Nonmanual Marking of Questions in Balinese Homesign Interactions: a Computer-Vision Assisted Analysis

Vadim Kimmelman[1], Ari Price[2], Josefina Safar[3], Connie de Vos[3], Jan Bulla[1,4]
[1]University of Bergen, [2]Western Norway University of Applied Sciences,
[3] Tilburg University, [4]University of Regensburg
[1,2]Bergen, Norway, [3] Tilburg, the Netherlands, [4]Regensburg, Germany,
{vadim.kimmelman, jan.bulla}@uib.no, julie.anna.price@hvl.no,
{C.L.G.deVos, J.Safar}@tilburguniversity.edu

Abstract

In recent years, both linguistic resources and computer-based tools have been developed that make it possible to investigate research questions that have not been studied before. In this study, we conduct a study of nonmanual question marking, using data from the Balinese Homesign Corpus – a unique resource documenting language use in several Balinese homesigners. We further demonstrate how using OpenFace, a Computer-Vision solution, allows for quantitative analysis of head tilts used by these signers in marking questions. We also showcase a pilot statistical analysis of the dynamic kinetic contours of the head movements.

Keywords: Balinese Homesign Corpus, questions, nonmanual marking, head tilt, OpenFace

1. Introduction

In recent years, both linguistic resources and computer-based tools have been developed that make it possible to investigate research questions that have not been studied before. In this study, we illustrate such a linguistic resource, namely the Balinese Homesign Corpus, containing unique data of several homesigners in conversations with their family members and other signers. This data set allows investigating nonmanual marking of questions in this special population, that is, deaf individuals who grew up without access to an already existing signed language. As we will discuss below, the homesigners are not in contact with each other, so their homesign systems are potentially completely distinct. At the same time, they co-create their homesign systems with their hearing family members, who are representatives of the local hearing community. Therefore, it is possible that the homesign systems will partially converge, also in the domain of question marking, due to the influence of the gestures of the hearing non-signers.

In addition, in this paper we explore the use of one of the relatively novel tools from the Computer Vision domain, namely OpenFace (Baltrusaitis et al., 2018). This tool allows identification and tracking of the head and the facial features, including measurements of the rotation of the head. Since, as we show, head pitch (up and down movement) are important markers of questions in our data set, we use the tool to measure pitch and correlate it with specific question types across the different signers in the corpus.

Finally, using a subset of the data, we show how OpenFace measurements can potentially be used to study dynamic kinematic properties of head movements in more detail, using various smoothing techniques. While this type of analysis will require extensive follow-up research, we demonstrate the promise that it has.

1.1. Question Marking in Sign Languages

Question marking has been studied for many sign languages (Zeshan, 2004; Cecchetto, 2012). Almost universally, nonmanual markers are employed in marking questions of different types, especially for polar questions (ibid.). For content questions, in many sign languages question words are used, often also accompanied by nonmanual markers. Quite strikingly, in a majority of sign languages, polar questions are accompanied by raised eyebrows, while content questions have more diverse patterns of marking. In addition, many studies report different types of head tilts marking for questions (Cecchetto, 2012). In her thesis about an emerging sign language in Brazil, Fusellier-Souza (2004, 304) specifically categorizes eyebrow raises and head tilts as modality-specific traits that function as non-manual features of question marking, in addition to expressing doubts and uncertainty, across sign languages. Very little is known about question marking in homesigners, with the exception of a case study on David, a child homesigner from the US, who has been reported to only use a manual flip gesture to mark wh-questions (Franklin et al., 2011). Based on the existing research, we thus decided to focus specifically on potential nonmanual markers in the homesign data described below.

Proceedings of the 11th Workshop on the Representation and Processing of Sign Languages, pages 168–177

1.2. Homesigners

Homesign is a visual-gestural communication system that is co-created by a deaf person who does not have full access to a conventionalized language and the attentive interlocutors in their proximity (De Vos, 2023). Observing homesign systems allows for an opportunity to view aspects of emerging linguistic systems that can provide insights into human language development. Each homesign system can have unique features, as with more conventionalized languages, but there are some elements that have been considered resilient properties of language (Brentari and Goldin-Meadow, 2017). What classifies these features as resilient is that they show up across different home sign systems, which all lack a distinct input from a preexisting language model. However, the form in which the functional feature may present itself can vary from one homesign system to another. This paper seeks to explore the forms that different homesigners, who are not in contact with each other but come from the same cultural background, use to mark question types with nonmanual markers typically found across sign languages. In particular, we observe the use of upward and downward head tilts across question types in conversational data between homesigners and their interlocutors.

1.3. Computer Vision Analysis of Nonmanuals

An important goal of this study is to test the applicability of Computer Vision tools to linguistic analysis of nonmanuals. In recent years, due to success of the Deep Learning approach to Computer Vision, several toolkits for detecting and tracking body landmakrs in video recording have appeared, including OpenPose (Cao et al., 2018) and MediaPipe (Lugaresi et al., 2019). Some of the tools also include automatic 3D reconstruction from 2D video recordings, and tracking head rotation, such as OpenFace (Baltrusaitis et al., 2018), which we use here for this reason.

First studies using Computer Vision tools to analyze sign languages have appeared over 10 years ago (Metaxas et al., 2012; Karppa et al., 2014). However, due to the relative user-friendliness of the new tools and their increased reliability and efficiency, in recent years, a large number of publications applying them to sign language and gesture data has appeared (see for example Östling et al. 2018; Trujillo et al. 2019; Fragkiadakis 2022; Börstell 2023), also for analyzing nonmanual markers (Kimmelman et al., 2020; Chizhikova and Kimmelman, 2022). While the use of these tools for sign language analysis is very promising, extensive testing and calibration of these tools is required (Kuznetsova et al., 2021); at this stage, it is necessary to combine these tools with manual annotations, as we do in the current study.

2. Methodology

2.1. The data

This data set consists of 5 videos from the Balinese Homesign Corpus. The videos contained recordings of 11 people who had experience using a homesign system. The participants were 5 prelingually deaf homesigners, who do not have input from an adult sign language model, 5 hearing interlocutors and 1 deaf interlocutor with knowledge of a conventionalized sign language (Indonesian Sign Language (BISINDO)). All homesigners and their interlocutors in our data set come from the Buleleng regency in Northern Bali, Indonesia. Due to their regional proximity, the homesigners and their interlocutors have a similar cultural background , which includes shared knowledge of locations, rituals, and traditional family systems among other norms. Having a common culture also gives these homesigners access to the gestural repertoire affiliated with the larger local community of hearing speakers of Balinese. The data was collected by a team of hearing and deaf research assistants in Bali. Each conversation was filmed with two Canon HF G50 cameras and conversations lasted from 10-50 minutes. This resulted in a total of 02:24:45 worth of video footage.

Figure 1: Homesigner HS01 (right) asking a polar question, in conversation with her mother.

Most of the deaf homesigners were filmed having a conversation with a hearing relative, such as their mother, sister, or sister-in-law, with one having two hearing relatives present (see Figure 1 for an illustration). One deaf homesigner was filmed signing with another deaf person, a man from a neighbouring village who he was not related to, but had met several times previously. This deaf man attended deaf school for a number of years, where he acquired BISINDO. Thus, all deaf homesigners without long-term formal schooling interacted with an interlocutor who knew a conventionalized language in addition to using homesign. However, participants differed in terms of their ages (27-53), professions, and marital statuses, which influenced conversation topics. For more detailed information, please see Safar and De Vos (2022), who used the same data set.

2.2. Annotation

The videos were annotated in ELAN 6.7 (Crasborn and Sloetjes, 2008). Previous annotations done by Safar and De Vos (2022) provided an English translation tier that acted as a baseline to pinpoint questions in the video data set. Expanding upon the original files from Safar and De Vos (2022), a tier was added to mark where questions came up in the interactions of homesigners. On this tier, question types were then marked as being 1 of 4 types: 'polar,' 'open,' 'content,' or 'huh'. While 'huh' is not necessarily a question in and of itself, it proved to have a fairly consistent form across homesigners and provided a similar function to an open question by prompting the interlocutor to give more information. The category of 'content' question was used when a manual sign (question word) was used, while 'open' question do not contain a manual question word, but instead a gap in place of one of the constituents, and presuppose the answer to fill in this gap.

After marking the question types, the 'NMM-annotation-template.etf' template created by Oomen et al. (2023) was imported into the original ELAN files to allow for the consistent annotation of nonmanual markers across homesigners. Following Oomen et al. (2023) , the nonmanual markers in each question were annotated, with special attention given to head position and eyebrows. In particular, up and down head movements were marked on the 'NMM.head-y' tier and more rapid movements were marked as nods on the 'NMM.head-move' tier. Raised, neutral and lowered eyebrow movements were also then marked on the 'NMM.eyebrows' tier. In order to explore the nonmanual markers of these signers as individuals, separate files were made for each signer in the data set, except for a 'third participant' in one video that did not actively participate in the conversation.

All the new annotations for this study were created by one of the authors, AP. Another author, VK, reviewed the annotations, and AP and VK discussed all the instances of disagreement.

2.3. Computer Vision Processing

We extracted the clips containing up and down head movements based on the annotation on the 'NMM.head-y' tier, using the `split_elan_videos` script (Börstell, 2022). Because the video recording contained two or three signers simultaneously, we cropped the clips to have only one signer in one clip with `ffmpeg`, (Tomar, 2006). The details can be found in the RMarkdown document in the repository linked below.

The clips were then analyzed in OpenFace (Baltrusaitis et al., 2018). OpenFace is a toolkit for face landmark detection, head pose estimation, and facial action unit recognition. Most relevant for this project is that OpenFace measures per frame head rotation along three axes (pitch, roll, and yaw) in radians. Up and down head movement is essentially pitch rotation, labeled as `pose_Rx` in OpenFace. We use the `pose_Rx` to measure head movements.[1] OpenFace also estimates confidence of the measurement (once per frame), and so we filter out the data points with confidence below 0.9.

2.4. Statistical Analysis

The full documentation of the statistical analysis and the data files used for the analysis can be found in this repository: `https://osf.io/5d7wu/`.

2.4.1. Analyzing the Annotations

As the first step, we graphically explore the relations between question type and the nonmanual markers (eyebrow movements, head movements, head pitch), both overall and for individual signers. We also compare the deaf signers to their hearing interlocutors. The analysis was conducted in R (R Core Team, 2022) with RStudio (Posit team, 2024), using `tidyverse` (Wickham et al., 2019) and `ggplot2` (Wickham and Chang, 2016).

2.4.2. Analyzing OpenFace Outputs

We used OpenFace to extract measurements of head pitch (`pose_Rx`). First, we investigated the relation between our annotations for head movement and the measures outputted by OpenFace in order to see whether they generally agree. After establishing that this is indeed the case, we investigated the relation between the OpenFace pitch

[1] OpenFace also tracks the eyebrows and even automatically detects eyebrow raise. However, as previous research has shown, these measures are very unreliable in the presence of head tilts (Kuznetsova et al., 2021). We have also tested them with this data set and came to the same conclusion: eyebrow measures from OpenFace cannot be used for linguistic analysis, at least not for question marking.

measurements and our question type annotations. We used the same tools as above, with addition of the lme4 package (Bates et al., 2015) for mixed effect regression.

2.4.3. Analyzing Specific Dynamic Movements

The head movements (both up and down) are dynamic movements, and our long-term goal is to investigate them as such, and not as average measures per movement as in the previous section. In order to start developing this approach, we have selected 24 up and down movements produced by HS01 and investigated them further.

The observations collected by the variable pose_Rx represent a continuous movement, but are of discrete nature. In addition, one has to assume that the recorded movements contain a certain amount of noise from the recording process. Last and maybe most importantly, head movements do not always follow a precise identical patterns, but may overlap with smaller movements which are – in principle – negligible.

We therefore process the head movement observations with three different non-parametric statistical methods that permit to detect general patterns in noisy data: locally estimated scatterplot smoothing (LOESS), kernel regression, and splines using the statistical software package R (R Core Team, 2021), version 4.05.

LOESS smoothing (Shyu and Cleveland, 1992) is based on local polynomial regression. It is available through the function loess, which uses polynomials of degree two by default. Moreover, LOESS requires input of the span parameter, which controls the degree of smoothing. We determined this parameter by 10-fold cross-validation with mean absolute error as criterion.

The core of Kernel regression is the Nadaraya–Watson estimate estimator (Watson, 1964; Nadaraya, 1964), available in R within the np package (Helwig, 2021). This estimator relies on input of an optimal bandwidth parameter, which determines the degree of smoothing. We chose Kullback-Leibler cross-validation (Hurvich et al., 1998) in the npregbw function for this task, because the default least-squares cross-validation turned out to be too wiggly.

A wide range of implementations exists for spline regression. All have in common that the shape of the resulting function mainly depends on the number (and placement) of knots, and a smoothing parameter. We considered i) the function ss from the np package with the default generalized cross-validation for choosing the smoothing parameter; ii) the gam function with default settings from the gam package (Hastie and Tibshirani, 1990; Chambers and Hastie, 1992); iii) p-splines via the function gam from the mgcv package.

3. Results

3.1. General findings

In total, we annotated 296 examples of questions in the data. However, the data is very unbalanced. First, 215 (73%) of the questions are polar questions. Second, different individuals produced drastically different numbers of examples. In fact, all but one examples of the *huh?* type were produced by a single hearing participant, and most examples of open questions were produced by another hearing participant. It is therefore difficult to make any generalization based on this data. However, it is important to remember that the individuals in the data set do not represent a population of users of a single language. Instead, each homesigner (possibly with their hearing family members) represents a completely unique system. If we discover some general tendency despite the skewed data sample and despite the potential differences between the homesign systems, it is even more surprising.

3.2. Annotation-Based Analysis

We explore the patterns of nonmanual marking by plotting the annotations for eyebrow movement, head movement and head pitch in relation to the type of question they overlap with.

In Figure 2 we can see the eyebrow movement patterns across the question types and the different signers.[2] It is clear that there is great variation between the signers. Focusing on the signers with the most data, HS01 deaf signer raised her eyebrows consistently for both polar and content questions; HS10's hearing conversational partner raised the eyebrows in open and content questions, but less frequently so for polar questions, and the HS17's deaf conversational partner (Deafb on the Figure) basically did not raise his eyebrows at all. One general pattern that emerges from these signers is that eyebrow marking is more varied for the polar questions than for the other types.

In Figure 3 we can see that again, there is a lot of variation between the signers, but something that is noticeable is that almost all the signers use head nods for polar questions, and less so for the other

[2]On this and the following Figures, the codes for individual signers consist of two parts. The first part refers to the conversation code in the corpus; the second part specifies whether the signer was deaf or hearing, with additional letters distinguishing the two deaf signers in one of the conversations.

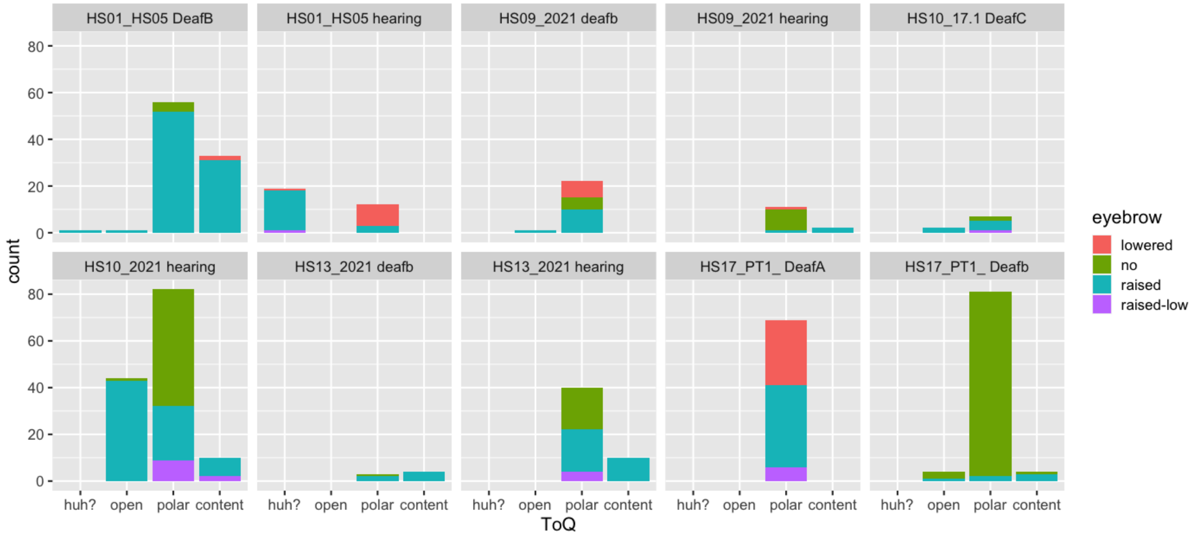

Figure 2: Eyebrow movements across question types, for each individual signer.

Figure 3: Head movements across question types, for each individual signer. Most relevant colors: brown: nods, pink: no movement, purple: headshake.

types.[3] One deaf signer (HS17) uses some headshakes in polar questions, which is explained by the fact that these are questions containing negation.

Finally, in Figure 4, despite the variation, we can see an interesting pattern emerging: the up movement is almost never used for polar questions (which use a lot of pitch down, or no pitch), but very dominantly for the other types. This is even more clear when the data is aggregated for all the signers in Figure 5.

In addition, we have explored whether there is a relation between the markers used by the deaf vs. hearing signers. Overall, we do not find clear differences. One noticeable difference is that, proportionally, the deaf signers had more neutral brow positions in polar questions and produced less down movements, but this is mostly driven by a single signer, as can be seen in Figure 4.

3.3. Computer-Vision-Based Analysis

After extracting the pose_Rx (pitch) measurements with OpenFace, as the first step we analyzed the relation between our annotated categories for pitch (up vs. down vs. neutral labels). The results are visualized in Figure 6.

Thus, as expected, the cases which we annotated as head pitched up have a higher average measurement of pitch in OpenFace than the cases annotated as pitch down, and the neutral cases are

[3]For the clarification of the other less frequent labels, see Oomen et al. (2023)

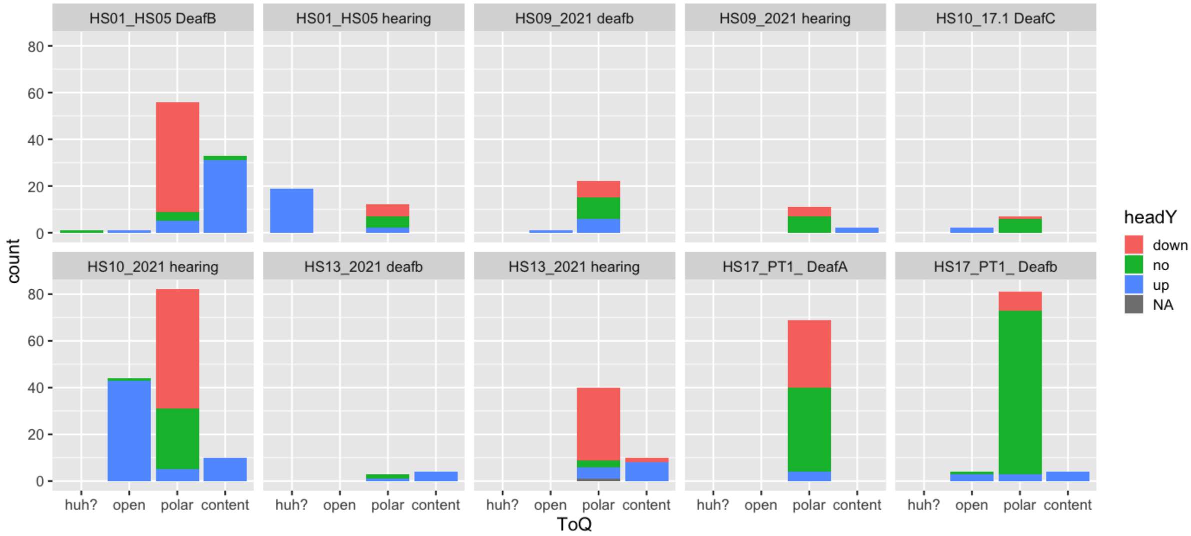

Figure 4: Head pitch across question types, for each individual signer.

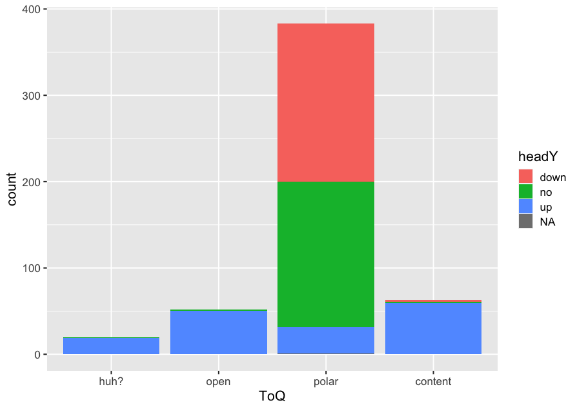

Figure 5: Head pitch across question types, aggregated.

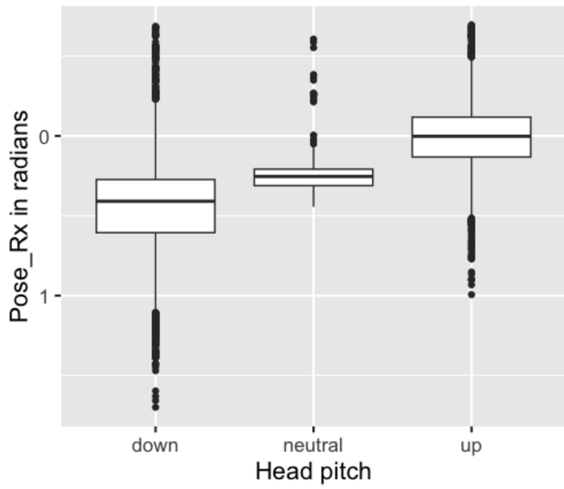

Figure 6: Relation between pose_Rx and manual annotations for head movement (pitch), aggregated over all the signers. Points beyond the ± 2SD removed for visualization purposes.

in the middle.

For a more insightful analysis, we also visualize the relation between our annotations for question types, and the overlapping measurements of pitch_Rx from OpenFace. This is represented in Figure 7.

It is clear that what we find based on our manual annotation is also very visible based on the OpenFace measurements: polar questions on average have a much lower head pitch (the head is moved down), while the other types have a higher pitch. The differences between polar vs. open and polar vs. content are highly significant (polar vs. open estimated difference 0.4 rad, $p < 0.001$, polar vs. content estimated difference 0.32 rad, $p < 0.001$), while the difference between polar and huh? is not significant (most likely because almost all instances of huh? are produced by a single signer).

Importantly, the same pattern is visible for the individual signers, modulo the fact that not all of them have all the question types present.

Thus, the measurements of head pitch from OpenFace produce results agreeing with our observations: polar questions are consistently marked by head down, while the other types of questions are marked with the opposite head movement.

3.4. Head Movements as Dynamic Patterns

We selected four typical head movements of varying duration to illustrate the performance of the three smoothing approaches described in Section 2.4.3. Figure 8 illustrates these four movements: the two top panels show a simple upward

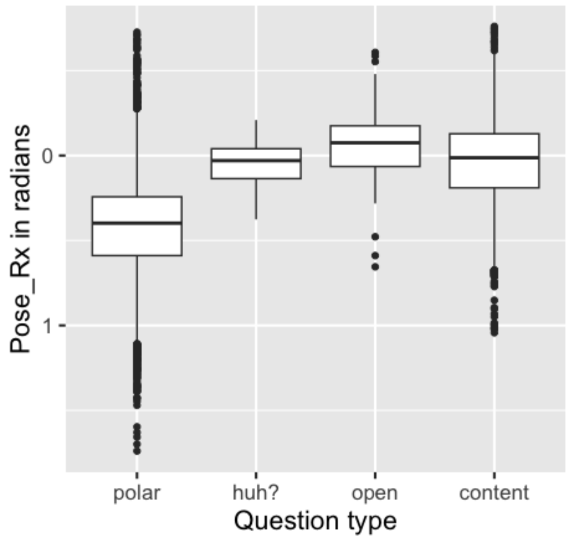

Figure 7: Relation between pose_Rx and manual annotations for type of question, aggregated over all the signers. Points beyond the ± 2SD removed for visualization purposes.

and downward nod, respectively. The two bottom panels present two slightly more complex movements consisting of a double and a multiple nods, respectively. For the simple movements, no large differences are visible between the smoothing methods. Splines obtained from the `gam` package exhibit the highest degrees of smoothing, while those resulting from the `npreg` adapt very (too) closely to the observations. Between these two cases lie the remaining methods, which visually do not differ substantially from each other. The more complex cases in the lower panel paint a more distinct picture: the splines from the `gam` and `mgcv` package do not capture the extent of the movement dynamics sufficiently, in particular for multiple tilt example. This example also illustrates a slight over-smoothing of the LOESS method. However, kernel regression and splines from the `npreg` package reproduce the movement dynamics well, where the latter again provides the highest fit to the observations.

4. Discussion

In this study, we aimed to investigate nonmanual marking of questions produced by Balinese homesigners and their family members in free conversation. We had three main goals: to provide a first description of such marking with attention to variation and similarities between the different signers, to test and explore using OpenFace as a tool to measure head movements in this type of data, and to start exploring analyzing head movements as dynamic patterns using these measurements.

Concerning the first goal, we found a rather interesting pattern. The signers, while not representing a single language, show some degree of convergence on the nonmanual strategies in marking different question types. The eyebrow movements show the most diversity between the signers. This is surprising given the prevalence of eyebrow raise used for polar question marking across different sign languages (Zeshan, 2004). However, the head movements, especially analyzed in terms of head pitch direction (up vs. down) show a surprisingly strong pattern which is similar between all the signers. Specifically, all the signers (both deaf and hearing) mark polar questions with downward pitch, while the other types are more characterized by upward pitch.

The most natural explanation that can be offered for this pattern is that head pitch is used for question marking in similar ways by the surrounding hearing community. This would naturally lead to the hearing family members using these nonmanuals also when signing with their deaf homesigner relatives. It is possible to hypothesize that, for the deaf homesigners and their relatives, the nonmanuals might undergo regularization and become partially obligatory due to their importance in communication. Note that it is clear that head tilt is not universally in other hearing communities, see for example Sze (2022) comparing head tilts in Cantonese speakers with Hong Kong Sign Language signers. So, further research on the nonmanual marking of questions among the surrounding hearing community is required to test this hypothesis.

As for the second goal, it turns out that using OpenFace for analyzing head pitch works very well, at least when averaging the pitch for individual instances of nods/tilts. The measurements agree with our pitch annotations, and there is a strong relation between our annotation for question type and the pitch measurements. Thus, we see an agreement between the manual annotation method and the Computer-Vision based method. Neither method can be considered fully reliable or the ground truth, but it can be a useful methodological improvement to compare and complement the two methods.

When inferring movement dynamics, we observed quite different degrees of smoothing by the various considered methods (see Figure 8). Hereby it should be noted that we mainly relied on default settings of the respective `R` packages, particularly for the spline regressions. The results suggest that the default number of knots is set too low in both the `gam` and the `mgcv` package, and potentially too high in the `npreg` package. In a couple of additional experiments (not shown) we investigated the effects of modifying various settings in the different packages. It turned out that the spline type has very little effect in our examples, whereas – as to

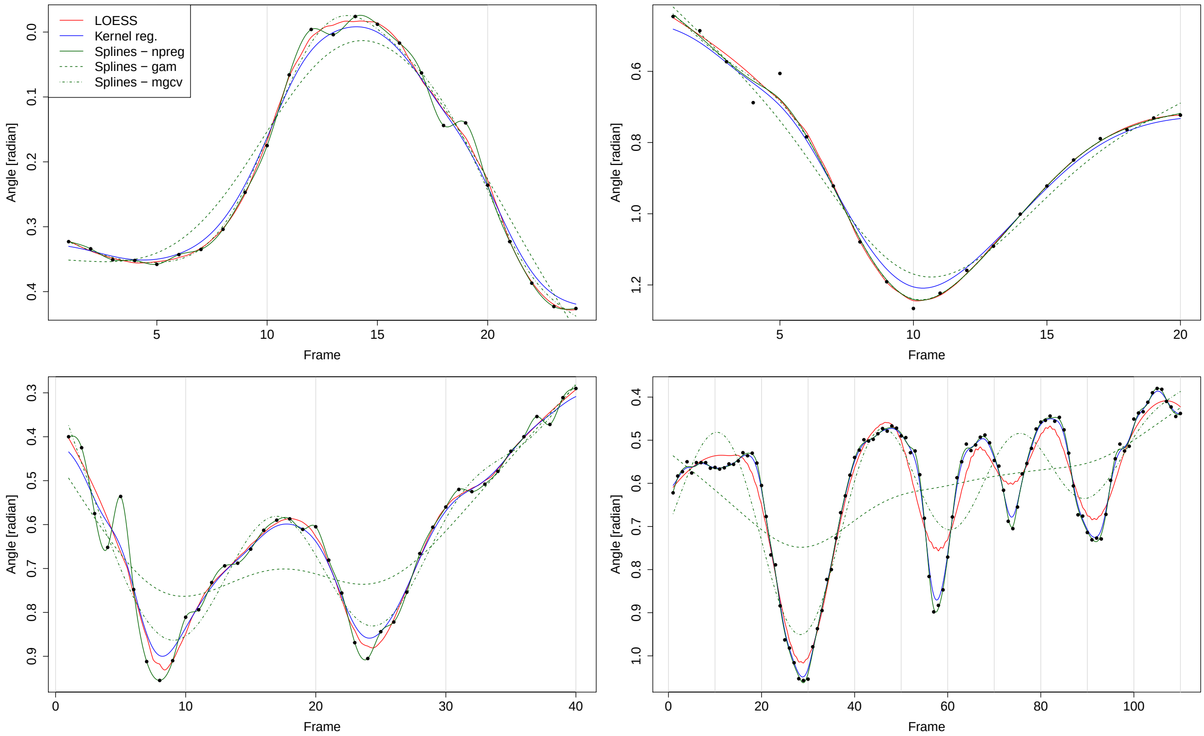

Figure 8: Typical head tilts / movements with inferred dynamics. The x- and y-axis show the frame and head tilt angle, respectively. The y-axis is mirrored for better interpretabilty. Black dots correspond to the observations, and lines result from fitted models (red: Loess, blue: kernel regression, green: spline regressions).

be expected – the knot number strongly affects the degree of smoothing. Hence, it remains to be investigated whether the performance of the spline regressions can be improved by e.g. optimizing the number of knots through cross-validation or model selection techniques. Furthermore, Kernel regression seems to satisfactorily capture the dynamics in all examples. Last, the cross-validation criteria of LOESS may also be improved, which could help to better describe the most dynamic movements.

Aside from these rather technical aspects, it also remains to investigate how the inferred movement dynamics should be post-processed. Analysis of the inferred curves from Figure 8 should be relatively straightforward by measures such as number of extreme points, duration of movements, or distances between extreme points, to name only a few. Challenges appear, however, from less clear sequences of movements such as displayed in Figure 9. This downward nod between approximately Frame 15 and Frame 35 constitutes the main dynamics of these observations. Problematic are several additional extreme points (Frames 13, 37, 47), which complicate the processing of such a sequence. Either these extremes are not real movements, or they are, but they should not be classified as nods in the linguistic sense. The development and testing of suitable methods constitute topics of ongoing research.

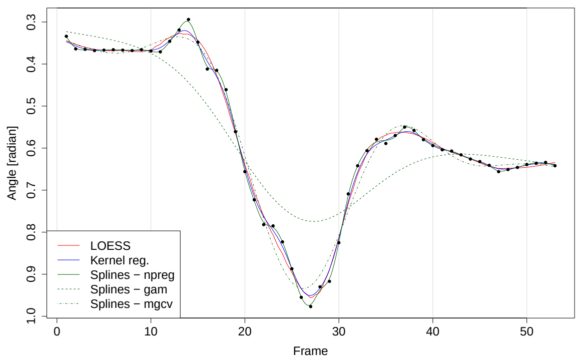

Figure 9: A more complex head tilt / movement with inferred dynamics. The x- and y-axis show the frame and head tilt angle, respectively. The y-axis is mirrored for better interpretability. Black dots correspond to the observations, and lines result from fitted models (red: Loess, blue: kernel regression, green: spline regressions).

Author Contributions

Vadim Kimmelman: Conceptualization, Funding Acquisition, Methodology, Investigation, Formal

Analysis, Visualization, Software, Writing. **Ari Price**: Methodology, Investigation, Writing. **Josefina Safar**: Methodology, Data Curation, Writing. **Connie De Vos**: Methodology, Data Curation, Writing. **Jan Bulla**: Formal Analysis, Visualization, Writing.

Acknowledgements

This project is funded by the European Union (ERC, Nonmanual, project number 101039378, awarded to V. Kimmelman). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

5. Bibliographical References

Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 59–66. IEEE.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48.

Diane Brentari and Susan Goldin-Meadow. 2017. Language Emergence. *Annual Review of Linguistics*, 3(1):363–388.

Carl Börstell. 2022. R functions for working with linguistic data (mapping, ELAN, video processing, etc.).

Carl Börstell. 2023. Extracting Sign Language Articulation from Videos with MediaPipe. In *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 169–178, Tórshavn, Faroe Islands. University of Tartu Library.

Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In *arXiv preprint arXiv:1812.08008*.

Carlo Cecchetto. 2012. Sentence types. In Roland Pfau, Markus Steinbach, and Bencie Woll, editors, *Sign language: An international handbook*, pages 292–315. De Gruyter Mouton, Berlin.

John M. Chambers and Trevor Hastie. 1992. *Statistical Models in S*. Wadsworth & Brooks/Cole Advanced Books & Software. Google-Books-ID: eKOaAQAACAAJ.

Anastasia Chizhikova and Vadim Kimmelman. 2022. Phonetics of Negative Headshake in Russian Sign Language: A Small-Scale Corpus Study. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 29–36, Marseille, France. European Language Resources Association (ELRA).

Onno Crasborn and Han Sloetjes. 2008. Enhanced ELAN functionality for sign language corpora. In *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 39–43. ELRA, Paris.

Connie De Vos. 2023. Cognitive pragmatics: Insights from homesign conversations. *Behavioral and Brain Sciences*, 46:e8.

Manolis Fragkiadakis. 2022. Assessing an Automated Tool to Quantify Variation in Movement and Location: A Case Study of American Sign Language and Ghanaian Sign Language. *Sign Language Studies*, 23(1):98–126.

Amy Franklin, Anastasia Giannakidou, and Susan Goldin-Meadow. 2011. Negation, questions, and structure building in a homesign system. *Cognition*, 118(3):398–416.

Ivani Fusellier-Souza. 2004. *Sémiogenèse des langues des signes. Etude de Langues de Signes Emergentes (LSE) pratiquées par des sourds brésiliens*. Theses, Université Paris 8 - École Doctorale "Cognition, Langage, Interaction" (ED 224). Tex.hal_id: tel-01701214 tex.hal_version: v1.

T. J. Hastie and R. J. Tibshirani. 1990. *Generalized Additive Models*. CRC Press. Google-Books-ID: qa29r1Ze1coC.

Nate Helwig. 2021. npreg: Nonparametric Regression via Smoothing Splines.

Clifford M. Hurvich, Jeffrey S. Simonoff, and Chih-Ling Tsai. 1998. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 60(2):271–293.

Matti Karppa, Ville Viitaniemi, Marcos Luzardo, Jorma Laaksonen, and Tommi Jantunen. 2014. SLMotion - an extensible sign language oriented video analysis tool. In *9th International Conference on Language Resources and Evaluation (LREC 2014)*, pages 1886–1891, Reykjavik, Iceland. European Language Resources Association (ELRA).

Vadim Kimmelman, Alfarabi Imashev, Medet Mukushev, and Anara Sandygulova. 2020. Eyebrow position in grammatical and emotional expressions in Kazakh-Russian Sign Language: A quantitative study. *PLOS ONE*, 15(6).

Anna Kuznetsova, Alfarabi Imashev, Medet Mukushev, Anara Sandygulova, and Vadim Kimmelman. 2021. Using Computer Vision to Analyze Non-manual Marking of Questions in KRSL. In *Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, pages 49–59, Virtual. Association for Machine Translation in the Americas.

Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-Teh Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. 2019. MediaPipe: A Framework for Building Perception Pipelines.

Dimitris Metaxas, Bo Liu, Fei Yang, Peng Yang, Nicholas Michael, and Carol Neidle. 2012. Recognition of nonmanual markers in American Sign Language (ASL) using non-parametric adaptive 2D-3D face tracking. In *8th International Conference on Language Resources and Evaluation (LREC 2012)*, pages 2414–2420, Istanbul, Turkey. European Language Resources Association (ELRA).

E. A. Nadaraya. 1964. On Estimating Regression. *Theory of Probability & Its Applications*, 9(1):141–142. Publisher: Society for Industrial and Applied Mathematics.

M. Oomen, Tobias de Ronde, Lyke Esselink, and Floris Roelofsen. 2023. ELAN template for annotation of non-manual markers.

Posit team. 2024. RStudio: Integrated development environment for R. manual, Posit Software, PBC, Boston, MA.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Josefina Safar and Connie De Vos. 2022. Pragmatic competence without a language model: Other-Initiated Repair in Balinese homesign. *Journal of Pragmatics*, 202:105–125.

William M. Grosse Shyu, Eric and William S. Cleveland. 1992. Local Regression Models. In *Statistical Models in S*. Routledge. Num Pages: 68.

Felix Sze. 2022. From gestures to grammatical non-manuals in sign language: A case study of polar questions and negation in Hong Kong Sign Language. *Lingua*, 267:103188.

Suramya Tomar. 2006. Converting video formats with FFmpeg. *Linux Journal*, 2006(146):10. Publisher: Belltown Media.

James P. Trujillo, Julija Vaitonyte, Irina Simanova, and Asli Özyürek. 2019. Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research. *Behavior Research Methods*, 51(2):769–777.

Geoffrey S. Watson. 1964. Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4):359–372. Publisher: Springer.

Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. 2019. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.

Hadley Wickham and Winston Chang. 2016. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.

Ulrike Zeshan. 2004. Interrogative Constructions in Signed Languages: Crosslinguistic Perspectives. *Language*, 80(1):7–39.

Robert Östling, Carl Börstell, and Servane Courtaux. 2018. Visual iconicity across sign languages: Large-scale automated video analysis of iconic articulators and locations. *Frontiers in Psychology*, 9(725).