CrossMark

Word order preference in sign influences speech in hearing bimodal bilinguals but not vice versa: Evidence from behavior and eye-gaze

Address for correspondence:
Francie Manhardt,
E-mail: francie.manhardt@ru.nl

Francie Manhardt[1] [iD], Susanne Brouwer[1] [iD], Eveline van Wijk[1]
and Aslı Özyürek[1,2,3]

[1]Centre for Language Studies, Radboud University, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands; [2]Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands and [3]Donders Center for Brain, Cognition and Behaviour, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

Abstract

We investigated cross-modal influences between speech and sign in hearing bimodal bilinguals, proficient in a spoken and a sign language, and its consequences on visual attention during message preparation using eye-tracking. We focused on spatial expressions in which sign languages, unlike spoken languages, have a modality-driven preference to mention grounds (big objects) prior to figures (smaller objects). We compared hearing bimodal bilinguals' spatial expressions and visual attention in Dutch and Dutch Sign Language ($N = 18$) to those of their hearing non-signing ($N = 20$) and deaf signing peers ($N = 18$). In speech, hearing bimodal bilinguals expressed more ground-first descriptions and fixated grounds more than hearing non-signers, showing influence from sign. In sign, they used as many ground-first descriptions as deaf signers and fixated grounds equally often, demonstrating no influence from speech. Cross-linguistic influence of word order preference and visual attention in hearing bimodal bilinguals appears to be one-directional modulated by modality-driven differences.

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 across modalities – that is, between speech and sign – occurs in hearing bimodal bilinguals, hearing individuals fluent in a spoken (vocal) and a sign (visuo-spatial) language. Previous research has provided first evidence for bi-directional cross-linguistic influence between categorical and arbitrary spatial expressions in speech (e.g., LEFT/RIGHT) and iconic expressions in sign (i.e., visual one-to-one mappings of the spatial configuration to sign; Manhardt, Brouwer & Özyürek, 2021) in hearing bimodal bilinguals. Here, we focused on another domain – that is, word order in spatial expressions where sign languages exhibit a modality driven order to mention grounds first and figures later. We investigated whether there is also bi-directional influence in this domain or whether this is constrained due to differences in modality.

In spatial expressions in sign languages, differences in modality are mostly visible with respect to the order in which objects are mentioned. Sign languages, studied so far, seem to universally prefer to mention grounds (bigger objects) before figures (smaller objects) to describe spatial relations between two objects (e.g., Emmorey, 2002 for American Sign Language; Kimmelman, 2012 for Russian Sign Language; Perniss, Zwitserlood & Özyürek, 2015 for German and Turkish Sign Language). This universal preference in visual expressions seems to be motivated by perceptual biases (rather than linguistic biases) based on principles of Gestalt perception (i.e., the human eye perceives visual elements prioritizing grounds in relation to figures, e.g., Rubin, 1915, 1958). Spoken languages, however, are more flexible and vary in the order in which objects are mentioned to describe spatial relations (Levinson, 1996, 2003). It is not known how these differences modulate cross-linguistic influences between a sign and spoken language.

In addition, previous research has shown evidence for parallels between order of mention of elements in an utterance and visual attention to objects (e.g., Griffin & Bock, 2000). Based on these previous findings, we explore whether cross-linguistic influence of word order in bilinguals, if any, can have cognitive consequences and also guides visual attention to objects. We adopt the approach of assessing visual attention before describing pictures which is found

CAMBRIDGE
UNIVERSITY PRESS

to reflect differences in conceptualization during message preparation across different spoken languages (see e.g., Flecken, Carroll, Weimar & Stutterheim, 2015; Manhardt, Özyürek, Sümer, Mulder, Karadölller & Brouwer, 2020; Papafragou, Hulbert & Trueswell, 2008; Trueswell & Papafragou, 2010). We ask whether sign and spoken language users allocate more visual attention during message preparation to those aspects of a visual scene that are mentioned first within an utterance.

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 focused 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 (e.g., prepositional dative structure) is reflected in the word order of a second sentence that is otherwise unrelated to the first.

However, 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) and not found in the absence of these typical priming paradigms. For instance, when Korean (ground-first)–English (figure-first) bilinguals described one out of four pictures based on its location (e.g., a picture of a cat that is above a piano) while trying to ignore written distractor words, the authors did not find any evidence for cross-linguistic influence[1] (Ahn, Gollan & Ferreira, 2019).

In the domain of visual modality, it has been found that word order can be primed within American Sign Language (i.e., prenominal vs. post-nominal sentence structure) in deaf signers (Hall, Ferreira & Mayberry, 2015). Furthermore, there is a recent study showing that if hearing non-signers are presented with a certain order of elements in silent gesture, this can prime word order preferences in their subsequent spoken utterances (Shurley, Schouwstra & Pickering, 2018). Therefore priming of order of elements can occur in the visual modality and across modalities and from a non-linguistic to a linguistic domain.

However, whether and how cross-linguistic influence of word order preference occurs in hearing bimodal bilinguals across speech and sign and whether this is constrained by modality driven word order preferences 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 and guided by message preparation.

The present study assesses the possibility of cross-linguistic influence in a special group of language users – namely, hearing bimodal bilinguals. Hearing bimodal bilinguals are often highly proficient in both a spoken and a sign language from birth as they are hearing children born to deaf parents. Thus, they are typically exposed naturally to sign language from early on as a home language. The sign language (i.e., 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, Lillo-Martin & Palmer, 2018; Quadros & Lillo-Martin, 2018) who have had 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.

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., Kimmelman, 2012; Perniss et al., 2015) and relates to the order in which two objects are mentioned. Following image-based Gestalt principles, the two objects (e.g., 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, 1915, 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. As grounds are bigger, they are assumed to have primacy in order in linguistic utterances compared to figures.

While word order preferences differ across sign languages in various non-spatial linguistic domains (e.g., Kimmelman, 2012; Sandler & Lillo-Martin, 2006), when describing such spatial relations, sign languages universally prefer establishing first the lexical sign for the ground (see Supplementary Materials, Figure S1, panel a) followed by introducing the lexical sign for the figure (see Supplementary Materials, Figure S1, panel b; 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, 1996, 2003). For one, 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 physical features of the objects as well as the actual spatial configuration (see Supplementary Materials, Figure S1, panel c). Similarly to the order of mentioning lexical signs for grounds and figures, in these so-called CLASSIFIER CONSTRUCTIONS the hand representing the ground is typically mapped onto the signing space first followed by the hand representing the figure (e.g., Emmorey, 2002; Perniss, 2007; Perniss et al., 2015; Sümer, 2015). Furthermore, when hearing non-signers are 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 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), which is in line with Gestalt and linguistic conceptual theories that identify grounds as bigger and more stable and permanent object (e.g., Rubin, 1915, 1958).

---

[1]In the original work the authors explicated whether sentence structure is co-activated and found no evidence of co-activation. To draw parallels to our work, we refer to the lack of influence when referring to this work.

The link between language production and visual attention

Previous research has shown that already prior to speaking, cross-linguistic variation between languages can influence message conceptualization and guides speakers' visual attention to different components of these visual scenes during message preparation 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, Von Stutterheim & Carroll, 2014; Flecken et al., 2015; Goller, Lee, Ansorge & Choi, 2017; Papafragou et al., 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). Based on these types of evidence, researchers have argued for a tight link between the way speakers linguistically encode visual scenes and how they visually attend to such scenes already during message preparation as well as during actual language production.

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 (Manhardt et al., 2020). In this work, another modality-specific aspect to describe spatial relations has been assessed regarding the use of iconic expressions in sign (see Supplementary Materials, Figure S1, panel c) versus categorical and arbitrary expressions in speech (e.g., LEFT/RIGHT), using the same materials as in the present study. This modality-specific difference has been found to guide deaf signers' visual attention differently to those spatial relations than that of hearing non-signers during message preparation showing evidence for THINKING FOR SIGNING. However, whether the influence of word order on eye-gaze, found for spoken languages, extends to sign productions during message preparation has not been explored yet, let alone in bilinguals.

Present study

In the present study, we investigated whether there is bi-directional cross-linguistic influence in hearing bimodal bilinguals, from sign to speech as well as vice versa by looking at spoken and signed descriptions of hearing bimodal bilinguals and comparing them to speech of hearing non-signers as well as to deaf signers' signed descriptions. We also assessed this at the level of visual attention by looking at eye-gaze during message preparation – that is, whether eye-gaze preferences to look more at grounds than figures depends on the language and whether it is sensitive to cross-linguistic influence.

We used a visual world language production eye-tracking paradigm (Manhardt et al., 2020) 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. The reason to investigate eye-gaze before signing/speaking, rather than during, was to control for differences in utterance production between hearing non-signers and deaf signers. Hearing non-signers are typically more flexible to speak while looking at the screen. Deaf signers, however, prefer eye contact with an addressee during signing, as

signing towards a screen is considered less appropriate. Moreover, signing (as well as gesturing) involves moving the head, hands and torso. This would lead to an increased loss of eye-gaze data, thus we assessed eye-gaze patterns prior to signing/speaking. Finally, this method of analyzing fixations before language production to understand language planning processes is also commonly used in previous studies (e.g., Manhardt et al., 2020; Papafragou et al., 2008; Trueswell & Papafragou, 2010). Overall, this paradigm allowed us to assess linguistic word order preferences during a (semi)naturalistic but controlled picture description task (i.e., refraining from using priming paradigms). At the same time, it gave us the opportunity to simultaneously measure 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. We then assessed cross-linguistic influence by comparing word order preferences of hearing 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 non-signers' and deaf NGT signers' eye-gaze during message preparation. Finally, we examined cross-linguistic influence of visual attention by assessing hearing 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 non-signers 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 hearing bimodal bilinguals, if there is cross-linguistic influence from the robust modality-driven word order in sign, we predicted that hearing bimodal bilinguals prefer less figure-first descriptions compared to their hearing non-signing peers. However, since NGT is a minority language in the Netherlands, influence from sign to speech might be absent 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 in spoken language bilingualism, 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 more resistant to change than the flexible word order in Dutch speech. Consequently, speech might not influence sign.

Finally, instead of experiencing cross-linguistic influence across modalities between speech and sign, hearing bimodal bilinguals might maintain their language-specific patterns as previously observed in such highly proficient bilinguals (e.g., Ahn et al., 2019; Azar, Özyürek & Backus, 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 what 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 non-signers over the time course of message preparation. This might reflect deaf-signers' modality-driven preference to produce predominantly ground-first descriptions. This would indicate that THINKING FOR SPEAKING extends to THINKING FOR SIGNING (Manhardt et al., 2020) in the domain of word order.

For hearing bimodal bilinguals, we predicted that cross-linguistic influence can go beyond language production and also influence message conceptualization. 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 conversely from speech to sign. In particular, we predicted that if there is cross-linguistic influence from sign to speech this might change hearing bimodal bilinguals' message conceptualization during spoken message preparation – that is, hearing 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 non-signers. In the reverse direction, if there is no cross-linguistic influence from speech to sign, then hearing 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 hearing bimodal bilinguals can be found only when there is cross-linguistic influence, thus from sign to speech but not from speech to sign.

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

The methods reported in this experiment were approved by the Humanities Ethics Assessment Committee of the Radboud University Nijmegen, The Netherlands. All data and analysis scripts are available at https://doi.org/10.17605/OSF.IO/86XP4.

Participants

The participants were the same as tested in Manhardt et al. (2020, 2021)[2]. This sample consisted of 21 hearing 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 hearing native Dutch non-signers (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). Three hearing bimodal bilinguals and two signers were excluded from the eye-tracking part of the study due to high eye-tracking loss (larger than 45%).

Crucially, hearing bimodal bilinguals participated twice, once in Dutch and once in NGT. The sessions took place three to five weeks apart and the order of the two sessions was counterbalanced to avoid priming effects. Therefore, one half of the hearing bimodal bilinguals (N = 10) performed their first session in Dutch followed by a second session in NGT, while the other half (N = 9) carried out the NGT session first followed by the Dutch session. The non-signers were tested once in Dutch and the deaf native signers once in NGT.
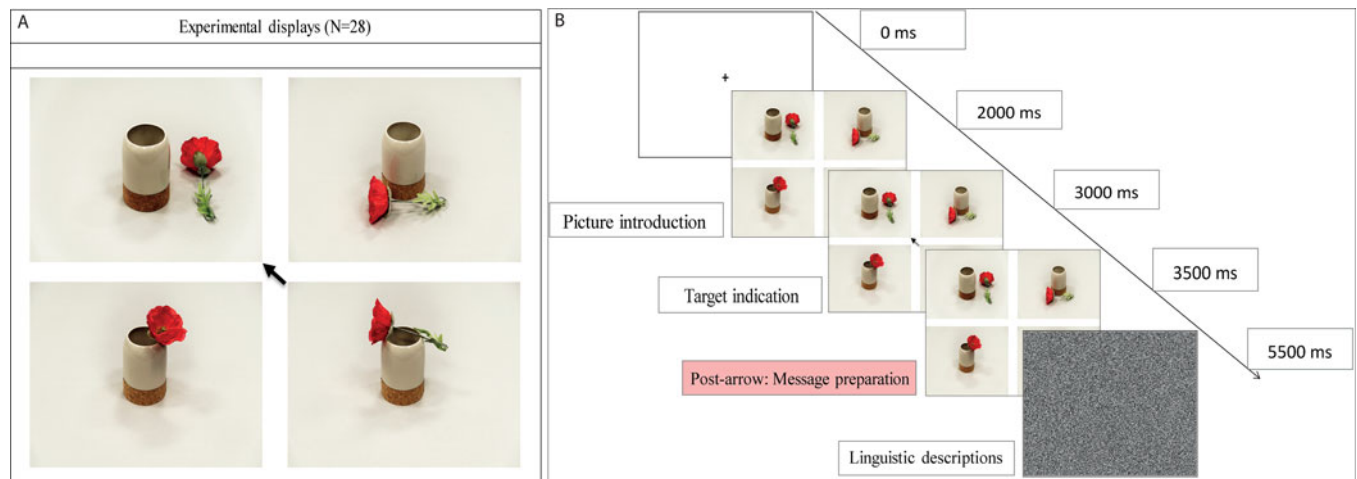
All deaf signers were born deaf and acquired NGT from birth from their deaf parents. Four of the deaf signers received a cochlear implant but only later in their lives (at age 12, 30, 37, 48). Thus, the signers had no access to auditory Dutch from birth, but had some knowledge of Dutch in its written form (formally instructed at school: M_{age}= 3.5, SD_{age} =2.8, for self-rated literacy skills in Dutch, see Appendix A).

All hearing non-signers were exposed to Dutch from birth and learned additional languages (mostly English, or German) later through instructional settings (for more information on hearing non-signer's language background, see Appendix B). The control groups were chosen on the basis of their native status and naturalistic acquisition (i.e., not instructional) of Dutch or NGT respectively, independent of whether they knew additional languages (as long as those were acquired later in life and through formal instruction).

Hearing bimodal bilinguals were born to at least one deaf parent, thus they simultaneously acquired NGT as minority language at home and Dutch as the majority environmental language from birth. We assessed fluency in Dutch and NGT by collecting self-ratings on a five-point Likert scale for language use (1 = never; 2 = rarely; 3 = sometimes; 4 = most of the time; 5 = all the time) as well as for proficiency for comprehension and production (1 = beginner, 2= intermediate, 3 = advanced, 4 = native-like, 5 = native). 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.

Regarding language use, scores indicated that hearing bimodal bilinguals use Dutch (M = 4.80, SD = 0.41) more often than NGT (M = 3.65, SD = .93) (paired samples t-test: $t(20)$ = -5.66, $p < .001$, Cohen's d = -1.81). Eight of the 21 hearing bimodal bilinguals reported to be professional sign language interpreters. Regarding language proficiency, ratings for production in NGT and Dutch indicated fluency levels somewhere between advanced and native-like, although scores were higher for Dutch (M = 4.55, SD = .51) than for NGT (M = 3.85, SD = .93) (paired samples t-test: $t(18)$ = -2.77, p = .01, Cohen's d = -.98). For comprehension, hearing bimodal bilinguals rated themselves as similarly native-like in both Dutch (M = 4.75, SD = .44) and NGT (M = 4.50, SD = .69) (paired samples t-test: $t(18)$ = -1.56, p = 0.21, Cohen's d = -.43).

Additionally, we assessed Dutch language fluency by measuring non-signers' and hearing bimodal bilinguals' articulation rate using speech analysis software Praat (Boersma & Weenink, 2001) (for the script, see de Jong & Wempe, 2009). To do so, articulation rate (number of syllables/time) was extracted from a elicited narrative (retelling of 3.41 min video narrative, for more narrative details see Herman et al., 2004). Articulation rate did not differ between the hearing bimodal bilinguals (M = 3.56, SD = .46) and the Dutch non-signers (M = 3.39, SD = .43) (independent samples t-test: $t(38)$ = -1.17, p = .25, Cohen's d = -.38), suggesting that hearing bimodal bilinguals were highly fluent in Dutch.

---

[2]We used the same sample and materials as in Manhardt et al. (2020, 2021). There is no additional overlap between the present study and these two previous studies in respect to the data.

Fig. 1. Example of an experimental display (panel A) and timeline (panel B) to describe "the flower is to the right of the vase". The arrow indicates the target picture.

To also assure that hearing bimodal bilinguals were highly proficient in NGT, we used an assessment tool for narrative production originally created to assess British Sign Language development (Herman et al., 2004). In particular, a deaf native NGT signer scored signed retellings of a 3.41 min video on two levels following a detailed and complex and objective protocol. The first level was on narrative structure, which included evaluations on mentioning all crucial events of the narrative in NGT appropriate structure. The second level was on NGT grammar, including evaluations on using spatial verbs and agreement verbs, aspect, classifiers, and role shift. Scores indicated no differences between hearing bimodal bilinguals and deaf native signers in both narrative structure (independent samples t-test: $t(38) = .09$, $p = .92$, *Cohen's d* = .03) and use of grammar (independent samples t-test: $t(38) = 1.61$, $p = .12$, *Cohen's d* = .05), suggesting that our hearing bimodal bilinguals were highly proficient in NGT.

Materials

We used the same stimuli set as in Manhardt et al. (2020, 2021), 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 1A). An arrow pointing at one of the pictures indicated the target picture participants had to describe. In 28 experimental displays the arrow pointed to a picture with left/right target relations, while the remaining three pictures in these items included other spatial relations (i.e., front, behind, in or on). We included 56 filler displays in which the arrow pointed at any other spatial relation (i.e., front, behind, in and on) to avoid emphasis on left/right relations during the whole experiment. We focused on left/right relations as these allowed us to assess eye-movement preferences to grounds and figures without overlapping locations or occlusions of the two objects (as is the case with in/on/behind relations). The distance between the ground and figure object was always kept equal across displays for the spatial relations respectively (i.e., for left/right, front/behind etc.). Irrespective of at which picture the arrow is pointing, ground objects were always located in the center of the pictures and figures were always placed to the left, right, front, behind, inside, or on top of the ground (Figure 1A). 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, 1978, 2003).

Procedure

Participants were individually tested on a SMI RED-250 mobile laptop. Before the actual experiment, participants performed a familiarization task. This task contained similar displays compared to those in the actual eye-tracking experiment to familiarize participants with the overall complexity and general arrangement of our displays – namely, a two-by-two grid in which each picture contained two objects in different spatial relations to each other. After answering some questions about the displays, we continued with the actual eye-tracking description task. The experiment was preceded by three practice trials and a five-point calibration and validation procedure. Each trial initiated with a fixation cross for 2000 ms (Figure 1B). After that, a four-picture display was introduced for 1000 ms followed by an arrow in the middle of the screen that indicated the target picture for a duration of 500 ms. The arrow then disappeared and the four pictures remained on the screen for 2000 ms until a gray, visual noise screen indicated the start of the picture description. This 2000 ms allowed us to measure eye-gaze during message preparation (for a similar approach, see e.g., Manhardt et al., 2020; Papafragou et al., 2008; Trueswell & Papafragou, 2010). During the gray noise screen, participants had to describe the target picture, thus the picture at which the arrow was pointing, to a trained confederate. After each picture description, the confederate pretended to select the described picture on a separate tablet. The confederates' four pictures were identical to those of the participant, except that they were spread differently on the tablet display (e.g., on the participant's screen the arrow pointed at the picture in the right upper corner, while the same picture could be located in the left lower corner on the confederates' tablet). After each picture description, participants initiated the next trial by pressing the ENTER button.

The timing of each trial element (e.g., fixation cross, introduction of four pictures) was always fixed to ensure that participants have equal viewing times of the visual displays before describing them. For this same reason, participants described the pictures after the

visual display disappeared (i.e., without seeing the visual display) to allow spontaneous selection of word order representative of the languages used, thus avoiding that different word orders might be a consequence of other experimental factors such as longer viewing times.

We used four pictures and a confederate to create face-to-face communicative situations. Hence, the confederate was always another person than the experimenter and importantly, participants were told that confederates were randomly selected participants. Participants did not receive feedback on their picture descriptions. In Dutch sessions, confederates were always Dutch native non-signers, while in NGT sessions, a deaf native NGT signing confederate and experimenter were present. Thus, hearing bimodal bilinguals were tested in a monolingual Dutch or NGT situation (i.e., not in a bilingual setting) to isolate unintentional transfer of word order preferences between NGT and Dutch. Furthermore, we used two sets of counterbalanced lists, thus hearing bimodal bilinguals did not describe the same pictures across the two sessions.

We used the software package Presentation NBS 16.4 (Neurobehavioral Systems, Albany, CA) to control and send triggers to the eye-tracker and present the stimuli. Eye-gaze was recorded binocular at a rate of 250 Hz (every 4 ms). Participants were always instructed orally/visually in form of a video. At the end of the session, participants received a language background questionnaire to assess language use, language proficiency, deafness in family, etc. Hearing bimodal bilinguals received the questionnaire always at the end of Dutch sessions (i.e., one half received it at the end of the first session, the other half received it at the end of the second session) to avoid that self-ratings for the less-dominant heritage language NGT are influenced by performing the communication task with a deaf NGT native signer. In total, the experimental session lasted approximately 45 minutes.

Data analysis

In this section, we will first describe how we analyzed language production in respect to preference of word order in Dutch across hearing non-signers and hearing bimodal bilinguals as well as in NGT across deaf signers and hearing bimodal bilinguals. Furthermore, we describe the analysis of eye-gaze

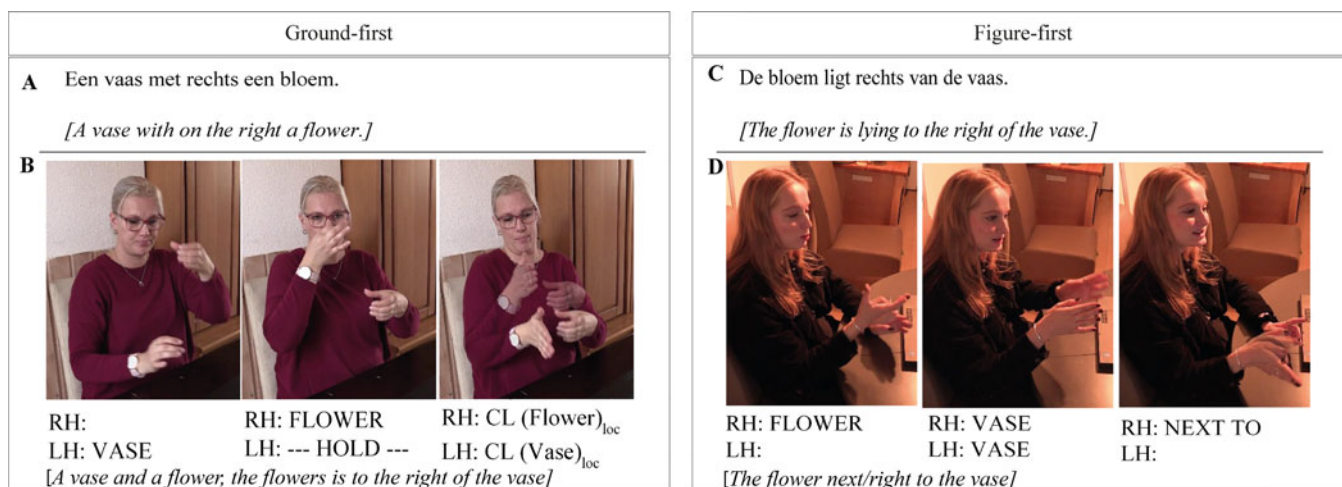preferences to look at grounds and figures during message preparation.

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, Brugman, Russel, Klassmann & Sloetjes, 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, such as grounds as bigger object placed in center of pictures, figures as smaller objects surrounding grounds. In Dutch, ground-first descriptions typically involved prepositional constructions using *met* ("with"; Figure 2A), while figure-first descriptions usually included verb constructions using *liggen/staan* ("lying/standing"; Figure 2C). 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 2B), while figure-first descriptions (albeit were very few) included lexical signs for the spatial relation (RELATION LEXEMES; Figure 2D). 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).

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, Davidson & Bates, 2008) 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 conducted two types of analyses: (1) we assessed whether word order preferences differ in Dutch and NGT by comparing descriptions between hearing non-signers and deaf

Ground-first	Figure-first
A Een vaas met rechts een bloem.	**C** De bloem ligt rechts van de vaas.
[A vase with on the right a flower.]	*[The flower is lying to the right of the vase.]*
B	**D**
RH: / RH: FLOWER / RH: CL (Flower)loc	RH: FLOWER / RH: VASE / RH: NEXT TO
LH: VASE / LH: --- HOLD --- / LH: CL (Vase)loc	LH: / LH: VASE / LH:
[A vase and a flower, the flowers is to the right of the vase]	*[The flower next/right to the vase]*

Fig. 2. An example for 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.

signers, and (2) we assessed whether hearing bimodal bilinguals differ in their word order preferences in Dutch compared to hearing non-signers and in NGT compared to deaf signing controls.

Visual attention

For each trial, eye movements were recorded from pre-arrow onset (0 ms) until the four-picture display disappeared (3500 ms). We analyzed fixation proportions (right eye only) across 50 ms continuous time bins. Our analyses focussed on a subset of the time course as we were only interested in examining the differences in eye movements linked to message preparation. We selected a 2000 ms post-arrow window initiating after target indication (1500 ms) until the onset of a production (3500 ms, Figure 1). This time window captures participants' message preparation phase linked to relational encoding (see Manhardt et al., 2020). This enabled us to assess whether there is cross-linguistic influence of visual attention in hearing bimodal bilinguals to grounds and figures respectively to the order in which they are being mentioned.

For the experimental trials, we defined two different rectangle-shaped Areas of Interests (AoIs): 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. Ground AoIs were larger capturing the ground object in the center 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 (version 3.3.1; R Core Team, 2013). 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 hearing bimodal bilinguals, as mentioned above in the participant section). Additionally, we excluded 3.6% of the trials in which track loss was higher than 50%.

We conducted two types of analyses on our binomial dependent variable (fixations to grounds (1) vs. figures (0)) 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 non-signers and deaf signers during message preparation, and (2) we assessed whether hearing 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. Due to multiple comparisons 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 language production data to assess word order preference in Dutch and NGT from hearing bimodal bilinguals compared to 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 hearing bimodal bilinguals' preference to look at grounds versus figures depending on the order in which they planned to mention them.

Language production

Figure 3 shows proportions of ground-first descriptions in Dutch and NGT between hearing bimodal bilinguals and controls, respectively. For plotting, data were averaged over trials and participants.

Word order preferences in Dutch and NGT (controls)

We assessed first whether word order preferences differed between the control groups by comparing Dutch hearing non-signers' and NGT deaf signers' picture descriptions. In particular, we investigated whether the ground was mentioned first (1) or not
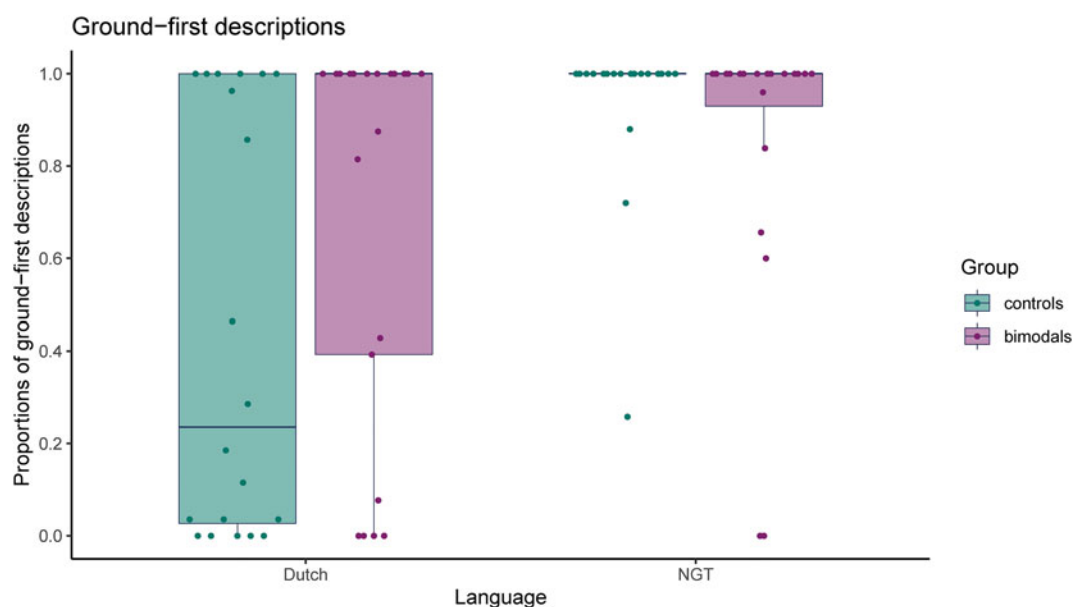
Fig. 3. Ground-first descriptions in Dutch and NGT. The left panel shows proportions of ground-first descriptions in Dutch across hearing non-signers (left) and hearing bimodal bilinguals (right). The right panel shows proportions of ground-first descriptions in NGT descriptions across deaf signers (left) and hearing bimodal bilinguals (right). Dots in the boxplot represent each data point (participant).
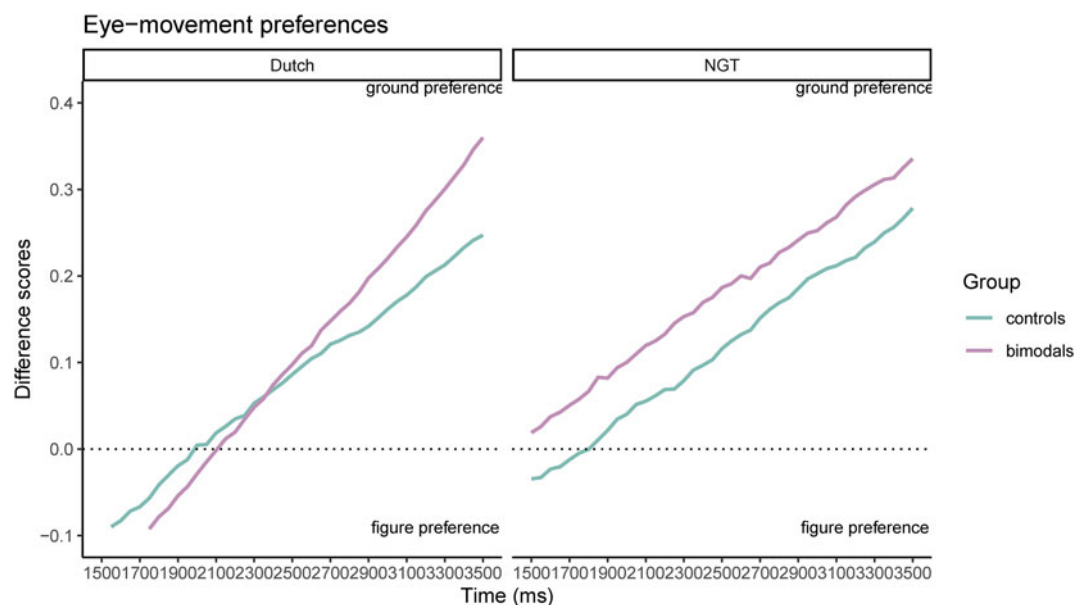
Fig. 4. Preference of fixating the ground vs. figure object during message preparation in Dutch and NGT. The left panel shows preferences in Dutch across hearing non-signers and hearing bimodal bilinguals. The right panel shows preferences in NGT, across deaf signers and hearing bimodal bilinguals. Y-axis 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 including 200ms saccade correction) until language production onset (3500 ms).

(0) using a general linear mixed-effects regression model with Group (hearing non-signers, 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 ($\beta = 9.57$, $SE = 2.45$, $z = 3.91$, $p < 0.001$), suggesting that deaf signers produced more ground-first descriptions in NGT ($M = 0.92$, $SD = 0.26$) than hearing non-signers in Dutch ($M = 0.56$, $SD = 0.50$; see Figure 3).

Cross-linguistic influence of word order preferences in hearing bimodal bilinguals

We compared hearing 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 non-signers, numerically contrast coded as -1/2 vs. hearing 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 ($\beta = 8.15$, $SE = 3.53$, $z = 2.31$, $p = 0.02$), suggesting that hearing bimodal bilinguals produced more ground-first descriptions ($M = 0.67$, $SD = 0.43$) than their hearing non-signing peers ($M = 0.43$, $SD = 0.50$; see Figure 3, left panel). No effect of Session Order on hearing bimodal bilinguals' ground-first preference was found (see Appendix C 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. hearing 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 ($\beta = -0.59$, $SE = 2.74$, $z = -0.22$, $p = .83$), revealing that hearing bimodal bilinguals did not differ from deaf signers in how often they produced ground-first descriptions when signing in NGT (hearing bimodal bilinguals: $M = 0.85$, $SD = 0.35$; deaf signers: $M = 0.92$, $SD = 0.27$; see Figure 3, right panel). Again, no effect of Session Order on hearing bimodal bilinguals' ground-first preference was found (see Appendix C for more information).

Visual attention

For plotting, we calculated difference scores by subtracting fixations to the ground AoI from the 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). Figure 4 illustrates these difference scores during message preparation in hearing bimodal bilinguals in Dutch (left panel) and NGT (right panel) compared to their hearing speaking (left panel) and deaf signing peers (right panel), respectively. The difference scores were plotted in successive 50 ms time bins initiating immediately after target indication (1500 ms plus 200ms saccade correction) until language production onset (3500 ms). A visualization of proportion of looks to both the ground and the figure can be found in the supplementary materials in Figure S2 (Supplementary Materials).

Eye-gaze preferences in Dutch and NGT (control groups)

We first examined whether eye-gaze to grounds versus figures differed in hearing non-signers 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 non-signers, numerically contrast coded as -1/2 vs. deaf signers, numerically contrast coded as +1/2) and Bin (continuous, centered 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 ($\beta = 0.11$, $SE = 0.14$, $z = 0.78$, $p = .44$), but a significant main effect of Bin ($\beta = 0.16$, $SE = 0.01$, $z = 12.74$, $p < .001$), and a significant interaction between Group by Bin ($\beta = -0.06$, $SE = 0.02$, $z = -2.69$, $p < .01$). This interaction suggests that during message preparation deaf signers preferred looking at grounds from the start, while for hearing non-signers a ground preference in eye-gaze emerged only later and instead started with a preference to fixate figures (Figure 4).

Cross-linguistic effects on visual attention in hearing bimodal bilinguals versus hearing non-signers

We assessed whether hearing bimodal bilinguals' eye-gaze patterns differed from hearing non-signers when planning Dutch descriptions. In particular, we investigated fixations to grounds (1) or figures (0) using a general linear mixed-effects regression model with Group (non-signers, numerically contrast coded as -1/2 vs. hearing bimodal bilinguals, numerically contrast coded as +1/2) and Bin (continues, centered 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 ($\beta = 0.23$, $SE = 0.25$, $z = 0.93$, $p = .35$), but a significant main effect of Bin ($\beta = 0.24$, $SE = 0.01$, $z = 19.51$, $p < .001$), and a significant interaction between Group by Bin ($\beta = 0.10$, $SE = 0.02$, $z = 4.10$, $p < .001$). This interaction suggests that during message preparation in Dutch, hearing bimodal bilinguals and hearing non-signers preferred looking at figures over grounds at the beginning of message preparation. However, when message preparation is unfolding, both groups preferred fixating grounds over figures. Crucially, hearing bimodal bilinguals' preference to look at grounds over figures increased more steeply over time compared to their hearing non-signing peers (Figure 4, left panel).

To further show that the relative looks over time to the grounds versus figures depend on the word order that the participants produced, we additionally analyzed whether hearing bimodal bilinguals and hearing non-signers look more to grounds when it is mentioned first than when figures are mentioned first (see Appendix D for more information). This analysis confirms that both hearing non-signers and hearing bimodal bilinguals look more to grounds over time when it is mentioned first than when figures are mentioned first ($\beta = 0.98$, $SE = 0.05$, $z = 20.21$, $p < .001$). Furthermore, hearing bimodal bilinguals look more often at grounds over time than mentioning grounds first than non-signers ($\beta = 0.11$, $SE = 0.05$, $z = 2.32$, $p < .025$).

Cross-linguistic effects on visual attention in hearing bimodal bilinguals versus deaf signers

We assessed whether hearing bimodal bilinguals' eye-gaze patterns differed from that of deaf signers when planning NGT descriptions. In particular, 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. hearing bimodal bilinguals, numerically contrast coded as +1/2) and Bin (continuous, centered 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 ($\beta = 0.18$, $SE = 0.16$, $z = 1.12$, $p = .26$), but a significant main effect of Bin ($\beta = 0.13$, $SE = 0.01$, $z = 9.88$, $p < .001$), and no significant interaction between Group by Bin ($\beta = 0.003$, $SE = 0.03$, $z = 0.14$, $p = .89$). This insignificant interaction suggests that the lack of a difference between groups remained over time (Figure 4, right panel).

To show that the relative looks over time to grounds versus figures depend on the word order that the participants produced, we additionally analyzed whether hearing bimodal bilinguals and deaf signers look more to grounds when it is mentioned first than when figures are mentioned first (see Appendix D for more information). This analysis confirmed that the order in which hearing bimodal bilinguals and deaf signers mention grounds or figures predicts where they look at most frequently ($\beta = 1.05$, $SE = 0.08$, $z = 12.92$, $p < .001$). This effect did not interact with time, thus the link between word order and eye-gaze progressed similarly over time for both groups ($\beta = 1.15$, $SE = 0.09$, $z = 1.59$, $p < .11$).

Discussion

The present study investigated whether and how different word order preferences in a sign and spoken language influence each other in hearing bimodal bilinguals in a domain where sign languages have a modality-driven word order. We further assessed whether influence of word order preferences between NGT and Dutch in hearing bimodal bilinguals has further cognitive consequences and influences visual attention during message preparation.

Language production

We found that in NGT, deaf signers produced mostly ground-first descriptions while with hearing non-signers there seems to be no clear preference for figure-first or ground-first order. For hearing bimodal bilinguals, in speech, they expressed more ground-first descriptions than hearing non-signers, showing influence from sign. In sign, they used as many ground-first descriptions as deaf signers, demonstrating no influence from speech. Cross-linguistic influence of word order preference in hearing bimodal bilinguals appears to be one-directional and might be modulated by modality-driven differences.

Word order differences in Dutch and NGT (controls)

Results revealed that deaf signers produced more ground-first descriptions than hearing non-signers. This confirms that NGT predominantly prefers ground-first order as found for all sign languages studied to date (e.g., Emmorey, 2002; Kimmelman, 2012; Morgan et al., 2008; Perniss, 2007; Sümer, 2015). Furthermore, we found that both signing groups (i.e., deaf signers and hearing bimodal bilinguals) showed a very strong and robust systematicity in mentioning grounds first in NGT. Taking these findings together, our results strengthen previous research suggesting that ground-first order so-far is likely to be a universal bias based on modality differences.

For hearing non-signers, results showed that in Dutch there is no clear preference for figure-first or ground-first order but, rather, half of the hearing non-signers 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, unlike in NGT but, rather, alternative orders are valid and acceptable (Hartsuiker et al., 1999).

Word order influence from sign to speech in hearing bimodal bilinguals' s descriptions

In Dutch, hearing bimodal bilinguals produced more ground-first descriptions than hearing speaking controls, suggesting an influence of word order preferences across modalities from sign to speech. This is in line with recent results on cross-modal influence in hearing bimodal bilinguals where the speech was influenced by specific iconic expressions in sign (Manhardt et al., 2021). Moreover, this finding aligns with previous research, showing influence of word order from silent gesture comprehension to spoken language production in a priming paradigm (Shurley et al., 2018). 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. Furthermore, our finding of word order influence from sign to speech 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). We show that word order preference can be influenced by language contact from another language (NGT) within a (bimodal) bilingual. It is possible that what is being influenced is driven from a non-linguistic cognitive bias towards perceiving grounds as more primary and salient than figures (based on principles of Gestalt principles, e.g., Rubin, 1915, 1958). This would also be in line with previous research showing that hearing non-signers preferred the same word order in silent gestures, suggesting that certain word orders in the visual modality might be a more general product of communicating in the visual-manual modality (Gershkoff-Stowe & Goldin-Medow, 2002).

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 (semi)naturalistic picture description setting without experimentally inducing 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 et al., 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 speech in our hearing bimodal bilingual sample seems to vary – that is, in Dutch, not all hearing 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.

No word order influence from speech to sign in hearing bimodal bilinguals' descriptions

When signing in NGT, hearing bimodal bilinguals did not differ from the deaf signing controls in their ground-first preference. This indicates no influence from sign to speech. This reveals that cross-linguistic influence of word order preference in hearing bimodal bilinguals is one-directional. This one-way influence occurred independent 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.

Interestingly, influence from speech to sign in hearing bimodal bilinguals has been evident in previous research (Manhardt et al., 2021). That study investigated the domain of iconicity where there is variation in linguistic choices in sign. However, in the present study word order preference in NGT might not be as variable as in Dutch due to the non-linguistic cognitive bias that seems to motivate ground-first order in sign. Hence, cross-linguistic influence from speech to sign 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 3 indicates that not all deaf signers produced ground-first utterances and that some of the hearing 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 hearing bimodal bilinguals might be modality-specific as cross-linguistic influence might be motivated by the modality-driven robust ground-first order rather than due to linguistic constraints of NGT.

Visual attention

For all three groups, we found that eye-gaze preferences to look at grounds or figures during message preparation aligns with the order they mention grounds and figures in their linguistic descriptions. This conforms with previous claims that during language production non-signers look first at the referent that is mentioned first (e.g., Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014). Even more, we show that such links between eye-gaze and word order can be also observed in deaf signers and hearing bimodal bilinguals and do not only arise during language production but already during message preparation.

Eye-gaze differences in Dutch and NGT (control groups)

Our results indicate that deaf signers preferred looking at grounds from the start of message preparation, while for hearing non-signers 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 non-signers. It also 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 is in line with primacy of grounds in their descriptions and reveals that THINKING FOR SPEAKING extends to THINKING FOR SIGNING (Manhardt et al., 2020), also in the domain of word order preferences.

Cross-linguistic influence affects visual attention in hearing bimodal bilinguals

For the hearing bimodal bilinguals, our results provide evidence for cross-linguistic influence of visual attention during message preparation but only when there was also cross-linguistic
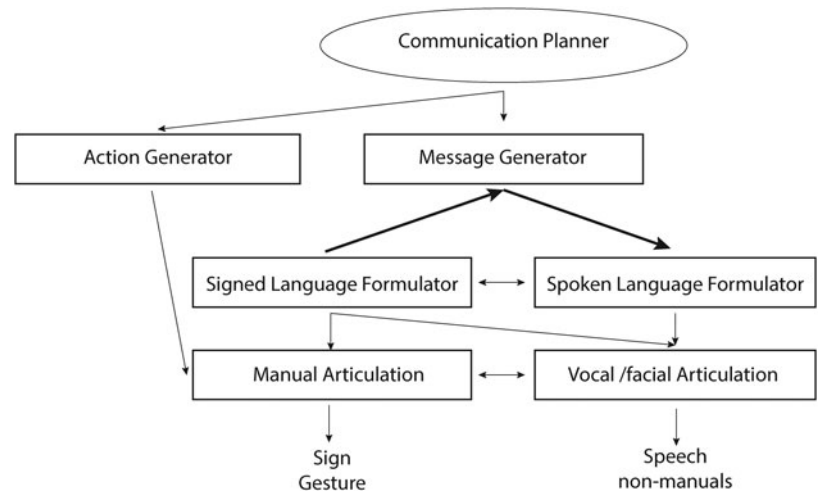
Fig. 5. A model of bilingual language production. Bold arrows indicate where our findings provide additional evidence for the model. Figure adapted from Emmorey et al. (2008).

influence at the level of language production – that is, we found cross-linguistic influence of visual attention 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 hearing bimodal bilinguals and hearing non-signers preferred looking at figures over grounds at the initial stages of message preparation (i.e., at the beginning of the timeline as shown in Figure 4). As message preparation unfolded, both groups developed a preference to look at grounds versus figures over time. However, hearing bimodal bilinguals' ground preference increased more over time compared to their hearing non-signing peers. However, during the time course of preparing NGT messages, hearing 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 center 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, however, differences in eye-gaze preferences in Dutch and NGT emerged from early on when message preparation begun.

Taken together, shifts in eye gaze patterns, motivated by influence from sign to speech, provide further evidence for an existing bimodal bilingual language production model (see Emmorey, Borinstein, Thompson & Gollan, 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 5). Additionally, it involves an Action Generator (a general mechanism for creating action plans) responsible for the production of gestures and which interacts with the Message Generator. Accordingly, we propose that cross-linguistic influences from sign to speech occur via the Message Generator (visualized by bold arrows in Figure 5, see also Manhardt et al., 2021 for a similar proposal for cross-linguistic influences between sign and speech in the domain of iconicity). Because the Message Generator is the place where preverbal 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 visual attention during message preparation. Specifically, ground-first order in the Message Generator used for modality-specific expressions in sign might make grounds more salient to hearing bimodal bilinguals than to hearing non-signers. Thus hearing bimodal bilinguals look at and produce grounds first more often. This then results in cross-linguistic influence from sign to speech as well as changes in visual attention when preparing to speak.

Conclusion

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 hearing bimodal bilinguals. It further showed that influence of word order to describe spatial language in hearing bimodal bilinguals is modality-specific and one-directional and has additional cognitive consequences that go beyond the level of language production modulating visual attention.

References

Ahn D, Gollan TH and Ferreira VS (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.

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

Baayen RH, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**, 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**, 307–340. doi: 10.1177/13670069050090030101

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

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

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

Cleland A and 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**, 214–230. doi: 10.1016/S0749-596X(03)00060-3

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

de Jong NH and Wempe T (2009) Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, **41**, 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**, 355–384. doi: 10.1353/sls.2018.0007

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

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

Emmorey K, Borinstein HB, Thompson R and Gollan TH (2008) Bimodal bilingualism. *Bilingualism*, **11**, 43–61. https://doi.org/10.1017/S1366728907003203

Flecken M, Carroll M, Weimar K and Stutterheim CV (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**, 100–122. doi: 10.1111/j.1540-4781.2015.12181.x

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

Gershkoff-Stowe L and Goldin-Medow S (2002) Is there a natural order for expressing semantic relations?. *Cognitive Psychology*, **45**, 375–412.

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

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

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

Griffin ZM (2004) Why look? Reasons for speech-related eye movements. In Henderson J and Ferreira F, *The integration of language, vision, and action:*

Eye movements and the visual world. New York, NY: Psychology Press. Scopus, pp. 213–248.

Griffin ZM and Bock K (2000) What the eyes say about speaking. *Psychological Science*, **11**, 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**, 3–15. doi: 10.1016/0093-934X(89)90048-5

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

Hall ML, Ferreira VS and Mayberry RI (2015) Syntactic priming in American Sign Language. *PloS One*, **10**, e0119611–e0119611. PubMed (25786230). doi: 10.1371/journal.pone.0119611

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

Hatzidaki A, Branigan HP and Pickering MJ (2011) Co-activation of syntax in bilingual language production. *Cognitive Psychology*, **62**, 123–150. doi: 10.1016/j.cogpsych.2010.10.002

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

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

Kita S and 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**, 16–32. https://doi.org/10.1016/S0749-596X(02)00505-3

Konopka AE and Meyer AS (2014) Priming sentence planning. *Cognitive Psychology*, **73**, 1–40. doi: 10.1016/j.cogpsych.2014.04.001

Kootstra G, van Hell J and 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

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

Kuznetsova A, Brockhoff PB and Christensen RHB (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**, 1–26. doi: 10.18637/jss.v082.i13

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

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

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

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

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

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

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

Manhardt F, Brouwer S and Özyürek A (2021) A tale of two modalities: Sign and speech influence in each other in bimodal bilinguals. *Psychological Science*, **32**, 424–436. doi:10.1177/0956797620968789.

Manhardt F, Özyürek A, Sümer B, Mulder K, Karadölller DZ and Brouwer S (2020) Iconicity in spatial language guides visual attention: A comparison between signers' and non-signers' eye gaze during message preparation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **46**, 1735–1753. doi: 10.1037/xlm0000843

Matin E, Shao KC and Boff KR (1993) Saccadic overhead: information-processing time with and without saccades. *Perception & psychophysics*, **53**, 372–380. https://doi.org/10.3758/bf03206780

Meyer AS, Sleiderink AM and Levelt WJM (1998) Viewing and naming objects: eye movements during noun phrase production. *Cognition*, **66**, B25–B33. doi: 10.1016/S0010-0277(98)00009-2

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

Morgan G, Herman R, Barrière I and Woll B (2008) The onset and mastery of spatial language in children acquiring British Sign Language. *Cognitive Development*, **23**, 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.

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

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

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

Pichler DC, Lillo-Martin D and Palmer JL (2018) A short introduction to heritage signers. *Sign Language Studies*, **18**, 309–327. doi: 10.1353/sls.2018.0005

Pickering MJ and Ferreira VS (2008) Structural priming: a critical review. *Psychological Bulletin*, **134**, 427–459. (18444704). doi: 10.1037/0033-2909.134.3.427

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

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

R Core Team. (2013) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rubin E (1915) *Synsoplevede Figurer*.

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

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

Sandler W and Lillo-Martin D (2006) *Sign language and linguistic universals*. Cambridge: Cambridge University Press.

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

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

Shurley J, Schouwstra M and Pickering MJ (2018, January) Structural Alignment in cross-modal priming: Linguistic representation is shared between and gesture and speech. In EvoLang XII: Evolution of Language: The 12th International Conference.

Slobin DI (2003) Language and thought online: Cognitive consequences of linguistic relativity. In Gentner D and Goldin-Meadow S (eds), *Language in mind: Advances in the study of language and thought*. MIT Press, pp. 157–191.

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.

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

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

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

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

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

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

Appendix

Appendix A: Information about deaf signing controls' language background

Self-rated scores (means, SD in parentheses) 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)

Appendix B: Information about hearing non-signing controls' language background

Language background from hearing Dutch non-signers 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 non-signers		
Most frequent other languages	AoA L2	Frequency of L2 use
English, German	13.56 (7.38)	3.10 (0.85)

Appendix C: Information about session order analysis in hearing bimodal bilinguals

We assessed possible effects of session order on the proportions of ground first-descriptions in hearing 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 ($\beta = -0.05$, $SE = 3.13$, $z = -0.02$, $p < 0.99$). For NGT, the model contained also no main effect of Session ($\beta = -0.65$, $SE = 3.51$, $z = -0.18$, $p < 0.85$).

Appendix D: Information about the analysis linking word order to eye-gaze preference

We used two general linear mixed-effects regression model, both assessing whether the relative looks to the ground vs. figure depend on the word order that the participants produced. For Dutch, we investigated fixations to grounds (1) or figures (0) using a general linear mixed-effects regression model with Group (hearing non-signers, numerically contrast coded as -1/2 vs. hearing bimodal bilinguals, numerically contrast coded as +1/2), Word Order (ground-first: 1, figure-first: 0) and Bin (continuous, centered 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 a significant main effect of Word Order ($\beta = 0.98$, $SE = 0.49$, $z = 20.21$, $p < .001$) and Bin ($\beta = 0.24$, $SE = 0.01$, $z = 20.95$, $p < .001$), significant interactions between Group by Word Order ($\beta = 0.36$, $SE = 0.97$, $z = 3.68$, $p < .001$), Group by Bin ($\beta = 0.14$, $SE = 0.02$, $z = 6.25$, $p < .001$) and Worder Order by Time ($\beta = 0.36$, $SE = 0.97$, $z = 3.68$, $p < .001$). There was also a significant interaction between Group by Word Order by Time ($\beta = -0.37$, $SE = 0.02$, $z = -15.92$, $p < .001$). For 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. hearing bimodal bilinguals, numerically contrast coded as +1/2), Word Order (ground-first: 1, figure-first: 0) and Bin (continuous, centered 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 a significant main effect of Word Order ($\beta = 1.05$, $SE = 0.08$, $z = 12.92$, $p < .001$) and Bin ($\beta = 0.21$, $SE = 0.02$, $z = 8.97$, $p < .001$), but no significant interactions between Group, Bin and Word Order.