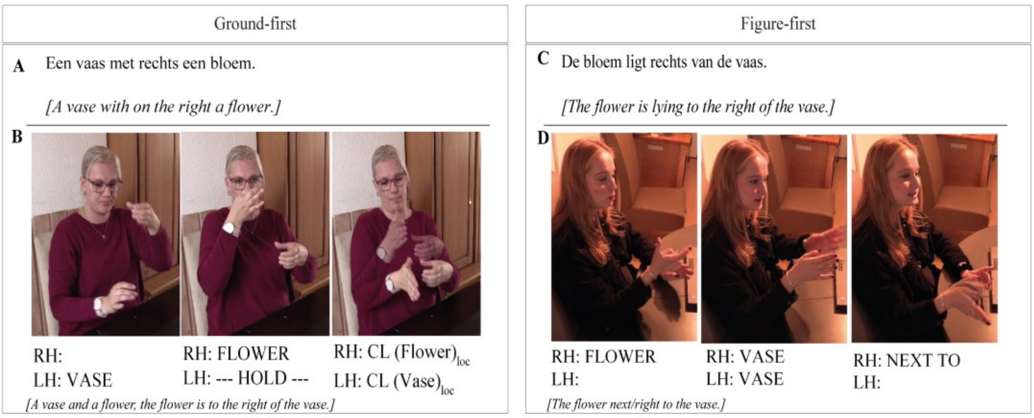
In this section, we will first describe how we analysed language production in respect to the preference of word order in Dutch across hearing speakers and bimodal bilinguals as well as in NGT across deaf signers and bimodal bilinguals. Furthermore, we describe the analysis of eye-gaze preferences to look at grounds and figures during message preparation.

Language production. We coded all picture descriptions using ELAN, a free annotation tool (http://tla.mpi.nl/tools/tla-tools/elan/) for multimedia resources developed by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands (Wittenburg et al., 2006). Trained, hearing native Dutch and deaf native NGT annotators performed annotation and coding of the data respectively. All coding was checked by an additional coder to find consensus. If no consensus could be reached, the trial was excluded from further analyses (5.81% of all descriptions).

For both Dutch and NGT, we coded for each picture description which object was mentioned first, namely the ground or the figure. This distinction was based on the arrangement of our stimuli, that is, grounds as the bigger object were placed in the centre of the pictures while figures as the smaller objects were surrounding grounds. In Dutch, ground-first descriptions typically involved prepositional constructions using *met* (“with”; Figure 26A), while figure-first descriptions usually included verb constructions using *liggen/staan* ("lying/standing"; Figure 26C). In NGT, ground-first descriptions typically involved CLs in which object properties and relations between them are mapped in a one-to-one relation on to the signing space (Figure 26B), while figure-first descriptions (albeit were very few) included lexical signs for the spatial relation (*relation lexemes;* Figure 26D). For both languages, descriptions in which only one object was mentioned (i.e. only figure, only ground) were omitted from further analyses (3.69% of all descriptions).

Figure 26

An example of mentioning the ground first and figure first in Dutch and NGT.



Panel (A) shows an example for mentioning the ground first "a vase with on the right a flower" in Dutch and panel (B) in NGT. Panel (C) shows an example for mentioning the figure first "the flower is to the right of the vase" in Dutch and panel (D) in NGT.

The coded data were analysed in R (version 3.6.2) (R Core Team, 2013). We used general and linear mixed-effects regression models (Baayen et al., 2008) using the packages *lme4* (version 1.1-19) (Bates et al., 2015) and *lmerTest* (version 3.0-1) (Kuznetsova et al., 2017) to retrieve p-values. We used binominal data as dependent variable and an optimizer to facilitate model convergence.

We conducted two types of analyses: (1) we assessed whether word order preferences differ in Dutch and NGT by comparing descriptions between hearing speakers and deaf signers, and (2) we assessed whether bimodal bilinguals differ in their word order preferences in Dutch compared to hearing speakers and in NGT compared to deaf signing controls.

Visual attention. Unlike Chapter 2 and 4, this chapter assess target fixations only. Consequently, the analysis reported in this chapter differs from that conducted in the previous eye-tracking chapters (Chapter 2 and 4)[8].

For each trial, eye movements were recorded from pre-arrow onset (0 ms) until the four-picture display disappeared (3500 ms). We analysed fixation proportions (right eye only) across 50 ms continuous time bins. Since we were interested in examining the differences in eye movements linked to message preparation, our analyses focussed on a subset of the time course. Namely, we selected a 2000 ms post-arrow window[9] initiating immediately after target indication (1500 ms) until linguistic production onset (3500 ms, Figure 8, Chapter 2, page 69). This time window captures participants' message preparation phase linked to relational encoding (shown in Chapter 2). This enabled us to assess whether there is cross-linguistic influence of visual attention in bimodal bilinguals to grounds and figures respectively to the order in which they are being mentioned.

We defined two different rectangle-shaped Areas of Interests (AoIs) for target pictures only: one for the ground object and one for the figure object. Eye-gaze to the remaining three pictures in visual displays was removed as they were not being described (27.26% of all fixations). The two AoIs did not overlap and differed slightly in size. Namely,

[8] Note that the effect of the second session on bimodal bilingual's visual attention to non-target pictures (reported in Chapter 4) did not affect the target fixations that we assessed in this chapter. Thus, here we were able to include eye-tracking data from 19 bimodal bilinguals.

[9] We focused only on the post-arrow window and did not include the pre-arrow window (as done in Chapter 2), because in the pre-arrow window the indication of the target picture did not take place yet and thus does not allow us to assess fixations to the target picture.

ground AoIs were larger capturing the ground object in the centre of the picture while figure AoIs were smaller capturing the figure object to the left or right side of the ground object. Fixation data were preprocessed in R. For each participant, we determined whether a fixation fell into one of the two AoIs in each of 40 consecutive bins of 50 ms. Participants with more than 45% track loss across all trials were excluded from the analysis (N = 4, of which 2 were deaf NGT signers and 2 were bimodal bilinguals, as mentioned above in the participant section). Additionally, we excluded trials in which track loss was higher than 50% (3.6%).

We conducted two types of analyses using general linear mixed-effects regression models: (1) we assessed whether preferences to fixate grounds versus figures differ in Dutch and NGT by comparing eye-gaze between hearing speakers and deaf signers during message preparation, and (2) we assessed whether bimodal bilinguals differ in their preferences to fixate grounds versus figures when preparing messages in Dutch compared to hearing speaking controls and in NGT compared to deaf signing controls. Our dependent variable was binomial (i.e. fixations to grounds vs. figures). Due to multiple comparisons of the two control groups, we conducted a Bonferroni correction on the p-values ($p < .025$). Fixation proportions were corrected in both time windows for 200 ms to plan a first saccade (Matin, Shao, & Boff, 1993).

Results

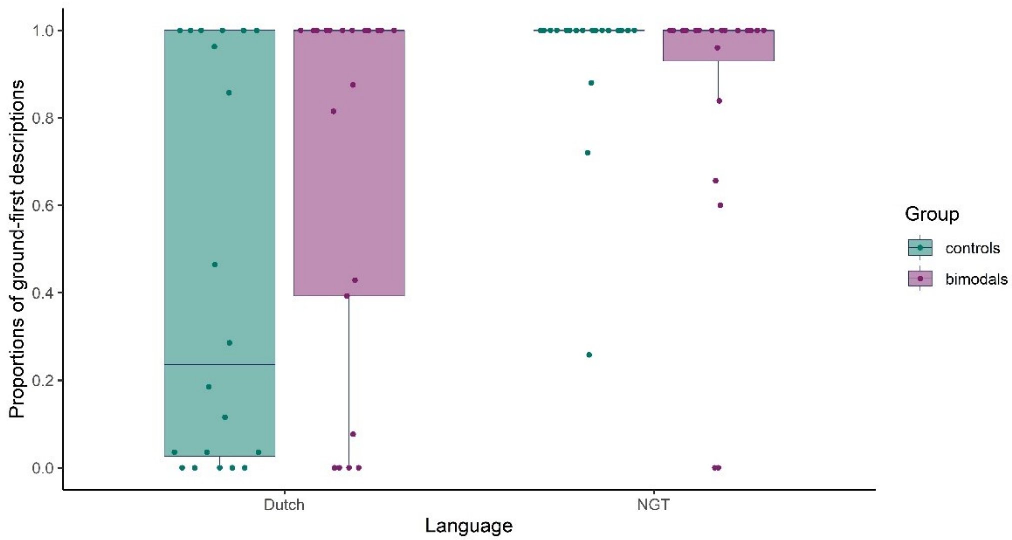
In this section, we will first report the linguistic description data to assess word order preference in Dutch and NGT from bimodal bilinguals as well as their hearing non-signing and deaf signing peers. After this, we will report the eye-gaze data from these three groups to assess whether possible cross-linguistic influence of visual attention modulates bimodal bilinguals' preference to look at grounds versus figures depending on the order in which they planned to mention them.

Language production

Figure 27 shows proportions of ground-first descriptions in Dutch and NGT between bimodal bilinguals and controls respectively. For plotting, data were averaged over picture descriptions (i.e. trials) and group (i.e. bimodal bilinguals, deaf signers, hearing speakers). To report means, data were averaged over picture descriptions and participants.

Figure 27

Ground-first descriptions in Dutch and NGT.



The left panel shows proportions of ground-first descriptions in Dutch across hearing speakers (left) and bimodal bilinguals (right). The right panel shows proportions of ground-first descriptions in NGT descriptions across deaf signers (left) and bimodal bilinguals (right). Dots in the boxplot represent each data point (participant).

Word order preferences in Dutch and NGT (controls). We assessed first whether word order preferences differed between the control groups by comparing Dutch hearing speakers' and NGT deaf signers' picture descriptions. In particular, we investigated whether the ground was mentioned first (1) or not (0) using a general linear mixed-effects regression model with Group (hearing speakers, numerically contrast coded as

-1/2 vs. deaf signers, numerically contrast coded as +1/2) as fixed effect. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model yielded a significant main effect of Group (*β* = 9.57, *SE* = 2.45, *z* = 3.91, *p* < 0.001), suggesting that deaf signers produced more ground-first descriptions in NGT (*M* = 0.92, *SD* = 0.26) than hearing speakers in Dutch (*M* = 0.56, *SD* = 0.50; see Figure 27).

Cross-linguistic influence of word order preferences in bimodal bilinguals. Next, we compared bimodal bilinguals' descriptions in each language to that of their hearing speaking and deaf signing peers. For Dutch, we investigated whether the ground was mentioned first (1) or not (0) using a general linear mixed-effects regression model with Group (hearing speakers, numerically contrast coded as -1/2 vs. bimodal bilinguals, numerically contrast coded as +1/2) as fixed effect. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model yielded a significant main effect of Group (*β* = 8.15, *SE* = 3.53, *z* = 2.31, *p* = 0.02), suggesting that bimodal bilinguals produced more ground-first descriptions (*M* = 0.67, *SD* = 0.43) than their hearing speaking peers (*M* = 0.43, *SD* = 0.50; see Figure 27, left panel). No effect of Session Order on bimodal bilinguals' ground-first preference was found (see the supplementary materials for more information), ruling out that this preference is due to priming of describing similar pictures in two testing sessions.

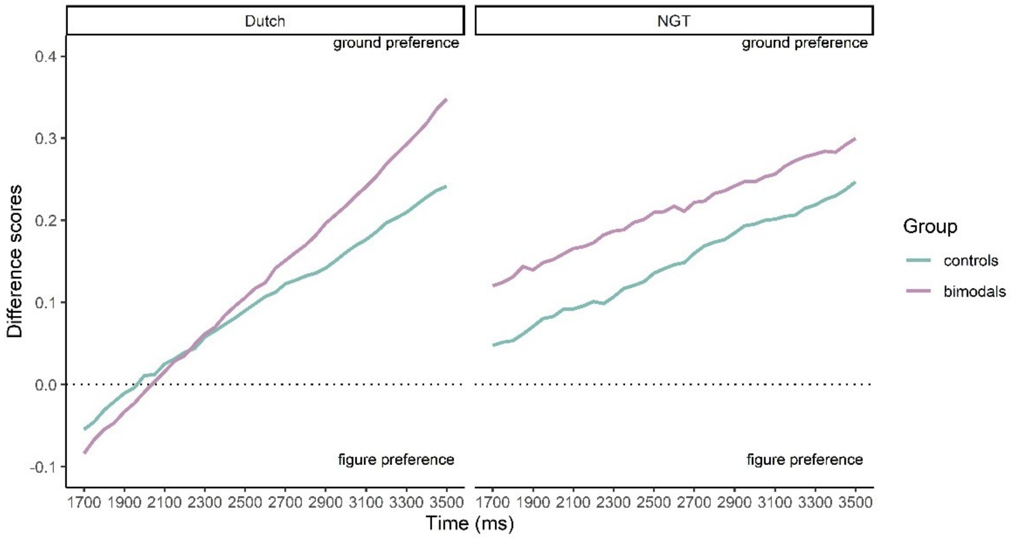
For NGT, we investigated whether the ground was mentioned first (1) or not (0) using a general linear mixed-effects regression model with Group (deaf signers, numerically contrast coded as -1/2 vs. bimodal bilinguals, numerically contrast coded as +1/2) as fixed effect. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model yielded no significant main effect of Group (*β* = -0.59, *SE* = 2.74, *z* = -0.22, *p* = .83), revealing that bimodal bilinguals did not differ from deaf signers in how often they produced ground-first descriptions when signing in NGT (bimodal bilinguals: $M = 0.85$, $SD = 0.35$; deaf signers: $M = 0.92$, $SD = 0.27$; see Figure 27, right panel). Again, no effect of Session Order on bimodal bilinguals' ground-first preference was found (see the supplementary materials for more information).

Visual attention

Figure 28 illustrates preferences to fixate grounds versus figures during message preparation in bimodal bilinguals in Dutch (left panel) and NGT (right panel) compared to their hearing speaking and deaf signing peers, respectively. For plotting only, we calculated difference scores between fixations to the ground AoI minus fixations to the figure AoI to illustrate a preference for looking at one object over the other (i.e. values above 0 indicate a ground preference and values below 0 indicate a figure preference). Plots show these difference scores averaged over picture descriptions (i.e. trials) across groups (i.e. hearing speakers, deaf signers, bimodal bilinguals) in successive 50 ms time bins initiating immediately after target indication (1500 ms plus 200 ms saccade correction) until language production onset (3500 ms).

Figure 28

Preference of fixating the ground vs. figure object during message preparation in Dutch and NGT.



The left panel shows preferences in Dutch across hearing speakers and bimodal bilinguals. The right panel shows preferences in NGT, across deaf signers and bimodal bilinguals. Y-axis values above 0 (dotted line) indicate a preference to look at the ground object. Values below 0 indicate a preference to look at the figure object. X-axis displays the time course of message preparation after target indication (1500 ms plus 200 ms saccade correction) until language production onset (3500 ms).

Eye-gaze preferences in Dutch and NGT (controls). We first examined whether eye-gaze to grounds versus figures differed in hearing speakers and deaf signers during message preparation. In particular, we investigated fixations to grounds (1) or figures (0) using a general linear mixed-effects regression model with Group (hearing speakers, numerically contrast coded as -1/2 vs. deaf signers, numerically contrast coded as +1/2) and Bin (continuous, centred and scaled) as fixed effects. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model yielded no significant main effect of Group (β = 0.11, *SE* = 0.14, *z* = 0.78, *p* = .44), but a significant main effect of Bin (β = 0.16, *SE* = 0.01, *z* = 12.74, *p*<.001), and a significant interaction between Group by Bin (β = -0.06, *SE* = 0.02, *z* = -2.69, *p*<.01). This interaction suggests that during message preparation deaf signers preferred looking at grounds from the start, while for hearing speakers a ground preference in eye-gaze emerged only later and instead started with a preference to fixate figures (Figure 28).

Cross-linguistic influence effects on visual attention in bimodal bilinguals. Next, we assessed how bimodal bilinguals' eye-gaze patterns might differ from that of hearing speakers when planning Dutch descriptions and from that of deaf signers when planning NGT descriptions. In particular, for Dutch, we investigated fixations to grounds (1) or figures (0) using a general linear mixed-effects regression model with Group (speakers, numerically contrast coded as -1/2 vs. bimodal bilinguals, numerically contrast coded as +1/2) and Bin (continues, centred and scaled) as fixed effects. The most parsimonious model included random intercepts for participants and items and a by-items random slope for Group. The model yielded no significant main effect of Group (β = 0.23, *SE* = 0.25, z = 0.93, p = .35), but a significant main effect of Bin (β = 0.24, *SE* = 0.01, z = 19.51, p<.001), and a significant interaction between Group by Bin (β = 0.10, *SE* = 0.02, z = 4.10, p<.001). This interaction suggests that during message preparation in Dutch, bimodal bilinguals and hearing speakers preferred looking at figures over grounds at the beginning of message preparation. However, when message preparation is unfolding, both groups increasingly preferred fixating grounds over figures. Crucially, bimodal bilinguals' preference to look at grounds over figures increased more steeply over time compared to their hearing speaking peers (Figure 28, left panel).

Finally, in NGT, we investigated fixations to grounds (1) or figures (0) using a general linear mixed-effects regression model with Group (deaf signers, numerically contrast coded as -1/2 vs. bimodal bilinguals, numerically contrast coded as +1/2) and Bin (continuous, centred and scaled) as fixed effects. The most parsimonious model included random intercepts for participants and items and a by-items random slope for

Group. The model yielded no significant main effect of Group (β = 0.18, *SE* = 0.16, *z* = 1.12, *p* = .26), but a significant main effect of Bin (β = 0.13, *SE* = 0.01, *z* = 9.88, *p* < .001), and no significant interaction between Group by Bin (β = 0.003, *SE* = 0.03, *z* = 0.14, *p* = .89). This lacking interaction between Group and Bin suggests that during message preparation in NGT, both bimodal bilinguals and deaf signers preferred looking more at grounds than figures. This ground preference in eye-gaze increased similarly over time for both groups and occurred right from the start of message preparation (Figure 28, right panel).

Discussion

The present study investigated how different word order preferences in a sign and spoken language influence each other in bimodal bilinguals in a domain where sign languages have a modality-driven word order. In particular, to describe spatial relations between entities, sign languages predominantly prefer ground-first order, while word order preferences differ across spoken languages. We assessed, whether there is a modality-specific influence of word order preferences between NGT and Dutch in bimodal bilinguals and whether this has further cognitive consequences and influences visual attention.

Language production

For NGT signers, results revealed that they produced more ground-first descriptions than hearing Dutch speakers. This confirms that NGT predominantly prefers ground-first order as found for many other sign languages (e.g. Emmorey, 2002; Kimmelman, 2012; Morgan et al., 2008; Perniss, 2007; Sümer, 2015) and indicates that ground-first is driven by the manual modality. In particular, both signing groups (i.e. deaf signers and bimodal bilinguals) show a very strong and robust systematicity in mentioning grounds first in NGT. Therefore, our results strengthen previous views that ground-first order might be based on modality-specific properties (e.g. Kimmelman, 2012; Perniss et al., 2015).

For Dutch speakers, results showed that in Dutch there is no clear preference for figure-first or ground-first order but rather, half of the speakers produced mostly figure-first descriptions while the other half preferred producing ground-first descriptions. This indicates, that there is no pre-set linguistic word order in Dutch for describing spatial relations but rather both figure-first and ground-first are valid and acceptable word orders (Hartsuiker, Kolk, & Huiskamp, 1999).

Influence from sign to speech in bimodal bilinguals. In Dutch, bimodal bilinguals produced more ground-first descriptions than hearing speaking controls, suggesting an influence of word order preferences across modalities from NGT sign to Dutch speech. This is in line with previous research assessing priming within a sign language (i.e. not at cross-linguistic influences) showing that word order can be primed within American Sign Language (i.e. pre-nominal vs. post-nominal sentence structure) in deaf signers (Hall et al., 2015). Our results extend these findings to cross-linguistic influence of word order from sign to speech even in absence of a priming paradigm and to spatial language. This is also in line with previous assumptions that word order preferences within a language might depend on other factors such as context, communicative pressure or language contact (Schouwstra & de Swart, 2014), namely, that word order preference can be influenced by language contact from another language (NGT) within a (bimodal) bilingual.

Interestingly, in similar descriptions of spatial relations, an influence of ground-first order was not observed in spoken English (figure-first)-Korean (ground-first) bilinguals (Ahn et al., 2019), suggesting that our observed influence is not related to conceptual groundedness of ground objects but rather to a modality-driven word order preference. It is possible that what is being influenced is a cognitive bias because the visuo-spatial modality adapts more easily to the visual perceptional principles of spatial scenes.

Our results go beyond previous findings on cross-linguistic influence of word order preferences in many ways. For one, we show here that effects of cross-linguistic influence emerged despite using a naturalistic picture description setting without experimentally forcing the mixing of bilinguals' languages as previously done in priming paradigms (e.g. Hatzidaki et al., 2011; Kootstra et al., 2012; Torres Cacoullos & Travis, 2016). Thus, in the present study, although only one language was relevant during the whole duration of the task, we still found cross-linguistic influence, while others have failed to show effects of word order influence in absence of priming paradigms (e.g. Ahn, Gollan, & Ferreira, 2019). Furthermore, our influence did not relate to the order in which the language session took place (i.e. first or second).

Nevertheless, word order preference in our bimodal bilingual sample seems to vary. That is, in Dutch, not all bimodal bilinguals showed a clear ground-first preference but a minority produced predominantly figure-first descriptions. This is in line with claims that cross-linguistic influences are intertwined and dynamic (Grosjean, 1989), resulting in weaker influences in some bilingual individuals and stronger influences in others.

Influence from speech to sign in bimodal bilinguals. Concerning the reverse direction from speech to sign, bimodal bilinguals did not differ from the deaf signing controls in their ground-first preference, indicating no influence from Dutch to NGT. These combined results provide first evidence for cross-linguistic influence of word order preference across modalities from sign to speech, but reversely, there was no influence from speech to sign, even though the spoken language was the majority language.

Taking both speech and sign results together, this study revealed that cross-linguistic influence of word order preference in bimodal bilinguals is one-directional, that is, it occurs from sign to speech but not vice versa. This one-way influence occurred independently of the language status, which contrasts with previous findings in proficient heritage bilinguals of two spoken languages, where cross-linguistic influence was typically evident from the majority to the minority language (e.g. Backus, 2005; Muysken, 2000; Polinsky, 2008) or where no cross-linguistic influences were found (e.g. Azar, 2020; Azar et al., 2019). This suggests that not only language status but also modality can be driving factor for cross-linguistic influence (for a more elaborate discussion on this, see Chapter 6).

One possibility for the lack of a bi-directional influence might be that word order preference in NGT might not be as variable as in Dutch. Hence, cross-linguistic influence might not take place since ground-first order might be robust and more resilient for change. Although NGT seems to have an invariant word order preference, Figure 27 indicates that not all deaf signers produced ground-first utterances and that some of the bimodal bilinguals in fact did produce figure-first descriptions in NGT. Thus, we argue that the one-way direction reveals that cross-linguistic influence of word order preferences in bimodal bilinguals might be modality-specific as cross-linguistic influence might be motivated by the modality-driven robust ground-first order rather than is due to linguistic constraints of NGT.

Visual attention

For the controls, our results further indicate that deaf signers preferred looking at grounds from the start of message preparation, while for hearing speakers a ground preference in eye-gaze emerged only later and instead started with a preference to fixate figures. This reflects that the modality-driven ground-first order in the language productions of deaf-signers also guides more attention to grounds right at the start of message preparation compared to speakers. This provides empirical evidence for the claim that what is mentioned first in a sentence is more conceptually foregrounded in the language users' mind (Gundel, 1985; Macwhinney, 1977). Furthermore, the fact that deaf signers prefer ground-first predominantly in their linguistic descriptions and also prefer looking at grounds over figures during the preparation of these descriptions reveals that *thinking for speaking* extends to *thinking for signing* (Chapter 2), also in the domain of word order preferences.

For the bimodal bilinguals, our results provide evidence for cross-linguistic influence of visual attention during message preparation but only when there was also cross-linguistic influence at the level of language production, that is, during spoken message preparation due to cross-linguistic influence from sign to speech but not during signed message preparation as there was no reverse cross-linguistic influence. In particular, when preparing Dutch descriptions, both bimodal bilinguals and hearing speakers preferred looking at figures over grounds at the initial stages of message preparation. As message preparation unfolded, both groups developed a preference to look at grounds versus figures over time. However, bimodal bilinguals' ground preference increased more over time compared to their hearing speaking peers. However, during the time course of preparing NGT messages, bimodal bilinguals preferred looking more at grounds than figures from early on and this preference did not differ from that of deaf signers.

Overall, during both language sessions, by the end of message preparation, all groups preferred looking at grounds over figures, which might be related to the arrangement of our visual displays. Namely, grounds were placed in the centre of the pictures while the location of figures varied in each picture (e.g. on the left, in the front). This might have attracted stationary gaze to grounds when messages were already largely prepared. Crucially, differences in eye-gaze preferences in Dutch and NGT emerged from early on when message preparation begun. This conforms with previous claims that during language production speakers look first at the referent that is mentioned first (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014). Even more, our results show that such eye-gaze language production links do not only arise during language production but already during message preparation.

Overall, our results reveal a one-way cross-linguistic influence from sign to speech not only in language production but also in visual attention. Thus the present study shows for the first time that cross-linguistic influence goes beyond language production and can at the same time shape the conceptualization of messages by influencing visual attention. This is in line with recent neuroimaging research showing different brain activation in hearing unimodal and bimodal bilingual infants when comprehending signed versus spoken language narratives. This suggests that neural substrates of language is also influenced by language modality already in infancy (Mercure et al., 2020).

To conclude, the current study revealed new insights into cross-linguistic influence by providing evidence from language production and visual attention. Particularly, our study revealed that cross-linguistic influence can occur across modalities in bimodal bilinguals in a naturalistic production setting. It further showed that the influence of word order to describe spatial language in bimodal bilinguals is modality-specific. Finally, the present study demonstrated that modality-specific influence has additional cognitive consequences that go beyond the level of language production.

Supplementary materials

Information about session order analysis in bimodal bilinguals

For Dutch, we assessed possible effects of session order on the proportions of ground first-descriptions in bimodal bilinguals. We used two general linear mixed-effects regression models, of which one assessed whether the ground was mentioned first (1) or not (0) in Dutch and the other assessed whether the ground was mentioned first (1) or not (0) in NGT. Both models included the categorical predictor Session (first or second) which was coded as a numeric contrast; that is, sessions conducted first were coded as -1/2 and sessions conducted as second as +1/2. For Dutch, the model contained no main effect of Session (β = -0.05, SE = 3.13, z = -0.02, p < 0.99). For NGT, the model contained also no main effect of Session (β = -0.65, SE = 3.51, z = -0.18, p < 0.85).

Chapter 6

General discussion and conclusion

In the current thesis, my goal was to enhance our understanding of how cross-modal diversity across and within populations can affect language production and cognition. I explored this by comparing diversity between a spoken and a sign language across hearing speaking and deaf signing populations as well as in hearing bilinguals who can both speak and sign, namely, bimodal bilinguals. Doing so, I provided not only new insights into how diversity in language production across populations can affect cognition but also revealed novel knowledge about how two languages can influence each other across modalities in bimodal bilinguals; that is, between speech and sign, which has not been examined before. Specifically, I tested hearing speakers of Dutch and deaf signers of NGT as well as bimodal bilinguals of Dutch and NGT. To do so, I chose the domain of spatial language, specifically into expressions of spatial relations, such as *the pen is to the right of the glass*, where previous research has identified clear modality-specific differences between spoken and sign languages in their linguistic patterning (e.g. Emmorey, 2002; Özyürek et al., 2010; Perniss et al., 2015; Zwitserlood et al., 2012). Furthermore, I focussed on two modality-specific differences between spoken and sign languages, namely iconicity and word order preference, in the domain of spatial language

I explored three overarching questions that I have investigated in four experimental studies in the current thesis and specifically for spoken Dutch and NGT:

(1) To what extent do differences in modality influence **production** across different speaking (hearing) and signing (deaf) populations?

(2) Do these modality-driven differences in production patterns found in (1) influence **cognition**, as measured by visual attention to referents during message preparation?

(3) How do patterns observed in (1) and (2) across different speaking (hearing) and signing (deaf) populations manifest themselves in bilinguals (hearing); that is, in the same individuals who can both speak and sign (i.e. bimodal bilinguals)?

By answering these questions and combining knowledge from three research areas, namely, language and cognition, sign languages, and bilingualism, this thesis enhanced our understanding of language diversity as well as whether and how this diversity can affect cognition. To explore these possibilities and answers to the questions addressed above, I combined elicited productions with an eye-tracking experiment to investigate possible relations between modality-driven production patterns and cognition.

In the remainder of this chapter, I will first outline the main findings described in the preceding experimental chapters and reflect on as well as discuss these findings within the broader literature and describe the implications of this work for current theory as well as limitations. Finally, I will speculate on how future studies can build on the experimental paradigm and findings presented in this thesis to understand language diversity and its effects on cognition.

6.1 Summary of main findings

In **Chapter 2,** I first examined the influence of iconicity on spoken and signed productions of left/right relations by comparing spatial descriptions between hearing speakers of Dutch and deaf signers of NGT (across populations). I also assessed within deaf signers to what extent they exploit different types of iconicity, such as those that map both spatial and physical properties of objects (CLs) and for those that map spatial properties only (RLs) onto the hands (see also, e.g. Emmorey, 2002; Emmorey & Herzig, 2003; Perniss et al., 2010). Secondly, I examined whether and how arbitrary (non-iconic in speech) versus iconic expressions (in sign) as well as different types of iconic expressions used to describe spatial relations influences message conceptualization by assessing how hearing speakers and deaf signers allocate their visual attention to pictures depicting spatial relations between objects during message preparation.

I predicted that, if iconicity is relevant for message conceptualization, then first, deaf signers, but not hearing speakers, would look more at all types of spatial relations. This might be because signing about spatial relations might require enhanced attention for the identification of the relative locations and/or the orientation, size and shape of the objects in relation to each other in all pictures in the display to map information onto the signing space iconically. Secondly, within signers, I predicted that depending on which features have been mapped, that is, spatial and physical features (CLs) versus spatial features only (RLs), deaf signers would attend to different pictures within the visual display. In particular, planning CLs in NGT might guide visual attention in addition to target pictures, to those non-target pictures that visually resemble the target picture (i.e. left/right/front/behind in which the figure object is placed to the sides of the ground object), whereas preparing RLs might guide visual attention to those non-target pictures that are only semantically relevant such as contrastive left/right pictures.

The production results of this chapter showed, as expected from previous research, that hearing speakers and deaf signers preferred modality-specific ways of expressing left/right spatial relations. Hearing speakers preferred categorical forms (LEFT/RIGHT) while deaf signers preferred iconic forms (CLs and/or RLs). Furthermore, within signers, they exploited different types of iconicity and preferred those expressions that resemble both spatial and physical properties of objects by using CLs over resembling only spatial properties by using RLs. Interestingly, using CLs was not the most preferred linguistic strategy, but instead, deaf signers overall preferred to use both CLs and RLs within one description instead of producing descriptions with either CLs or RLs alone. They also used RLs more frequently than what would have been expected from previous literature.

The results on message conceptualization showed that, as predicted, the modality-specific way of expressing spatial descriptions iconically resulted in increased visual attention to non-target pictures for deaf signers, but not for hearing speakers during message preparation. Moreover, within deaf signers, when preparing to resemble both spatial and physical properties of objects by using CLs signers allocated more visual attention to other pictures that visually resembled the target picture (left/right/front/behind) compared to pictures that did not share visual resemblance (in/on). Finally when preparing to resemble only spatial properties by using RLs signers allocated visual attention to only those pictures that were semantically relevant (left/right), that is, to those in which figures were located on the same lateral axis and therefore those that were contrastively related to the spatial location only (i.e. left/right). This demonstrates that differences in visual attention between hearing speakers and deaf signers cannot be simply attributed to different hearing statuses (hearing vs. deafness) but is rather specific to using iconic expressions. In sum, with **Chapter 2** I revealed that *thinking for signing* differs from *thinking for speaking* due to the iconic expressions in spatial language and that linguistic diversity even within a sign language can shape visual attention and thus conceptualization during message preparation.

In **Chapter 3,** I examined cross-linguistic diversity within populations to see whether bimodal bilinguals show similar production patterns as found across different speaking and signing individuals in Chapter 2, or if they speak and sign differently providing evidence for cross-linguistic influence across modalities. I tested whether cross-linguistic influence is present bi-directionally, namely, across modalities from sign to speech and from speech to sign, as well as within the same modality, that is, between co-speech gestures and sign as found in previous research (e.g. Casey & Emmorey, 2009; Gu et al., 2018). To explore this, I compared bimodal bilinguals' language production in each of their languages to that of their hearing speaking and deaf signing peers (who are also reported in Chapter 2). I hypothesised that arbitrary (non-iconic) and categorical expressions in speech and iconic expressions in sign might shape each other in a way that bilinguals' productions will differ from that of the hearing speaking and deaf signing controls. More specifically, as expressions in sign are more semantically specific due to the use of iconic forms (i.e. as in the use of CLs), I predicted that this might increase semantically specific expressions in speech. In return, using arbitrary and categorical forms might decrease iconicity in signed descriptions and increase the use of less semantically specific forms such as RLs.

Confirming these predictions, results revealed that speech and sign shape each other bi-directionally in bimodal bilinguals. In particular, using commonly iconic expressions in sign increased semantically specific information about physical features of objects in speech (cross-modal influence) and the use of iconic co-speech gestures alongside speech (within-modality influence), that is, sign influenced speech. At the same time, using categorical expressions in speech decreased iconicity in sign and increased the use of RLs, that is, speech influenced sign. Therefore, in this chapter, I revealed that cross-linguistic influence across modalities can shape bilingual diversity that differs from that found across populations demonstrating that bimodal bilinguals do no resemble two monolinguals in one but are instead shaped by two languages from different modalities.

In **Chapter 4**, I examined to what extent visual attention during message conceptualization can be changed within bimodal bilinguals as a consequence of bi-directional cross-linguistic influence shown in Chapter 3. To explore this, I assessed bimodal bilinguals' visual attention during spoken versus signed message preparation and then compared those to that of hearing speakers and deaf signers (similar to and reported in Chapter 2). I predicted that since bimodal bilinguals' speak and sign about spatial relations differently in Chapter 3 due to cross-linguistic influences, this might also affect how bimodal bilinguals visually attend to these spatial relations when preparing to speak versus to sign compared to that of the controls. Alternatively, cross-linguistic influence might not have cognitive consequences that can go beyond language production. Consequently, bimodal bilinguals might maintain their modality-specific conceptualization patterns and thus would not differ from that of hearing speakers when preparing spoken messages and from that of deaf signers when preparing signed messages.

In line with the first set of predictions, results indicated that during spoken message preparation, bimodal bilinguals, but not hearing speakers, allocated increased visual attention to non-target pictures in similar ways that were found in deaf signers of NGT during message preparation. Thus, cross-linguistic influence from sign in the form of increased semantic specificity (in both speech and co-speech gestures, Chapter 3) shaped bimodal bilinguals' visual attention to spatial relations differently to that of hearing speakers. Furthermore, during signed message preparation, both bimodal bilinguals and deaf signers experienced increased eye-gaze competition over time from non-target pictures. However, bimodal bilinguals' increase in eye-gaze competition was less steep than that of deaf signers. Thus, cross-linguistic influence from speech to sign in the form of decreased iconicity (Chapter 3) also shaped bimodal bilinguals' visual attention to spatial relations differently to that of deaf signers. Overall, this chapter revealed that cross-linguistic influence across modalities has further consequences for cognition, and more specifically, influences message conceptualization in bimodal bilinguals.

Finally, **Chapter 5,** investigated whether and how another modality-driven difference across spoken and sign languages; that is, a fixed word order preference, influences production patterns and conceptualization during message preparation across hearing and signing populations and in hearing bimodal bilinguals. This was to show that the effects found in Chapter 3 and 4 are not restricted to iconicity, but can also occur in another domain where modality-specific patterning occurs. Rather than comparing across populations and within bilinguals (as in Chapters 2, 3 and 4), in this chapter I compared all 3 populations in one study in the domain of word order preference.

For production, I compared bimodal bilinguals' descriptions of e target pictures in each of their languages to that of their hearing speaking and deaf signing peers. Descriptions were first assessed with respect to whether grounds or figures were mentioned first to examine which word order all three populations preferred. Comparing across populations, I hypothesised that deaf signers would consistently use ground-first but hearing Dutch speakers would use both types of word order. For bimodal bilinguals, I predicted that the modality-specific ground-first order in sign might influence speech, as I have found for the influence of iconic expressions in sign on speech in Chapter 3. For the reverse effect, if ground-first word order in sign is the modality-driven order and grounded in cognitive perceptual biases, it might be resistant to be influenced by the spoken language. Thus when signing, bimodal bilinguals might not differ from their deaf signing peers, unlike found in Chapter 3 for the effect of iconicity on speech.

Results indicated that deaf signers indeed predominantly preferred ground-first order while hearing speakers produced both types of word order. Bimodal bilinguals, when speaking produced more ground-first descriptions than their hearing speaking peers; however, when signing they produced as much ground-first order than their deaf signing peers. For bimodal bilinguals, thus, results indicated a one-way cross-linguistic influence from sign to speech, but there was no influence from speech to sign.

Next, I examined the influence of these word order preferences on message conceptualization by comparing visual attention to grounds and figures within the target picture during spoken versus signed message preparation in all the three groups. Previous research suggested that language users allocate more visual attention to those aspects of a visual scene that are mentioned first within an utterance (e.g. Griffin, 2004; Griffin & Bock, 2000; van de Velde, Meyer, & Konopka, 2014). Based on these findings, when comparing hearing speakers to deaf signers, I expected deaf signers to attend more to grounds than hearing speakers. For bimodal bilinguals, I predicted that the one-way cross-linguistic influence from sign to speech, will change how bimodal bilinguals visually attend to grounds versus figures when preparing to speak but not when preparing to sign.

Results confirmed these predictions and showed that deaf signers allocated more attention to grounds than figures during message preparation compared to hearing speakers. Bimodal bilinguals allocated more visual attention to grounds than to figures compared to their hearing speaking peers when preparing to speak. However, when preparing to sign bimodal bilinguals' visual attention did not differ from those of the deaf signers. Therefore, in this chapter, I revealed that modality-driven differences in word order preferences can influence language production and message conceptualization in bimodal bilinguals and can at the same time also hinder cross-linguistic influences.

6.2 Discussion and implications

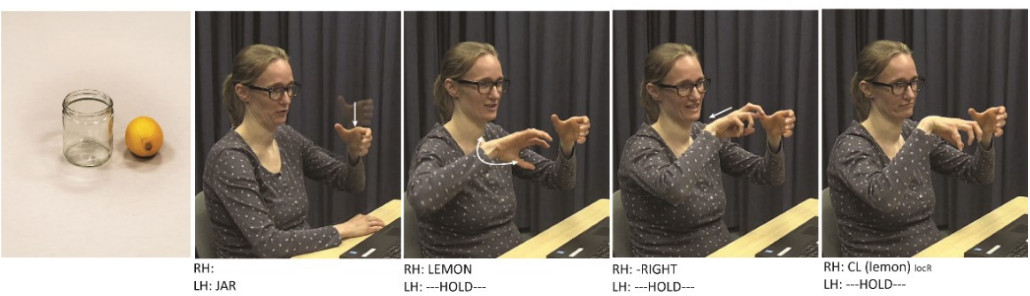
6.2.1 Differences in modality shape language production across hearing speakers and deaf signers in the domain of spatial language

The first core question of this thesis was to what extent do differences in modality across different speaking (hearing) and signing (deaf) populations influence production in general and in the language pair of spoken Dutch and NGT in particular. In this thesis, I have provided concrete quantitative data about how frequently different types of linguistic strategies are preferred by spoken and sign languages confirming that modality-specific patterning in a sign language arises in the domain of spatial language (e.g. Emmorey, 2002; Özyürek et al., 2010; Perniss et al., 2015; Zwitserlood et al., 2012). I have demonstrated this for the specific language pair Dutch and NGT for left/right spatial expressions in two domains, namely, iconicity and word order preference. Results provide novel insight into the nature of iconic expressions and modality-specific word order in sign as well as confirm how differences in modality shapes cross-modal diversity. I will discuss these contributions in detail below.

The nature of iconic expressions in sign. Initially, sign language researchers have assumed that CLs are the most common and universal strategy across different sign languages for expressing space (see Cuxac, 1996; Sallandre & Cuxac, 2002; Talmy, 2003). From this assumption, it has been suggested that signers achieve the highest degree of iconicity possible (Perniss et al., 2015). More recent comparative work has shown that there might be variation and that other linguistic strategies such as RLs are used more often than originally assumed (e.g. Perniss et al., 2015; Sümer, 2015), however, overall quantitative studies were missing. In **Chapter 2**, I have provided first of all evidence that while CLs are a more preferred strategy then RLs, using them alone is not the most preferred strategy in deaf signers of NGT. Instead, deaf signers overall favoured using both CLs and RLs in one description instead of producing descriptions that contained either CLs or RLs alone (see Figure 29). This occurrence of double strategies has been also previously observed in Turkish Sign Language during spatial descriptions of front/behind (Sümer, 2015; Sümer et al., 2014). The preference of double strategies confirms that iconicity is more than just a binary property (i.e. iconic vs. non-iconic) and rather manifests itself in various forms and degrees and involves different types of form-to-meaning mappings (e.g. Bellugi & Klima, 1976; Dingemanse et al., 2020; Ortega et al., 2017). Moreover, combining both CLs and RLs also suggest that CLs seem to serve the function of highlighting and depicting (Ferrara & Hodge, 2018) rather than obligatory marking and classifying locations and object properties (Zwitserlood, 2003).

Figure 29

Example of a double strategy for "the lemon is to the right of the jar" in NGT using both RL (RIGHT) and CL within one description.



These findings suggest that signers do not simply aim to exploit the highest degree of iconicity at any given movement but other factors seem to determine how often signers use linguistic strategies, such as communicative context and informativeness. Based on the frequent use of double strategies it seems that deaf signers preferred exploiting the most informative strategy over the most iconic strategy only. In particular, by producing CLs, deaf signers deploy the highest level of iconicity, through which objects and relations are depicted and highlighted in a one-to-one fashion. At the same time, by adding categorical RLs signers emphasise the location of the figure object in respect to the ground object. Thus by combining CLs and RLs in one description deaf signers exploit enhanced informativeness and demonstrate high efficiency in their choice of linguistic strategies (for more insights into signers' communicative efficiency through using a variety of iconic forms in Italian Sign Language, see Slonimska et al., 2020). Overall, in this thesis, I have revealed that informativeness in addition to being fully iconic drives spatial expressions (see also Perniss, 2007; Sümer, 2015) in certain communicative contexts.

One reason for this high level of informativeness might be the nature and complexity of the experimental paradigm used in this thesis. In particular, I presented four-picture displays containing two objects in different spatial relations. In fact, hearing speakers were also highly efficient by preferring concrete spatial expressions, such as LEFT/RIGHT, over more ambiguous and vaguer general expressions, such as NEXT TO. Crucially, these trends occurred irrespectively of how complex the presented displays were (i.e. more complex displays: both left and right present vs. less complex displays: either left or right). Thus, my Visual World Language Production Paradigm seems to simulate a communicative context in which deaf signer's preferred achieving a high level of informativeness. Overall, these findings reveal that the communicative context and informativeness might also determine how often a strategy is used in sign languages.

Modality specific word order preference in sign. In **Chapter 4,** I confirmed that there is no pre-set linguistic word order in Dutch for describing spatial relations but rather both figure-first and ground-first are valid and acceptable word orders. While this flexibility in word order preference for spatial relations has been previously suggested (Hartsuiker, Kolk, & Huiskamp, 1999) this thesis provides empirical and quantitative evidence based on an elicited corpus in Dutch.

Furthermore, I have revealed that NGT predominantly prefers ground-first order, as has been found in many other sign languages (see among others Emmorey, 1996 for American Sign Language; Morgan, Herman, Barrière, & Woll, 2008 for British Sign Language; Sümer, 2015 for Turkish Sign Language; Perniss, 2007 for German Sign Language; and Kimmelman, 2012 for Russian Sign Language). This confirms the notion of ground-first order as a universal and modality-specific word order for the visuo-spatial modality.

Overall, the findings on hearing speakers' and deaf signers' language production of **Chapter 2** and **Chapter 5** confirm that there is cross-modal diversity in the way spoken and sign languages express spatial relations in terms of both iconicity and word order preferences, driven by affordances in each modality. Furthermore, the fact that the preferred forms found in NGT match those found in many sign languages around the globe (i.e. iconic forms and ground-first word order preference) indicates not only that these favoured forms are a result of the affordance of the visuo-spatial modality of sign languages but also that signers might use such forms for purposes other than being fully iconic (Perniss et al., 2015).

6.2.2 Differences in modality shape message conceptualization across hearing speakers and deaf signers

The second core question of this thesis was to find out to what extent the differences in modality found across hearing and deaf populations discussed above influence cognition. Going beyond previous investigations that have focussed on differences between spoken languages only, in this thesis I have revealed that differences in modality (vocal vs. visuo-spatial) can shape cognition for language production between speaking (*thinking for speaking*) and signing (*thinking for signing*). These findings offer important contributions to our understanding of how iconicity and modality might influence cognition as well as more generally how language production can influence cognition, which will be discussed below.

Effects of iconicity and modality on cognition. In **Chapter 2** I showed that iconicity can shape message conceptualization differently for signers than for speakers during message preparation of spatial relations. In particular, due to using iconic expressions of spatial relations deaf signers seem to allocate more attentional resources to the relative locations and/or physical features of the objects in relation to each other compared to hearing speakers that use arbitrary and categorical expressions when using a spoken language. This demonstrates for the first time that message preparation for language production affects how hearing speakers and deaf signers conceptualise spatial relations differently. Thus, *thinking for speaking* (Slobin, 2003) differs *from thinking for signing*. Furthermore, I demonstrated for the first time that sign production and eye-gaze are linked as speech production is linked to eye-gaze. Furthermore, I showed links between eye-gaze and sign language production not only at the lexical level (shown previously for sign language comprehension, e.g. Lieberman et al., 2015) but also at the sentence level (as shown for spoken language production; e.g. Griffin, 2004; Griffin & Bock, 2000; Koneke & Meyer, 2014; van de Velde et al., 2014) and during message preparation.

Moreover, I demonstrated that deaf signers' visual attention was also modulated by the different types of iconicity that they exploited in their spatial expressions. Depending on which features have been mapped, that is, spatial and physical features (CLs) versus spatial features only (RLs), they attended to different aspects of the visual display. This reveals that within a sign language, different types of iconic expressions such as those that are iconic to both the object properties and spatial location influence signer's conceptualization of messages differently than when using signs that iconically depict spatial locations only.

This supports the idea that iconicity involves semiotic relations which manifests itself in various forms and degrees (e.g. Bellugi & Klima, 1976; Dingemanse et al., 2020; Ortega et al., 2017) and reveals that these types of iconicity can shape signers' cognition relevant for language production and already during message formulation. These results can be considered to confirm the earlier mentioned argument that CLs can be used optionally and flexibly rather than obligatorily to highlight and depict object locations and their physical properties (see e.g. Ferrara & Hodge, 2018). If CLs were frozen and obligatory forms then they might not have changed deaf signers' message conceptualization as strongly. That is, highlighting the size, shape and orientation of the objects as well as the locations of these entities might have modulated the conceptualization of messages depending on the type of iconicity even before they are produced. These results can be considered to be in line with viewing CLs as *depictive signs* rather than as *classifier predicates* with morphological structures.

Overall, these effects of iconicity on message conceptualization demonstrate that iconicity as well as the type of iconicity matters for language processing and shed therefore more light in the ongoing debate whether iconicity matters or not (for evidence that iconicity does not matter, see e.g. Anderson & Reilly, 2002; Bosworth & Emmorey, 2010; Orlansky & Bonvillian, 1984; Poizner et al., 1981; for evidence that iconicity matters, see e.g. Campbell et al., 1992; Grote & Linz, 2003; Navarrete et al., 2017; Sümer, 2015; Thompson et al., 2010, 2012). Furthermore, the thesis goes beyond these previous findings that are based on the lexical level for language comprehension and learning and reveals that iconicity on the sentence level can affect cognition for language production.

These modality-driven effects on message conceptualization across spoken and sign languages seem not to be restricted to iconicity but can also occur in a different domain, such as word order as shown in **Chapter 5**. In particular, deaf signers' modality-driven preference to mention grounds first shaped their message conceptualization during message preparation and required more attentional resources to grounds than figures. Hearing speakers, however, preferred to fixate figures at the beginning of message preparation. This finding too, demonstrates that *thinking for speaking* extends to *thinking for signing* in the domain of word order preferences. Furthermore, this link between preferences in visual attention and the linguistic order in which these aspects are mentioned is in line with previous research on spoken languages (e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014). These findings extend our knowledge from speech eye-gaze links to sign eye-gaze links in the domain of word order preference and eye-gaze. Even more, the results reveal that such eye-gaze links do not only arise during language production (*ibid.*) but already occur during message preparation.

Considering the key findings from **Chapter 2** and **5** together, I have demonstrated that not only language-specific (e.g. Bock, 1995; Flecken et al., 2015; Levelt, 1989; Papafragou et al., 2008) but also modality-specific ways in which signed messages are formulated become automatized and shape visual attention during the preparation of messages before they are produced. Now I want to highlight the differences in time course in which the effects on message conceptualization occurred across **Chapter 2** and **5**, which provide further insights into message conceptualization during signed message preparation.

In particular, in **Chapter 2,** the effects of iconic expressions on deaf signers' visual attention (i.e. increased visual attention to non-target pictures) were largest at the end of message preparation, that is, right before deaf signers described the spatial relations. In contrast, in **Chapter 5,** where I investigated the order of first mentions of grounds versus figures, effects were largest at the very initial stages of message preparation (i.e. preference to visually attend to grounds over figures), that is, right at the beginning of linguistic planning. These differences in time course might be related to the linguistic order in which grounds/figures and iconic forms are mentioned within a signed sentence during sign language production (first vs. later). The production model by Levelt (1989) suggests that pre-linguistic formulation begins with the activation of lexical items. In line with this, **Chapter 5** indicates that during signed message preparation the lexical items of the objects involved are being activated first and in the order there are mentioned. Further in Levelt's view, all other linguistic aspects of an utterance are generated based on the linguistic properties of these activated items. In line with this, **Chapter 2** suggests that the planning of iconic expressions (such as CLs and RLs) seem to follow the formulation of the lexical items based on the linguistic order in which they are mentioned as visual attention to non-target pictures increased up to the point of language production. However, I had noticed that iconic forms (CLs/RLs) can occur at different positions in a sentence (e.g. after both the ground and figure are mentioned or after the ground and before the figure is mentioned), although I have no quantitative reports for this observation. Based on this, the fact that effects of iconic expressions were largest at the end of message preparation might also indicate the temporal dynamics of formulating iconic expressions. That is, the concepts and lexical items of the objects might need to be activated first to identify the spatial and/or physical features of these objects to map them onto the signing space iconically. This difference in the time course of visual attention provides first and novel insights into message conceptualization during signed message preparation reflecting how different modality-specific aspects are packaged and represented in the signers' mind.

General implications for the relationship between language and cognition. The results of this thesis also contribute at a more general level to the debate on the language-cognition interface, showing that at the moment prior not only speaking but also signing, language can serve as a toolkit to highlight and represent certain aspects of the world which then change the conceptualization of this world (e.g. Flecken et al., 2015;

Papafragou et al., 2008; Slobin, 2003). By going beyond previous investigations that have focussed on differences between spoken languages only, I have demonstrated that differences in modality can shape cognition for language production. More specifically, I have revealed that the relationship between visual attention and language production is not only linked to discrete and categorical aspects of language production but also to iconic expressions.

In studies assessing how classifiers in spoken languages can shape speakers' cognition, it has been debated whether such grammatical systems can affect the way speakers think about objects and referents (e.g. Saalbach & Imai, 2007; Schmitt & Zhang, 1998) or whether they simply reflect conceptual structures (e.g. Speed, Chen, Huettig, & Majid, in press). While this has been typically assessed within categorical language structures here I introduced iconicity where language forms are more analogue and grounded in semiotic relations. This investigation revealed that using iconic expressions (that are in some ways akin to classifiers but see my point on depictions above) can, in fact, shape rather than just reflect conceptual structures as the planning of these CLs in sign shaped message conceptualization differently than the categorical and arbitrary forms in speech.

6.2.3 Cross-modal influence in bilingual diversity and its effects on message conceptualization

The last core question of this thesis asked how the language patterns, as well as their effects on cognition, discussed so far across different speaking (hearing) and signing (deaf) populations manifest themselves in bilinguals (hearing); that is, in the same individuals who can both speak and sign (i.e. bimodal bilinguals). In this thesis I revealed that cross-linguistic influence can occur between speech and sign across modalities in bimodal bilinguals and also changes bilinguals' message conceptualization compared to that of hearing speakers and deaf signers. That is, (bimodal) bilingual diversity patterns go uniquely above and beyond diversity found across populations and shape bilinguals' cognition differently. Below I will discuss this diversity and its implication for cognition in more detail as well as implications for bimodal bilinguals' production.

Bilingual diversity can be shaped by languages from different modalities. Cross-linguistic influence can occur in bimodal bilinguals between sign and speech across different domains, namely, iconicity (**Chapter 3**) and word order preference (**Chapter 5**). These influences occurred independently of the language status, which contrasts with previous findings in heritage bilinguals of two spoken languages, where cross-linguistic influence was typically evident from the majority to the minority language (e.g. Backus, 2005; Muysken, 2000; Polinsky, 2008). The findings in the current thesis indicate that cross-linguistic influence can occur between the minority (NGT) and majority language (Dutch) and not necessarily only from a majority to a minority language. Even more, influence from a majority to a minority language can be even constrained as shown in **Chapter 3**. This provides evidence that modality might be a driving factor for cross-linguistic influence and confirm our results about bimodal bilingualism as a special form of bilingual diversity. Furthermore, my findings also contrast with previous claims that if spoken heritage bilinguals are highly proficient in both their majority and minority language there is less or no cross-linguistic influence (e.g. Azar, 2020; Azar et al., 2019). The bimodal bilinguals assessed in this thesis are highly proficient in both Dutch and NGT as they did not differ from their hearing speaking and deaf signing peers in the extensive objective proficiency measures that I have conducted in this thesis. Under which exact circumstances cross-linguistic influence can occur is unclear, however, might not always depend on language proficiency but also on other factors such as the domain or language modality which I will discuss more below.

Interestingly, I found some differences in the modality-specific domains, such as iconicity versus word order preference in which cross-linguistic influences occurred. On the one hand, cross-modal influence resulted in the creation of a new conceptual system (**Chapter 3**). On the one hand, in the domain of iconicity cross-modal influence resulted in bi-directional influence and the creation of a new conceptual system (**Chapter 3**). However, in the domain of word order, cross-modal influence was more constrained and driven by modality-specific affordances of the visuo-spatial modality of sign languages, as in the robust and cognitively driven ground-first word order of NGT (**Chapter 5**). Below I will compare each influence in detail and will then discuss why these differences in cross-linguistic influence might have arisen.

In **Chapter 3** I confirmed previous evidence for bi-directional cross-linguistic influence, as found for spoken language bilinguals (e.g. Brown & Gullberg, 2011; Jarvis, 2011; Pavlenko & Jarvis, 2002). Crucially, I extended these findings to cross-linguistic influence across modalities. Arbitrary and categorical expressions in speech and iconic expressions in sign shaped each other in a way that bilinguals' productions differed from that of their hearing speaking and deaf signing peers. This influence between speech and sign crossed the modality boundary and gives rise for the existence of what I term *cross-modal influence*. Despite the differences in language formats, cross-linguistic influence found its way into speech by transforming iconic expressions from NGT into adverbial and prepositional phrases containing increased semantically specific information about the physical properties of the objects. At the same time, using arbitrary and categorical expressions in speech reduced iconicity and semantic specificity in signed expressions. In the current thesis, I revealed for the first time that such a cross-modal influence can occur in bimodal bilinguals and showed this way that bimodal bilinguals do not resemble two monolinguals in one (see Grosjean, 1989). Instead, their language production can be modulated by two intertwined languages from different modalities. This reveals that cross-modal influence can shape (bimodal) bilingual diversity which differs from that found across populations.

In addition, my results revealed cross-linguistic influence within a modality, that is, from sign to co-speech gestures which support previous findings (e.g. Casey &

Emmorey, 2009; Casey, Emmorey, & Larrabee, 2012; Gu et al., 2018; Weisberg et al., 2020). In detail, bimodal bilinguals produced more semantically specific manual productions such as iconic gestures/signs (e.g. CLs) accompanying their speech than their non-signing peers. This within-modality influence from sign and co-speech gestures occurred more frequently than the across-modality influence from sign to speech. Thus, even though sign can influence speech, this influence seems to be more constrained than influence from sign to co-speech gestures. This suggests that cross-linguistic influences are stronger within the same modality, that is, from sign to manual productions than across modalities, that is, from sign to speech. This can be due to the lack of corresponding forms between sign and speech and points out to the unique diversity found in bimodal bilinguals.

Interestingly, the cross-modal bi-directional influence on language production found for iconicity in **Chapter 3** did not directly replicate for word order in **Chapter 5**. That is, in the domain of word order preferences I could not observe a bi-directional influence between speech and sign, but rather a one-way influence; that is, from sign to speech but not from speech to sign. In particular, word order preferences in sign influenced speech, but in turn, word order preference in speech did not influence sign. I argue that this one-way direction reflects modality-specific mechanisms that drive this cross-modal influence because the modality-driven ground-first order in sign is more resilient for change.

One of the reasons for this difference in cross-linguistic influence across the domain of iconicity and word order might be differences in their variability in the language. In the case of word order preference, speech does not influence sign because ground-first order in sign appears to be invariable and more resilient to change due to the modality-specific nature of ground-first and its groundedness in cognitive perceptual biases. For iconicity, using CLs versus RLs seems to allow more variability than word order for speech to influence the choices in sign for bimodal bilinguals to describe spatial relations and thus enables cross-linguistic influence. Thus, these findings provide new insights into language variability in NGT, showing that the type of iconicity is open for change and allows variability while ground-first order is more robust and constrained. Thus the domain of language determines what can be influenced or not.

Bilingual diversity changes message conceptualization during message preparation. In this thesis, I also showed that cross-linguistic influences at the behavioural level can also affect bimodal bilinguals' cognition. In the case of **Chapter 4**, I revealed that languages can not only converge at the linguistic level but bilinguals' cognition can also converge into a new system when preparing to speak or sign, which differs to that of the controls. Similar effects have been found previously showing that spoken language bilinguals conceptualise room sizes differently than their monolingual peers, however, these findings have been not directly linked to language production (Wolter et al., 2020). To my knowledge, cross-linguistic influence in language production and differences in conceptualization has been always studied separately. Combining these two approaches, the findings of **Chapter 4** and **5** reveal for the first time that cross-linguistic influence can go beyond language production and also have cognitive consequence, such as changing message conceptualization during message preparation. Importantly, **Chapter 5** demonstrates further that *only* where there is cross-linguistic influence there are effects on message conceptualization. Overall, these results provide evidence also from the cognitive level that bimodal bilinguals can accommodate two languages from two modalities in a single system and extend the validity of this claim from spoken to bimodal bilinguals (e.g. Brown & Gullberg, 2011; Jarvis, 2011; Pavlenko & Jarvis, 2002).

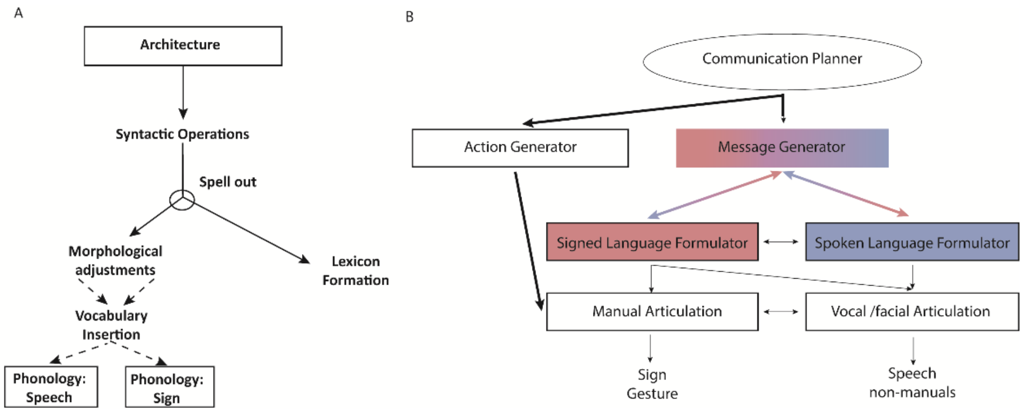
Implications for bimodal bilingual language production models. Results on the production of both languages in bimodal bilinguals in different domains provide novel insight into the way two languages from different modalities can influence and shape each other and result in a bilingual diversity that differs from that found across populations. It moreover reflects that bimodal bilinguals activate the sign language while speaking the spoken language and vice-versa. This supports previous claims on co-activation of sign languages while using spoken languages shown (during language production, see e.g. Emmorey et al., 2008; during language comprehension, see e.g. Giezen et al., 2015; Shook & Marian, 2012) and extends evidence from the lexical to the sentence level of spatial descriptions.

Finally, my results on cross-linguistic influences at the level of language production as well as during message conceptualization provide further evidence for the existing two bimodal bilingual language production models by Lillo-Martin, de Quadros, and Pichler (2016) and by Emmorey et al. (2008) which will be discussed further below.

First, the Language Synthesis model (Figure 30A; Lillo-Martin, de Quadros, & Pichler, 2016) claims that cross-linguistic influence is simply generated by one architecture that allows features from different languages to enter into one single source as long as it is not violating any language constraints (for a similar claim for spoken languages, see MacSwan, 2000). Thus, in line with the models' claims, I found that cross-linguistic influence occurs between linguistic features despite differences in non-corresponding formats, but only for those domains that allow variability and do not violate language constraints. Even though I found cross-linguistic influence in the domain of iconicity, in the domain of word order I found no influence from speech on sign because it may have violated the modality-specific and robust word order preference in sign.

Figure 30

Contribution to the bimodal bilingual language production models.



Panel (A) illustrates an adapted figure from Lillo-Martin et al.'s (2016) model. Panel (B) illustrates an adapted figure from Emmorey et al.'s (2008) model.

Second, in respect to the model by Emmorey et al. (2008), within-modality influences between languages (sign to co-speech gestures) have been suggested to occur at the level of the Message Generator (see bold black lines in Figure 30B; e.g. Casey & Emmorey, 2009). In line with this, I propose that cross-linguistic influences across modalities (between speech and sign) in both domains, namely, iconicity and word order, also occur via the Message Generator (visualized by the coloured boxes and arrow lines in Figure 30B). Because the Message Generator is the place where pre-verbal messages are formulated, an influence between the Spoken and Signed Formulator via the Message Generator reflects not only possible influences on language production but also on message conceptualization during message preparation. Thus, the thesis provides first evidence for the existing routes between the Message Generator and both Signed and Spoken Language Formulators. It also supports previous claims that commonly used expressions in a spoken and sign language guide attention to certain aspects of states and events at the level of the Message Generator, supporting previous claims on the role of the Message Generator in multimodal utterance production (i.e. for speech and gesture; Kita & Özyürek, 2003), extending it to the mind of bimodal bilinguals. Specifically, semantically specific information and ground-first order in the Message Generator used for modality-specific expressions in sign, could influence bimodal bilinguals to produce more semantically specific expressions and more ground-first order in speech through the Spoken Language Formulator. This then results in cross-linguistic influence from sign to speech as well as changes in message conceptualization when preparing to speak. Similarly, speech can influence sign in the domain of iconicity but not word order (as discussed above) through the Message Generator, which is affected by categorical forms in speech. The Message Generator influenced by speech might then influence the Signed Language Formulator, resulting in cross-linguistic influence from speech to sign as well as changes message conceptualization when preparing to sign.

6.3 Methodological contributions of the thesis

Having discussed the theoretical contributions of this thesis, I would also like to highlight the methodological contributions of this thesis, that is, the use of the Visual World Language Production Eye-Tracking Paradigm. With this paradigm, I created a platform to study sign language production and its links to message conceptualization in the same experiment. Through the deployment of a visual cue to indicate the target picture, the paradigm is usable for both speaking and signing populations as well as for both adults and children. As a whole, the paradigm is the first of its kind and an important contribution to the sign language research field were many aspects of language processing are still understudied compared to the spoken language research field and where eye-tracking is still not used very frequently. This paradigm provides hopefully a powerful tool for sign language researchers enabling the investigation of various aspects of different stages of online sign language production and its cognitive underpinnings (for ideas for future research, see section 6.4 below).

Further and beyond the field of sign language research, the paradigm is also a first to involve displays which each picture included two objects in different spatial configurations. Typically, pictures used in Visual World Paradigms involve a single referent or object. Thus as such, the paradigm provides many possibilities to expand our knowledge onto the relationship between spatial language production, both spoken as well as signed and eye-gaze.

Additionally, the paradigm resembles a relatively naturalistic language production experiment. Many of the theoretical developments on which I have based my research, come from the domain of eye-gaze language production links. In these studies, participants are typically instructed to describe pictures or animated scenes to a computer screen while their speech and eye-movements are being recorded (e.g. Griffin & Bock, 2000; Papafragou et al., 2008; van de Velde et al., 2014). This method is noticeably very useful and has provided many important breakthroughs in understanding how eye-gaze and language production are connected. However, I believe that the method I have presented in this thesis provides an additional useful contribution and opens up new avenues to assess not only isolated speech but rather language production as a multimodal construct. In particular, I wanted to move towards a more multi-person design and therefore involved a confederate to resemble a more naturalistic communication setting with a concrete communicative goal: constructing an utterance with content so the interlocutor can identify the correct picture. Consequently, descriptions are more naturalistic than those addressed to a computer screen and allow the communicative use of co-speech gestures. Thus, the paradigm as such provides an important contribution to the field because it enables the investigation of eye-gaze language production links in a naturalistic setting that is akin to a communicative director-matcher task, while at the same time experimental control, necessary for eye-tracking research, is ensured.

In addition to the Visual World Language Production Eye-Tracking Paradigm, the thesis is also the first to bring together cross-linguistic influence and message conceptualization in bilinguals. Previously, these two domains have been assessed separately. By assessing both language and conceptualization patterns and by investigating those across populations and within bilinguals as well as considering both of bilinguals' languages the thesis provides rich insides into how two languages can influence each other behaviourally and cognitively. This approach might hopefully set a first stepping stone for assessing bilingual diversity and the bilingual mind combined.

6.4 Future directions

6.4.1 Going beyond left/right spatial relations

In this thesis, I have revealed how differences in modality can shape spoken versus signed production and how this can affect message conceptualization during spoken versus signed message preparation across speaking and signing population as well as how these manifest themselves in bimodal bilinguals. The effects that I have demonstrated and discussed so far, all relate to the expressions of left/right relations. I focussed on left/right relations because they can be mapped onto and relative to the body and space and are contrastive, which allows studying visual attention effects most clearly. Furthermore, left/right relations do not necessarily involve occlusion of the figure objects in contrast to in/on/behind relations, where the figure object touches or partly hides behind the ground object. Thus, they are especially suitable for measuring visual attention prior to language production.

Nevertheless, it would be worthwhile to explore whether similar effects revealed in this thesis can be observed for other spatial relations as well. In fact, in **Chapter 2** deaf signers' eye-movements indicated increased eye-gaze competition for the sagittal competitor (i.e. front/behind) when preparing descriptions with CLs. Further, the fact that signers experienced increased eye-gaze competition from all types of spatial relations supports the idea that signing about spatial relations requires the identification of the relative locations and/or the orientation, size and shape of the objects in relation to each other to map information onto the signing space iconically. Consequently, disambiguating where to move or place the hands in space might require more effort for signers during production planning. Thus, I would predict to find similar eye-gaze effects of iconicity during the planning of for example sagittal relations as well. Overall, I expect that the finding from **Chapter 2**, that *thinking for speaking* differs from *thinking for signing*, is not depending on the type of spatial relation that is being conceptualised.

Lastly, it would be interesting to ask which other aspects of iconicity could affect message conceptualization, such as iconicity found in the domain of motion events (Emmorey, 2002; Perniss, 2007; Sümer, 2015). To describe, for instance, a scene in which a woman is running towards a car, signers would express the size and shape of the woman's legs, and the position and shape of the car in signing space. With a wiggling motion of the fingers moving towards the handshape that represents the car, signers can then resemble the running manner towards the car. Thus, in the domain of motion events sign languages can also exploit rich iconic expressions that visually resemble the motion event in an ono-to-one fashion due to the use of the visuo-spatial modality.

Based on this, it would be worthwhile to investigate whether such complex and iconic expression would also guide signers' visual attention during message preparation to describe motion events and also perhaps increase semantic specificity in bimodal bilinguals' motion expressions in speech. According to the findings of **Chapter 2** and **4**, I would expect that signers experience increased visual attention to the manner of motion during signed message preparation as well as bimodal bilinguals during spoken message preparation due to cross-linguistic influence from sign. Choosing handshapes and movements of the hands and/or fingers to represent a woman running, might require enhanced attentional resources to those aspects of the scene.

6.4.2 Implications for developing systems in children

The work presented in this thesis also raises interesting questions on whether differences in *thinking for speaking* versus *thinking for signing* would also extend to across populations in children, whose languages and ability to talk about space is still developing. Previous research has suggested that due to the affordances of the visuo-spatial modality of sign languages deaf children of Turkish Sign Language acquire left/right relations earlier than their hearing speaking peers of Turkish (Sümer, 2015).

Based on these modality-specific differences in language development it would be interesting to investigate whether deaf signing children would also conceptualise their messages differently during message preparation than hearing speaking children. Even though deaf signing children are found to use CLs less often than their adult peers, their expressions are still more iconic compared to those of deaf children (Sümer, 2015). Thus, I would expect deaf signing children to also allocate more visual attention to non-target pictures during message preparation than hearing speaking children.

Furthermore, it would be interesting to look at the effect of delayed language input in deaf children based on recent findings showing that delayed sign language acquisition hinders the acquisition of descriptions of viewpoint-dependent relations (Karadölller et al., in press). In this study, deaf children acquiring Turkish Sign Language late between the age of 6 and 7 used CLs less frequently than their deaf native signing peers. Based on this reduced use of CLs I would expect that late signing children would also allocate less visual attention to non-target pictures during message preparation than native signing children.

6.4.3 Implications for L2 learners of a sign language

My findings for cross-linguistic influences between speech and sign in bimodal bilinguals raise also interesting questions about whether and how these findings are specific to bimodal bilinguals or whether they also extend to L2 learners of a sign language. This might have important implications for sign language interpreters who often learn a sign language later in life when being enrolled in interpreter programmes (as it is the case in the Netherlands). That is, possible influence effects might be important to consider for their speech structure as well as their signing.

Previous research has indicated that already after one year of sign language instruction, these learners produced more iconic co-speech gestures compared to before being exposed to sign language (Casey et al., 2012; Weisberg et al., 2020). This suggests that the newly learnt sign language influenced their spoken language already a year after sign language exposure. This raises questions on whether the iconic expressions in a sign language can also lead to increased semantically specific speech as I have found in **Chapter 3** in bimodal bilinguals. Typically and as mentioned above, CLs are harder to learn for children acquiring a sign language than for instance RLs, because they require detailed and complex mappings between the real referents and the hands onto the signing space (Karadölller et al., in press; Sümer, 2015). However, in sign language L2 teaching, especially for teaching NGT, training on using CLs take up a large part of the teaching syllabus. Consequently, after being enrolled in an interpreter study for two years, sign language L2 learners are found to actually produce more CLs than their deaf signing peers to express motion events (Visker, 2020). Based on this it would be interesting to see if this over-use of CLs also influences expressions in their spoken L1 in form of semantic specificity as well as influences their message conceptualization when preparing to speak and sign. Alternatively, these cross-over effects might be specific to only bimodal bilinguals who are born to deaf parents and have had early input to a sign language.

Another interesting question would be to ask whether the results from **Chapter 5,** would be also visible in beginning L2 learners, in particular, those that are native in a spoken language that allows both ground-first and figure-first. Specifically, could ground-first preference in sign influence their word order preference in speech and when? It is possible, that the influence from sign to speech shown in this thesis is also specific to bimodal bilinguals studied in this thesis. That is, sign language learners might need a certain amount of sign language exposure to switch from mixing ground-first or figure-first in their spoken language to using ground-first consistently in sign. Cross-linguistic influence of ground-first from sign to speech might thus take more time and experience with a sign language.

Finally, if any of these cross-linguistic influences occur in more proficient sign language users such as sign language interpreters, then this suggests that interpreters might co-activate their spoken and sign language in both directions at the level of semantic specificity and word order when speaking or signing. Findings of this thesis could then be used to improve their training by increasing awareness of these influences and using this knowledge to enhance their interpreting skills.

6.5 Societal implications

The type of work presented in this thesis can have a profound impact on the debate over using sign language with deaf and hard of hearing children who have received cochlear implants, high-tech medical devices that help the deaf perceive sound. Often, medical professions advice against the use of sign language after receiving such implants, arguing that sign language would hinder the development of speech (e.g. Humphries et al., 2012, 2013). The findings on bimodal bilinguals' language use provide evidence that signing does not harm speech development. Importantly, cross-linguistic influences can occur naturally in bilinguals and are not an indication that bilinguals are confused (e.g. Unsworth, 2016). Instead, cross-linguistic influences provide even evidence that bilinguals are sensitive to grammatical and sociolinguistic constraints in both languages. Based on the findings in **Chapter 3,** acquiring sign language and spoken language together from early on (i.e. bilingual education) might even make speech and communication richer. I want to emphasise that sign languages do not harm the acquisition of speech but that a spoken and a sign language support and shape each other. This is in line with recent views suggesting that bilingual development prevents language deprivation and poor health outcomes in deaf children (Wilkinson & Morford, 2020). Thus, with this thesis, I also would like to increase awareness of language practices in the deaf community. Bimodal bilingual language development should be accepted as much as bilingual development of two spoken languages.

6.6 Conclusion

Languages differ fundamentally from one another in how they package and represent different aspects of our experiences. This language diversity influences how speakers perceive and conceptualise the world differently. In the current thesis, I hope to have enhanced our understanding of how language diversity can affect cognition for language production by comparing users who use a spoken language and/or a sign language. I showed that the modality-specific affordances of the visuo-spatial modality of sign languages influence production differently in signing (deaf) populations compared to speaking (hearing) populations. These modality-driven differences in production can influence cognition; that is, *thinking for speaking* differs from *thinking for signing*. Lastly, these patterns observed across different speaking (hearing) and signing (deaf) populations manifest themselves in bimodal bilinguals (hearing), in form of cross-linguistic influence between speech and sign across modalities, which changes their message conceptualization compared to that of different populations. By showing that not only using the vocal but also using the visuo-spatial modality affects bilingual diversity and cognition I hope to have provided an avenue to increase awareness of the overall acceptance and equality of sign languages to those of spoken languages in (language) research in general.

References

Ahn, D., Gollan, T. H., & Ferreira, V. S. (2019). Not keeping another language in mind: How bilinguals represent sentence structure in two languages. *Poster Presented at Psychonomic Society's 60th Annual Meeting*. Presented at the Montréal, Québec, Canada. Montréal, Québec, Canada.

Aikhenvald, A. (2006). Classifiers and noun classes: Semantics. In K. Brown, *Encyclopedia of language & linguistics* (pp. 463–471). Oxford, UK: Elsevier. doi: 10.1016/B0-08-044854-2/01111-1

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439. doi: https://doi.org/10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition*, *111*(1), 55–71. doi: 10.1016/j.cognition.2008.12.005

Anderson, D., & Reilly, J. (2002). The MacArthur Communicative Development Inventory: Normative data for American Sign Language. *Journal of Deaf Studies and Deaf Education*, *7*(2), 83–106. doi: 10.1093/deafed/7.2.83

Arik, E. (2013). Space, time, and iconicity in Turkish Sign Language (TID). *Trames. Journal of the Humanities and Social Sciences*, *16*(4), 345–358. doi: 10.3176/tr.2012.4.03

Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2008). Language is shaped by the body. *Behavioral and Brain Sciences*, *31*(5), 509–511. doi: 10.1017/S0140525X08005001

Assaneo, M. F., Nichols, J. I., & Trevisan, M. A. (2011). The anatomy of onomatopoeia. *PLOS ONE*, *6*(12), e28317. doi: 10.1371/journal.pone.0028317

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., … Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, *26*(4), 518–526.

Azar, Z. (2020). *Effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands* (Unpublished doctoral dissertation). Max Planck Institute for Psycholinguistics.

Azar, Z., Özyürek, A., & Backus, A. (2019). Turkish-Dutch bilinguals maintain language-specific reference tracking strategies in elicited narratives. *International Journal of Bilingualism*, *24*(2), 376–409. doi: 10.1177/1367006919838375

Bååth, R. (2014). *BayesianFirstAid: A package that implements bayesian alternatives to the classical *.test functions in R*.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi: 10.1016/j.jml.2007.12.005

Backus, A. (2005). Codeswitching and language change: One thing leads to another? *International Journal of Bilingualism*, *9*(3–4), 307–340. doi: 10.1177/13670069050090030101

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474. doi: 10.1016/j.jml.2007.09.002

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281–1289. doi: 10.1098/rstb.2008.0319

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). doi: 10.18637/jss.v067.i01

Bellugi, U., & Klima, E. S. (1976). Two faces of sign: Iconic and abstract. *Annals of the New York Academy of Sciences*, *280*(1 Origins and E), 514–538. doi: 10.1111/j.1749-6632.1976.tb25514.x

Blumenfeld, H. K., & Marian, V. (2005). Covert bilingual language activation through cognate word processing: An eye-tracking study. *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society*, 286–291. Mahwah, NJ: Erlbaum.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, *22*(5), 633–660. doi: 10.1080/01690960601000746

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*(5), 547–567. doi: 10.1080/20445911.2013.812093

Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, *5*, 341–345.

Bosworth, R. G., & Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in American Sign Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1573–1581. doi: 10.1037/a0020934

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, *30*(2), 225–251. doi: 10.1017/S0272263108080327

Brown, A., & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of Path among Japanese learners of English. *Bilingualism: Language and Cognition*, *14*(1), 79–94. doi: 10.1017/S1366728910000064

Brown, A. L. (1994). The advancement of learning. *Educational Researcher*, *23*(8), 4–12. doi: 10.3102/0013189X023008004

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: implications for language production and attention. *Language, Cognition and Neuroscience*, *31*(8), 1015–1037. doi: 10.1080/23273798.2016.1175649

Campbell, R., Martin, P., & White, T. (1992). Forced choice recognition of sign in novice British Sign Language learners. *Applied Linguistics*, *13*(2), 185–201. doi: 10.1093/applin/13.2.185

Casey, S., & Emmorey, K. (2009). Co-speech gesture in bimodal bilinguals. *Language and Cognitive Processes*, *24*(2), 290–312. doi: 10.1080/01690960801916188

Casey, S., Emmorey, K., & Larrabee, H. (2012). The effects of learning American Sign Language on co-speech gesture. *Bilingualism: Language and Cognition*, *15*(04), 677–686. doi: 10.1017/S1366728911000575

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620539

Cleland, A., & Pickering, M. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*(2), 214–230. doi: 10.1016/S0749-596X(03)00060-3

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107. doi: 10.1016/0010-0285(74)90005-X

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* (Unpublished doctoral dissertation). McGill University, Montreal, Canada.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. de Groot, *Handbook of bilingualism: Psycholinguistic approaches.* (pp. 289–307). NY, US: Oxford University Press.

Cuxac, C. (1996). *Fonctions et structures de l'iconicité des langues des signes* (Unpublished doctoral dissertation). Université René Descartes, Paris V.

Daller, M. H., Treffers-Daller, J., & Furman, R. (2011). Transfer of conceptualization patterns in bilinguals: The construal of motion events in Turkish and German. *Bilingualism: Language and Cognition*, *14*(1), 95–119. Cambridge Core. doi: 10.1017/S1366728910000106

Daller, M. H., Yıldız, C., de Jong, N. H., Kan, S., & Başbaĝi, R. (2011). Language dominance in Turkish-German bilinguals: methodological aspects of measurements in structurally different languages. *International Journal of Bilingualism*, *15*(2), 215–236. doi: 10.1177/1367006910381197

Davies, C., & Kreysa, H. (2017). Looking at a contrast object before speaking boosts referential informativeness, but is not essential. *Acta Psychologica*, *178*(Supplement C), 87–99. doi: 10.1016/j.actpsy.2017.06.001

Davies, C., & Kreysa, H. (2018). Look before you speak: children's integration of visual information into informative referring expressions. *Journal of Child Language*, *45*(5), 1116–1143. doi: 10.1017/S0305000918000120

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390. doi: 10.3758/BRM.41.2.385

De Quadros, R. Mü. (2018). Bimodal bilingual heritage signers: A balancing act of languages and modalities. *Sign Language Studies*, *18*(3), 355–384. doi: 10.1353/sls.2018.0007

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot, *Handbook of bilingualism: Psycholinguistic approaches.* (pp. 179–201). NY, US: Oxford University Press.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, *6*(10), 654–672. doi: 10.1002/lnc3.361

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615. doi: 10.1016/j.tics.2015.07.013

Dingemanse, M., Perlman, M., & Perniss, P. (2020). Construals of iconicity: experimental approaches to form–meaning resemblances in language. *Language and Cognition*, *12*(1), 1–14. doi: 10.1017/langcog.2019.48

Emmorey, K. (1996). The confluence of space and language in signed languages. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 171–209). Cambridge, MA: MIT Press.

Emmorey, K. (2002). The effects of modality on spatial language: How signers and speakers talk about space. In D. Quinto-Pozos, K. Cormier, & R. P. Meier (Eds.), *Modality and structure in signed and spoken languages* (pp. 405–421). New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511486777.019

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1651), 20130301. (25092669). doi: 10.1098/rstb.2013.0301

Emmorey, K., Borinstein, H. B., Thompson, R., & Gollan, T. H. (2008). Bimodal bilingualism. *Bilingualism: Language and Cognition*, *11*(01), 43–61. doi: 10.1017/S1366728907003203

Emmorey, K., Borinstein, H., & Thompson, R. (2005). Bimodal bilingualism: Code-blending between spoken English and American Sign Language. In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 663–673). Somerville, MA: Cascadilla Press.

Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In *Perspectives on classifier constructions in signed languages* (pp. 222–246). Mahwah, NJ: Lawrence Erlbaum.

Emmorey, K., McCullough, S., Mehta, S., Ponto, L. B., & Grabowski, T., J. (2013). The biology of linguistic expression impacts neural correlates for spatial language. *Journal of Cognitive Neuroscience*, *25*(4), 517–533. doi: 10.1162/jocn_a_00339

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: Signum Press. Retrieved from https://books.google.nl/books?id = LMrxAAAAMAAJ

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences*, *32*(5), 429–448. doi: 10.1017/S0140525X0999094X

Ferrara, L., & Halvorsen, R. P. (2017). Depicting and describing meanings with iconic signs in Norwegian Sign Language. *Gesture*, *16*(3), 371–395. doi: 10.1075/gest.00001.fer

Ferrara, L., & Hodge, G. (2018). Language as description, indication, and depiction. *Frontiers in Psychology*, *9*, 716. doi: 10.3389/fpsyg.2018.00716

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, *69*(3), 165–182. doi: 10.1016/j.jml.2013.06.001

Flecken, M., Carroll, M., Weimar, K., & Stutterheim, C. V. (2015). Driving along the road or heading for the village? Conceptual differences underlying motion event encoding in French, German, and French–German L2 users. *The Modern Language Journal*, *99*(S1), 100–122. doi: 10.1111/j.1540-4781.2015.12181.x

Flecken, M., Von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: findings from a cross- linguistic non-verbal recognition task. *Language and Cognition*, *6*(01), 45–78. doi: 10.1017/langcog.2013.2

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003). *Language in mind: advances in the study of language and thought*. Cambridge, Mass: MIT Press.

Giezen, M. R., Blumenfeld, H. K., Shook, A., Marian, V., & Emmorey, K. (2015). Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition*, *141*, 9–25. doi: 10.1016/j.cognition.2015.04.009

Giezen, M. R., & Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. *Bilingualism: Language and Cognition*, *19*(2), 264–276. doi: 10.1017/S1366728915000097

Gleitman, L., January, D., Nappa, R., & Trueswell, J., C. (2007). On the give and take between event apprehension and utterance formulation. *Language-Vision Interaction*, *57*(4), 544–569. doi: 10.1016/j.jml.2007.01.007

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*(27), 9163. doi: 10.1073/pnas.0710060105

Goller, F., Lee, D., Ansorge, U., & Choi, S. (2017). Effects of language background on gaze behavior: A crosslinguistic comparison between Korean and German speakers. *Advances in Cognitive Psychology*, *13*(4), 267–279. doi: 10.5709/acp-0227-z

Griffin, Z. M. (2004). Why look? Reasons for speech-related eye movements. In J. Henderson & F. Ferreira, *The integration of language, vision, and action: Eye movements and the visual world* (pp. 213–248). New York, NY: Psychology Press. Scopus.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279. doi: https://doi.org/10.1111/1467-9280.00255

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Bilingualism and Neurolinguistics*, *36*(1), 3–15. doi: 10.1016/0093-934X(89)90048-5

Grote, K., & Linz, E. (2003). The influence of sign language iconicity on semantic conceptualization. In W. G. Müller & O. Fisher (Eds.), *From sign to signing: iconicity in language and literature* (pp. 23–40). Amsterdam/Philadelphia: John Benjamins.

Gu, Y., Zheng, Y., & Swerts, M. (2018). Having a different pointing of view about the future: The effect of signs on co-speech gestures about time in Mandarin–CSL bimodal bilinguals. *Bilingualism: Language and Cognition*, *22*(4), 1–12. doi: 10.1017/S1366728918000652

Gundel, J. K. (1985). 'Shared knowledge' and topicality. *Journal of Pragmatics*, *9*(1), 83–107. doi: 10.1016/0378-2166(85)90049-9

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. (2015). Syntactic priming in American Sign Language. *PloS One*, *10*(3), e0119611–e0119611. PubMed (25786230). doi: 10.1371/journal.pone.0119611

Hartsuiker, R., Kolk, H., & Huiskamp, P. (1999). Priming word order in sentence production. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*, *52*(1), 129–147. doi: 10.1080/027249899391250

Hatzidaki, A., Branigan, H. P., & Pickering, M. J. (2011). Co-activation of syntax in bilingual language production. *Cognitive Psychology*, *62*(2), 123–150. doi: 10.1016/j.cogpsych.2010.10.002

Herman, R., Grove, N., Holmes, S., Morgan, G., Sutherland, H., & Woll, B. (2004). *Assessing BSL development: Production test*. City University Publication.

Hintz, F., Meyer, A. S., & Huettig, F. (2019). Visual context constrains language-mediated anticipatory eye movements. *Quarterly Journal of Experimental Psychology*, *73*(3), 458–467. doi: 10.1177/1747021819881615

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23–B32. doi: 10.1016/j.cognition.2004.10.003

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. doi: 10.1016/j.actpsy.2010.11.003

Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D. J., Padden, C., Rathmann, C., & Smith, S. R. (2012). Language acquisition for deaf children: Reducing the harms of zero tolerance to the use of alternative approaches. *Harm Reduction Journal*, *9*, 16–16. PubMed (22472091). doi: 10.1186/1477-7517-9-16

Indefrey, P., şAhin, H., & Gullberg, M. (2017). The expression of spatial relationships in Turkish–Dutch bilinguals. *Bilingualism: Language and Cognition*, *20*(03), 473–493. doi: 10.1017/S1366728915000875

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. *Language and Space.*, 1–30.

Jarvis, S. (2011). Conceptual transfer: Crosslinguistic effects in categorization and construal. *Bilingualism: Language and Cognition*, *14*(1), 1–8. Cambridge Core. doi: 10.1017/S1366728910000155

Karadölller, D. Z., Sümer, B., & Özyürek, A. (in press). Effects and non-effects of late language exposure on spatial language development: Evidence from deaf adults and children. *Language Learning and Development*.

Kellerman, E. (1979). Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition*, *2*(1), 37–57. Cambridge Core. doi: 10.1017/S0272263100000942

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, *139*(4), 638–653. doi: 10.1037/a0020507

Kimmelman, V. (2012). Word order in Russian Sign Language. *Sign Language Studies*, *12*(3), 414–445. doi: 10.1353/sls.2012.0001

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32. doi: 10.1016/S0749-596X(02)00505-3

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40. doi: 10.1016/j.cogpsych.2014.04.001

Kootstra, G., van Hell, J., & Dijkstra, T. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency. *Bilingualism: Language and Cognition*, *15*, 1–23. doi: 10.1017/S136672891100068X

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: the case of grammatical gender. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 843–858. doi: 10.1037/0278-7393.34.4.843

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(2), 119–135. doi: 10.1017/S1366728906002483

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Chapter seven - Juggling two languages in one mind: What bilinguals tell us about Language processing and its consequences for cognition. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 56, pp. 229–262). Academic Press. doi: 10.1016/B978-0-12-394393-4.00007-8

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In V. Ferreira, M. Goldrick, & M. Miozzo, *The oxford handbook of language production.* (pp. 165–181). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199735471.013.001

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. doi: 10.18637/jss.v082.i13

Lai, V. T., Rodriguez, G. G., & Narasimhan, B. (2014). Thinking-for-speaking in early and late bilinguals. *Bilingualism: Language and Cognition*, *17*(01), 139–152. doi: 10.1017/S1366728913000151

Laudanna, A., & Volterra, V. (1991). Order of words, signs, and gestures: A first comparison. *Applied Psycholinguistics*, *12*(2), 135–150. doi: 10.1017/S0142716400009115

Length, R. (2019). *emmeans: Estimated marginal means, aka least-squares means*. Retrieved from https://CRAN.R-project.org/package = emmeans

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA, US: The MIT Press.

Levinson, S., C. (1996). Language and space. *Annual Review of Anthropology*, *25*(1), 353–382. doi: https://doi.org/10.1146/annurev.anthro.25.1.353

Levinson, S. C. (2003). *Space in language and cognition: explorations in cognitive diversity*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511613609

Levinson, S. C., & Wilkins, D. (Eds.). (2006). *Grammars of space: explorations in cognitive diversity*. Cambridge, UK ; New York: Cambridge University Press.

Liddell, S. K. (1995). Real, surrogate and token space: Grammatical consequences in ASL. In K. Emmorey & J. S. Reilly, *Language, gesture, and space* (pp. 19–41). Hillsdale, NJ: Lawrence Erlbaum Associates.

Liddell, S. K. (1998). Grounded blends, gestures, and conceptual shifts. *Cognitive Linguistics*, *9*(3), 283–314.

Liddell, Scott K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press. Cambridge Core. doi: 10.1017/CBO9780511615054

Lieberman, A. M., Borovsky, A., Hatrak, M., & Mayberry, R. I. (2015). Real-time processing of ASL signs: Delayed first language acquisition affects organization of the mental lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1130–1139. doi: 10.1037/xlm0000088

Lieberman, A. M., Borovsky, A., & Mayberry, R. I. (2017). Prediction in a visual language: real-time sentence processing in American Sign Language across development. *Language, Cognition and Neuroscience*, *0*(0), 1–15. doi: 10.1080/23273798.2017.1411961

Lillo-Martin, D., de Quadros, R. M., Chen Pichler, D., & Fieldsteel, Z. (2014). Language choice in bimodal bilingual development. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01163

Lillo-Martin, D., de Quadros, R. M., & Pichler, D. C. (2016). The development of bimodal bilingualism: Implications for linguistic theory. *Linguistic Approaches to Bilingualism*, *6*(6), 719–755. (28603576). doi: 10.1075/lab.6.6.01lil

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, *41*(5), 791–824. doi: 10.1515/ling.2003.026

MacSwan, J. (2000). The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Bilingualism: Language and Cognition*, *3*(1), 37–54. Cambridge Core. Retrieved from Cambridge Core.

MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C. R., … Brammer, M. J. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain*, *125*(7), 1583–1593. doi: 10.1093/brain/awf153

Macwhinney, B. (1977). Starting points. *Language*, *53*(1), 152–168. doi: 10.2307/413059

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–380. doi: 10.3758/BF03206780

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago, IL, US: University of Chicago Press.

Meier, R. P. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In R. P. Meier, K. Cormier, & D. Quinto-Pozos, *Modality and structure in signed and spoken languages.* (pp. 1–25). New York, NY, US: Cambridge University Press. doi: 10.1017/CBO9780511486777.001

Mercure, E., Evans, S., Pirazzoli, L., Goldberg, L., Bowden-Howl, H., Coulson-Thaker, K., … MacSweeney, M. (2020). Language experience impacts brain activation for spoken and signed language in infancy: Insights from unimodal and bimodal bilinguals. *Neurobiology of Language*, *1*(1), 9–32. doi: 10.1162/nol_a_00001

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition*, *66*(2), B25–B33. doi: 10.1016/S0010-0277(98)00009-2

Michael, L. (2014). Social dimensions of language change. In C. Bowern & B. Evans, *The routledge handbook of historical linguistics* (pp. 484–502). Abingdon: Routledge. doi: 10.4324/9781315794013.ch22

Morgan, G., Herman, R., Barrière, I., & Woll, B. (2008). The onset and mastery of spatial language in children acquiring British Sign Language. *Cognitive Development*, *23*(1), 1–19. doi: 10.1016/j.cogdev.2007.09.003

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge, UK: Cambridge University Press.

Navarrete, E., Peressotti, F., Lerose, L., & Miozzo, M. (2017). Activation cascading in sign production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 302–318. doi: 10.1037/xlm0000312

Orlansky, M. D., & Bonvillian, J. D. (1984). The role of inconicity in early sign language acquisition. *The Journal of Speech and Hearing Disorders*, *49*(3), 287–292. doi: 10.1044/jshd.4903.287

Ortega, G., Sümer, B., & Özyürek, A. (2017). Type of iconicity matters in the vocabulary development of signing children. *Developmental Psychology*, *53*(1), 89–99. doi: 10.1037/dev0000161

Özyürek, A., & Woll, B. (2019). Language in the Visual Modality: Co- speech Gesture and Sign Language. In P. Hagoort, *Human language: From genes and brain to behavior* (pp. 67–83). Cambridge, MA: MIT Press.

Özyürek, A., Zwitserlood, I., & Perniss, P. (2010). Locative expressions in signed languages: a view from Turkish Sign Language (TİD). *Linguistics*, *48*(5), 1111–1145. doi: 10.1515/ling.2010.036

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155–184. doi: 10.1016/j.cognition.2008.02.007

Pavlenko, A., & Jarvis, S. (2002). Bidirectional transfer. *Applied Linguistics*, *23*(2), 190–214. doi: 10.1093/applin/23.2.190

Pavlenko, Aneta, & Malt, B. C. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, *14*(1), 19–45. Cambridge Core. doi: 10.1017/S136672891000026X

Perniss, P. (2007). *Space and iconicity in German Sign Language (DGS)* (Unpublished doctoral dissertation). Max Planck Institute for Psycholinguistics, The Netherlands.

Perniss, P., Özyürek, A., & Morgan, G. (2015). The influence of the visual modality on language structure and conventionalization: Insights from sign language and gesture. *Topics in Cognitive Science*, *7*(1), 2–11. doi: 10.1111/tops.12127

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 227. doi: 10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 20130300. doi: 10.1098/rstb.2013.0300

Perniss, P., Zwitserlood, I., & Özyürek, A. (2015). Does space structure spatial language?: A comparison of spatial expression across sign languages. *Language*, *91*(3), 611–641. doi: http://dx.doi.org/10.1353/lan.2015.0041

Pichler, D. C., Lillo-Martin, D., & Palmer, J. L. (2018). A short introduction to heritage signers. *Sign Language Studies*, *18*(3), 309–327. doi: 10.1353/sls.2018.0005

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, *134*(3), 427–459. (18444704). doi: 10.1037/0033-2909.134.3.427

Pinker, S. (1995). *The language instinct: The new science of language and mind*. Penguin Adult.

Poizner, H., Bellugi, U., & Tweney, R. D. (1981). Processing of formational, semantic, and iconic information in American Sign Language. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1146–1159. doi: 10.1037/0096-1523.7.5.1146

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, *6*, 40–71.

Pyers, J. E., & Emmorey, K. (2008). The face of bimodal bilingualism: Grammatical markers in American Sign Language are produced when bilinguals speak to English monolinguals. *Psychological Science*, *19*(6), 531–535. doi: 10.1111/j.1467-9280.2008.02119.x

Quadros, M. R., & Lillo-Martin, D. (2018). Brazilian bimodal bilinguals as heritage signers. *Languages*, *3*(3), 32. doi: 10.3390/languages3030032

Rubin, E. (1958). Figure-ground perception. In D. C. Beardslee & M. Wertheimer, *Readings in perception* (pp. 194–203). Princeton, NJ: Van Nostrand.

Russo, T. (2004). Iconicity and Productivity in Sign Language Discourse: An Analysis of Three LIS Discourse Registers. *Sign Language Studies*, *4*(2), I64–I97. doi: 10.1353/sls.2004.0006

Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*(3), 485–501. doi: 10.1037/0096-3445.136.3.485

Salamoura, A., & Williams, J. N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual lexicon. *Applied Psycholinguistics*, *28*(4), 627–660. doi: 10.1017/S0142716407070348

Sallandre, M. A., & Cuxac, C. (2002). Iconicity in sign language: A theoretical and methodological point of view. In I. Wachsmuth & T. Sowa, *Gesture-based communication in human-computer interaction* (pp. 171–180). Berlin: Springer Verlag.

Sapir, E. (1912). Language and environment. *American Anthropologist*, *14*(2), 226–242.

Sauter, M., Uttal, D. H., Alman, A. S., Goldin-Meadow, S., & Levine, S. C. (2012). Learning what children know about space from looking at their hands: The added value of gesture in spatial communication. *Journal of Experimental Child Psychology*, *111*(4), 587–606. doi: 10.1016/j.jecp.2011.11.009

Schembri, A. C. (2003). Rethinking “classifiers” in signed languages. In K. Emmorey, *Classifier constructions in signed languages* (pp. 3–34). Mahwah, NJ:Lawrence Erlbaum Associates.

Schembri, A. C., Jones, C., & Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners’ gestures without speech. *Journal of Deaf Studies and Deaf Education*, *10*(3), 272–290. doi: https://doi.org/10.1093/deafed/eni029

Schmitt, B. H., & Zhang, S. (1998). Language structure and categorization: A study of classifiers in consumer cognition, judgment, and choice. *Journal of Consumer Research*, *25*(2), 108–122. doi: 10.1086/209530

Schoonbaert, S., Hartsuiker, R., & Pickering, M. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, *56*(2), 153–171. doi: 10.1016/j.jml.2006.10.002

Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, *131*(3), 431–436. doi: 10.1016/j.cognition.2014.03.004

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician*, *34*(4), 216–221. doi: 10.1080/00031305.1980.10483031

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147. doi: 10.1016/s0010-0277(99)00025-6

Sherman Wilcox. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language. *Cognitive Linguistics*, *15*(2), 119–147. doi: 10.1515/cogl.2004.005

Shook, A., & Marian, V. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition*, *124*(3), 314–324. doi: 10.1016/j.cognition.2012.05.014

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. *Language in Mind: Advances in the Study of Language and Thought*, *157*(1), 178.

Slonimska, A., Özyürek, A., & Capirci, O. (2020). The role of iconicity and simultaneity for efficient communication: The case of Italian Sign Language (LIS). *Cognition*, *200*, 104246. doi: 10.1016/j.cognition.2020.104246

Sorensen, D. W., & Bailey, K. G. D. (2007). The world is too much: Effects of array size on the link between language comprehension and eye movements. *Visual Cognition*, *15*(1), 112–115.

Speed, L., Chen, J., Huettig, F., & Majid, A. (in press). Classifier categories reflect, but do not affect conceptual organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Stokoe, W. C., Jr. (2005). Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. *The Journal of Deaf Studies and Deaf Education*, *10*(1), 3–37. doi: 10.1093/deafed/eni001

Sümer, B. (2015). *Acquisition of spatial language by signing and speaking children: a comparison of Turkish Sign Language (TİD) and Turkish* (Unpublished doctoral dissertation). Radboud University Nijmegen, The Netherlands.

Sümer, B., Perniss, P., Zwitserlood, I., & Özyürek, A. (2014). Learning to express left-right & front-behind in a sign versus spoken language. In P. Bello, M. Guarini, & M. McShane, *Proceedings of the 36th annual meeting of the Cognitive Science Society* (pp. 1550–1555). Austin,Tx : Cognitive Science Society.

Talmy, L. (1978). The relation of grammar to cognition: a synopsis. *Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing*, 14–24. Association for Computational Linguistics.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen, *Language typology and syntactic description,Vol. 3: Grammatical categories and the lexicon* (Vol. 3, pp. 225–282). Cambridge, UK: Cambridge University Press.

Talmy, L. (2003). The representation of spatial structure in spoken and signed language. In K. Emmorey, *Perspectives on classifier constructions in sign languages* (pp. 169–195). Mahwah, NJ: Erlbaum.

Talmy, Len. (2003). The Representation of Spatial Strucutre in Spoken and Signed Language. In *Perspectives on Classifier Constructions in Sign Language*.

Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movements and spoken language comprehension. In M. Traxler & M. A. Gernsbacher, *Handbook of psycholinguistics* (pp. 863–900). Oxford, UK: Elsevier. doi: 10.1016/B978-012369374-7/50023-7

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. doi: 10.1126/science.7777863

Taub, S. F. (2000). Iconicity in American Sign Language: Concrete and metaphorical applications. *Spatial Cognition and Computation*, *2*(1), 31–50. doi: 10.1023/A:1011440231221

Thompson, R. L., Vinson, D. P., Fox, N., & Vigliocco, G. (2013). Is lexical access driven by temporal order or perceptual salience? Evidence from British Sign Language. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*, 1450–1455.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 550–557. doi: 10.1037/a0014547

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British sign language: effects of iconicity for phonological decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 1017. doi: 10.1037/a0019339

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, *23*(12), 1443–1448. doi: 10.1177/0956797612459763

Torchiano, M. (2019). *Effsize: Efficient effect size computation* [R package version 0.7.6]. Retrieved from https://CRAN.R-project.org/package = effsize

Torres Cacoullos, R., & Travis, C. E. (2016). Two languages, one effect: Structural priming in spontaneous code-switching. *Bilingualism: Language and Cognition*, *19*(4), 733–753. doi: 10.1017/S1366728914000406

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64–82. doi: 10.1016/j.jml.2010.02.006

Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations: Language and cognition. *Language Learning*, *66*(3), 554–580. doi: 10.1111/lang.12188

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari, *Lifespan perspectives on bilingualism* (pp. 136–196). Mouton de Gruyter/APA.

van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing “easy” and “hard” events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*(1), 124–144. doi: 10.1016/j.jml.2013.11.001

Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130292. doi: 10.1098/rstb.2013.0292

Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W. ., & Woll, B. (2005). Language and imagery: Effects of language modality. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1574), 1859–1863. doi: 10.1098/rspb.2005.3169

Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language. *Journal of Memory and Language*, *82*, 56–85. doi: 10.1016/j.jml.2015.03.002

Visker, E. . (2020). *Learning to use space: A study into the SL2 acquisition process of adult learners of sign language of the Netherlands* (Unpublished doctoral dissertation). LOT.

Weber, K., & Indefrey, P. (2009). Syntactic priming in German–English bilinguals during sentence comprehension. *NeuroImage*, *46*(4), 1164–1172. doi: 10.1016/j.neuroimage.2009.03.040

Weisberg, J., Casey, S., Sevcikova Sehyr, Z., & Emmorey, K. (2020). Second language acquisition of American Sign Language influences co-speech gesture production. *Bilingualism: Language and Cognition*, *23*(3), 473–482. doi: 10.1017/S1366728919000208

Whorf, B. L. (1997). The relation of habitual thought and behavior to language. In N. Coupland & A. Jaworski (Eds.), *Sociolinguistics: A Reader* (pp. 443–463). London: Macmillan Education UK. doi: 10.1007/978-1-349-25582-5_35

Wilkinson, E., & Morford, J. P. (2020). How bilingualism contributes to healthy development in deaf children: A public health perspective. *Maternal and Child Health Journal*. doi: 10.1007/s10995-020-02976-6

Wilson, K. M. (1999). Validity of global self-ratings of ESL speaking proficiency based on an FSI/ILR-referenced scale. *ETS Research Report Series*, *1999*(2), i–64. doi: 10.1002/j.2333-8504.1999.tb01811.x

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 1556–15594). Genoa: European Language Resources Association.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265. doi: 10.1002/wcs.104

Wolter, B., Yamashita, J., & Leung, C. Y. (2020). Conceptual transfer and lexical development in adjectives of space: Evidence from judgments, reaction times, and eye tracking. *Applied Psycholinguistics*, *(advanced online publication)*, 1–31. doi: 10.1017/S0142716420000107

Zwitserlood, I. (2003). *Classifying hand configurations in Nederlandse Gebarentaal (Sign Language of the Netherlands) /* (Unpublished doctoral dissertation). University of Utrecht, The Netherlands.

Zwitserlood, I. (2012). Classifiers. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 158–186). Berlin, Germany: Mouton de Gruyter.

Zwitserlood, I., Perniss, P., & Özyürek, A. (2012). An empirical investigation of expression of multiple entities in Turkish Sign Language (TİD): Considering the effects of modality. *Lingua*, *122*(14), 1636–1667. doi: 10.1016/j.lingua.2012.08.010

English summary

Our ability to communicate with each other using language is at the heart of what makes us human. However, languages around the world differ in their sounds, words and grammar. Consequently, there are sometimes big differences between languages in the meanings that they express. This diversity has motivated researchers to explore whether language diversity causes speakers of different languages to see the world with different eyes.

Especially between spoken and sign languages, there are fundamental differences in how meaning is expressed. Sign languages express meaning in a different modality than spoken languages: not with the vocal cords but with the visible parts of the body, using hand, face and body movements. These real and complex languages are used by deaf communities. Does this language diversity in modality (vocal versus visual) result in that hearing speakers and deaf signers look at the world differently? This is a question that no one has yet investigated before. Furthermore, previous studies have mainly asked whether groups of people that speak different languages look at the world differently. But what happens if one person knows two different languages (such as a sign language and a spoken language)? How do they express meaning? That is, do they talk and sign differently because the two languages might influence each other and how does this affect the way they look at the world?

In this thesis, I addressed these open questions by comparing users of a spoken (Dutch) and a sign language (Sign Language of the Netherlands, NGT). I looked at how they express meaning in each language and how this affects the way they look at the world. I also investigated these same things in hearing bilinguals who can speak and sign, so-called bimodal bilinguals.

To do so, I studied descriptions of pictures that show two objects, such as pen and glass. These objects were in a spatial relation to each other: the pen was for example to the right of the glass. I chose this spatial domain because it has been previously shown that spoken and sign languages differ immensely from another when it comes to describing such relations. There are two main differences: (1) Due to using the hands and other visible parts of the body signers can depict objects and the spatial relations between them by mirroring these aspects with their hands and placing them in front of the body. This visual mirroring of spatial relations, which I call “iconicity” in this thesis, is not possible by using speech. (2) Signers of many sign languages use a consistent order in which they mention the two objects (e.g. pen and glass), that is, bigger objects such as the glass first and smaller objects such as the pen second. This seems to be driven by using the hands to describe spatial relations. In contrast, spoken languages do not use a consistent order in which they mention objects to describe these relations. Looking at these two differences, I created an experiment for which I have combined picture descriptions with an eye-tracking experiment. This way, I investigated possible relations between the way hearing speakers, deaf signers and bimodal bilinguals describe pictures of spatial relations and how they look at these pictures.

In **Chapter 2**, I was interested in how hearing speakers and deaf signers look at the world differently because their descriptions are non-iconic (speech) versus iconic (sign). In sign languages descriptions often visually mirror the spatial scene, but in spoken languages descriptions often contain non-iconic words like "left". Does this difference cause signers to look at these spatial relations differently than speakers?

To experimentally test this, hearing speakers and deaf signers were seated in front of a laptop and saw always 4 pictures on the screen. Each picture showed the same two objects (e.g. glass and pen) but always in different spatial relations (e.g. left/right/front/on). The participants’ task was to describe one of the pictures (indicated by an arrow) to a trained interlocutor whose hearing status matched that of the participant, that is, hearing speakers described pictures to a hearing speaking addressee and deaf signers described them to a deaf signing addressee. With a little camera in front of them, I tracked with infrared light where on the screen they look at. I wanted to see at which picture participants look and how often they look at it and compare this between hearing speakers and deaf signers based on the different strategies they describe those pictures in their own language.

I found that deaf signers, but not hearing speakers, looked more at all four pictures despite that they only had to describe one picture. So why did deaf signers look also at the other pictures that they did not had to describe? I argued that this is because signing about spatial relations might require extra attention to identify the spatial relation and/or the orientation, size and shape of the objects in all four pictures to mirror information using the hands. Thus, in this chapter, I have shown that deaf signers see the world with different eyes than hearing speakers due to the iconic descriptions they use when describing spatial relations.

In **Chapter 3**, I wanted to see how bimodal bilinguals express meaning. More concretely, I was interested in how a spoken and a sign language can influence each other within one person looking at the difference in how spatial relations are described across spoken and sign languages, as investigated in Chapter 2. From previous research that looked at bilinguals speaking two spoken languages we know that languages can influence each other within an individual, also called cross-linguistic influence. I tested for the first time whether languages from different modalities can influence each other by asking whether bimodal bilinguals talk and sign differently than their hearing speaking and deaf signing peers tested in Chapter 2. I used the same experiment as described in Chapter 2 but only looked at participants' picture descriptions, not their eye-movements. I compared bimodal bilinguals' descriptions in each of their languages to that of their hearing speaking and deaf signing peers.

I found that non-iconic expressions in speech and iconic expressions in sign shape each other in a way that bilinguals' descriptions differ from that of the hearing speaking and deaf signing controls. More specifically, as expressions in sign are more iconic and contain elaborate information about how objects are placed (e.g. horizontally, vertically) and shaped (e.g. round, long elongated), I found that this made speech more elaborate about these features as well. In return, using non-iconic forms in speech decreased iconicity in signed descriptions. Overall, with this chapter, I showed that bimodal bilinguals do no resemble two monolinguals in one but how they use each language is instead shaped by the language from the different modality.

In **Chapter 4**, I explored how bimodal bilinguals look at the world. Thus, does the fact that bimodal bilinguals talk and sign differently than their hearing speaking and deaf signing peers also cause bimodal bilinguals to see spatial relations with different eyes? Using the same experiment again I looked this time at bimodal bilinguals' eye-movements while they prepared those picture descriptions that I have explored in Chapter 3 and then compared those eye-movements to those of hearing speakers and deaf signers (similar to and reported in Chapter 2).

I found that when preparing to speak, bimodal bilinguals, but not hearing speakers, looked more at the other pictures they did not have to describe, similarly to deaf signers as shown in Chapter 2. Thus, iconic sign did not only make bimodal bilinguals' speech more elaborate but it also made them pay more attention to spatial relations. When preparing to sign, both bimodal bilinguals and deaf signers paid a lot of attention to those pictures they did not have to describe. However, signers paid more attention to those pictures than bimodal bilinguals. Thus, non-iconic speech did not only make sign descriptions less iconic but it also caused bimodal bilinguals to look at spatial relations differently than deaf signers. Overall, this chapter revealed that a spoken and sign language can influence each other not only in form of different spatial descriptions but it can also affect how bimodal bilinguals view these spatial scenes before they start describing them.

Finally**, Chapter 5**, I looked at descriptions and eye-movements together to explore how bimodal bilinguals express meaning in each language and how this affects the way they look at the world. This time, I looked at another difference in how spoken and sign languages describe spatial relations; that is, sign languages prefer to mention bigger objects first to describe spatial relations while spoken languages differ in their preference for which object is mentioned first. I used again the same experiment and looked at pictures descriptions as well as eye-movements and tested the same bimodal bilinguals as well as the hearing speaking and deaf signing controls as in Chapter 2-4. I looked at this order difference to make sure that the effects I found in the previous chapters are not only limited to iconicity but are broader effects that stem from the different modalities in which the languages are used.

I found that there is indeed a difference in respect to the order in which spoken and sign languages mention the small and bigger objects: deaf signers predominantly preferred mentioning the bigger objects first, not the small objects, while hearing speakers mentioned sometimes bigger objects first and sometimes smaller objects first. When bimodal bilinguals described the pictures using speech, they mentioned bigger objects first more often than their hearing speaking peers. However, when describing the pictures using sign they mentioned bigger objects first as much as their deaf signing peers. Thus in contrast to Chapter 3, I showed that sign influences speech in bimodal bilinguals, but there was no influence from speech to sign. This suggests that differences in modality can limit cross-linguistic influences.

Next, I also looked at eye-movements. In detail, I focused on only the picture the participant had to describe and compared eye-movements to the smaller and the bigger object. I expected that participants would look most to the object they would describe first, whether this would be big or small.

I found that deaf signers not only mention bigger objects first more often but they also look at them more often than hearing speakers. For bimodal bilinguals, they not only mentioned bigger objects first more often than hearing speakers when speaking but they also looked more at bigger objects when preparing to speak about the bigger object first. However, when preparing to sign bimodal bilinguals' eye-movement did not differ from those of the deaf signers. Thus, here, we found an influence from sign to speech, but not from speech to sign.

To summarize all these findings, I have shown that language diversity in modality (vocal versus visual) cause hearing speakers and deaf signers to look at the world differently. I have also shown that bilinguals who know both a spoken and a sign language use both languages differently from their hearing speaking and deaf signing peers, which in turn makes them look at the world with different eyes when preparing to speak or sign.

Nederlandse samenvatting

Ons vermogen om met elkaar te communiceren door middel van taal vormt de kern van wat ons menselijk maakt. Talen over de hele wereld verschillen echter in hun klanken, woorden en grammatica. Hierdoor zijn er soms grote verschillen tussen talen in de betekenis die ze uitdrukken. Deze diversiteit heeft onderzoekers gemotiveerd om te onderzoeken of taaldiversiteit ertoe leidt dat mensen die verschillende talen spreken de wereld ook met andere ogen zien.

Vooral tussen gesproken talen en gebarentalen zijn er fundamentele verschillen in de manier waarop betekenis wordt uitgedrukt. Gebarentalen drukken betekenis uit in een andere modaliteit dan gesproken talen: niet met de stembanden maar met de zichtbare delen van het lichaam, dus met hand-, gezichts- en lichaamsbewegingen. Deze echte en complexe talen worden gebruikt door dovengemeenschappen. Heeft deze taaldiversiteit in modaliteit (vocaal versus visueel) tot gevolg dat horende mensen die een gesproken taal spreken (bijvoorbeeld Nederlands) en dove mensen die een gebarentaal gebruiken (bijvoorbeeld Nederlandse Gebarentaal (NGT)) anders naar de wereld kijken? Dit is een vraag die nog niemand eerder heeft onderzocht. Bovendien hebben eerdere studies vooral gevraagd of groepen mensen die verschillende talen spreken anders naar de wereld kijken. Maar wat gebeurt er als iemand twee verschillende talen kent (zoals een gebarentaal en een gesproken taal)? Hoe drukken deze tweetaligen betekenis uit? Dat wil zeggen: praten en gebaren ze anders omdat de twee talen elkaar kunnen beïnvloeden en hoe beïnvloedt dit dan de manier waarop ze naar de wereld kijken?

In dit proefschrift heb ik deze open vragen behandeld door horende Nederlands-sprekers en dove NGT-gebaarders te vergelijken. Ik heb gekeken naar hoe ze betekenis uitdrukken in elke taal en hoe dit de manier waarop ze naar de wereld kijken beïnvloedt. Ik heb dezelfde dingen ook onderzocht bij zogenaamde bimodale tweetalige mensen die vloeiend Nederlands spreken en vloeiend gebaren in NGT.

Hiervoor heb ik beschrijvingen bestudeerd van afbeeldingen die twee objecten toonden, zoals pen en glas. Deze objecten stonden in een ruimtelijke relatie tot elkaar: de pen lag bijvoorbeeld rechts van het glas. Ik koos voor dit ruimtelijke domein omdat eerder is aangetoond dat gesproken talen en gebarentalen enorm verschillen als het gaat om het beschrijven van dergelijke ruimtelijke relaties. Er zijn twee belangrijke verschillen: (1) Door het gebruik van de handen en andere zichtbare delen van het lichaam kunnen gebaarders objecten en de ruimtelijke relaties daartussen weergeven door ze met hun handen te imiteren en voor het lichaam te plaatsen. Zulke visuele imitaties van ruimtelijke relaties, die ik in dit proefschrift 'iconiciteit' noem, is niet mogelijk met gesproken taal. (2) Gebaarders van veel gebarentalen gebruiken een consequente volgorde waarin ze twee objecten noemen (bijv. een pen en een glas): ze noemen grotere objecten zoals het glas eerst en kleinere objecten zoals de pen als tweede. Dit lijkt voort te komen uit het gebruik van de handen om ruimtelijke relaties te beschrijven. Gesproken talen daarentegen gebruiken verschillende volgorden waarin ze objecten noemen om de bijbehorende ruimtelijke relaties te beschrijven. Met deze twee verschillen in gedachten heb ik een experiment ontworpen waarin het beschrijven van afbeeldingen werd gecombineerd met eye-tracking. Zo heb ik mogelijke relaties onderzocht tussen hoe beelden van ruimtelijke relaties worden beschreven en bekeken door horende Nederlands-sprekers, dove NGT-gebaarders en horende bimodale tweetaligen die Nederlands en NGT beheersen.

In **hoofdstuk 2** was ik geïnteresseerd in of Nederlands-sprekers en NGT-gebaarders anders naar de wereld kijken, omdat hun beschrijvingen niet-iconisch (spraak) versus iconisch (gebaren) zijn. In gebarentaal gebruikt men iconische beschrijvingen die de ruimtelijke scène visueel imiteren, maar in gesproken taal gebruikt men niet-iconische woorden zoals "links". Zorgt dit verschil ervoor dat NGT- anders naar deze ruimtelijke relaties kijken dan Nederlands-sprekers?

Om dit experimenteel te testen, plaatste ik Nederlands-sprekers en NGT-gebaarders voor een laptop. De deelnemers zagen altijd 4 afbeeldingen op het scherm. Elke afbeelding liet dezelfde twee objecten zien (bijv. een glas en een pen), maar altijd in verschillende ruimtelijke relaties (bijv. links/rechts/voor/op). De taak van de deelnemers was om een van de plaatjes (aangegeven door een pijl) te beschrijven aan een getrainde gesprekspartner. De gehoorstatus van de gesprekspartner kwam altijd overeen met die van de deelnemer: horende Nederlands-sprekers beschreven plaatjes aan een horende Nederlands-sprekende gesprekspartner en dove NGT-gebaarders beschreven ze aan een dove NGT-gebarende gesprekspartner. Met een kleine camera op de laptop, volgde ik met infrarood licht waar op het scherm de deelnemers keken. Ik wilde zien naar welke foto de deelnemers keken en hoe vaak ze ernaar keken, en ik wilde dit vergelijken tussen Nederlands-sprekers en NGT-gebaarders op basis van de verschillende strategieën die ze gebruiken om deze foto's in hun eigen taal te beschrijven.

Ik ontdekte dat NGT-, maar niet Nederlands-sprekers, meer naar alle vier fotos keken, hoewel er maar één moest worden beschreven. Waarom keken NGT-gebarders naar de afbeeldingen die ze niet hoefden te beschrijven? Ik heb in mijn proefschrift betoogd dat dit komt doordat het gebaren over ruimtelijke relaties extra aandacht zou kunnen vragen om de ruimtelijke relatie en/of de oriëntatie, grootte en vorm van de objecten in alle vier de afbeeldingen te identificeren om informatie met de handen te imiteren. Samenvattend heb ik in dit hoofdstuk laten zien dat NGT-gebaarders de wereld met andere ogen zien dan Nederlands-sprekers vanwege de iconische beschrijvingen die ze gebruiken om ruimtelijke relaties te beschrijven.

In **hoofdstuk 3** wilde ik zien hoe bimodale tweetaligen betekenis uitdrukken. Ik was concreet geïnteresseerd in hoe een gesproken taal en een gebarentaal elkaar kunnen beïnvloeden binnen één persoon. Hiervoor onderzocht ik weer hetzelfde verschil in de beschrijving van ruimtelijke relaties tussen gesproken en gebarentaal als in hoofdstuk 2. Uit eerder onderzoek naar tweetaligen van twee gesproken talen weten we dat talen elkaar binnen één persoon kunnen beïnvloeden, ook wel cross-linguïstische invloed genoemd. Ik heb voor de eerste keer getest of talen uit verschillende modaliteiten elkaar kunnen beïnvloeden door te vragen of bimodale tweetaligen anders praten en gebaren dan de Nederlands-sprekers en NGT-gebaarders die in hoofdstuk 2 zijn getest. Ik heb hetzelfde experiment gebruikt als beschreven in hoofdstuk 2, maar heb alleen gekeken naar de beeldbeschrijvingen van de deelnemers, niet naar hun oogbewegingen. Ik vergeleek de beschrijvingen van bimodale tweetaligen in het Nederlands met die van de Nederlands-sprekers en in NGT met die van de NGT-gebaarders.

Ik ontdekte dat niet-iconische uitdrukkingen in spraak en iconische uitdrukkingen in gebarentaal elkaar zodanig vormen dat de beschrijvingen van bimodale tweetaligen verschillen van die van de Nederlands-sprekers en NGT-gebaarders. In het bijzonder, ik ontdekte dat aangezien uitdrukkingen in gebarentaal vaak iconischer zijn en gedetailleerdere informatie bevatten over hoe objecten worden geplaatst (bijv. horizontaal, verticaal) en zijn gevormd (bijv. rond, lang, langwerpig), maakte dit de spraak van de tweetaligen ook gedetailleerder over deze kenmerken. Andersom zorgde het gebruik van niet-iconische vormen in spraak voor een vermindering in het gebruik van iconische vormen in gebarentaal. In het algemeen heb ik met dit hoofdstuk laten zien dat bimodale tweetaligen niet twee eentaligen in één persoon zijn, maar dat hun taalgebruik in elke taal gevormd wordt door het taalgebruik in de andere taal (met de andere modaliteit).

In **hoofdstuk 4** heb ik onderzocht hoe bimodale tweetaligen naar de wereld kijken. Dus, veroorzaakt het feit dat bimodale tweetaligen anders praten en gebaren dan de Nederlands-sprekers en NGT-gebaarders ook dat bimodale tweetaligen ruimtelijke relaties zien met andere ogen? Met hetzelfde experiment onderzocht ik dit keer de oogbewegingen van bimodale tweetaligen terwijl ze de beeldbeschrijvingen planden die ik in hoofdstuk 3 heb onderzocht. Vervolgens heb ik deze oogbewegingen vergeleken met die van Nederlands-sprekers en NGT-gebaarders (vergelijkbaar met Hoofdstuk 2).

Ik ontdekte dat bij de voorbereiding om Nederlands te spreken, bimodale tweetaligen, maar niet Nederlands-sprekers, meer naar de andere plaatjes keken die ze niet hoefden te beschrijven, net als NGT-gebaarders deden in hoofdstuk 2. De invloed van iconische vormen in NGT zorgde er niet alleen voor dat de Nederlandse beschrijvingen van de bimodale tweetaligen uitgebreider werden, maar het zorgde er ook voor dat ze meer aandacht gingen besteden aan ruimtelijke relaties. Bij de voorbereiding om te gebaren, besteedden zowel bimodale tweetaligen als NGT-gebaarders veel aandacht aan de foto's die ze niet hoefden te beschrijven. De gebaarders besteedden echter nog meer aandacht aan die afbeeldingen dan bimodale tweetaligen. Dit betekent dat niet-iconische vormen in gesproken taal niet alleen beschrijvingen in gebarentaal minder iconisch maakten, maar ze zorgden er ook voor dat bimodale tweetaligen anders naar ruimtelijke relaties keken dan NGT-gebaarders. In het algemeen onthulde dit hoofdstuk dat een gesproken taal en gebarentaal elkaar niet alleen kunnen beïnvloeden in hoe ruimtelijke relaties worden beschreven, maar ook in hoe deze worden bekeken ter voorbereiding op het beschrijven.

Ten slotte heb ik in **hoofdstuk 5** naar beschrijvingen en oogbewegingen samen gekeken om te onderzoeken hoe bimodale tweetaligen betekenis uitdrukken in elke taal en hoe dit de manier waarop ze naar de wereld kijken beïnvloedt. Deze keer keek ik naar een ander verschil in de manier waarop gesproken talen en gebarentalen ruimtelijke relaties beschrijven; dat wil zeggen, gebarentalen geven er de voorkeur aan om eerst grotere objecten te noemen, terwijl gesproken talen verschillen in hun voorkeur voor welk object als eerste wordt genoemd. Ik gebruikte weer de beschrijvingen en oogbewegingen uit hetzelfde experiment en testte dezelfde bimodale tweetaligen, Nederlands-sprekers en NGT-gebaarders als in hoofdstuk 2-4. Ik heb naar dit verschil in volgorde gekeken om er zeker van te zijn dat de effecten die ik in de vorige hoofdstukken aantrof niet alleen beperkt zijn tot iconiciteit, maar dat het bredere effecten zijn die voortkomen uit de verschillende modaliteiten waarin de talen worden gebruikt.

Ik bevestigde dat er inderdaad een verschil is in de volgorde waarin gesproken talen en gebarentalen de kleine en grotere objecten noemen: NGT-gebaarders noemden vaak eerst de grotere objecten, terwijl Nederlands-sprekers soms eerst grotere objecten eerst noemden en soms eerst de kleineree. Toen bimodale tweetaligen de afbeeldingen in het Nederlands beschreven, noemden ze grotere objecten vaker eerst dan hun horende sprekende leeftijdsgenoten. Bij het beschrijven van de afbeeldingen in NGT noemden ze echter even vaak grotere objecten eerst als NGT-gebaarders. Ik liet dus zien dat voor bimodale tweetaligen de volgorde in gebarentaal de volgorde in gesproken taal beïnvloedt, maar andersom was er geen invloed van gesproken taal naar gebarentaal.

Vervolgens keek ik ook naar oogbewegingen. Specifiek onderzocht ik alleen de foto die de deelnemer moest beschrijven en vergeleek ik de oogbewegingen naar het kleinere en grotere object. Ik verwachtte dat de deelnemers het meest zouden kijken naar het object dat ze het eerst zouden beschrijven, of dit nu groot of klein was.

Ik ontdekte dat NGT-gebaarders niet alleen grotere objecten vaker als eerst noemen, maar er ook vaker naar kijken dan Nederlands-sprekers. Als de bimodale tweetaligen Nederlands spraken, dan noemden ze niet alleen grotere objecten vaker als eerst dan Nederlands-sprekers, maar keken ze ook meer naar grotere objecten terwijl ze zich voorbereidden om als eerst over die grotere objecten te praten. Tijdens de voorbereiding op het gebaren in NGT, verschilden de oogbewegingen van de bimodale tweetaligen echter niet van die van NGT-gebaarders. Dit alles betekent dat ik in dit hoofdstuk een invloed heb gevonden van gebarentaal naar gesproken taal, maar andersom niet van gesproken taal naar gebarentaal. Dus, taaldiversiteit in modaliteit kan cross-linguïstische invloed bemoeilijken.

Om alle resultaten uit de vier hoofdstukken samen te vatten, ik heb laten zien dat taaldiversiteit in modaliteit (vocaal versus visueel) ervoor zorgt dat horende gebruikers van een gesproken taal en dove gebruikers van een gebarentaal anders naar de wereld kijken. Ik heb ook aangetoond dat tweetaligen die zowel een gesproken taal als een gebarentaal kennen, beide talen anders gebruiken dan hun horende sprekende en dove gebarende leeftijdsgenoten, waardoor ze ook met andere ogen naar de wereld kijken terwijl ze zich voorbereiden om te spreken of te gebaren.

Deutsche Zusammenfassung

Unsere Fähigkeit mit Sprache miteinander zu kommunizieren ist das Herzstück dessen was uns menschlich macht. Sprachen auf der ganzen Welt unterscheiden sich jedoch in ihren Lauten, Wörtern und ihrer Grammatik. Folglich gibt es manchmal große Unterschiede zwischen Sprachen in den Bedeutungen, die sie ausdrücken. Diese Vielfalt hat Forschende motiviert zu untersuchen, ob Sprachvielfalt dazu führt, dass Personen mit unterschiedlichen Sprachen die Welt auch mit unterschiedlichen Augen sehen.

Besonders zwischen gesprochenen Sprachen und Gebärdensprachen gibt es grundlegende Unterschiede in der Art und Weise wie Bedeutung verwirklicht wird. Gebärdensprachen drücken die Bedeutung in einer anderen Modalität aus als Lautsprachen, also nicht mit dem vokalen Sprechapparat, sondern mit den sichtbaren Körperteilen, indem sie Hand-, Gesichts- und Körperbewegungen verwenden. Diese realen und komplexen Sprachen werden von gehörlosen Gemeinschaften verwendet. Führt diese sprachliche Andersartigkeit in der Modalität (vokal oder visuell) dazu, dass hörende Personen die eine Lautsprache sprechen (z. B. Niederländisch) und gehörlose Personen die eine Gebärdensprache sprechen (z.B. Niederländische Gebärdensprache (NGT)) die Welt unterschiedlich betrachten? Dies ist eine Frage, die bis jetzt noch wenig erforscht wurde. Darüber hinaus haben frühere Studien hauptsächlich gefragt, ob Menschen unterschiedlicher Sprachgemeinschaften die Welt unterschiedlich sehen. Aber was passiert, wenn eine einzelne Person zwei verschiedene Sprachen beherrscht (z. B. eine Gebärdensprache und eine Lautsprache)? Wie drücken diese Zweisprachigen Bedeutung aus, das heißt, sprechen und gebärden sie anders als Einsprachige, weil sich die beiden Sprachen gegenseitig beeinflussen könnten, und wie wirkt sich dies auf ihre Sichtweise auf die Welt aus?

In dieser Dissertation gehe ich auf diese offenen Fragen ein. Dazu habe ich hörende Niederländisch-Sprechende und gehörlose NGT-Gebärende dahingehend verglichen, wie sie in ihrer Sprache Bedeutung ausdrücken und wie sich dies auf ihre Sichtweise auf die Welt auswirkt. Ich habe diese Dinge in sogenannten bimodalen Zweisprachigen untersucht, die fließend Niederländisch sprechen und in NGT gebärden.

Zu diesem Zweck habe ich Beschreibungen von Bildern erforscht, die zwei Objekte wie zum Beispiel Stift und Glas zeigen, die in einem räumlichen Verhältnis zueinander stehen. Zum Beispiel: „der Stift liegt rechts neben dem Glas“. Ich habe mich für diesen räumlichen Bereich entschieden, weil andere Studien gezeigt haben, dass sich Lautsprachen und Gebärdensprachen bei der Beschreibung solcher Beziehungen erheblich voneinander unterscheiden. Es wurden zwei Hauptunterschiede vorgeschlagen: (1) Gebärdende können Objekte und ihre räumlichen Beziehungen darstellen, indem sie sie mit ihren Händen spiegeln und vor den Körper platzieren. Diese Art von visueller Spiegelung von räumlichen Beziehungen, die ich in dieser Arbeit als „Ikonizität“ bezeichne, ist mit Lautsprachen nicht möglich. (2) Gebärdende vieler Gebärdensprachen verwenden eine feste Reihenfolge, in der sie die beiden Objekte (z. B. Stift und Glas) erwähnen. Größere Objekte wie zum Beispiel das Glas werden in der Regel zuerst erwähnt und kleinere Objekte wie zum Beispiel den Stift später. Diese Reihenfolge scheint darauf zurückzuführen zu sein, dass mit den Händen räumliche Beziehungen beschrieben werden. Im Gegensatz dazu bevorzugen verschiedene Lautsprachen verschiedene Reihenfolgen in denen sie Objekte erwähnen um räumliche Beziehungen zu beschreiben. Mit Blick auf diese beiden Unterschiede habe ich Bildbeschreibungen mit einem Eye-Tracking-Experiment kombiniert, um mögliche Beziehungen zwischen der Beschreibung und Betrachtung von Bildern räumlicher Beziehungen bei hörenden Niederländisch-Sprechenden, gehörlosen NGT-Gebärdenden und bimodalen Zweisprachigen zu untersuchen.

In **Kapitel 2** war ich daran interessiert ob Niederländisch-Sprechende und NGT-Gebärdende die Welt unterschiedlich sehen, weil ihre Beschreibungen nicht ikonisch (Lautsprache) oder ikonisch (Gebärden) sind. Mit anderen Worten, führt die visuelle Spiegelung von räumlichen Szenen mit den Händen im Gegensatz zur Verwendung nicht-ikonischer Wörter wie "links" dazu, dass NGT-Gebärdende diese räumlichen Beziehungen anders betrachten als Niederländisch-Sprechende?

Um dies experimentell zu testen, habe ich Niederländisch-Sprechende und NGT-Gebärdende vor einen Laptop gesetzt. Die Teilnehmenden sahen immer 4 Bilder auf dem Bildschirm, die dieselben zwei Objekte (z. B. Glas und Stift) zeigten, jedoch immer in unterschiedlichen räumlichen Beziehungen (z. B. links/rechts/davor/auf). Die Aufgabe der Teilnehmenden bestand darin, einem heimlich geschulten Gegenüber eines der Bilder (durch einen Pfeil gekennzeichnet) zu beschreiben. Der Hörstatus des Gegenüber entsprach dem des Teilnehmenden: hörende Niederländisch-Sprechende beschrieben Bilder einem hörenden Niederländisch-sprechenden Gegenüber und gehörlose NGT-Gebärdende beschrieben sie einem gehörlosen NGT-gebärdenden Gegenüber. Mit einer kleinen Kamera verfolgte ich mit Infrarotlicht, wohin sie auf dem Bildschirm schauen. Ich wollte sehen, auf welches Bild die Teilnehmenden schauen und wie oft sie es betrachten und dies dann zwischen Niederländisch-Sprechenden und NGT-Gebärdenden vergleichen, basierend auf den verschiedenen Strategien, mit denen sie diese Bilder in ihrer eigenen Sprache beschreiben.

Ich stellte fest, dass NGT-Gebärdende, nicht jedoch Niederländisch-Sprechende, alle vier Bilder genauer betrachteten, obwohl nur eins davon beschrieben werden musste. Warum betrachteten NGT-Gebärdende die Bilder, die sie nicht beschreiben mussten? Ich argumentierte in meiner Dissertation, dass dies darauf zurückzuführen ist, dass das Gebärden über räumliche Beziehungen möglicherweise besondere Aufmerksamkeit erfordert, um die räumliche Beziehung und/oder die Ausrichtung, Größe und Form der Objekte in allen vier Bildern zu identifizieren und Informationen mit den Händen zu spiegeln. Zusammenfassend habe ich in diesem Kapitel gezeigt, dass NGT-Gebärdende die Welt aufgrund der ikonischen Beschreibungen, die sie bei der Beschreibung räumlicher Beziehungen verwenden, mit anderen Augen sehen als Niederländisch-Sprechende.

In **Kapitel 3** wollte ich sehen, wie bimodale Zweisprachige Bedeutung ausdrücken. Konkreter war ich daran interessiert, wie sich eine Laut- und eine Gebärdensprache innerhalb einer Person gegenseitig beeinflussen können. Dazu habe ich den gleichen Unterschied in der Beschreibung räumlicher Beziehungen zwischen Laut- und Gebärdensprachen untersucht, wie in Kapitel 2. Aus früheren Untersuchungen, die sich mit Zweisprachigen von Lautsprachen befassten, wissen wir, dass sich Sprachen innerhalb einer Person gegenseitig beeinflussen können, was auch als wechselseitiger Einfluss bezeichnet wird. Ich habe zum ersten Mal getestet, ob Sprachen aus verschiedenen Modalitäten sich gegenseitig beeinflussen können, indem ich gefragt habe, ob bimodale Zweisprachige anders sprechen und gebärden als die Niederländisch-Sprechenden und NGT-Gebärdenden aus Kapitel 2. Ich habe das gleiche Experiment wie in Kapitel 2 verwendet, aber diesmal nur die Bildbeschreibungen der Teilnehmer untersucht, nicht ihre Augenbewegungen. Ich habe die Beschreibungen der bimodalen Zweisprachigen in Niederländisch mit denen von den Niederländisch-Sprechenden und in NGT mit denen von den NGT-Gebärdenden verglichen.

Ich fand heraus, dass sich nicht-ikonische Ausdrücke in Lautsprache und ikonische Ausdrücke in Gebärdensprache gegenseitig beeinflussen, sodass sich die Beschreibungen der Zweisprachigen von denen der Niederländisch-Sprechenden ebenso wie von denen der NGT-Gebärdenden unterscheiden. Da Beschreibungen in Gebärdensprache ikonischer sein können und ausführliche Informationen darüber enthalten, wie Objekte platziert (z. B. horizontal, vertikal) und geformt (z. B. rund, länglich) werden, war besonders auffällig, dass die Beschreibung dieser Merkmale bei den Zweisprachigen in Lautsprache ausführlicher waren. Im Gegenzug verringerte die

Verwendung nicht-ikonischer Formen in Lautsprache die Nutzung von Ikonizität in der Gebärdensprache. Insgesamt habe ich mit diesem Kapitel gezeigt, dass bimodale Zweisprachige nicht zwei einsprachigen Personen in einem sind, sondern ihre Ausdrucksweise von der Kombination von Sprachen verschiedener Modalitäten geprägt ist.

In **Kapitel 4** habe ich untersucht, wie bimodale Zweisprachige die Welt betrachten. Führt die Tatsache, dass bimodale Zweisprachige anders sprechen und gebärden als ihre hörenden und gehörlosen Gleichgesinnten, auch dazu, dass bimodale Zweisprachige räumliche Beziehungen mit anderen Augen sehen? Mit demselben Experiment habe ich diesmal die Augenbewegungen von bimodalen Zweisprachigen untersucht, während sie die Bildbeschreibungen aus Kapitel 3 vorbereitet haben. Diese Augenbewegungen habe ich dann mit denen von Niederländisch-Sprechenden und NGT-Gebärdenden verglichen.

Ich stellte fest, dass bimodale Zweisprachige, anders als Niederländisch-Sprechende, bei der Vorbereitung des niederländisch Sprechens mehr auf die anderen Bilder schauten, die sie nicht beschreiben mussten. Damit ähneln sie den NGT-Gebärdenden in Kapitel 2. Das heißt, dass der Einfluss ikonischer Formen in NGT nicht nur die niederländischen Beschreibungen der bimodalen Zweisprachigen ausführlicher machte. Vielmehr brachte es sie auch dazu, den räumlichen Beziehungen mehr Aufmerksamkeit zu schenken. Bei der Vorbereitung des NGT-gebärdens haben sowohl bimodale Zweisprachige als auch NGT-Gebärdende den Bildern, die sie nicht beschreiben mussten, große Aufmerksamkeit geschenkt. Allerdings widmeten die NGT-Gebärdenden diesen Bildern noch mehr Aufmerksamkeit als bimodale Zweisprachige. Das bedeutet, dass die nicht-ikonische Formen der Lautsprache Beschreibungen in der Gebärdensprache nicht nur weniger ikonisch machte, sondern auch dazu führte, dass bimodale Zweisprachige räumliche Beziehungen anders betrachteten als NGT-Gebärdende. Insgesamt hat dieses Kapitel gezeigt, dass eine Lautsprache und eine

Gebärdensprache sich nicht nur in Form unterschiedlicher räumlicher Beschreibungen gegenseitig beeinflussen können, sondern auch die Art und Weise, wie bimodale Zweisprachige räumliche Szenen betrachten, bevor sie Bilder in der jeweiligen Sprache beschreiben.

Zum Schluss habe ich in **Kapitel 5** Beschreibungen und Augenbewegungen zusammen betrachtet, um zu untersuchen, wie bimodale Zweisprachige in jeder Sprache Bedeutung ausdrücken und wie sich dies auf die Art und Weise auswirkt, wie sie die Welt betrachten. Dieses Mal habe ich mir einen weiteren Unterschied angesehen, wie Laut- und Gebärdensprachen räumliche Beziehungen beschreiben. Wie oben beschrieben, bevorzugen es Gebärdensprachen, zuerst größere Objekte zu erwähnen, um räumliche Beziehungen zu beschreiben, während Lautsprachen sich in ihrer Präferenz , für welches Objekt zuerst erwähnt wird unterscheiden. Ich habe wieder das gleiche Experiment durchgeführt und Bildbeschreibungen sowie Augenbewegungen bei denselben bimodalen Zweisprachigen, Niederländisch-Sprechenden und NGT-Gebärendedn wie in Kapitel 2-4 untersucht. Ich habe mir diesen Reihenfolgeunterschied angesehen, um sicherzustellen, dass die Effekte, die ich in den vorherigen Kapiteln gefunden habe, nicht nur auf Ikonizität beschränkt sind, sondern auch allgemeine Effekte sind, die sich aus den verschiedenen Modalitäten ergeben, in denen die Sprachen verwendet werden.

Ich fand heraus, dass es tatsächlich einen Unterschied in Bezug auf die Reihenfolge gibt, in der die Laut- und Gebärdensprachen kleine und größere Objekte erwähnen: NGT-Gebärdende zogen es vor, zuerst die größeren Objekte zu erwähnen, während die Niederländisch-Sprechenden mal größere Objekte und mal kleinere Objekte zuerst erwähnten. Wenn bimodale Zweisprachige die Bilder auf Niederländisch beschrieben, erwähnten sie häufiger zuerst größere Objekte als die Niederländisch-Sprechenden. Bei der Beschreibung der Bilder in NGT erwähnten sie jedoch stets zuerst größere Objekte ebenso wie die NGT-Gebärdenden. So zeigte ich für bimodale Zweisprachige, dass die

Reihenfolge in der Gebärdensprache die Lautsprache beeinflusst, aber es gab keinen Einfluss der Lautsprache auf die Gebärdensprache.

Als nächstes habe ich mir auch die Augenbewegungen angesehen. Im Detail konzentrierte ich mich diesmal nur auf das Bild, das die Teilnehmer beschreiben mussten, und verglich die Augenbewegungen zu dem kleineren und dem größeren Objekt. Ich hatte erwartet, dass die Teilnehmer am meisten auf das Objekt schauen würden, das sie zuerst beschreiben würden, egal ob dies groß oder klein wäre.

Ich fand heraus, dass NGT-Gebärdende größere Objekte nicht nur häufiger zuerst erwähnen, sondern sie auch häufiger ansehen als Niederländisch-Sprechende. Die bimodalen Zweisprachigen erwähnten nicht nur zuerst größere Objekte als Niederländisch-Sprechende beim Sprechen, sondern schenkten auch größeren Objekten mehr Aufmerksamkeit, wenn sie sich darauf vorbereiteten zuerst über das größere Objekt zu sprechen. Während der Vorbereitung NGT zu gebärden unterschieden sich die Augenbewegungen der bimodalen Zweisprachigen jedoch nicht von denen der gehörlosen NGT-Gebärdenden. Dies alles bedeutet, dass ich in diesem Kapitel einen Einfluss von Gebärdensprache zu Lautsprache gefunden habe, aber im Gegensatz zu Kapitel 3, nicht von Lautsprache zu Gebärdensprache. Sprachliche Andersartigkeit kann also wechselseitige Einflüsse hindern.

Zusammenfassend habe ich in diesen vier Studien gezeigt, dass Sprachvielfalt in der Modalität (vokal versus visuell) dazu führt, dass Niederländisch-Sprechende und NGT-Gebärdende die Welt unterschiedlich betrachten. Ich habe auch gezeigt, dass Zweisprachige, die sowohl eine Laut- als auch eine Gebärdensprache beherrschen, beide Sprachen anders verwenden als die einsprachigen Niederländisch-Sprechenden und NGT-Gebärdenden. Dieser Umstand bringt sie dazu , die Welt bereits während der Vorbereitung auf das Sprechen oder Gebärden mit anderen Augen zu betrachten.

Acknowlegdements

Having received the official approval for this dissertation today, this is a good time for me to sit down, cherish this moment and think about all the amazing people that have contributed to this major accomplishment.

First and foremost, it was an honour to work with and learn from two extremely smart and inspiring mentors and supervisors, **Asli & Susanne**. Over the years, the three of us turned into the three musketeers, taking every hurdle that has faced us (especially those from certain editors). Once corona is over I can hopefully give you the group hug I wanted to give you both since my second paper has been accepted in Psychological Science.

Asli, thank you for your support and patience and for trusting me enough to be part of your VICI project. Even more, thanks for believing in me to conduct the cross-linguistic influence projects, even though these weren't part of your original VICI project plan. Thank you for pushing me to take chances and submit to such high journals - something I would have never dared myself. You were always available day and night no matter how busy you were. I highly valued this. Thank you also for trusting me in the role as research assistant/ lab manager for the MLC group, which has opened already many career paths and has shown me that my strength and passion not only lie in research but also in supporting others to make their research lives better. Thank you for providing such a warm and friendly lab environment by opening your home so many times for unforgettable Christmas and summer gatherings. It was a privilege to work with you and being part of your research group.

Susanne, thank you for believing in me especially at times in which I myself couldn't. You have been my emotional rock, someone I always looked up to and the (academic) big sister I never had. Thank you for investing your time, resources and efforts in making sure that I succeed, even though the empirical topic of the thesis was a bit out of your comfort zone. Thank you especially for making sure that I stay happy and sane along the way. Thank you for all your helpful and quick feedback, for your kind and encouraging words, for your helpful and clever analysis suggestions, for your professional as well as personal advice and your encouragement. You introduced me to the fabulous world of R – a world that I don't want to miss anymore. Most importantly, thank you for genuinely caring about me and checking up on me hourly in times of sickness and giving health advice from a Dutch perspective. I really appreciate and value our conversations in Norway and elsewhere about life, science and everything in between. Thank you for always being there for me and for giving me virtually or physical hugs when I needed some. I will be forever grateful.

Many thanks to my manuscript committee, **Falk Huettig, Pamela Perniss, Marianne Gullberg, Rain Bosworth & Chloë Marshall** for reading and evaluating my thesis, especially considering these hard times we currently live in. I'm very grateful you took the time to read through all of my pages and I look forward to having deep discussions with you during the defense.

A special thanks goes to **Tom.** Je hebt me van dag één tot het einde bij dit project gesteund. Dankzij jou kon ik alle NGT data begrijpen en gebruiken die je in de loop der jaren hebt gecodeerd en bedankt voor al je steun tijdens excursies door heel Nederland. Jouw aanwezigheid maakte het verzamelen van data bij de deelnemers thuis een leuke en onvergetelijke ervaring. Dit project zou niet mogelijk geweest zijn zonder jou!

Big thanks to all my participants, particularly, to the deaf signers and bimodal bilinguals. Without you, this dissertation would not have been possible. You opened up your homes to a total stranger (and sometimes even two or three when I had help from Tom or another confederate), allowed me to use all your sockets and to move all your furniture to get nice looking recordings and made me feel at home. Thanks for all the drinks and snacks, for the quality cuddling times with your pets and kids and for all the great stories you have shared.

I would also like to thank the **MPI,** especially the **library team** and the **Technical Group of the MPI**. Although my project was part of the CLS I had the privilege to receive the amazing support from the MPI. Many thanks to **Peter, Reiner, Gert-Jan, Tobias & Johan** for your patience and supporting me during my technical journey within the MPI and via phone at participant's homes and finding my "lost" data files. Thanks to **Jeroen** and **Nick** for all your help with processing the massive amount of video data that I have collected. Nick, you are greatly missed. Thanks also to **Jan & Rober** for basically everything and Jan, for always greeting me with my name when passing by on the MPI corridors even when I just started working there. Thanks to **Alex** for fixing all my technical problems and for always turning on my PC when I shut it down remotely over and over again. I would also like to thank IMPRS coordinator **Kevin**. Thanks to you the IMPRS feels like a big family. Thank you for your care, dedication and open ear at every given moment. Thanks for fun and inspiring photography adventures and thanks for making the IMPRS a place where I learned so much. I couldn't have done most of my work without all of you. Thank you.

I was lucky to always have an intern army around me. Thank you **Marcel, Jessica, Lianne & Anique** for being my NGT confederate and travelling around the country for me. Without all of you, I couldn't have collected the NGT data. Thank you **Marlijn M.** for being my addressee, memorable fieldtrips, safe car rides, untangling my wool chaos during long (very long) train rides and joyful addressee moments with the bimodal bilinguals and their pets. Thank you **Eveline** for being my first MA student and for making cold, snowy and creepy moments during field trips in Zoetermeer fun and unforgettable. Thank you **Els** for being such a hero and handling and coding so much of my data. Thank you **Madeleine** for being such an independent, talented and efficient intern and helping me so much in setting up the bimodal bilingual data collection.

Thanks to **Julia M.** to be such a burst of sunshine and being the addressee for all my speakers. Thank you **Linda L., Ruth and Kaya** for all your work with processing, handling and coding data.

I'd also like to thank all present and past members of the MLC lab: **Linda, Tom, Renske, Marlou, Louise, Dilay, Ezgi, James, Anita, Annika, Beyza, Erce, Gerardo, Zeynep, Junfei, Clarissa, Marlijn, Vicky, Huku, Christina, Kazuki & Kimberley** for all the stimulating discussion we had during lab meetings or lunch and for all the fun moments, creative Sinterklaas poems and delicious homemade food during lab events. Special thanks goes to the VICI team – **Dilay, Beyza, Erce, Tom, Kimberley & Susanne** for your feedback and support and for designing the experiments together.

Many thanks to my MLC friends. **Renske,** for all your help and support getting me started as a PhD student, for helping me with the ethical approval, for coding my data and for being there for me academically and non-academically. **Marlou,** thanks for introducing me to the groentepakket (I wouldn't know how to live without it anymore), for listening to my worries and for going on walks with me when I felt corona-lonely. **Dilay**, thanks for starting this PhD journey with me and providing emotional support when and where needed. **Ezgi**, thanks for being the best Paris tour guide ever and for fun Waal swimming, 4-daagse and Meringue baking moments. **Anita,** thank you for sharing the troubles and hard moments at the end of both our PhD journeys. I also want to thank my office mates - **Dilay, Christina, Marlijn, Rehana and Ezgi**. Thanks for your emotional support, gezellige coffee breaks and snack sharing.

There are also so many more MPI people I want to thank. **Eirini, Greta, Julia M., Birgit** and many more for uplifting corridor chats. Thank you **Marlijn** for becoming such a good friend. Thanks for sharing my plant love, our hip hop adventures and Michelin star-worthy food ;). Thanks especially for helping me with thesis summaries and for designing with me my thesis and thesis cover and making its creation so unforgettable. Thank you **Anne** for starting with me at the MPI as an RA and for ending this PhD not

alone by submitting the thesis at the same time. Thank you for our cooking adventures, endless conversations about love and life and being my East-German heritage mate here in Nijmegen. **Limor,** thank you for your everlasting energy and positivity, for female power uniting against men and for making me write sexactly into my acknowledgements. **Hannah**, thank you for being always so welcoming and for providing me with a safe haven when I really needed it. Thank you **Joe**, for allowing me to make a cutting of your office Monstera; you made me a very happy owner of a (by now very giant) Monstera. And thank you **Ksenija** for all the pep talks on the hallway and inside the 2^{rd} floor MPI bathroom.

And thanks to my non-MPI friends as well. Danke **Antonia, Lotti & Alena.** Alles begann in Leipzig und hat mich bis hierher gebracht. Meine Freundschaft mit euch bedeutet mir unglaublich viel und ich bin mir sicher, dass diese noch ewig anhalten wird. **Annika,** thank you for being always there for me, for listening to my worries, for great Netflix parties, for long sign language discussions and for your help with my German thesis summary. I hope our friendship will continue for many years to come. **Lisette**, thanks for your everlasting positivity, for your caring and emotional support, for helping me move houses on such short notice and still keeping a cool head and for making moving my 35 plants (now it's 85+, but who is counting?) such an unforgettable experience.

And now to my paranymphs. **Tom,** we hadden een min of meer moeilijke start omdat ik je taal niet kende, maar je werd een waardevolle vriend en een belangrijk persoon in mijn leven. Dankzij jou kwam ik in de visuele wereld van NGT, een taal waar ik inmiddels echt van hou. En ook bedankt om me zo enthousiast te maken voor visuele muziek en me een visuele muziektolk te laten worden. **Linda,** you have been with me from day one of my PhD when approaching me on my first day for coffee when everyone else from the group wasn't back from their Christmas holidays. Thanks for all your advice, your wisdom, your support, for always believing in me and for being there for me both academically and non-academically, no matter when or what. **Julia,** thank you for being

basically my best friend in Nijmegen. Thank you for always being there for me, for providing me shelter when I needed it, for the best holiday and spa moments ever, for sharing my video making love and for so much more. There are just not enough words to express my love and gratitude for you. Thank you for really everything.

Danke an meine Familie und Familienfreunde. **Opa Helmut, Ute, Opa Arthur, Oma Elsbeth & Oma Helga** - danke, dass ihr immer an mich glaubt auch wenn ich so weit weg bin und es manchmal schwer nachvollziehbar ist, was ich hier eigentlich tue. Danke **Netti,** für alles. Du bist immer da für mich und warst schon immer so stolz auf mich. Dies gab und gibt mir immer so viel Kraft. **Sandro**, zu dir habe ich mich schon immer unglaublich verbunden gefühlt. Du warst mein Vorbild schon von klein an und ich weiß, dass ich immer auf dich zählen kann auch wenn die Zeit und Distanz uns oft ein paar Kommunikationspausen eingeräumt hat. Danke für alles. **Mama & Papa** ohne euer Vertrauen, eure Unterstützung und Liebe hätte ich es nie so weit geschafft. Ich bin euch so unglaublich dankbar für alles. Ihr seid die besten Eltern, die man sich wünschen kann.

And **Mathieu**, I have been thinking a lot about whether it is appropriate to mention you here last given that we don’t know each other for that long. But you deserve no other place than this one here. I’m so glad that I gave you my permission to steal my heart in the most stressful time ever, right at the end of my PhD journey. I couldn’t have made it through these past months without you. Thank you so much for your support, for your concern, for caring, for believing in me and for your unconditional love. Let’s continue to make each other stronger and happier. I can’t wait to share many more unforgettable moments with you.

Curriculum vitae

Francie Manhardt was born on April 11, 1991, in Erfurt, Germany. She obtained a bachelor's degree in Linguistics from Universität zu Leipzig (Germany) and a master's degree in Linguistics from Radboud University. In her master's thesis, Francie used eye-tracking to study the link between cross-language activation and cognitive control in bimodal bilinguals under supervision of dr. Ellen Ormel. Francie then continued her academic studies and joined the Multimodal Language and Cognition group to work with prof. dr. Asli Özyürek and dr. Susanne Brouwer on the studies presented in this doctoral thesis. In addition to her research, Francie worked as a lecturer at the Department of Language and Communication. She also fulfilled the role as the lab manager for the MLC group for several years. Francie has been involved in various science outreach projects such as science festivals and writing blog articles for the science blog Donders Wonders. In September 2020, Francie joined the Research Services group to work as Data Steward at the Institute for Management Research to support the ethical committee, support developing ICT data infrastructure and coordinate research data management for the Nijmegen School of Management.

Author publications

Manhardt, F., Brouwer, S., & Özyürek, A. (in press). A tale of two modalities: Signs and speech influence each other in bimodal bilinguals. *Psychological Science.*

Manhardt, F., Özyürek, A., Sumer, B., Mulder, K., Karadöller, D. Z., & Brouwer, S. (2020). Iconicity guides visual attention: A comparison between signers' and speakers' eye-gaze during message preparation. *Journal for Experimental Psychology: Learning, Memory, and Cognition*. 46(9), 1735–1753.

Manhardt, F., Brouwer, S., van Wijk, E., & Özyürek, A. (under review). Word order preference in sign influences speech in bimodal bilinguals but not vice versa: Evidence from language production and eye-visual attention.

Ercenur, Ü., **Manhardt, F.**, & Özyürek, A. (under review). Speaking and gesturing guide event perception during message conceptualization: Evidence from eye movements.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*

21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*

65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*

107. Situational variation in non-native communication. *Huib Kouwenhoven*

108. Sustained attention in language production. *Suzanne Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*

110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*

111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*

114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*

116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*

117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*

118. Language processing in a conversation context. *Lotte Schoot*

119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*

120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*

122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*

125. Lapse organization in interaction. *Elliott Hoey*

126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*

128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*

129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*

130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*

131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*

132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*

133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*

134. Indirect request comprehension in different contexts. *Johanne Tromp*

135. Envisioning Language - An Exploration of Perceptual Processes in Language Comprehension. *Markus Ostarek*

136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*

137. Let the agents do the talking: on the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*

138. Age and hearing loss effects on speech processing. *Xaver Koch*

139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*

140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*

141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*

142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*

143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas-Berscia*

144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*

145. Understanding temporal overlap between production and comprehension. *Amie Fairs*

146. The role of exemplars in speech comprehension. *Annika Nijveld*

147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*

148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*