Person Identification Based On Sign Language Motion: Insights From Human Perception And Computational Modeling

Félix Bigand
LIMSI - CNRS
Paris-Saclay University
Orsay, France
felix.bigand@limsi.fr

Elise Prigent
LIMSI - CNRS
Paris-Saclay University
Orsay, France
elise.prigent@limsi.fr

Annelies Braffort
LIMSI - CNRS
Paris-Saclay University
Orsay, France
annelies.braffort@limsi.fr

ABSTRACT

Previous research has shown that human perceivers can identify individuals from biological movements, such as walking or dancing. It remains to be investigated whether sign language motion, which obeys to other constraints than pure biomechanical ones, also allows for person identification. The present study is the first to investigate whether deaf perceivers recognize signers based on motion capture (mocap) data only. Point-light displays of 4 signers producing French Sign Language utterances were presented to a group of deaf participants. Results revealed that participants managed to identify familiar signers above chance level. Computational analysis of the mocap data provided further evidence that morphological cues were unlikely to be sufficient for signer identification. A machine learning approach aiming to evaluate the motion features that can account for human performance is currently being developed. First results of the model reveal high accuracy for signer identification based on the same stimulus material, even after having normalized for size and shape. The present behavioral and computational findings suggest that mocap data contain sufficient information to identify signers, and this beyond simple cues related to morphology.

CCS CONCEPTS

• **Computing methodologies** → **Motion capture**; **Machine learning**; **Perception**;

KEYWORDS

Motion Analysis, Perception, Machine Learning, Sign Language, Motion Capture

ACM Reference Format:
Félix Bigand, Elise Prigent, and Annelies Braffort. 2020. Person Identification Based On Sign Language Motion: Insights From Human Perception And Computational Modeling. In *7th International Conference on Movement and Computing (MOCO '20), July 15–17, 2020, Jersey City/ Virtual, NJ, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3401956.3404187

MOCO '20, July 15–17, 2020, Jersey City/ Virtual, NJ, USA

ACM ISBN 978-1-4503-7505-4/20/07…$15.00
https://doi.org/10.1145/3401956.3404187

1 INTRODUCTION

Despite numerous advantages, the use of virtual signers (or signing avatars) is still limited to date. One of the main reasons is that most models designed for signing animations hardly generate natural motion. The advent of motion capture (mocap) systems brought new potentials along this line, enabling animators to replay movements that were recorded with persons. These accurate systems provide more natural and comprehensible motion. They raise, however, another unexpected problem, notably related to person identification. As in the auditory domain, where voice parameters inform about a speaker's identity, a signer's identity could be conveyed by motion. The present study investigates how signer identity can be extracted from mocap.

Recognizing the identity of signers is an important issue in the domain of sign language. Deaf people would like to be able to produce messages anonymously for some applications, such as sharing anonymous testimony on TV or on the web. Moreover, even though written content might allow for some anonymity, it is hardly accessible for deaf persons as it is only a second language for many of them, which is not always mastered [1]. Up to now, current research about person identification in sign language motion remains sparse (if any). It is unclear which motion features are responsible for signer identification from mocap and how these features could be manipulated in sign language animations.

In this study, we (1) evaluated to which extent deaf perceivers actually manage to identify signers from sign language mocap, (2) assessed the features that could be used for identification using computational methods and (3) developped computational models of the involved mechanisms. A summary of these 3 approaches is given in Table 1.

2 THEORETICAL BACKGROUND

2.1 Perception of identity from motion

The recognition of identity from motion information has been initially addressed for biological motion. In Johansson's studies [12] [13], point-light (PL) stimuli have separated information given by dynamic cues from information given by other characteristics such as shape or aspect of the person's body. Using this device, Johansson showed that humans were able to recognize a set of moving dots as a human walker. Based on mocap recordings, different studies then demonstrated that PL displays contain enough information to recognize familiar people from their gaits [8] [15] [28]. These results have then been extended to other movements such as dancing [3],

[1] "American deaf students around age 18 have a reading level more typical of 10-year-old hearing pupils" [10]

Table 1: Three analysis approaches for signer identification from mocap in the present paper, combining (1) human behavior, (2) computational analysis and (3) machine learning, each based on the same motion features with different normalizations and analyzed variables.

	Analysis	Motion features	Normalizations	Analyzed variables
(1)	Perceiving signer identity from mocap	3D global coordinates	Natural motion	Correct identification of perceivers
(2)	Feature analysis of signers' mocap	3D global coordinates	Natural motion	Morphological similarity using PCA
(3)	Computational model for signer identification from mocap	3D global coordinates	- Natural motion - Size-normalized motion - Shape-normalized motion	Correct identification of the model

[15], or clapping in synchrony with music [20]. Recently, Carlson et al. (2018) explored the role of individual differences in music-induced movements [5].

Some research has also aimed for a better understanding of the underlying mechanisms occurring in these kinds of identification. Some studies questioned the role of motor and visual experience in motion perception, reporting higher identification scores when participants viewed their own actions in comparison to those of their friends [15]. Some studies have assessed the respective role of different factors in recognition, such as viewpoint, revealing more accurate identifications for the frontal view than for the profile view [28]. Troje et al. (2005) also evaluated the role of specific features in gait identification by manipulating structural and kinematic cues [28]. They demonstrated that even when participants were deprived of information about shape and gait frequency, their identification performance was still five to six times above chance level. Removing size information did not affect recognition. The low impact of size and shape suggests that most of the information used for identification is contained in motion kinematics. However, gait frequency, which was the only parameter of this kind that Troje et al. (2005) tested, did not seem to play a major role. Manipulations of numerous other parameters are still needed to better understand the role of kinematic cues in identification. Other approaches than human perception measurements can be used to address the question, notably using computational methods, including machine learning.

2.2 Motion analysis using machine learning

Machine learning has been used in several domains aiming to classify categories from a given input. Automatic identification has been developed for various potential applications, such as security, but mainly using voice, images or videos. However, much less is known about how machines could extract identity from sign language mocap data. In combination with basic classifiers or more complex neural networks, mocap data can be used to discriminate between categories or individuals. Troje (2002) investigated the underlying mechanisms of gait perception using Principal Component Analysis (PCA) to classify gender from walking patterns [27]. Candidate features were assessed using a pattern recognition approach. The developed framework also allowed for the synthesis of new gait patterns for which users were able to manipulate the gender attribute. Tilmanne & Dutoit (2010) used a similar approach to learn and then generate stylistic gait, analyzing 11 styles recorded on mocap with PCA and gaussian modeling [23]. This PCA approach takes its inspiration from previous research investigating face classification. These studies have revealed that some eigenvectors extracted from a full face dataset could account for face identification [29]. A recent study presented at *MOCO'*19 on functional PCA analysis of mocap data on different rowing techniques also revealed significant differences between individuals [1].

This modeling framework has also been successfully used to analyze motion in artistic domains. Tits et al. (2015) showed that components emerging from PCA analysis of fingers gesture enable predicting the degree of expertise of piano players [25]. Machine learning was also used to classify emotions from dance movements [4], or to analyse synchronicity between music-induced body movements and music periodic structure [26]. A more recent study revealed that individuals could be classified by this kind of machine learning approach when dancing to music, whatever its genre [6]. Using a Support Vector Machine model, the authors reported that individual classification was achieved at higher scores than genre classification, despite the lower chance level (73 individuals versus 8 genres). These results suggest that a machine could extract motion features that discriminate individuals dancing freely. The aim of the present study is to address this hypothesis for sign language motion, investigating whether a machine could detect and qualitatively define inter-individual differences in signing.

2.3 Motion capture and sign language

Point-light displays also inspired studies investigating sign language (SL) perception, evaluating comprehensibility [18] [22]. Poizner et al. (1981) presented PL stimuli of American Sign Language (ASL) while manipulating specific aspects of the movements. Their findings suggest that the more distal the joint of the body (e.g. fingers), the more information it carries. In Tartter et al. (1981), pairs of participants managed to have discussions in ASL by the only means of 27 lighting spots attached to hand articulations. These moving dots were sufficient to understand one another despite the reduced information and the increased difficulty of the task.

More recent studies have used mocap to investigate motion in sign language from a linguistic approach. The collection of LSF poetic sequences in mocap allowed comparing prosodic variations in spoken and signed languages [7]. Based on a parallel corpus with a LSF mocap version and several French audio versions of the poems, Catteau et al. (2016) have been able to outline the strategies of interpreters to convey prosodic variation. SL mocap corpora have also been used to study how kinematic features of signed movements can be affected by semantics and prosody. Differences in kinematics in ASL have been revealed as being a way to convey specific semantic properties of verb classes [16].

To the best of our knowledge, mocap corpora have not been used yet to investigate person identification in LSF and sign languages in general. The human ability to recognize individuals from motion has only been shown for biological movements, such as

walking or dancing. Compared to biological motion, SL motion is constrained by both biomechanic and linguistic rules. On one hand, the combination of greater constraints could make the perception of identity more difficult than for biological movements. On the other hand, numerous evidence from neuroscience suggest that human language evolved from a cortical system understanding movement. Mirror neurons in the premotor cortex of the macaque monkey were found to respond both when performing an action and when observing the same action performed by another monkey [19]. Imaging data of the human brain revealed that comparable mechanisms can be located in the inferior frontal cortex, notably Broca's area, which has been shown to be involved in spoken language [11] and sign language processing [17]. Person identification might thus be processed differently by deaf people when perceiving signing movements than when perceiving purely biological ones.

3 PERCEIVING SIGNER IDENTITY FROM MOCAP

The present study evaluated the ability of deaf perceivers to identify signing individuals that were presented as PL displays. Point lights were computed from a motion capture dataset in which four different signers freely described pictures in French Sign Language. Short excerpts of these descriptions were randomly presented to the participants. For each excerpt, participants were asked to identify the signer with a four-alternative forced choice.

3.1 Methods

Participants

24 participants (mean age = 42.0, SD of age = 11.0) took part in the study. The Research Ethics Committee of Paris-Saclay University validated the experiment. All participants were deaf and were users of French Sign Language. Language level was self-assessed in the beginning of the experiment. Participants mainly reported the highest levels C1-C2 [2] (79,17%). 12% reported advanced levels (B1-B2). 8.33% reported intermediate level A2.

Stimuli

Major joints of the body were displayed with white dots on a black background. We used MOCAP1 [2], a 3D corpus of mocap data where eight signers each had freely described the content of 24 different pictures in French Sign Language. From this collection, we selected 16 different descriptions performed by four signers. They were chosen so that we could expect familiarity with signers to be gradual. (see Section 3.2).

For the corpus construction [2], a set of 40 reflective markers had been attached to their body and face. Signers had been wearing suits to record markers of the shoulders, elbows, wrists, hands and chest. Four markers had been attached to a cap and had recorded the head movements. No marker had recorded finger movements. We did not include the 13 facial markers in the present study.

Recordings had been done with an Optitrack S250e, equipped with 10 cameras with a spatial resolution of 0.7 Mpixels and a temporal resolution of 250 Hz. From the 27 body markers, we derived 19 virtual markers that optimally describe the major joints of the body for the PL displays. All sensors were presented as 3D global positions in a reference with the pelvis as the origin (see Figure 1). All the stimuli were displayed in front view.

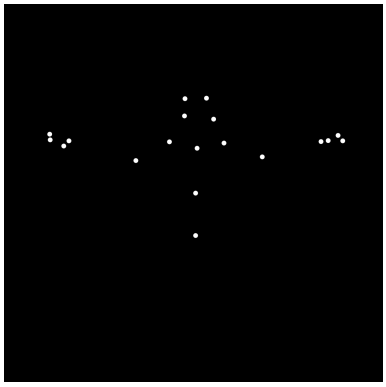

Figure 1: Example of the point-light displays used in the experiment (all in front view).

Design and procedure

The participants took part in the experiment via an online survey (mean duration = 11.04 min). Most participants used a computer (70.83%) rather than a tablet or smartphone.

Before the test session, each of the four signers were presented on the screen with their names and three photos taken from public content (TV, Youtube, conferences...). Participants reported their familiarity with each signer by answering this question : "Have you ever seen this person?". Four levels could be reported : "No, never" (0), "Yes, occasionally" (1), "Yes, often" (2) and "Yes, very regularly" (3). After that, the test session consisted of 16 trials (4 signers x 4 picture descriptions). In each trial, the PL stimulus (mean duration = 10.8s, SD = 2.6) was presented, followed by the presentation of four buttons, illustrated by each signer's photo. Participants were asked to identify the signer in this four-alternative forced choice. All of the 16 stimuli were presented in random order. No response feedback was given to the participants.

3.2 Results

One sample t-tests revealed that identification performance was significantly above chance level (25%) for signer 1 ($t(23)= 8.21$, $p<.001$), signer 2 ($t(23)=4.05$, $p<.001$), signer 3 ($t(23)=2.30$, $p<.05$) but not for signer 4 ($t(23)=1.16$, $p=.26$). Identification scores as a function of the four signers are shown in Figure 2.

A repeated-measure one-way ANOVA was performed with signer (four levels) as within-subjects factor and correct identification as dependent variable. A significant main effect of signer was found on correct identification ($F(3,69) = 12.36$, $p<.001$, $\eta^2=.25$). Bonferroni-adjusted post-hoc tests were performed to test for differences between signers. They revealed a significant increase ($p < .001$) in performance between signer 4 (M=30.2%) and signer 1 (M=65.6%). There is an increase between all four signers (30.2%, 36.5%, 45.8%, 65.6%) but no significant differences were found between signer 4 and both signers 2 and 3.

[2] http://www.visuel-lsf.org/les-niveaux-de-competences-cecrl/

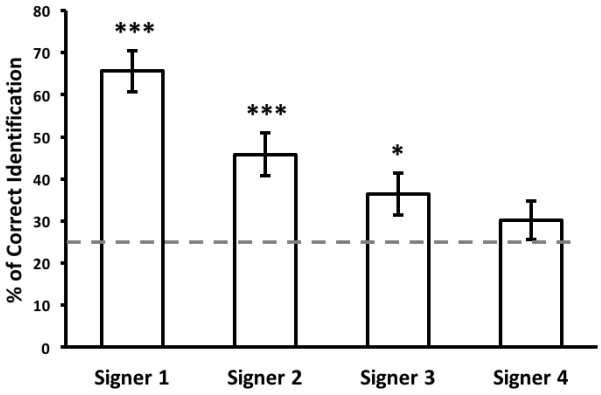

Figure 2: Performance scores from the four-alternative forced choice identification task. Dashed horizontal line indicates chance performance levels. Error bars indicate standard errors. Significant differences from chance level : * (p<.05), * (p<.001).**

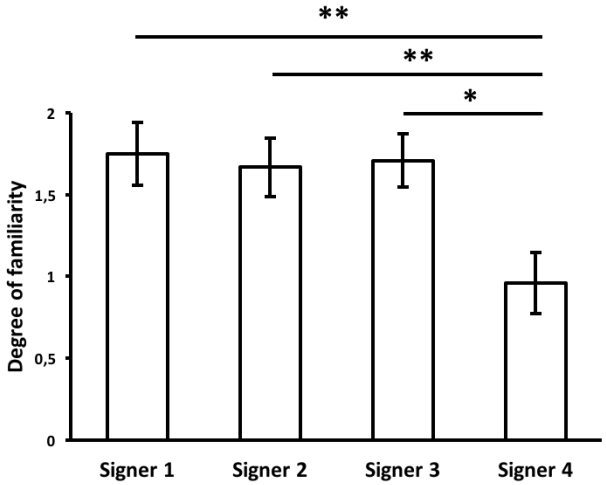

Figure 3: Self-reported familiarity for each signer. Error bars indicate standard errors. Significant differences between signers : *(p<.05), ** (p<.01).

Familiarity with signers was expected to be variable because of their different exposure to deaf people. Signer 1 (M=65.6%) has been a popular story-teller for children and producer for deaf TV shows, being the first deaf person seen on TV in France, in 1979. Signer 2 (M=45.8%) is an LSF translator and actor. He notably worked as a translator for Websourd, a highly popular deaf media. Signer 3 (M=36.5%) is a young LSF journalist working for different media. We were expecting that she would get lower recognition rates as she only recently appeared in the field. Signer 4 (M=30.2%) is involved in some projects on LSF but with lower exposure. She has worked as an LSF translator and trainer but mainly in a local environment. To verify these background differences, we measured self-reported familiarity for each signer and participant. Measures, averaged over participants, are shown in Figure 3.

A repeated-measure one-way ANOVA was performed with signer (four levels) as within-subjects factor and self-reported familiarity as dependent variable. A significant main effect of signer was found on self-reported familiarity ($F(3,69)=6.65$, $p<.001$, $\eta^2=.13$). Bonferroni-adjusted post-hoc tests revealed that familiarity was significantly lower for signer 4 (M=0.96) than for signer 1 (M=1.75, p<.01), signer 2 (M=1.67, p<.01) and signer 3 (M=1.71, p<.05). No significant differences were found between the 3 first signers.

3.3 Discussion

Even though it is not an easy task for participants to evaluate the signers' familiarity, the analysis revealed significant differences, discriminating signer 4 from the three others. This finding provides a partial account for the correct identification scores, which differed from chance level for all signers except for signer 4. This lower value could be explained by the lower exposure of Signer 4 to the general public. In other words, participants managed to identify familiar signers above chance level. It suggests that the mocap data we used in the experiment included sufficient information for participants to identify familiar signers. Further work is needed to explain the differences in correct identification between the 3 first signers, despite equal self-reported familiarity.

The first study by Cutting & Kozlowski (1977) reported performance scores above chance level (16%) but only reaching 38% [14]. Troje et al. (2005) reported 76% correct identification but involving extensive pretraining for the participants [28]. Our results (Figure 2) include all participants' responses, whatever their familiarity with signers. In addition, limitations of the online survey can be discussed. The design of the survey ensured that participants were not able to pause the video or to stop it before the end, but the conditions under which each participant responded to the survey could still vary. Therefore, reaching identification scores such as 65.6% or 45.8% for some signers suggest that their movements provided critical information for identification. Computational feature analysis enabled us to better understand the nature of this information.

4 FEATURE ANALYSIS OF SIGNERS' MOCAP

The excerpts that were presented to the participants were rather short (mean duration = 10.8s, SD = 2.6) and were randomly extracted from the original recordings, regardless of the linguistic content. The data we used did not include facial nor finger sensors. Prior studies have demonstrated that a precise display of the fingertips was needed to ensure comprehensibility in sign language, especially for lexical signs [18]. Other studies pointed out the crucial role of the signer's facial movements during comprehension of ASL [9], notably mouthing and eye gaze. None of these informations was present in the point-light displays presented here to the participants. It is therefore assumed that the identification of signers was achieved beyond simple differences in linguistics.

4.1 Morphological cues

As presented in section 3.1, skeletons of the four signers were displayed as global coordinates for which the pelvis was the origin.

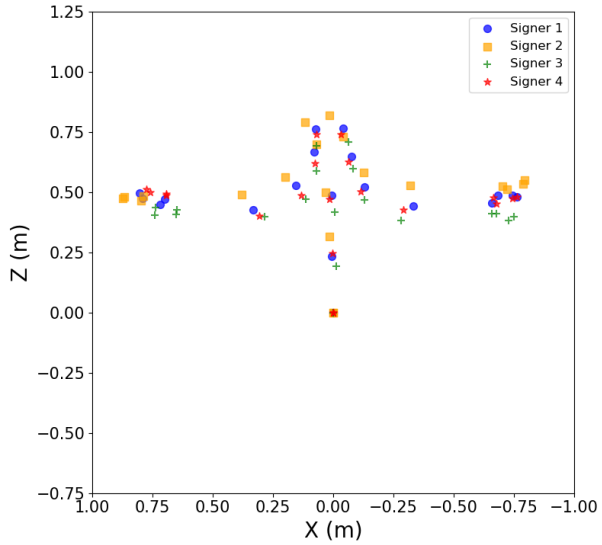

Figure 4: Reference posture of the four signers.

This coding allowed for consistent comparison of the signers who had been placed and oriented differently in the motion capture space. Nevertheless, different morphologies of the four signers were represented in this coordinate system, as shown in Figure 4. To evaluate the role of these differences in identification, we used a PCA approach to compute morphological similarities and compared it with participants' confusion errors between signers.

4.2 Morphology definition using PCA

Our aim was to assess to what extent morphology could account for participants' performance in the experiment. As it can be defined with various indices such as height or shoulder width, we ran a Principal Component Analysis (PCA) to find the most relevant variables to represent morphology. Similarly to Tits (2018) [24], PCA was performed on distance from head to pelvis, distance from hand to hand (in extension), shoulder width, and individual segment lengths (trunk, arm and forearm).

The first principal component accounted for 72% of variance in the data, and it was highly correlated with the distance from head to pelvis ($r(2) = .99, p < .05$). Consequently, this variable was chosen to define morphology. Figure 5 illustrates morphological differences between all signers, using this index. This emphasizes an important gap between signer 2 (highest) and signer 3 (lowest) indices, and specifies a higher index for signer 1 than for signer 4, while both fit in a similar range.

4.3 Results

Based on this morphology factor, a similarity matrix was computed among all signers. Using the Euclidean distance, similarity is computed as follows :

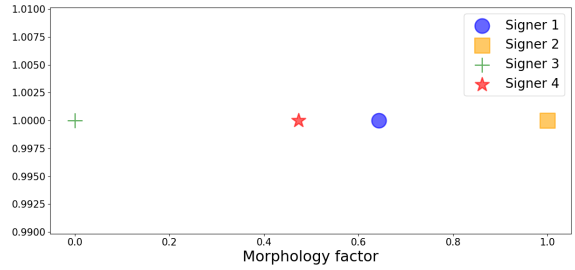

Figure 5: Ranking of the four signers as a function of the normalized morphology factor.

$$s_{i,j} = \frac{1}{1 + \sqrt{(m_i - m_j)^2}} \tag{1}$$

m_i is the normalized morphology factor of signer i.

The computed similarity matrix is shown in Figure 6 (left). Each row represents the amount of similarity with the 3 other signers. According to this representation, signer 2 is more likely to be confused with signer 1 (39%) than with signer 3 (26%). Signer 4 would have equal chances to be confused with both signers 2 and 3. This measure based on computational and statistical analysis allows us to predict confusions in the identification of signers, based on morphological cues. We can compare it to the real confusions of participants in the experiment, as shown in Figure 6 (right). 43% of the confusions for signer 2 are made with signer 3. The less morphologically similar two signers are, the more confused they seem to be.

The Pearson's correlation between similarity matrix and confusion matrix was measured. The resulting coefficient reveals that correlation is not significant ($r(10) = -.36, p = .26$). Taking into account only familiar signers (i.e. the 3 first signers, see Figure 3), correlation between morphological similarities and participants' confusions is also not significant ($r(7) = -.20, p = .61$).

Figure 6: Morphological similarity among signers (left). Participants confusions between signers (right).

4.4 Discussion

The role of morphology in motion perception has been a matter of debate for several types of movements. Sie et al. (2014) proposed a simple skeleton scaling method, by placing the coordinate system on a reference node of the body (i.e. on the pelvis), and dividing all

node coordinates by the torso height [21]. Troje et al. (2005) used normalizations in size or/and shape, using linear regression models [28]. In the specific context of LSF motion, our computational analysis based on PCA revealed that morphological similarities between signers were not correlated with participants' confusions. This suggests that morphology alone cannot account for correct signer identifications. A second human experiment based on normalized stimuli is ongoing, in order to test this hypothesis with behavioral data. We now present a computational model that aims to investigate other candidate features for identification.

5 COMPUTATIONAL MODEL FOR SIGNER IDENTIFICATION FROM MOCAP

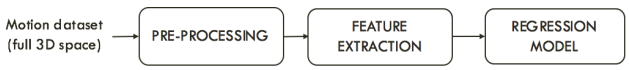

Figure 7: General workflow of the computational model.

Figure 7 illustrates the general workflow of the proposed model for identification. The aim is to extract low-dimensional features from the input, and then discriminate between the signers. The model is designed in a way that allows for generalization to various types of motion, extracted features and targeted task. This complementary approach supports the presented outcomes given by the perceptual experiment and allows us to manipulate different parameters of motion in order to better understand the mechanisms underlying signer identification.

5.1 Methods

A first model was tested with the dataset used for the behavioral experiment (see Section 3.1), but extending it to 24 descriptions for each signer (instead of 4). 10-sec excerpts were extracted from each original description. Each input variable thus contained 250 samples (fps = 25) and 57 coordinates (3D global positions of 19 sensors, origin placed on the pelvis).

As presented in section 4.2, mocap coordinates were relative to the pelvis marker. Similarly to Troje et al. (2005), these data were able to undergo two steps of normalization :

- Size-normalization : Reference postures were recorded for each signer (Figure 4). An overall reference posture was computed by averaging across the four signers. Slope of the regression between each reference posture and the overall reference posture was then computed. The slope defined relative sizes [28] for each signer : 0.999, 1.075, 0.924, 1.003. After dividing signers data by its size, they all had the same size but relative positions of the articulations still differ, keeping shape intact.
- Shape-normalization : New references were computed from each signer's size-normalized data. New overall reference was defined. Shape-normalized data were obtained by subtracting individual reference postures from each frame then adding the overall reference posture. After that transformation, all signers had the same reference "T" posture.

Then, PCA was performed, allowing for a reduction of the number of variables while keeping the largest variance possible. Using PCA, the first k principal components can be kept to describe the input data, provided that they represent a sufficient amount of the variance in the data. Finally, a regression model was trained for signer classification, based on the reduced number of principal components. A leave-one-observation-out cross validation was employed. It means that the model was trained on N-1 (23) observations for each signer, and the remaining 1 observation was used as test exemplar. All observations were used as test exemplars so the model was tested 24 times and performance was computed as an average across these iterations. This enabled us to see how well the model learns individual signature, idiosyncratic movement patterns that generalize to new observations.

To that end, multinomial logistic regression was used. It is a multi-class generalization of logistic regression that models a binary dependent variable as a logistic function. Equations for the probabilities of each class are :

$$P(Y = c) = \frac{e^{\beta_c . \mathbf{X}}}{\sum_{k=1}^{4} e^{\beta_k . \mathbf{X}}} \tag{2}$$

$\mathbf{X}$ is the input vector (PCs in our case), $\beta_\mathbf{k}$ are the regression coefficients optimized for class k, Y is the class prediction, i.e. output of the model.

As mentioned above, this model can be generalized (see Figure 7). Work is ongoing on testing potential alternatives at each step of the workflow. First results relying on PCA and logistic regression are presented, training the model on natural and shape-normalized motion.

5.2 Results

PCA revealed that 40 components were required to cover 90% of the variance in the natural data (including morphological cues). More interestingly, the two first PCs were able to account for some differentiation between signers (explaining 39% of the variance). Figure 8 illustrates this, showing projections of the input data over the two components. We also note that the different observations for each signer were consistently distributed across the 2 axes.

The first component enabled differentiating signer 2 from all other signers, while the second component enabled differentiating signer 3 from signers 1 and 4. Next components (>2) did not seem to enable any differentiation. This phenomenon and the fact that the first 2 PCs only accounted for 39% of data variance might be a consequence of the nature of the corpus. MOCAP1 provides free signing movements elicited by the description of pictures, thus signers were able to use a wide variety of movements. In other words, first PCs differentiated between individuals but were not sufficient to account for the complex variety of movement patterns in the corpus.

The computational model was also trained on shape-normalized data. The extracted PCs were compared with those obtained using natural data. The two first PCs only accounted for a lower percentage (27%) of the data variance but still enabled differentiation between signers, as shown in Figure 9. The first component discriminated between signer 3 and signer 4, while the second component discriminated between signer 1 and all other signers. The lower percentage of information given by these 2 components was expected as data were normalized. The main outcome of these first results is

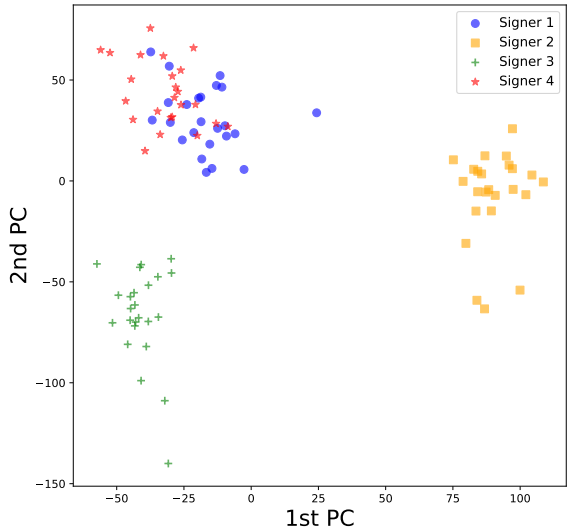

Figure 8: PCA on the natural data. Projection of the dataset over the first 2 PCs

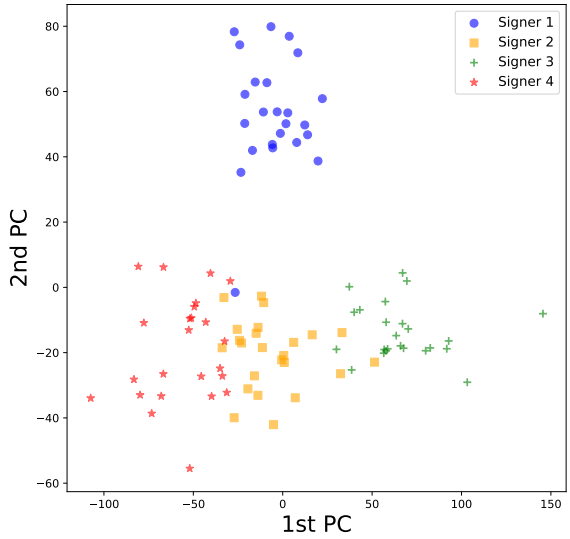

Figure 9: PCA on the shape-normalized data. Projection of the dataset over the first 2 PCs

that PCA still extracted substantial information that discriminates between signers, even though size and shape information were removed.

Results given by the regression model confirmed these observations. Figure 10 shows the model accuracy from the classification task, given the shape-normalized input. Despite the lack of size and shape information, the model classified the four signers above 80% accuracy.

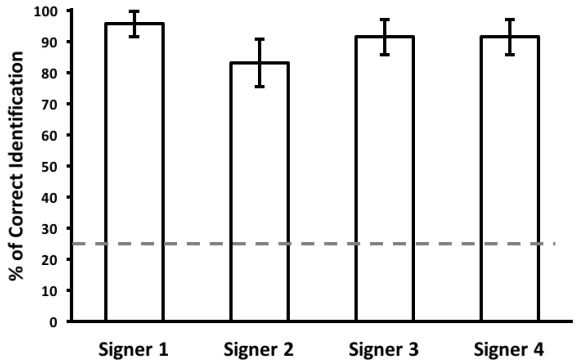

Figure 10: Performance scores from the computational model, trained for identification with shape-normalized data. Dashed horizontal line indicates chance performance levels.

5.3 Discussion

A first model based on PCA and logistic regression was presented. When trained with size and shape-normalized data, the model revealed that the first 2 principal components were able to account for correct identification above 80%. This computational framework is still in progress but these results suggest that differentiation is possible between gestural identities, relying on other cues than size and shape.

A main advantage of this framework is that it can be generalized. It can be used with any type of mocap data (local or global coordinates, angles etc), any method for feature extraction, and any classification model. In the presented results, all PCs (including PC 1&2) that were extracted included joint global coordinates of the signers. It means that separation was possible with low-level features. Further work is ongoing on analyzing which specific characteristics of motion the two first PCs described. More than optimizing computational models for signer identification purposes, our approach aims at using these models in order to qualitatively define motion parameters that are responsible for signer identification.

6 GENERAL DISCUSSION

The present study provides the first evidence that deaf perceivers managed to identify familiar signers, shown as point-light displays, above chance level, as demonstrated for walking [28] or dancing [15]. The second outcome is that morphology was not sufficient to identify the signers. A computational analysis based on PCA revealed a non-significant correlation between the participants' confusion errors and the morphological similarities among signers. This is also consistent with prior studies on the perception of identity from gait [28]. Even after having removed size and shape information, the different walkers were still identified with high accuracy (about five to six times higher than chance level).

Given that the present point-light stimuli were as short as 10s, randomly selected in the original recordings and as it was demonstrated that fingers were needed for SL comprehensibility [18], linguistics were unlikely to play a major role in the identification.

Participants thus may have used other cues. To address this question, we developed a machine learning workflow. Based on mocap data, a regression model was trained to classify LSF utterances by individual signer. First results revealed that the model was able to identify signers, as it was reported for dance mocap data [6]. Using PCA as a means for feature extraction, our results are also in accordance with earlier findings on gender classification from gait [27]. We also trained our first computational model on normalized mocap data, using the same methods as Troje et al. (2005). In line with the human data reported by this study [28], performance of the model was above 80% even after having normalized size and shape of the signers. As mentioned above, we did not test the role of morphology directly in a perceptual experiment. For that goal, we have now designed a second experiment using normalized PL displays, which is currently in progress. We expect participants' correct identifications to be above chance level, even after normalization for size and shape.

Combining human data and computational modeling, the main findings of the present study suggest that SL mocap data contain enough information for signers to be identified, and this beyond morphology-related cues. This outcome calls for additional research further investigating the contribution of other cues, such as kinematic cues in particular. The first machine learning workflow that we presented enables us to evaluate these cues and their potential role in identification. Using PCA as a feature extraction step allows for the evaluation of these cues without a priori hypotheses. Such a data-driven approach is particularly interesting in the case of identification as the discriminant features are mainly idiosyncratic and thus hardly predictable for each individual. Another benefit of methods like PCA is its invertibility. It is therefore possible to compute new motion patterns based on a linear combination of the principal components. On one hand, the PCA modeling approach enables us to interpret components that allow for differentiating between signers, by exaggerating their weight in the linear combination. Manipulations of that kind have been used by Troje et al. (2002) in order to exaggerate the differences in male and female walking patterns [27]. On the other hand, being able to control 'identity features' in motion synthesis will provide promising perspectives into how to anonymize SL motion for virtual signers.

7 ACKNOWLEDGMENTS

This work has been funded by the Bpifrance investment project "Grands défis du numérique", as part of the ROSETTA2 project (Subtitling RObot and Adapted Translation).

REFERENCES

[1] Artur Becker, Henrik Herrebrøden, Victor E González Sánchez, Kristian Nymoen, Carla Maria Dal Sasso Freitas, Jim Torresen, and Alexander Refsum Jensenius. 2019. Functional Data Analysis of Rowing Technique Using Motion Capture Data. In *Proceedings of the 6th International Conference on Movement and Computing*. 1–8.

[2] Mohamed-el-Fatah Benchiheub, Bastien Berret, and Annelies Braffort. 2016. Collecting and Analysing a Motion-Capture Corpus of French Sign Language. In *Workshop on the Representation and Processing of Sign Languages*. Portoroz, Slovenia. https://hal.archives-ouvertes.fr/hal-01633625

[3] Bettina E Bläsing and Odile Sauzet. 2018. My action, my self: Recognition of self-created but visually unfamiliar dance-like actions from point-light displays. *Frontiers in psychology* 9 (2018), 1909.

[4] Antonio Camurri, Ingrid Lagerlöf, and Gualtiero Volpe. 2003. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International journal of human-computer studies* 59, 1-2 (2003), 213–225.

[5] Emily Carlson, Birgitta Burger, and Petri Toiviainen. 2018. Dance Like Someone is Watching: A Social Relations Model Study of Music-Induced Movement. *Music & Science* 1 (2018), 2059204318807846.

[6] Emily Carlson, Pasi Saari, Birgitta Burger, and Petri Toiviainen. 2020. Dance to your own drum: Identification of musical genre and individual dancer from motion capture using machine learning. *Journal of New Music Research* (2020), 1–16.

[7] Fanny Catteau, Marion Blondel, Coralie Vincent, Patrice Guyot, and Dominique Boutet. 2016. Variation prosodique et traduction poétique (LSF/français): Que devient la prosodie lorsqu'elle change de canal?. In *Journées d'Étude sur la Parole*, Vol. 1. 750–758.

[8] James E Cutting and Lynn T Kozlowski. 1977. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the psychonomic society* 9, 5 (1977), 353–356.

[9] Karen Emmorey, Robin Thompson, and Rachael Colvin. 2009. Eye gaze during comprehension of American Sign Language by native and beginning signers. *Journal of Deaf Studies and Deaf Education* 14, 2 (2009), 237–243.

[10] Judith A Holt. 1993. Stanford Achievement TestâĂŤ8th edition: Reading comprehension subgroup results. *American Annals of the Deaf* 138, 2 (1993), 172–175.

[11] Marco Iacoboni, Roger P Woods, Marcel Brass, Harold Bekkering, John C Mazziotta, and Giacomo Rizzolatti. 1999. Cortical mechanisms of human imitation. *science* 286, 5449 (1999), 2526–2528.

[12] Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Perception & psychophysics* 14, 2 (1973), 201–211.

[13] Gunnar Johansson. 1976. Spatio-temporal differentiation and integration in visual motion perception. *Psychological research* 38, 4 (1976), 379–393.

[14] Lynn T Kozlowski and James E Cutting. 1977. Recognizing the sex of a walker from a dynamic point-light display. *Perception & psychophysics* 21, 6 (1977), 575–580.

[15] Fani Loula, Sapna Prasad, Kent Harber, and Maggie Shiffrar. 2005. Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance* 31, 1 (2005), 210.

[16] Evie Malaia, Ronnie B Wilbur, and Marina Milković. 2013. Kinematic parameters of signed verbs. *Journal of Speech, Language, and Hearing Research* 56, 5 (2013), 1677–1688.

[17] Kayoko Okada, Corianne Rogalsky, Lucinda O'Grady, Leila Hanaumi, Ursula Bellugi, David Corina, and Gregory Hickok. 2016. An fMRI study of perception and action in deaf signers. *Neuropsychologia* 82 (2016), 179–188.

[18] Howard Poizner, Ursula Bellugi, and Venita Lutes-Driscoll. 1981. Perception of American sign language in dynamic point-light displays. *Journal of experimental psychology: Human perception and performance* 7, 2 (1981), 430.

[19] Giacomo Rizzolatti, Leonardo Fogassi, and Vittorio Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature reviews neuroscience* 2, 9 (2001), 661–670.

[20] Vassilis Sevdalis and Peter E Keller. 2009. Self-recognition in the Perception of Actions Performed in Synchrony with Music. *Annals of the New York Academy of Sciences* 1169, 1 (2009), 499–502.

[21] Mu-Syuan Sie, Yu-Chuan Cheng, and Cheng-Chin Chiang. 2014. Key motion spotting in continuous motion sequences using motion sensing devices. In *2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. IEEE, 326–331.

[22] Vivien C Tartter and Kenneth C Knowlton. 1981. Perception of sign language from an array of 27 moving spots. *Nature* 289, 5799 (1981), 676.

[23] Joëlle Tilmanne and Thierry Dutoit. 2010. Expressive gait synthesis using PCA and Gaussian modeling. In *International Conference on Motion in Games*. Springer, 363–374.

[24] Mickaël Tits. 2018. *Expert Gesture Analysis through Motion Capture using Statistical Modeling and Machine Learning*. Ph.D. Dissertation.

[25] Mickaël Tits, Joëlle Tilmanne, Nicolas D'Alessandro, and Marcelo M Wanderley. 2015. Feature Extraction and Expertise Analysis of PianistsâĂŹ Motion-Captured Finger Gestures. In *ICMC*.

[26] Petri Toiviainen, Geoff Luck, and Marc R Thompson. 2010. Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Perception: An Interdisciplinary Journal* 28, 1 (2010), 59–70.

[27] Nikolaus F Troje. 2002. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of vision* 2, 5 (2002), 2–2.

[28] Nikolaus F Troje, Cord Westhoff, and Mikhail Lavrov. 2005. Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics* 67, 4 (2005), 667–675.

[29] Dominique Valentin, Hervé Abdi, and Alice J O'TOOLE. 1994. Categorization and identification of human face images by neural networks: A review of the linear autoassociative and principal component approaches. *Journal of biological systems* 2, 03 (1994), 413–429.