

РАЗДЕЛ II. ЭКСПЕРИМЕНТАЛЬНЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ ЯЗЫКОВОГО
ПОВЕДЕНИЯ В МАШИННО-ГЕНЕРИРУЕМЫХ СРЕДАХ
SECTION II. EXPERIMENTAL METHODS FOR STUDYING HUMAN LANGUAGE
BEHAVIOUR IN MACHINE-GENERATED ENVIRONMENTS

UDC 811.93

DOI: 10.18413/2313-8912-2024-10-4-0-7

Konstantin I. Belousov¹
Aleksandar Taleski²
Artem R. Agaev³

Technosemantics of gesture: on the possibilities of using
Perm sign notation in software-generated environments

¹Perm State University,
15 Bukirev St., Perm, 614068, Russia
E-mail: belousovki@gmail.com
ORCID: 0000-0003-4447-1288

²Perm State University,
15 Bukirev St., Perm, 614068, Russia
E-mail: taleski87@yahoo.com
ORCID: 0000-0002-1901-8572

³Human Semantics LLC
15 Bukirev St., Perm, 614068, Russia
E-mail: agaev-artem1@yandex.ru
ORCID: 0009-0001-8948-2143

Received 07 September 2024; accepted 15 December 2024; published 30 December 2024

Abstract: This paper is dedicated to the development of a concept and software solution for generating human movements based on a semantically-oriented language notation created by the authors. The language notation is presented as a formula with a flexible structure of concepts and rules for their implementation, allowing easy adaptation of movement parameter changes to match an ideal or real sample.

To model movements represented in the language notation, a cross-platform application was developed using the Blender 4.2 for the visualization and generation of gestures for anthropomorphic models. The movement control system consists of the following stages: translating it into a gesture notation record; parsing this record; constructing an internal representation of the movement, which is a sequence of frames. Frames contain information about which bone (body part) of the performer they refer to, how they affect its position, and at what time from the start of the movement this frame is relevant. In the final stage, the internal representation of the movement is transformed into the performer's movement, which can be either virtual anthropomorphic 3D models or physical software-hardware systems in the form of anthropomorphic robots.

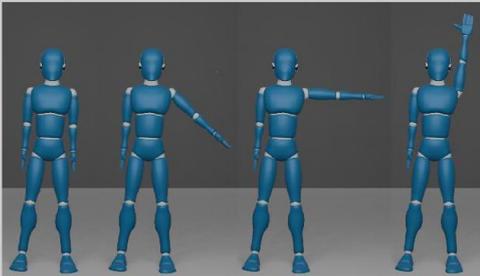
To better correspond to anthropomorphic behavior, in addition to "ideal samples" of human movement, models of human gesture behavior presented in a multimodal corpus specially created for this purpose by the team were used. The material

consists of audiovisual recordings of spontaneous oral texts where recipients describe a wide range of their own emotional states. The data obtained during the experimental research confirmed the flexibility, enhanced controllability, and modularity of the language notation, as well as the ability to model the continuous space of human motor activity.

Keywords: Technosemantics; Movement generation; Language notation of gestures; Gesture; Visualization; Multimodal corpus; 3D graphics; Interpretability

How to cite: Belousov, K. I., Taleski, A., Agaev, A. R. (2024). Technosemantics of gesture: on the possibilities of using Perm sign notation in software-generated environments, *Research Result. Theoretical and Applied Linguistics*, 10 (4), 126–146. DOI: 10.18413/2313-8912-2024-10-4-0-7

TECHNOSEMANTICS OF GESTURE: ON THE POSSIBILITIES OF USING PERM SIGN NOTATION IN SOFTWARE-GENERATED ENVIRONMENTS

Problem	Solution	Results
The absence of a universal language for recording human and anthropomorphic robotic models' movements in human-machine communication	A concept has been developed for modeling and generating human and anthropomorphic robotic models' movements based on a semantically oriented linguistic notation	<ul style="list-style-type: none">• For describing movements, the Perm Gesture Notation was created• A cross-platform application has been developed for visualizing and generating gesture animations of anthropomorphic models• To achieve plausible anthropomorphic gestural behavior, audiovisual recordings of humans describing emotional states were used
		
Generating 3D-Model Movements		Modeling Gesture Behavior

NETWORK SCIENTIFIC JOURNAL
RESEARCH RESULT
НАУЧНЫЙ РЕЗУЛЬТАТ

УДК 811.93

DOI: 10.18413/2313-8912-2024-10-4-0-7

Белоусов К. И.¹
Талески А.²
Агаев А. Р.³

Техносемантика жеста: о возможностях использования пермской жестовой нотации в программно-генерируемых средах

¹ Пермский государственный национальный исследовательский университет,
ул. Букирева, 15, Пермь, 614068, Россия
E-mail: belousovki@gmail.com
ORCID: 0000-0003-4447-1288

²Пермский государственный национальный исследовательский университет,
ул. Букирева, 15, Пермь, 614068, Россия
E-mail: taleski87@yahoo.com
ORCID: 0000-0002-1901-8572

³ООО «ХьюменСемантикс»
15 Bukireva St., Perm, 614068, Russia
E-mail: agaev-artem1@yandex.ru
ORCID: 0009-0001-8948-2143

*Статья поступила 07 сентября 2024 г.; принята 15 декабря 2024 г.;
опубликована 30 декабря 2024 г.*

Аннотация: Статья посвящена разработке концепции и программного решения генерации движений человека на основе созданной коллективом семантико-ориентированной жестовой нотации. Нотация представляется в виде формулы, отличающейся гибкой структурой понятий и правил их реализации, что позволяет легко адаптировать изменения параметров движения для соответствия идеальному или реальному образцу.

Для моделирования движений, представленных в языковой нотации, разработано кроссплатформенное приложение на базе игрового движка Blender 4.2 для визуализации и генерации анимации жестов антропоморфных моделей. Система управления движений состоит из следующих стадий: запись жестовой нотации; разбор этой записи; построение внутреннего представления движения, которое является последовательностью ряда кадров. Кадры имеют информацию о том, к какой кости (части тела) исполнителя они относятся, как влияют на её положение, в какой момент времени от начала движения этот кадр актуален. На последнем этапе осуществляется преобразование внутреннего представления движения в движение исполнителя, в качестве которого могут выступать как виртуальные антропоморфные 3D-модели, так и материальные программно-аппаратные комплексы в виде антропоморфных роботов.

Для большего соответствия антропоморфному поведению помимо «идеальных образцов» движения человека были использованы модели жестового поведения людей, представленные в специально создаваемом для этой цели коллективом мультимодальном корпусе. Материал представляет собой аудиовизуальные записи устных спонтанных текстов описания реципиентами широкого спектра собственных эмоциональных состояний. Полученные в ходе экспериментального исследования данные подтвердили гибкость, повышенную контролируемость и модульность языковой нотации, возможность моделирования непрерывного пространства человеческой двигательной активности.

Ключевые слова: Техносемантика; Генерация движений; Языковая нотация жестов; Жест; Визуализация; Мультимодальный корпус; 3D-графика; Интерпретируемость

Информация для цитирования: Белоусов К. И., Талески А., Агаев А. Р. Техносемантика жеста: о возможностях использования пермской жестовой нотации в программно-генерируемых средах // Научный результат. Вопросы

теоретической и прикладной лингвистики. 2024. Т. 10. № 4. С. 126–146.
DOI: 10.18413/2313-8912-2024-10-4-0-7

ТЕХНОСЕМАНТИКА ЖЕСТА: О ВОЗМОЖНОСТЯХ ИСПОЛЬЗОВАНИЯ ПЕРМСКОЙ ЖЕСТОВОЙ НОТАЦИИ В ПРОГРАММНО ГЕНЕРИРУЕМЫХ СРЕДАХ

Проблема

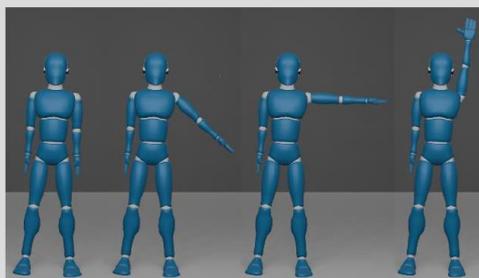
Отсутствие универсального языка для записи движений человека и антропоморфных роботизированных моделей в человеко-машинной коммуникации

Решение

Разработана концепция для моделирования и генерации движений человека и антропоморфных роботизированных моделей на основе семантически ориентированной жестовой нотации

Результаты

- Для описания движений создана Пермская жестовая нотация
- Разработано кроссплатформенное приложение для визуализации и генерации анимации жестов антропоморфных моделей
- Для достижения правдоподобности антропоморфного жестового поведения использовались аудиовизуальные записи описания людьми эмоциональных состояний



Генерация движений 3D-моделей



Моделирование жестового поведения

Introduction

The development of systems for describing human movements is a relevant task across various fields, such as choreography, robotics, animation, and many others. Just as a musical score represents musical notes, movement notation is designed to provide precise instructions regarding specific movements, actions, as well as their combinations and/or sequences. The creation of notations that describe movements across different domains faces several challenges. The first challenge is the absence of a universal language for recording movements. Different fields use distinct methods and languages to record movements, leading to a “Tower of Babel” situation where communication between disciplines becomes difficult (Laumond, Abe, 2016). The second

challenge arises from the established terminological systems, paradigms, and goals in different fields (e.g., choreography vs. robotics), which further complicates interdisciplinary communication and hinders scientific integration (Calvert, 2015).

At the same time, the metalinguistic and conceptual fragmentation of disciplinary fields contradicts the qualitative changes in the nature of human communication and activity, driven by the increasing prominence of interface-mediated communication and the digitization of various forms of activity. This type of communication demands universal “seamless” tools (meta-languages and technologies) for transitioning between different forms of human activity. Hence, there is a need to explore new, potentially universal approaches to describing human-

machine interaction (HMI) and human-computer interaction (HCI).

Electronic computing devices and virtual anthropomorphic 3D models, such as robots and virtual avatars, imitate human behavior, communication, emotions, movements, gestures, and other human traits and behavioral patterns. As these anthropomorphic models become more integrated into everyday life, they must be able to handle a wide range of human actions and linguistic expressions. When discussing human movements, for effective interaction with people and the extraction of information from linguistic and gestural data, these intelligent models need to understand the referential relationships between actions and language, as well as interpret actions conveyed through language.

The development of tangible human-machine interfaces is based on the principles of human communication, where information is perceived, processed, and transmitted using multiple modalities (Belousov et al., 2024). Semantics in this context covers not only auditory or visual messages but also motor actions. For example, van Elk et al. describe the semantics of actions as a hierarchically organized system, whose subsystems are selectively activated depending on the subject's intentions and task context (van Elk et al., 2009). In natural communication, information is transmitted through several channels of different modalities to "narrow" and focus the scope of meaning. For instance, spatial deixis in face-to-face interaction requires gestural involvement: "Look there!" (accompanied by a hand gesture or a nod).

Typically, the problem of describing movements relates to the discussion of the differences between movements and gestures. Several approaches distinguish between movement and gesture. For example, movement can be viewed as a concept encompassing any physical activity of the body, including gestures, which, in turn, carry symbolic meaning in communication (Bull, Doody, 2013; Kendon, 1997; Kilpatrick, 2020; Murillo et al., 2021; Novack,

Wakefield, 2016: 340; Stults-Kolehmainen, 2023). According to the approach of A.M. Novack et al., a gesture is not directly related to altering the world (e.g., opening a jar by twisting it) or representing movement patterns (e.g., performing dance steps). Instead, it represents movement that may change the world (e.g., a gesture showing how to open a jar by twisting it) or represents movement that exists on its own (e.g., a gesture showing how to perform a dance) (Novack, Wakefield, 2016: 340). Both examples are generally correct; however, one can find contexts where dance movements or opening a jar can be viewed as gestures. For example, professional dance movements (e.g., in ballet) are loaded with communicative and semantic meaning, and when directed at an audience, they are interpreted similarly to any other sign. The example involving material transformation of the world ("opening a jar", etc.) can also be contextualized as a sign of professional skill.

We prefer an understanding of the relationship between gesture and movement, where movement is not only a general concept for gesture but also unites them in sequential chains (Kendon, 1997; Kilpatrick, 2020; Streeck, 2010; Trujillo et al., 2018). If a movement fragment is communicatively directed and semantically coherent, it can be considered a gesture. However, the same movement might not be a gesture if it lacks an addressee capable of interpreting the sign (with the addressee potentially being the communicator themselves, i.e., in the case of self-communication).

In the modern technological world, gestures are forms of non-verbal communication used not only for interacting with people but also with anthropomorphic virtual or robotic models. Recording human movements, developing models and algorithms for effective motion recognition, creating unified movement notation, and improving human-machine communication are pressing challenges in science, robotics, machine learning, and other fields.

The goals of movement notation – whether in dance, animation, or robotics –

include providing clear, interpretable instructions that can be executed by performers, machines, or digital characters. Accurate movement notation is essential for their present and future use, study, or reproduction, ensuring consistency across different situations and among different performers (both human and anthropomorphic models). In this paper, we present a new concept and software solution for modeling and generating movements based on a semantically-oriented linguistic notation, which we have named the Perm Gesture Notation (PGN). The primary goal of this paper is to present a methodology for modeling human movements based on the developed linguistic notation. This includes the presentation of the developed notation, a description of an application that implements a movement control system based on the Blender 4.2 engine. In the developed application, visualization and generation of movements for anthropomorphic models, as well as compositional and sequential control of the generated movements based on the developed notation, are carried out. To achieve plausible anthropomorphic behavior in the 3D model, we use a multimodal corpus of audiovisual recordings of people describing various emotional states. This confirms the possibility of modeling a wide range of human motor actions.

Theoretical review

The development of movement notation systems has a history spanning several centuries. Since the 15th century, at least 87 different movement notation systems have been used across Europe and North America (Abbie, 1974; Farnell, 1996; Guest, 1984; Key, 1977). These systems often arose to document specific dance styles or gesture languages, and many fell out of use as the associated movement practices evolved or disappeared (Farnell, 1996).

In the field of non-verbal communication research, any attempt to develop a movement alphabet must inevitably take into account historically significant systems of movement notation, including the

Beauchamp-Feuillet system that thrived at the French court in the 17th and 18th centuries, Gilbert Austin's gesture system for public speaking from 1806, Ray Birdwhistell's kinesics from the mid-20th century, William Stokoe's system for recording American Sign Language, as well as adaptations for other sign languages by Adam Kendon and La Mont West Jr. Generalized systems like Labanotation, Laban Movement Analysis, Benesh Choreology, and Eshkol-Wachman notation are designed for adaptation to a wide range of movements, functioning as a "phonetic alphabet" for body actions, capable of recording any physical action while preserving its semantic content (Abbie, 1974; Bashan et al., 2024; Benesh, Benesh, 1956; Bernardet et al., 2019; Birdwhistell, 1952; Dell, 1977; El Raheb, Ioannidis, 2014; El Raheb et al., 2023; Eshkol, Wachmann, 1958; Farnell, 1996; Frishberg, 1983; Grushkin, 2017; Guest, 1984; 2005; Harrigan, 2008; Kendon, 1997; Key, 1977; Laban, Lawrence, 1974).

Several key systems of movement description have emerged from related fields such as anthropology, ethology, and psychology. The Facial Action Coding System (FACS) (Ekman, Friesen, 1978) is designed for analyzing and interpreting facial expressions by identifying specific facial actions. The Bernese System (Frey et al., 1982) offers a methodology for analyzing movement in psychological and anthropological studies. The most recent system, the Body Action and Posture Coding System (BAP) (Dael et al., 2012), was developed for detailed analysis of body movements and postures using 141 behavioral variables that capture complex temporal movement segments. These systems provide a comprehensive toolkit for studying non-verbal communication and human behavior in various scientific and applied contexts (Bernardet et al., 2019).

Movement classification, coding, and interpretation systems differ across several parameters, including their intended purpose, origin, level of detail, and reliability. First,

these systems vary in their objectives: some are designed for reproducing or analyzing movements, while others aim to draw conclusions about the individual performing the movements. For example, systems originating in dance, such as Labanotation and Laban Movement Analysis, focus on the precise transmission and analysis of dancers' movements. In contrast, systems from anthropology, ethology, and psychology, like FACS and kinesics, are aimed at studying and interpreting human behavior in a broader context (Birdwhistell, 1952; Bernardet et al., 2019; Duprey et al., 2017; Ekman, Friesen, 1978; Laban et al., 1974).

Second, the level of subjective inference varies among different systems. Some systems rely on subjective judgments by coders, while others use systematic labels and observational coding, or direct measurement of muscle activity. This distinction is particularly important in scientific research, where high precision and data reliability are required (Bernardet et al., 2019).

The focus on movement quality or type is another distinguishing feature of various systems. For instance, Labanotation and the Bernese System emphasize movement quality, while kinesics focuses on types of movements. The level of system detail, including parameters such as the minimum time unit for coding, the size of the alphabet (number of distinct codes), and the number of simultaneously coded tracks (body parts), plays a key role in determining the system's level of detail. The smaller the time unit, the larger the alphabet, and the more tracks coded, the more detailed the system is. This, in turn, affects coding time: detailed systems require significantly more time for data coding than less detailed ones (Birdwhistell, 1952; Bernardet et al., 2019; Dael et al., 2012; Ekman, Friesen, 1978; Farnell, 1996; Laban, Lawrence, 1974).

Reliability also varies across systems. Some, like FACS and BAP, have undergone empirical reliability evaluation, while others, like LMA, have been evaluated to a limited extent, and some systems have not been

assessed for reliability at all (Bernardet et al., 2019; Dael et al., 2012; Ekman, Friesen, 1978; Laban, Lawrence, 1974; Shafir et al., 2016).

Thus, differences among movement classification, coding, and interpretation systems include their intended purpose, origin, degree of subjective inference, focus on movement quality or type, level of detail, coding time, and reliability. M. Karg et al. (2013) emphasize that movements are typically represented as joint angle trajectories or kinematic features, which, while detailed, can be computationally intensive and may overlook the expressive qualities of movements. Movement notation systems offer a more efficient and expressive representation.

Due to the conceptual proximity to our notation system, we pay particular attention to the Common Morphokinetic Alphabet (CMA) for transcribing and analyzing body movements, which seeks to provide a comprehensive, precise, and user-friendly method for recording physical movements, establishing grammatical rules based on Noam Chomsky's phrase grammar (Izquierdo, Anguera, 2018). Like our notation, the CMA system incorporates linguistic principles, particularly the structure and syntax of movement phrases, to create a notation system capable of capturing the complexity of human movement in three-dimensional space over time (Izquierdo, Anguera, 2018). The CMA system builds on previous systems, such as Ray Birdwhistell's kinesic notation system (KNS) (Birdwhistell, 1952) and the Facial Action Coding System (FACS) (Ekman, Friesen, 1978), but positions itself as a more comprehensive and adaptable system capable of keeping pace with technological advancements (Izquierdo, Anguera, 2018).

It should be noted that the research conducted by Izquierdo, Anguera (2018) is closely related to our work in certain aspects, which will be discussed in the next section.

Methodology. Description of the Perm Gesture Notation (PGN)

The movement notation system, as we conceptualize it, must satisfy several key criteria.

First, the movement notation should be semantically oriented and based on natural language. This would allow subject matter experts, when writing or interpreting a movement formula, to visualize how the action will be performed. The language for recording movements should be represented as a formalized structure of concepts and rules for their implementation. This would facilitate the use of computational technologies, particularly allowing for encoded messages to be interpreted by an intelligent interpreter.

The language for recording movements must be flexible, extendable through the integration of new objects, concepts, and rules, and easily adaptable to changes in movement parameters. The notation should be comprehensive, covering all aspects of movement, and universal, capable of encoding all types of movements. Furthermore, the notation must exhibit anatomical and physiological accuracy, representing the mechanics of the body correctly. There should also be a possibility for quantitative evaluation of the generated movement models.

Additionally, the notation must be practical and easily integrated with modern technologies. It is crucial that the system of notation can document observed movements without relying on contextual assumptions or subjective interpretations, ensuring objectivity and minimizing coder bias. We have endeavored to account for these criteria as

thoroughly as possible in the creation of our movement notation system.

Unlike other movement notation systems, particularly the Common Morphokinetic Alphabet (CMA) (Izquierdo, Anguera, 2018), which is the closest to ours, where the movement notation includes not only position/state but also complex movement patterns (e.g., both hands together), with repetitions of entities (e.g., right and left sides for position and body parts), and where some indicators are described using numbers, our system models movement as a sequential transition from one stationary state of the body part performing the movement to another stationary state. The key distinction in our approach is that instead of describing the movement itself, which requires the inclusion of numerous variables, we define only the initial and final states of the body part performing the movement (intermediate states can be included, as discussed below) and the temporal characteristics of executing the movement as a transition from one state of the body part to another.

Each state has specific indicators associated with certain entities (movement system, numerical indicators, etc.), body parts, and their nature and functions. The entities, nature, and functions of body parts have different types, allowing them to be represented in a structured and symbolic form. Latin is chosen as the natural language for description. A fragment of the system is provided in Table 1.

Table 1. Fragment of the Perm Gesture Notation (PGN)
Таблица 1. Фрагмент Пермской жестовой нотации (ПЖН)

Entities	Types	Symbol	Latin Name	
Movement	Transitional	[]	transitorius	motus
	Basic	[[]]	basicus	
State		{ }	status	
Event		#	eventus	
Motion System	Sequence	+	sequentia	motus ratio
	Simultaneity	*	simultaneus	
	Cyclicity	◊	repetitio	
Numerical Indicator	Time Interval (ms)		tempus intervallum	secundum numerum indicator
Nature and Functions	Name	Symbol	Latin Name	
Body Parts	Hand (overall)	M	manus	pars corporis humani
	Palm	P	palma	
	Wrist	MO	manus	
	Forearm	AB	antibrachium	
Movement Direction (Trajectory)	Up	su	sursum	directio (trajectoria) motus
	Down	de	deorsum	
	Middle	med	medium	
	Right	dx	dextra	
	Left	si	sinistra	
	Clockwise	shm	secundum horologii motum	
Hand Position	Extended	ex	extensus	manus positio
	Bent	an	ancus	
Movement Form	Circle	c	circulus	forma motus
	Semicircle	sc	semicirculus	

The notation allows for the description of both complete movements and movement fragments. An example of a movement formula using our notation is:

:1000:{*si Ma* } (1)

The description of movement using a specific notation system can be illustrated by the example of raising a hand, as shown in Figure 1. For the precise description of a body

part's state, the following model is used (demonstrated for the left hand):

- M (manus) – denotes the hand.
- Left lower index (si): Specifies the selected hand, where “si” means the left hand (similarly, “dx” refers to the right hand).
- Right upper index (ex): Indicates the hand's position, where “ex” denotes an extended hand.

- Right lower index (de): Specifies the direction, where “de” means downward.

The use of curly braces {} around the notation indicates that this part of the movement represents a state, while :1000: specifies the duration of this state in milliseconds. Thus, formula (1) conveys the initial state of the movement.

This can be translated into a more machine-readable form as follows:

:1000: {_si_M^ex_de_} (2)

Some positions and states of the hand are assumed by default. For instance, the fully extended arm state is considered the default. Therefore, such states may not need to be explicitly marked in the notation. Consequently, formula (2) simplifies to:

:1000: {_si_M_de_} (3)

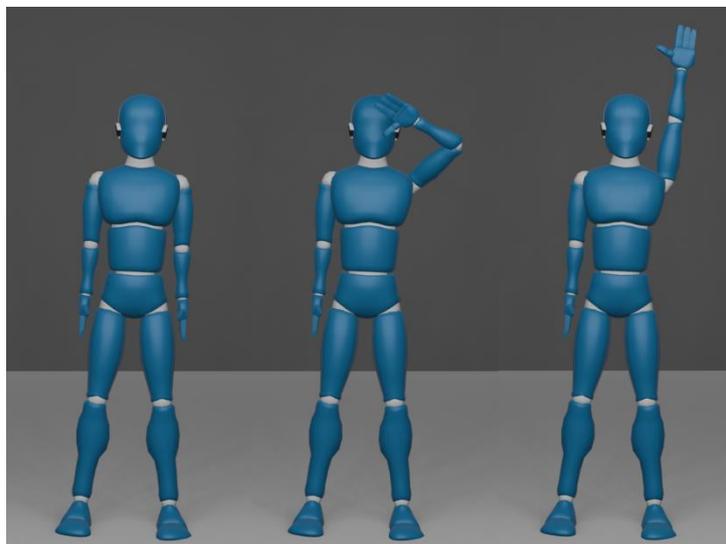
Using these notation rules, we can describe the movement shown in Figure 1 as follows:

:1000: {_si_M_de_}:1000: [_si_M_su_]:1000: {RE} (4)

From this formula, it is clear that it does not describe intermediate states: the left hand moves directly from a downward extended position to an upward extended position. It may seem that the movement is described separately (e.g. 1000: [_si_M_su_]). However, this is not the case: this notation captures the final state as the result of the movement, and its placement within the movement code serves to: a) indicate the movement itself, b) specify its execution time, and c) allow for the possibility of simultaneous execution with another action (an example is provided below). At the same time, since the final state is already presented in the formula, there is no need to duplicate its notation, which is why the formula ends with the fragment 1000: {RE}, indicating that this state, lasting 1000 ms, is described in the movement block, i.e., as {_si_M_su_}.

Figure 1. Raising the Arm in Blender (three motion slices)

Рисунок 1. Поднятие руки в приложении Blender (три среза движения)



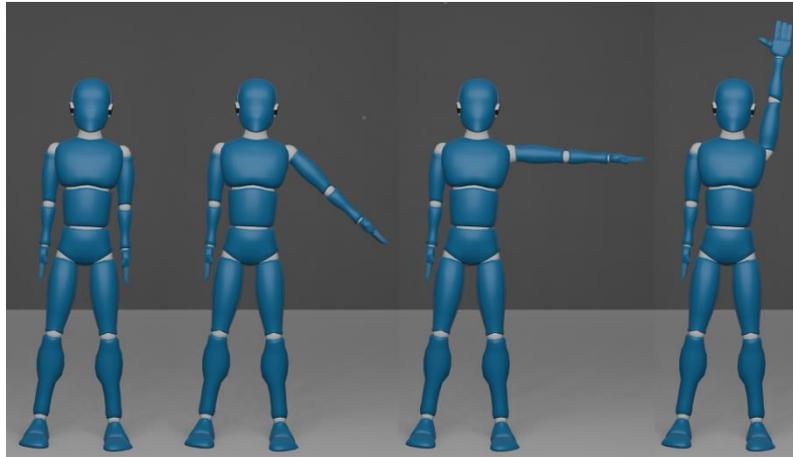
In this case, the movement is performed in the most efficient manner. However, if we want to execute the movement differently, we need to define intermediate states for the arm's position. This will result in the arm moving along a different trajectory. For

example, raising the arm through an intermediate position, where the arm is extended to the side, followed by a pause of 1 millisecond (1: {RE} denotes a resting state for 1 millisecond), can be described as follows:

:200:{_si_M_de_}:1000:[_si_M_si_]:1: {RE}:1000:[_si_M_su_]:200:{RE}

Figure 2. Raising an arm in Blender (four motion slices)

Рисунок 2. Поднятие руки в приложении Blender (четыре среза движения)



In Figure 2, it is evident that the final position of the arm coincides with the position shown in Figure 1: the arm is extended with the palm facing forward. To specify a different final position, we need to use additional notation to clarify the precise orientation, e.g.:

:200:{_si_M_de_}:1000:[_si_M_si_]:1:

{RE}:1000:[_si_M_su*MO_dx_]:200:{RE}

Let's examine the construction [_si_M_su*MO_dx_]:

- The final position of the left arm is indicated (upwards).
- The '*' symbol denotes simultaneity, meaning another action occurs at the same time as the arm is in the upward position.
- "MO" represents the palm, and "dx" indicates that the palm is turned to the right. For the left arm, this means the palm is turned towards the body.

This example illustrates that the directions and positions of body parts, as described in Table 1, can be applied to various body parts depending on their anatomical capabilities. For instance, for the palm, "MO_at" (forward) indicates the arm is rotated so that the palm faces forward. Therefore, the position of the arm in its final state, as shown in Figures 1 and 2, can be

precisely described using this notation system.

Since the notation is designed for modeling movements in computerized systems, there are preset configurations (settings) through which the movements of body parts in an anthropomorphic 3D model are executed. For example, raising the arm upward is performed through an intermediate state of elbow bending, thus such movement does not require special marking. At the same time, this approach raises the issue of determining the most human-like configurations, which can be achieved through statistical analysis of a wide range of human movements modeled with varying psychobiological characteristics (e.g., age, gender, temperament, etc.), a promising area for further research.

Up to this point, movement descriptions have been focused on a single body part. But how does the notation work with simultaneous movements of multiple body parts, such as both arms?

For modeling simultaneous-successional movements (simultaneous execution of actions by multiple body parts in their temporal dynamics), the following concepts are introduced:

1. Independence (autonomy) of the 3D model's body parts, limited only by the

anatomical and physiological capabilities of a human, as fixed in the 3D model settings.

2. “Recording tracks”, which are used for independent description of each body part’s movement (each body part is assigned its own “track”).

3. A movement event, understood as either: a) any change in state/movement of a body part caused by external factors (another body part or object, e.g., the hand striking another body part or object), or b) a complex movement of body parts, where mandatory states of these parts are defined relative to each other over time. It is worth noting that despite the autonomy of a body part’s movement, its motion can influence another body part due to connectivity and inertia (e.g., arm rotation affecting the torso). Such inertial changes are natural and are not considered events.

An event is tied to time but occurs within a certain interval Δt . The Δt parameters can be used to calculate and assess the precision of modeled movements. The event is denoted by the symbol “#” in the upper left index of the body part. For example, the description of the event (#1) where two hands make contact in front of the chest: for the left hand[_si_#1^M_dx*at_]and for the right hand[_dx_#1^M_si*at_]. Events have their indices, allowing distinction between different events and linking formula fragments describing various body parts.

It should be noted that complex movements (such as “hands together” in CMA) can be defined separately in our case and stored as formula variants. In other words, certain patterns of states or movements can be implemented as units in a dictionary, with their formal representations.

Implementation of Motion Control Systems

The development of a motion control system based on linguistic notation involves several stages: gesture notation, construction of an internal representation of motion, and its translation into executable movements, demonstrated using 3D modeling and animation software.

Stage One: Parsing Notation

The first stage is parsing the notation (formula parsing). Parsing involves analyzing the notation string, extracting terms, and creating an abstract syntax tree (AST). The parsing process is illustrated with the following formula example:

:400: {_dx_M_2la-de-2at_}:500:[_dx_M_la-su-2at_]:400:{RE}

1. Extraction of Individual Terms:

○ First Term: :400: {_dx_M_2la-de-2at_}

■ Time: 400 milliseconds.

■ Type: State ({})

■ Position Details:

■ dx_M: Right hand.

■ 2la: Movement left with coefficient 2.

■ de: Movement down.

■ 2at: Movement forward with coefficient 2.

■ Term Object: A State Term is created with attributes:

■ time = 400.

■ node = dx_M.

■ details = {'2la', 'de', '2at'}.

○ Second Term: :500:[_dx_M_la-su-2at_]

■ Time: 500 milliseconds.

■ Type: Movement ([])

■ Direction Details (Final State):

■ dx_M: Right hand.

■ la: Movement left.

■ su: Movement up.

■ 2at: Movement forward with coefficient 2.

■ Term Object: A Move Term is created with attributes:

■ time = 500.

■ node = dx_M.

■ details = {'la', 'su', '2at'}.

○ Third Term: :400:{RE} ■ Time: 400 milliseconds.

■ Type: State ({})

■ Details: RE (State fixation).

■ Term Object: A State Term is created with attributes:

■ time = 400.

■ details = {'RE'}.

Stage Two: Construction of Internal Motion Representation

In the second stage, an internal representation of motion is constructed based on the AST. This representation consists of a series of frames that include information about which body part of the performer is associated with each frame, how they affect its position, and the specific moment in time from the start of the motion when each frame is relevant. Only key frames are recorded in this series – extreme positions of primitive movements. E.g. if a performer needs to trace the letter “L” with their hand, the series will include key frames for the hand’s positions at the bottom, top, and top right. Motion interpolation is the responsibility of the performer.

Stage Three: Transformation of Internal Representation into Final Motion

The third stage involves transforming the internal representation into final motion. The internal motion representation is handed over to the performer, who converts it into the final result visible to the user. The performer can be a robot or any program working with skeletal animation, including game engines. In this study, Blender 3D, a 3D modeling and animation program, is used as the performer. Blender 3D allows for the creation of custom extensions using Python and provides a comprehensive set of graphical interfaces and software utilities for working with 3D graphics.

Results and discussions

In this paper, the results are presented as a demonstration of the notation’s ability to describe natural human movements extracted from video segments of a multimodal corpus depicting various emotional states described by individuals. The primary aim of the developed gesture notation is to facilitate seamless transitions from one form of human activity to another in different interface-mediated communications by providing a universal and precise system for encoding and describing human movements, including gestures, postures, and complex interactions of body parts.

At this stage, we do not aim for complete and precise tempo-rhythmic and spatial reproduction of human movements by machines due to existing technical limitations. Therefore, we will define the formal description boundaries: notation will be applied only to hand movements.

The complex of movements is described as a series of states for each hand, represented in two recordings (referred to as “tracks”). In formal notation, the transition from one track to another is separated by a comma (this symbol is used exclusively for this function).

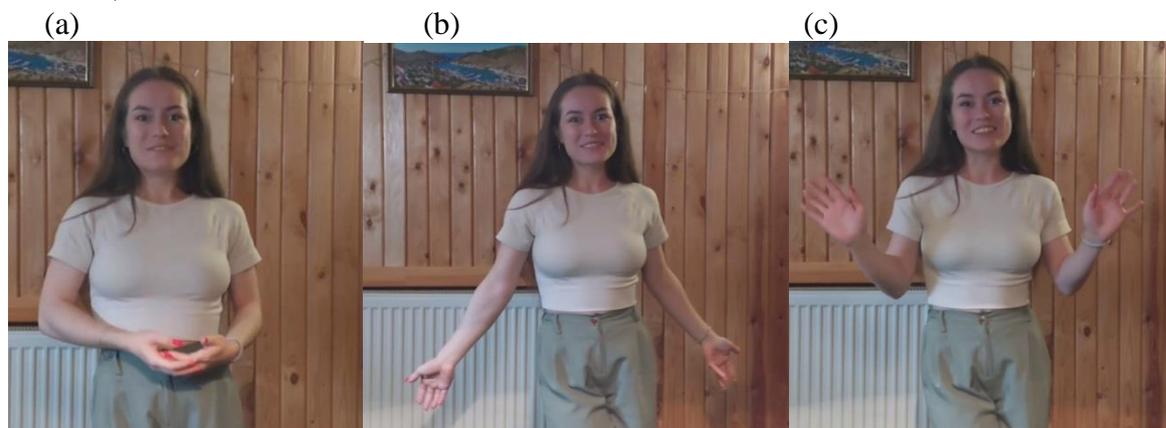
The formal notation of the movement ultimately shapes our understanding of the gesture, which, from a semantic perspective, is communicatively directed and semantically coherent, and, from a formal perspective, represents a movement from one state to another (including intermediate states if present).

We examine fragments of movements from a specialized multimodal corpus of emotions and gestures. Emotions in this corpus can be broadly categorized as positive and negative. The quantitative and qualitative characteristics of the movements accompanying these emotions may vary. As an example, we consider two sets of movements: one associated with the positive emotion “Delight” (Figure 3) and another associated with the negative emotion “Anger” (Figure 4), as well as the corresponding linguistic contexts of their use (see Example 1 and Example 2). Illustrations will use screenshots of hand movement states from informants, which serve as model representations in the form of movement formulas.

Each of the movement complexes selected for illustration differs in several parameters. For the first complex, which occurs in the context of the emotion “Delight”, it is noted for its simplified structure. Additionally, this movement is an example of synchronous movement of both hands.

Figure 3. Illustration of the complex of movements within the framework of the emotion “Delight” and its components

Рисунок 3. Иллюстрация комплекса движений в рамках эмоции «Восторг» и его составляющие



This movement complex is observed in the following linguistic context:

Example 1. ...И вот (**Bent arms are folded with one palm placed on the back of the other hand palm**) полет на парашуте (**Arms extended outward and downward**), я была в восторге (**Bent arms raised on either side**).

In describing the emotion “Delight”, the informant uses this movement complex (Figure 3), which can be divided into three main states: a) bent arms and one palm placed on the back of the other hand palm; b) arms extended outward and downward; c) bent arms raised on either side.

A detailed description of these states reveals the following characteristics: the initial state (a) represents the starting position where the arms are bent at the elbows at a 90-degree angle, with the left forearm directed forward and to the right, the right forearm forward and to the left, and the palms pressed together at the level of the abdomen (palms up). This component can be represented by the terms ($\{_{dx_B_de_+dx_AB_ad-at-si_}\}$ and $\{_{si_B_de_+si_AB_ad-at-dx_}\}$). In state (b), the right arm is extended downward and to the right, and the left arm is extended downward and to the left ($[_{dx_AB_de-la_}]$ and $[_{si_AB_de-la_}]$). Finally, the last state (c), which is the concluding stage of the

complex, is characterized by the right arm being bent with the forearm directed upward and to the right, and the left arm also bent with the forearm directed upward and to the left ($[_{dx_AB_su-la_}]$ and $[_{si_AB_su-la_}]$). Each part of the movement component is followed by a fixation with a specific time measurement, e.g. (:190:{RE}).

This movement complex can be clearly encoded as the following formula:

```
:415:{_dx_B_de_+dx_AB_ad-at-si_}:375:[_dx_AB_de-la_]:765:{RE}:690:[_dx_AB_su-la_]:190{RE};:415:{_si_B_de_+si_AB_ad-at-dx_}:375:[_si_AB_de-la_]:765:{RE}:690:[_si_AB_su-la_]:190{RE}
```

The structure of this formula consists of six terms, three for each hand. These terms are distributed across two tracks, one for the right hand and one for the left hand, respectively. Comparing the terms and their sequences across both tracks reveals their complete identity, indicating the synchronicity of the movements of both hands. Furthermore, each term can be associated with a specific linguistic context, as illustrated in Table 2.

The following section presents an illustration of the movement complex accompanying the negative emotion “Anger”.

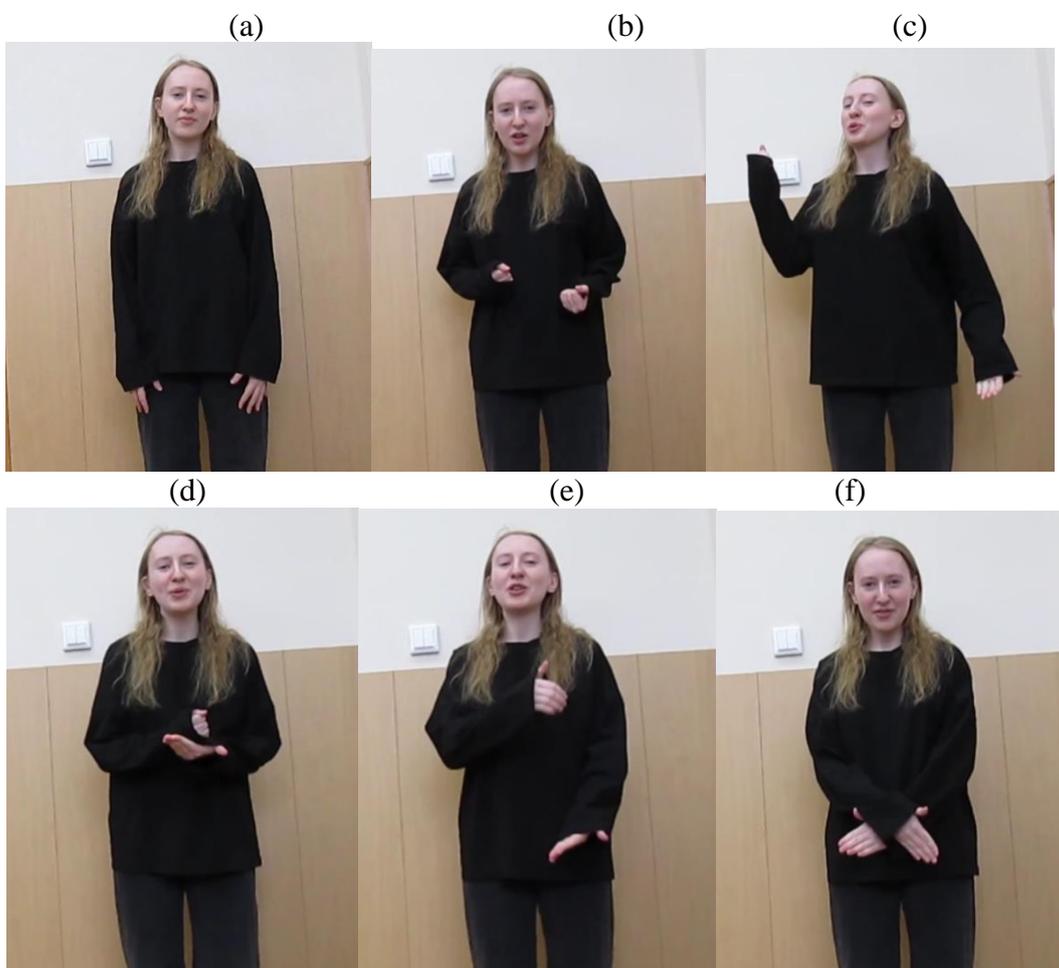
Table 2. The ratio of terms and contexts of the emotion “Delight”

Таблица 2. Соотношение термов и контекстов эмоции «Восторг»

Terms	Track 1	:415: {_dx_B_de_+d x_AB_ad-at-si_}	:375:[_dx_AB_de- la_]:765:{RE}	:690:[_dx_AB_su- la_]:190:{RE}
	Track 2	:415: {_si_B_de_+si _AB_ad-at-dx_}	:375:[_si_AB_de- la_]:765:{RE}	:690:[_si_AB_su- la_]:190:{RE}
Linguistic Context		И вот	полет на парашуте	я была в восторге

Figure 4. Illustration of the complex of movements within the framework of the emotion “Anger” and its components

Рисунок 4. Иллюстрация комплекса движений в рамках эмоции «Злость» и его составляющие



The complex movement patterns associated with the emotion of “Anger” are accompanied by the following linguistic context:

Example 2. **(Both hands are lowered)**
 Злость – это когда постоянно хочется прямо

подражаться (**Hands bent, the forearms of the right and left arms are lifted upwards and come together – right to the left, left to the right**) или ты хочешь (**Right hand bent, forearm upward, left hand extended downward and slightly to the side**) стукнуть

по столу (**Bent hands are stacked, one palm in a fist on top of the other**), сказать (**Right hand bent, right forearm to the left and slightly up, left hand extends forward, to the right, and then drops down with a five-fold amplitude**), что нет (**Hands extended downward and diverge – right to the left, left to the right, and cross**), ты не прав...

The interpretation of movements in this sequence, as opposed to the previous one, is more complex, involving unsynchronized and additional actions for the right (dx) and left (si) hands, structured by terms in two main tracks.

For the right hand, movement begins from the initial position with the hand lowered down ($\{dx_M_de\}$). This is followed by several components. The second component (b) ($[dx_AB_at-ad-si_]$) is observed when the right forearm extends forward, bends towards the shoulder, and moves left, after which the final position is fixed ($\{RE\}$). Next is the third component (c) – the forearm raises up and moves left ($[dx_AB_su-la_]$), also followed by fixation ($\{RE\}$).

The first event (d) ($[dx^{\#1}AB_at-ad-2si_]$) involves the forearm extending forward, bending towards the shoulder, and moving left with double amplitude. After fixing the position ($:200:\{RE\}$), we encounter the next component (e) ($[dx_AB_at-ad-2si-su_]$), where the forearm again extends forward, bends towards the shoulder, moves left with double amplitude, and then rises. The stage ends with fixation ($\{RE\}$). The final event (f) includes lowering the hand down and moving left with triple amplitude ($[dx^{\#2}M_de-3si_]$), followed by fixation ($\{RE\}$).

The left hand repeats a similar sequence of movements. In the initial position (a) it is also lowered down ($\{si_M_de\}$). The next component (b) includes extending the forearm forward, bending towards the shoulder, moving right, and lowering down ($[si_AB_at-ad-dx-de_]$), after which the position is fixed ($\{RE\}$). Next (c), the forearm lowers down with double amplitude and moves left ($[si_M_2de-la_]$), with fixation ($\{RE\}$).

The first event (d), where the forearm extends forward, bends towards the shoulder, and moves right, can be noted as ($[si^{\#1}AB_at-ad-dx_]$), with subsequent fixation ($\{RE\}$). Then follows the movement component (e), where the forearm extends forward, moves right, and lowers down with five-fold amplitude ($[si_AB_at-ad-dx-5de_]$), after which the position is fixed ($\{RE\}$). The final event (f) includes lowering the hand down and moving right with triple amplitude ($[si^{\#2}M_de-3dx_]$), ending with fixation ($\{RE\}$).

The overall description of this movement complex, considering its temporal range, can be encoded as follows:

:240: $\{dx_M_de\}$:1535: $[dx_AB_at-ad-si_]$:3365: $\{RE\}$:335: $[dx_AB_su-la_]$:175: $\{RE\}$:405: $[dx^{\#1}AB_at-ad-2si_]$:685: $\{RE\}$:240: $[dx_AB_at-ad-2si-su_]$:130: $\{RE\}$:100: $[dx^{\#2}M_de-3si_]$:1170: $\{RE\}$;
:240: $\{si_M_de\}$:1535: $[si_AB_at-ad-dx-de_]$:3365: $\{RE\}$:335: $[si_M_2de-la_]$:175: $\{RE\}$:405: $[si^{\#1}AB_at-ad-dx_]$:685: $\{RE\}$:240: $[si_AB_at-ad-dx-5de_]$:130: $\{RE\}$:100: $[si^{\#2}M_de-3dx_]$:1170: $\{RE\}$

Based on the presented formula, it is noted that the examined sequence represents a complex hand movement pattern, which includes not only elementary manipulations but also more complex forms of interaction between the hands, indicating a high level of coordination and synchronization of movements within this sequence. Thus, it can be emphasized that the proposed gesture notation is an effective tool for systematizing and displaying not only primitive but also complex movements. This notation allows for accurate and detailed recording of any gesture interactions, ensuring a precise representation of their structure and sequence.

Table 3. The ratio of terms and contexts of the emotion “Anger”
Таблица 3. Соотношение термов и контекстов эмоции «Злость»

	Term Number	1	2	3
Tracks	Track 1	:240:[_dx_M_de_]	:1535:[_dx_AB_at-ad-si_]:3365:{RE}	:335:[_dx_AB_sula_]:175:{RE}
	Track 2	:240:[_si_M_de_]	:1535:[_si_AB_at-ad-dx-de_]:3365:{RE}	:335:[_si_M_2de-la_]:175:{RE}
Linguistic Context			Злость – это когда постоянно хочется прямо подраться	или ты хочешь

	Term Number	4	5	6
Tracks	Track 1	:405:[_dx^#1^AB_at-ad-2si_]:685:{RE}	:240:[_dx_AB_at-ad-2si-su_]:130:{RE}	:100:[_dx^#2^M_de-3si_]:1170:{RE}
	Track 2	:405:[_si^#1^AB_at-ad-dx_]:685:{RE}	:240:[_si_AB_at-dx-5de_]:130:{RE}	:100:[_si^#2^M_de-3dx_]:1170:{RE}
Linguistic Context		стукнуть по столу	сказать	что нет ты не прав

The notations presented in Tables 2 and 3 can be enriched with additional parameters. The first set consists of dynamically variable behavioral parameters, based on the already identified parameters (sequence of terms in the formula, speech material, duration of gestures), and includes metrics derived from these parameters, as well as optional physiological characteristics that can be measured in experimental situations (pulse, tidal volume, saturation, etc.). The second set consists of optional constants (for ongoing communication) describing socio-demographic, psychological, and physiological parameters of the individual (gender, age, education, height, arm span, etc.).

The construction of a hierarchy of metrics for gesture analysis and their constraints relies on extensive research materials. Recent studies (Liu et al., 2022; Qi et al., 2024; Tonoli et al., 2024; Yang et al., 2023; Yoon et al., 2020; Zhi et al., 2023) discuss metrics such as Mean Squared Error (MSE), Mean Average Error (MAE), Fréchet Gesture Distance (FGD), Gesture Area

Coverage (GAC), among others, which compare movements to a reference, assess naturalness and similarity, analyze dynamic characteristics (speed, acceleration, etc.), and consider spatial coverage of gestures.

It is proposed that some metrics might also be used for evaluating speech-gesture behavior. Assuming that the integral behavioral unit is the term (as a final state of movement, represented in notation by a transition from the previous state to the next via movement, i.e., as {termK} → [termK+1]), all measurements should be tied to terms. This approach leads to the following metrics:

1. **L** – Length of the hand movement trajectory. Calculations can be performed for a segment (termK → termK+1) as well as for the entire chain of movement for one or both hands: term1 → term2 → ... → termN.

2. **S** – Area under the curve of the hand movement trajectory within a single gesture, i.e., the designated behavioral unit termK → termK+1, and for the entire movement.

3. **V** – Volume of space captured by the gesture (term $K \rightarrow$ term $K+1$). Measurement methods may vary from the space of a single hand or both hands to the space considered between two hands in the gesture.

Parameters **L**, **S**, and **V** require normalization in the description of real communication (without using 3D models); we propose introducing a normalizing parameter = $\frac{L}{N^2}$, where (l_1 – height, l_2 – arm span) for **L**, N^2 for **S**, and N^3 for **V**.

4. **speech** – Volume of speech material associated with the duration of the gesture, i.e., term $K \rightarrow$ term $K+1$.

5. **temp** – Metric characterizing the dynamics of the space captured by gestures: = $\frac{V}{t}$. The temp metric can also be computed differently, such as the ratio of the number of terms or volume of speech material to the unit of time.

Overall, the recording of the function arguments for Behavior can be represented as follows:

Behavior = {(term $_1$, speech $_1$, t $_1$, L $_1$, S $_1$, V $_1$, ...),..., (term $_N$, speech $_N$, t $_N$, L $_N$, S $_N$, V $_N$, ...); (pulse $_1$, tidal volume $_1$, ...)*, ..., (pulse $_N$, tidal volume $_N$, ...)*; (gender, age, N, temp...)*}, where the asterisk (*) denotes groups of optional parameters.

Conclusions and Future Work

The creation of a formalized system for encoding human movements for interface-mediated communication in various machine-generated environments, as we have attempted to demonstrate, is a complex yet solvable scientific problem. In the course of the study, the Perm Gesture Notation system was introduced, and the unit for describing human behavior – the term – was identified. A term represents the record of a transition from a preceding fixed state of a body part within its movement to a subsequent fixed state. As a unit of movement discretization, a term aligns with other forms of activity (i.e., it functions as a cognitive-behavioral unit), such as speech (we discussed the volume of speech material, though other measurable parameters, such as acoustic measures, may also be relevant), as

well as facial expressions or data from wearable devices that convey information about pulse, breathing volume, and other factors.

Given that each human movement likely has semantic meaning, i.e., is not random and/or chaotic in the sequence of human behavior acts, we should ultimately require machines to achieve similar semantic understanding of motor activity and synchronize it with the speech semantics of the message. Complete and accurate tempo-rhythmic and spatial reproduction of human movements by a machine (i.e., essentially replicating the human chronotope by its virtual avatar) can be achieved by combining computer vision methods with modern neural network transformer models. The ultimate goal of this approach is to generate a movement formula in the developed formal language notation based on the analysis of video sequences of human movements input into the model. The generated formula can then be transmitted to an executor (either a virtual 3D model or a physical robotic device). Reproducing the encoded movements by the executor is necessary for analyzing the accuracy and realism of the generated animations through comparison with actual movements, followed by formula adjustments and model retraining.

It is evident that the quality of the model's reproduction of human speech-gesture behavior by the machine should ultimately be assessed by humans themselves. It is important that machines not only replicate physical movements but also appropriately convey the semantic and emotional values of these movements, in accordance with human expectations and cultural norms. In this context, there is a need to develop a separate branch of general semantics – technosemantics, as semantics generated by anthropomorphic machines in human-machine interaction and directed towards the communicating human (for issues related to technosemantics, see (Belousov et al., 2024)).

The synthesis of movements, which is oriented not only towards mechanical reproduction but also towards meaningful content, is a crucial component in creating more effective and human-centered communication systems in various machine-generated environments. It is important to consider not only visual and kinematic parameters but also the hidden meanings carried by gestures and movements in the context of communication. This task becomes increasingly relevant in the context of the development of anthropomorphic machines and large language models, which strive not only to imitate humans but also to communicate with them at a level close to human interaction.

In conclusion, further development of movement encoding systems, including those presented in this work, and their application in various fields such as the film industry, computer games, education, medicine, sports, and robotics, could lay the foundation for new forms of communication where machines not only execute commands but also actively engage in semantically meaningful interactions, maintaining a comprehensive and intuitive dialogue with humans.

References

Abbie, M. (1974). Movement notation, *The Australian journal of physiotherapy*, 20 (2), 61–69. [https://doi.org/10.1016/S0004-9514\(14\)61177-6](https://doi.org/10.1016/S0004-9514(14)61177-6) (In English)

Bashan, M., Einbinder, H., Harries, J., Shosani, M. and Shoval, D. (2024). *Movement Notation: Eshkol and Abraham Wachmann*, Verlag der Buchhandlung Walther König, Köln, Germany. (In English)

Belousov, K. I., Sazina, D. A., Ryabinin, K. V. and Brokhin, L. Yu. (2024). Sensory Technolinguistics: On Mechanisms of Transmitting Multimodal Messages in Perceptual-Cognitive Interfaces, *Automatic Documentation and Mathematical Linguistics*, 58 (2), 108–116. <https://doi.org/10.3103/s0005105524700079> (In English)

Benesh, R. and Benesh, J. (1956). *An Introduction to Benesh Dance Notation*, A & C Black, London, UK. (In English)

Bernardet, U., Fdili Alaoui, S., Studd, K., Bradley, K., Pasquier, P. and Schiphorst, T. (2019). Assessing the reliability of the Laban Movement Analysis system, *PLoS ONE*, 14 (6): e0218179. <https://doi.org/10.1371/journal.pone.0218179> (In English)

Birdwhistell, R. L. (1952). *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture*, Foreign Service Institute, Washington, DC, USA. (In English)

Bull, P. and Doody, J. P. (2013). 8 Gesture and body movement, *De Gruyter eBooks*, 205–228. <https://doi.org/10.1515/9783110238150.205> (In English)

Calvert, T. (2015). Approaches to the Representation of Human Movement: Notation, Animation and Motion Capture, *Dance Notations and Robot Motion, Springer Tracts in Advanced Robotics*, 111, 49–68. https://doi.org/10.1007/978-3-319-25739-6_3 (In English)

Dael, N., Mortillaro, M. and Scherer, K. R. (2012). The Body Action and Posture Coding System (BAP): Development and Reliability, *Journal of Nonverbal Behavior*, 36 (2), 97–121. <https://doi.org/10.1007/s10919-012-0130-0> (In English)

Dell, C. (1977). *A Primer for Movement Description: Using Effort-shape and Supplementary Concepts*, Dance Notation Bureau Press, New York, USA. (In English)

Duprey, S., Naaim, A., Moissenet, F., Begon, M. and Chèze, L. (2017). Kinematic models of the upper limb joints for multibody kinematics optimisation: An overview, *Journal of Biomechanics*, 62, 87–94. DOI: 10.1016/j.jbiomech.2016.12.005 (In English)

Ekman, P. and Friesen, W. V. (1978). *Facial Action Coding System*, Consulting Psychologists, Palo Alto, CA, USA. (In English)

El Raheb, K. and Ioannidis, Y. (2014). From dance notation to conceptual models: a multilayer approach, *Proceedings of the 2014 International Workshop on Movement and Computing, MOCO, ACM, New York*, 25–30. (In English)

El Raheb, K., Buccoli, M., Zanoni, M., Katifori, A., Kasomoulis, A., Sarti, A. and Ioannidis, Y. (2023). Towards a general framework for the annotation of dance motion sequences, *Multimed Tools Appl*, 82, 3363–3395. <https://doi.org/10.1007/s11042-022-12602-y> (In English)

- Eshkol, N. and Wachmann, A. (1958). *Movement Notation*, Weidenfeld and Nicolson, London, UK. (In English)
- Farnell, B. M. (1996). Movement Notation Systems, *The World's Writing Systems*, in Daniels P. T. (ed.), 855–879. (In English)
- Frishberg, N. (1983). Writing systems and problems for sign language notation, *Journal for the Anthropological Study of Human Movement*, 2 (4), 169–195. (In English)
- Frey, S., Hirsbrunner, H-P. and Jorns, U. (1982). *Time-Series Notation: A Coding Principle for the Unified Assessment of Speech and Movement in Communication Research*, Gunter NarrVerlag, Tübingen, Germany. (In English)
- Grushkin, D. A. (2017). Writing Signed Languages: What For? What Form?, *American Annals of the Deaf*, 161 (5), 509–527. <https://doi.org/10.1353/aad.2017.0001> (In English)
- Guest, A. H. (1984). *Dance Notation: The Process of Recording Movement on Paper*, Dance Horizons, New York, USA. (In English)
- Guest, A. H. (2005). *Labanotation: The System of Analyzing and Recording Movement (4th ed.)*, Routledge, New York, USA. <https://doi.org/10.4324/9780203823866> (In English)
- Harrigan, J. A. (2008). Proxemics, Kinesics, and Gaze, *The New Handbook of Methods in Nonverbal Behavior Research*, 136–198. <https://doi.org/10.1093/acprof:oso/9780198529620.003.0004> (In English)
- Izquierdo, C. and Anguera, M. T. (2018). Movement notation revisited: syntax of the common morphokinetic alphabet (CMA) system, *Front. Psychol.*, 9:1416. <https://doi.org/10.3389/fpsyg.2018.01416> (In English)
- Karg, M., Samadani, A.-A., Gorbet, R., Kuhnlenz, K., Hoey, J. and Kulic, D. (2013). Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation, *IEEE Transactions on Affective Computing*, 4 (4), 341–359. <https://doi.org/10.1109/t-affc.2013.29> (In English)
- Kendon, A. (1997). Gesture, *Annual Review of Anthropology*, 26 (1), 109–128. <https://doi.org/10.1146/annurev.anthro.26.1.109> (In English)
- Key, M. R. (1977). *Nonverbal communication: a research guide and bibliography*, The Scarecrow Press, Metuchen, N.J., USA. (In English)
- Kilpatrick, C. E. (2020). Movement, Gesture, and Singing: A Review of Literature. Update: Applications of Research in Music Education, 38 (3), 29–37. DOI: 10.1177/8755123320908612 (In English)
- Laban, R. von and Lawrence, F. C. (1974). *Effort: Economy of Human Movement 2nd ed.*, Macdonald & Evans, London, UK. (In English)
- Laumond, J. and Abe, N. (2016). *Dance Notations and Robot Motion*, Springer International Publishing AG, Cham (ZG), Switzerland. <https://doi.org/10.1007/978-3-319-25739-6> (In English)
- Liu, H., Zhu, Z., Iwamoto, N., Peng, Y., Li, Zh., Zhou, Y., Bozkurt, E. and Zheng, B. (2022). BEAT: A Large-Scale Semantic and Emotional Multi-modal Dataset for Conversational Gestures Synthesis, *Computer Vision – ECCV 2022*, 612–630. <https://doi.org/10.48550/arXiv.2203.05297> (In English)
- Murillo, E., Montero, I. and Casla, M. (2021). On the multimodal path to language: The relationship between rhythmic movements and deictic gestures at the end of the first year, *Frontiers in Psychology*, 12, 1–8. <https://doi.org/10.3389/fpsyg.2021.616812> (In English)
- Novack, A. M. and Wakefield, E. M. (2016). Goldin-Meadow S. What makes a movement a gesture?, *Cognition*, 146, 339–348. <https://doi.org/10.1016/j.cognition.2015.10.014> (In English)
- Qi, X., Liu, C., Li, L., Hou, J., Xin, H. and Yu, X. (2024). Emotion Gesture: Audio-Driven Diverse Emotional Co-Speech 3D Gesture Generation, *IEEE Transactions on Multimedia*, 1–11. <https://doi.org/10.1109/TMM.2024.3407692> (In English)
- Streeck, J. (2010). The Significance of Gesture: How it is Established, *Papers in Pragmatics*; 2 (1-2). <https://doi.org/2.10.1075/iprapip.2.1-2.03str> (In English)
- Shafir, T., Tsachor, R. and Welch, K. B. (2016). Emotion Regulation through Movement: Unique Sets of Movement Characteristics are Associated with and Enhance Basic Emotions, *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.02030> (In English)

Stults-Kolehmainen, M. A. (2023). Humans have a basic physical and psychological need to move the body: Physical activity as a primary drive, *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1134049> (In English)

Tonoli, R. L., Costa, P. D. P., Marques, L. B. d. M. M. and Ueda, L. H. (2024). Gesture Area Coverage to Assess Gesture Expressiveness and Human-Likeness', *International Conference on Multimodal Interaction (ICMI Companion '24)*, 4–8 November 2024, San Jose, Costa Rica. ACM, New York, NY, USA. <https://doi.org/10.1145/3686215.3688822> (In English)

Trujillo, J. P., Vaitonyte, J., Simanova, I. and Özyürek, A. (2018). Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research, *Behavior Research Methods*, 51 (2), 769–777. <https://doi.org/10.3758/s13428-018-1086-8> (In English)

Van Elk, M., van Schie, H. T. and Bekkering, H. (2009). Short-term action intentions overrule long-term semantic knowledge, *Cognition*, 111 (1), 72–83. <https://doi.org/10.1016/j.cognition.2008.12.002> (In English)

Yang, S., Wu, Z., Li, M., Zhang, Z., Hao, L., Bao, W. and Zhuang, H. (2023). QPGesture: Quantization-Based and Phase-Guided Motion Matching for Natural Speech-Driven Gesture Generation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2321–2330. <https://doi.org/10.48550/arXiv.2305.11094> (In English)

Yoon, Y., Cha, B., Lee, J.-H., Jang, M., Lee, J., Kim, J. and Lee, G. (2020). Speech Gesture Generation from the Trimodal Context of Text, Audio, and Speaker Identity, *ACM Transactions on Graphics*, 39 (6). <https://doi.org/10.1145/3414685.3417838> (In English)

Zhi, Y., Cun, X., Chen, X., Shen, X., Guo, W., Huang, S. and Gao, S. (2023). LivelySpeaker: Towards Semantic-Aware Co-Speech Gesture Generation, *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 20807–20817. <https://doi.org/10.1109/ICCV51070.2023.01902> (In English)

The authors have read and approved the final manuscript.

Авторы прочитали и одобрили окончательный вариант рукописи.

Conflicts of interests: the authors have no conflicts of interest to declare.

Конфликты интересов: у авторов нет конфликтов интересов для декларации.

Konstantin I. Belousov, Doctor of Philology, Professor, Department of Theoretical and Applied Linguistics, Perm State University, Perm, Russia.

Белусов Константин Игоревич, доктор филологических наук, профессор кафедры теоретического и прикладного языкознания, Пермский государственный национальный исследовательский университет, Пермь, Россия.

Aleksandar Taleski, Candidate of Philological Sciences, Associate Professor, Department of Theoretical and Applied Linguistics, Perm State University, Perm, Russia.

Талески Александр, кандидат филологических наук, доцент кафедры теоретического и прикладного языкознания, Пермский государственный национальный исследовательский университет, Пермь, Россия.

Artem R. Agaev, software developer, Human Semantics LLC, Perm, Russia.

Агаев Артем Русланович, разработчик программного обеспечения, ООО «Хьюмен Семантикс», Пермь, Россия.

УДК 811.93

DOI: 10.18413/2313-8912-2024-10-4-0-7

Белоусов К. И.¹
Талески А.²
Агаев А. Р.³

Техносемантика жеста: о возможностях использования пермской жестовой нотации в программно-генерируемых средах

¹ Пермский государственный национальный исследовательский университет,
ул. Букирева, 15, Пермь, 614068, Россия
E-mail: belousovki@gmail.com
ORCID: 0000-0003-4447-1288

² Пермский государственный национальный исследовательский университет,
ул. Букирева, 15, Пермь, 614068, Россия
E-mail: taleski87@yahoo.com
ORCID: 0000-0002-1901-8572

³ООО «ХьюменСемантикс»
15 Bukireva St., Perm, 614068, Russia
E-mail: agaev-artem1@yandex.ru
ORCID: 0009-0001-8948-2143

*Статья поступила 07 сентября 2024 г.; принята 15 декабря 2024 г.;
опубликована 30 декабря 2024 г.*

Аннотация: Статья посвящена разработке концепции и программного решения генерации движений человека на основе созданной коллективом семантико-ориентированной жестовой нотации. Нотация представляется в виде формулы, отличающейся гибкой структурой понятий и правил их реализации, что позволяет легко адаптировать изменения параметров движения для соответствия идеальному или реальному образцу.

Для моделирования движений, представленных в языковой нотации, разработано кроссплатформенное приложение на базе игрового движка Blender 4.2 для визуализации и генерации анимации жестов антропоморфных моделей. Система управления движений состоит из следующих стадий: запись жестовой нотации; разбор этой записи; построение внутреннего представления движения, которое является последовательностью ряда кадров. Кадры имеют информацию о том, к какой кости (части тела) исполнителя они относятся, как влияют на её положение, в какой момент времени от начала движения этот кадр актуален. На последнем этапе осуществляется преобразование внутреннего представления движения в движение исполнителя, в качестве которого могут выступать как виртуальные антропоморфные 3D-модели, так и материальные программно-аппаратные комплексы в виде антропоморфных роботов.

Для большего соответствия антропоморфному поведению помимо «идеальных образцов» движения человека были использованы модели жестового поведения людей, представленные в специально создаваемом для этой цели коллективом мультимодальном корпусе. Материал представляет собой аудиовизуальные записи устных спонтанных текстов описания реципиентами широкого спектра собственных эмоциональных состояний. Полученные в ходе

экспериментального исследования данные подтвердили гибкость, повышенную контролируемость и модульность языковой нотации, возможность моделирования непрерывного пространства человеческой двигательной активности.

Ключевые слова: Техносемантика; Генерация движений; Языковая нотация жестов; Жест; Визуализация; Мультимодальный корпус; 3D-графика; Интерпретируемость

Информация для цитирования: Белоусов К. И., Талески А., Агаев А. Р. Техносемантика жеста: о возможностях использования пермской жестовой нотации в программно-генерируемых средах // Научный результат. Вопросы теоретической и прикладной лингвистики. 2024. Т. 10. № 4. С. 126–146. DOI: 10.18413/2313-8912-2024-10-4-0-7

ТЕХНОСЕМАНТИКА ЖЕСТА: О ВОЗМОЖНОСТЯХ ИСПОЛЬЗОВАНИЯ ПЕРМСКОЙ ЖЕСТОВОЙ НОТАЦИИ В ПРОГРАММНО ГЕНЕРИРУЕМЫХ СРЕДАХ

Проблема

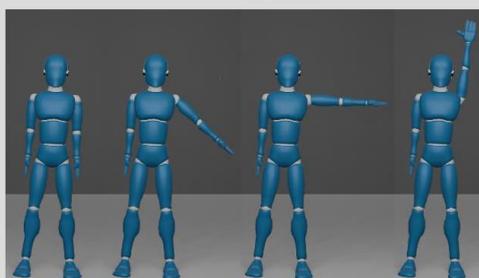
Отсутствие универсального языка для записи движений человека и антропоморфных роботизированных моделей в человеко-машинной коммуникации

Решение

Разработана концепция для моделирования и генерации движений человека и антропоморфных роботизированных моделей на основе семантически ориентированной жестовой нотации

Результаты

- Для описания движений создана Пермская жестовая нотация
- Разработано кроссплатформенное приложение для визуализации и генерации анимации жестов антропоморфных моделей
- Для достижения правдоподобности антропоморфного жестового поведения использовались аудиовизуальные записи описания людьми эмоциональных состояний



Генерация движений 3D-моделей



Моделирование жестового поведения



UDC 811.93

DOI: 10.18413/2313-8912-2024-10-4-0-7

Konstantin I. Belousov¹
Aleksandar Taleski²
Artem R. Agaev³

Technosemantics of gesture: on the possibilities of using Perm sign notation in software-generated environments

¹ Perm State University,
15 Bukirev St., Perm, 614068, Russia
E-mail: belousovki@gmail.com

ORCID: 0000-0003-4447-1288

²Perm State University,
15 Bukirev St., Perm, 614068, Russia
E-mail: taleski87@yahoo.com
ORCID: 0000-0002-1901-8572

³Human Semantics LLC
15 Bukirev St., Perm, 614068, Russia
E-mail: agaev-artem1@yandex.ru
ORCID: 0009-0001-8948-2143

Received 07 September 2024; accepted 15 December 2024; published 30 December 2024

Abstract: This paper is dedicated to the development of a concept and software solution for generating human movements based on a semantically-oriented language notation created by the authors. The language notation is presented as a formula with a flexible structure of concepts and rules for their implementation, allowing easy adaptation of movement parameter changes to match an ideal or real sample.

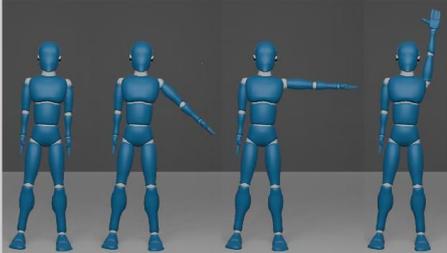
To model movements represented in the language notation, a cross-platform application was developed using the Blender 4.2 for the visualization and generation of gestures for anthropomorphic models. The movement control system consists of the following stages: translating it into a gesture notation record; parsing this record; constructing an internal representation of the movement, which is a sequence of frames. Frames contain information about which bone (body part) of the performer they refer to, how they affect its position, and at what time from the start of the movement this frame is relevant. In the final stage, the internal representation of the movement is transformed into the performer's movement, which can be either virtual anthropomorphic 3D models or physical software-hardware systems in the form of anthropomorphic robots.

To better correspond to anthropomorphic behavior, in addition to "ideal samples" of human movement, models of human gesture behavior presented in a multimodal corpus specially created for this purpose by the team were used. The material consists of audiovisual recordings of spontaneous oral texts where recipients describe a wide range of their own emotional states. The data obtained during the experimental research confirmed the flexibility, enhanced controllability, and modularity of the language notation, as well as the ability to model the continuous space of human motor activity.

Keywords: Technosemantics; Movement generation; Language notation of gestures; Gesture; Visualization; Multimodal corpus; 3D graphics; Interpretability

How to cite: Belousov, K. I., Taleski, A., Agaev, A. R. (2024). Technosemantics of gesture: on the possibilities of using Perm sign notation in software-generated environments, *Research Result. Theoretical and Applied Linguistics*, 10 (4), 126–146. DOI: 10.18413/2313-8912-2024-10-4-0-7

TECHNOSEMANTICS OF GESTURE: ON THE POSSIBILITIES OF USING PERM SIGN NOTATION IN SOFTWARE-GENERATED ENVIRONMENTS

Problem	Solution	Results
<p>The absence of a universal language for recording human and anthropomorphic robotic models' movements in human-machine communication</p>	<p>A concept has been developed for modeling and generating human and anthropomorphic robotic models' movements based on a semantically oriented linguistic notation</p>	<ul style="list-style-type: none">• For describing movements, the Perm Gesture Notation was created• A cross-platform application has been developed for visualizing and generating gesture animations of anthropomorphic models• To achieve plausible anthropomorphic gestural behavior, audiovisual recordings of humans describing emotional states were used
 <p>Generating 3D-Model Movements</p>		 <p>Modeling Gesture Behavior</p>
		

Введение

Разработка систем описания движений человека является актуальной задачей в различных областях деятельности, таких как хореография, робототехника, анимация и мн.др. Подобно тому как музыкальная партитура представляет музыкальные ноты, нотация движений призвана передать точные инструкции в отношении конкретных движений, действий, а также их комплексов и/или последовательностей. Создание нотаций, позволяющих описывать движения в разных предметных областях, сталкивается с рядом проблем. Первая проблема заключается в отсутствии универсального языка для записи движений. Различные предметные области используют отличающиеся друг от друга методы и языки для записи движений, что приводит к ситуации «Вавилонской башни», когда общение между дисциплинами становится затруднительным (Laumond, Abe, 2016). Вторая проблема возникает из-за уже

сложившихся в разных предметных областях своих терминосистем, парадигм и целей (ср. хореографию и робототехнику), что еще больше усложняет междисциплинарное общение и препятствует интеграции наук (Calvert, 2015).

В то же время метаязыковая и концептуальная фрагментарность дисциплинарных областей входит в противоречие с качественными изменениями в природе человеческой коммуникации и деятельности, обусловленными выдвиганием на передний план интерфейсно-опосредованной коммуникации и цифровизации многочисленных форм деятельности. Такая коммуникация требует универсальных «бесшовных» инструментов (метаязыка и технологий) для перехода от одной формы человеческой активности к другой. Отсюда вытекает потребность в поиске новых, возможно универсальных, подходов к описанию взаимодействия человека и

машины (НМИ) и взаимодействия человека и компьютера (НСИ).

Электронные вычислительные устройства и виртуальные антропоморфные 3D-модели, такие как роботы и виртуальные аватары, имитируют поведение человека, общение, эмоции, движения, жесты и другие человеческие черты и поведенческие паттерны. Поскольку эти антропоморфные модели становятся все более интегрированными в повседневную жизнь, они должны уметь оперировать множеством человеческих действий и языковых выражений. Если мы говорим о движениях человека, то для эффективного взаимодействия с людьми и извлечения информации из лингвистических и жестовых данных, эти интеллектуальные модели должны понимать референциальные отношения между действиями и языком, а также интерпретировать действия, передаваемые через язык.

Разработка осязаемых человеко-машинных интерфейсов основывается на принципах человеческого общения, где информация воспринимается, обрабатывается и передается с использованием нескольких модальностей (Belousov et al., 2024). Семантика в этом контексте охватывает не только слуховые или визуальные сообщения, но и моторные действия. Например, ван Элк и коллеги описывают семантику действий как иерархически организованную систему, подсистемы которой избирательно активируются в зависимости от намерений субъекта и контекста задачи (van Elk et al., 2009). В естественной коммуникации передача информации осуществляется сразу по нескольким каналам разной модальности для «сужения» и фокусирования области смысла, например, пространственный дейксис в ситуации непосредственного общения требует жестового участия: «Посмотри там!» (указание рукой или кивком головы).

Обычно проблема описания движений связана с обсуждением различий между движениями и жестами. Существует несколько подходов к различению движения

и жеста. Так, движение можно рассматривать как понятие, которое охватывает любую физическую активность тела, в том числе и жесты, которые, в свою очередь, несут символическое значение в коммуникации (Bull, Doody, 2013; Kendon, 1997; Kilpatrick, 2020; Murillo et al., 2021; Novack, Wakefield, 2016: 340; Stults-Kolehmainen, 2023). Согласно подходу А.М. Новака и коллег жест не связан с непосредственным изменением мира (например, открытие банки путем ее поворота) или не отображает образцы движения (например, выполнение шагов в танце), а вместо этого представляет движение, которое может изменить мир (например, жест, показывающий, как можно открыть банку, повернув ее), или представляет движение, которое существует само по себе (например, жест, показывающий, как следует исполнять танец) (Novack, Wakefield, 2016: 340). Оба приведенные примера, на наш взгляд, в целом корректны, однако можно обнаружить контексты, в которых танцевальные движения или открытие банки можно рассматривать в качестве жестов. Так, профессиональные танцевальные движения (например, в балете) коммуникативно и семантически нагружены и будучи обращены к зрителю интерпретируются, подобно тому, как интерпретируется любой другой знак. Пример священным преобразованием мира («открытие банки» и под.), также можно погрузить в контекст, в котором такое преобразование может рассматриваться как знак профессионального мастерства.

Нам ближе понимание взаимосвязи жеста и движения, в котором движение является не только родовым понятием для жеста, но и объединяет их в следующих друг за другом последовательностях (Kendon, 1997; Kilpatrick, 2020; Streeck, 2010; Trujillo et al., 2018). При этом если фрагмент движения коммуникативно направлен и семантически целостен, то его можно рассматривать в качестве жеста. Но заметим, что точно такое же движение может и не являться жестом, если оно не имеет

адресата, способного интерпретировать данный знак (в качестве адресата может выступать сам адресант, т.е. речь идет об автокоммуникации).

В современном технологичном мире жесты являются формами невербального общения, используемыми не только для взаимодействия с людьми, но также с антропоморфными виртуальными или роботизированными моделями. Запись человеческих движений, разработка моделей и алгоритмов для эффективного распознавания движений, создание единой нотации движений и улучшение человеко-машинной коммуникации являются актуальными задачами в науке, робототехнике, машинном обучении и многих других областях.

Цели нотации движений включают – будь то в танце, анимации или робототехнике, – и предоставление четких, интерпретируемых инструкций, которые могут быть выполнены исполнителями, машинами или цифровыми персонажами. Точная нотация движений необходима для их настоящего и будущего использования, изучения или воспроизведения, обеспечивая согласованность в различных ситуациях и среди различных исполнителей (как людей, так и антропоморфных моделей). В этой статье мы представляем новую концепцию и программное решение для моделирования и генерации движений на основе семантически ориентированной языковой нотации, названной нами Пермской жестовой нотацией. Основная задача этой статьи – представить методологию моделирования человеческих движений на основе разработанной языковой нотации. Это включает представление разработанной нотации, описание приложения, реализующего систему управления движениями, основанного на движке Blender 4.2. В разработанном приложении осуществляется визуализация и генерация движений для антропоморфных моделей, а также композиционное и последовательное управление сгенерированными движениями на основе разработанной нотации. Для достижения правдоподобности

антропоморфного поведения 3D-модели мы используем мультимодальный корпус аудиовизуальных записей описания людьми различных эмоциональных состояний. Это подтверждает возможность моделирования широкого спектра человеческих моторных действий.

Теоретический обзор

Создание системы нотации движений имеет многовековую историю. С XV века в Европе и Северной Америке использовалось как минимум 87 различных систем нотации движений (Abbie, 1974; Farnell, 1996; Guest, 1984; Key, 1977). Эти системы часто возникали для документирования определенных стилей танца или жестов (языков жестов), и многие из них выходили из употребления по мере эволюции или исчезновения связанных с ними практик движения (Farnell, 1996).

В области исследований невербального общения любая попытка разработать алфавит движений неизбежно должна учитывать значимые системы из истории нотации движений, включая систему Бошана-Фёйе, которая процветала при французском дворе в XVII и XVIII веках, систему жестов для публичных выступлений Гилберта Остина 1806 года, кинесику Рэя Бёрдвистелла середины XX века, систему записи американского языка жестов Уильяма Стокое и адаптации для других языков жестов Адама Кендона и Ла Монта Веста младшего, обобщенные системы нотации движений, такие как Лабанотация, анализ движений Лабана, хореология Бенеша и нотация Эшкола-Вахмана, которые предназначены для адаптации к различным формам движений и функционируют как Международный фонетический алфавит для движений, способный записывать любое телесное действие, сохраняя его семантическое содержание (Abbie, 1974; Bashan et al., 2024; Benesh, Benesh, 1956; Bernardet et al., 2019; Birdwhistell, 1952; Dell, 1977; El Raheb, Ioannidis, 2014; 2023; Eshkol, Wachmann, 1958; Farnell, 1996; Frishberg, 1983; Grushkin, 2017; Guest, 1984; 2005; Harrigan, 2008;

Kendon, 1997; Key, 1977; Laban, Lawrence, 1974).

Несколько ключевых систем описаний движений возникли из смежных областей антропологии, этологии и психологии. Система кодирования лицевых действий (FACS) (Ekman, Friesen, 1978) предназначена для анализа и интерпретации мимических выражений путем идентификации конкретных лицевых действий. Бернская система (Frey et al., 1982) предлагает методологию для анализа движений в психологических и антропологических исследованиях. Самая последняя система – система кодирования телесных действий и поз (BAP) (Dael et al., 2012) разработана для детального анализа движений и поз тела с использованием 141 поведенческой переменной, фиксирующей сложные временные сегменты движения. Эти системы предоставляют обширный инструментарий для изучения невербального общения и человеческого поведения в различных научных и прикладных контекстах (Bernardet et al., 2019).

Системы классификации, кодирования и интерпретации движений различаются по ряду параметров, определяющих их целевое назначение, происхождение, уровень детализации и надежность. Во-первых, эти системы различаются по своему назначению: одни предназначены для воспроизведения или анализа движений, в то время как другие направлены на создание заключения/выводов о субъекте, выполняющем движения. Например, системы, происходящие из области танца, такие как Лабанотация и анализ движений Лабана, сосредоточены на точной передаче и анализе движений танцоров. Напротив, системы, возникшие из антропологии, этологии и психологии, такие как FACS и кинесика, направлены на изучение и интерпретацию человеческого поведения в более широком контексте (Birdwhistell, 1952; Bernardet et al., 2019; Duprey et al., 2017; Ekman, Friesen, 1978; Laban et al., 1974).

Во-вторых, уровень субъективного вывода различается среди разных систем.

Некоторые системы полагаются на субъективные суждения кодировщиков, другие используют систематические метки и наблюдательное кодирование, а третьи – на прямое измерение мышечной активности. Это различие особенно важно при выборе системы для научных исследований, где требуется высокая точность и надежность данных (Bernardet et al., 2019).

Фокус на качестве или типе движения также является отличительной чертой различных систем. Например, Лабанотация и Бернская система подчеркивают качество движения, тогда как кинесика концентрируется на типах движений. Уровень детализации системы, включающий такие параметры, как минимальная временная единица для кодирования, размер алфавита (количество различных кодов) и количество одновременно кодируемых дорожек (частей тела), играет ключевую роль в определении уровня детализации системы. Чем меньше временная единица, больше алфавит и больше дорожек, тем более детальной является система. Это, в свою очередь, влияет на время кодирования: детализированные системы требуют значительно больше времени для кодирования данных, чем менее детализированные (Birdwhistell, 1952; Bernardet et al., 2019; Dael et al., 2012; Ekman, Friesen, 1978; Farnell, 1996; Laban, Lawrence, 1974).

Надежность систем также различается: некоторые системы, такие как FACS и BAP, прошли эмпирическую оценку надежности, тогда как другие, такие как LMA, были оценены в ограниченной степени, и существуют системы, надежность которых вообще не была оценена (Bernardet et al., 2019; Dael et al., 2012; Ekman, Friesen, 1978; Laban, Lawrence, 1974; Shafir et al., 2016).

Таким образом, различия между системами классификации, кодирования и интерпретации движений включают их целевое назначение, происхождение, степень субъективного вывода, фокус на качестве или типе движения, уровень детализации, время кодирования и надежность. М. Карг с соавторами (Karg et al., 2013) подчеркивают,

что движения обычно представляются траекториями углов суставов или кинематическими признаками, которые, хотя и детализированы, могут быть вычислительно затратными, а также могут игнорировать качества выразительности движений. Системы нотации движений предлагают более эффективное и выразительное представление.

В связи с концептуальной близостью к нашей системе нотации, мы уделяем особое внимание системе Common Morphokinetic Alphabet (CMA) для транскрипции и анализа движений тела, которая стремится обеспечить комплексный, точный и удобный метод записи физических движений, устанавливая грамматические правила на основе фразовой грамматики Ноама Хомского (Izquierdo, Anguera, 2018). Как и наша нотация, система CMA включает принципы лингвистики, особенно структуру и синтаксис фраз движений, для создания системы нотации, способной захватить сложность человеческого движения в трехмерном пространстве во времени (Izquierdo, Anguera, 2018). Система CMA основана на предыдущих системах, таких как кинесическая система записи движений Рэя Бёрдвистелла (KNS) (Birdwhistell, 1952) и система кодирования лицевых действий (FACS) (Ekman, Friesen, 1978), но, в отличие от них, эта система позиционирует себя как более комплексную и адаптируемую систему, способную идти в ногу с технологическими достижениями (Izquierdo, Anguera, 2018).

Следует подчеркнуть, что исследования, проведенные Izquierdo, Anguera (2018), очень тесно связаны с нашей работой в некоторых аспектах, которые будут обсуждены в следующем разделе.

Методология.

Система нотации движений в нашем понимании должна удовлетворять нескольким ключевым критериям.

Нотация движений должна быть семантико ориентированной, основанной на естественном языке. В этом случае предметный эксперт на этапе написания формулы движения (или ее

чтении) представлял бы, как будет совершаться действие.

Язык записи движений должен быть представлен в виде формализованной структуры понятий и правил их реализации. Это позволит использовать компьютерные технологии, в частности, чтение интеллектуальным интерпретатором закодированного сообщения.

Язык для записи движений должен быть гибким, расширяемым за счет внедрения новых объектов, понятий и правил, а также легко адаптируемым к возможности изменения параметров движения.

Нотация должна обладать полнотой (охватыванием всех аспектов движения) и универсальностью (способность кодировать все виды движений)

Нотация должна обладать анатомической и физиологической точностью (правильным представлением механики тела).

Должна существовать возможность количественной оценки создаваемых моделей движения.

Кроме того, нотация должна быть практичной и легко интегрируемой с современными технологиями. Также важно, чтобы система нотации могла документировать наблюдаемые движения, не полагаясь на контекстуальные предположения или субъективные интерпретации, чтобы обеспечить объективность и минимизировать предвзятость кодировщиков. Мы постарались максимально учитывать эти критерии при создании нашей системы нотации движений.

В отличие от других систем нотации движений, в частности системы CMA (Izquierdo, Anguera, 2018), которая наиболее близка к нашей, где нотация движения включает не только положение/состояние, но и комплексные модели движений (например, руки вместе), и где наблюдаются повторения сущностей (например, правая и левая стороны для положения и частей тела), а некоторые индикаторы описываются цифрами, наша система нотации моделирует

движение как последовательный переход от одного неподвижного состояния части тела, выполняющего движение, к другому неподвижному состоянию. Таким образом ключевое отличие в описании движения – вместо описания самого движения, требующее внедрения многочисленных переменных, мы задаем только начальное и конечное состояния части тела, выполняющего движение (можно включать и промежуточные состояния, об этом ниже), а также временные характеристики исполнения самого движения как перехода

от одного состояния части тела к другому его состоянию.

Каждое состояние имеет конкретные индикаторы, связанные с определенными сущностями (система движения, числовые индикаторы и т.д.), частями тела, и их природой и функциями. Сущности, природа и функции частей тела имеют различные типы, что позволяет представлять их в структурированной и символической форме. Латинский язык выбран как естественный язык описания. Фрагмент системы приведен в таблице 1.

Table 1. Fragment of the Perm Gesture Notation (PGN)

Таблица 1. Фрагмент Пермской жестовой нотации (ПЖН)

Сущности	Типы	Символ	Латинское название	
Движение	переходное	[]	transitorius	motus
	основное	[[]]	basicus	
Состояние		{ }	status	
Событие		#	eventus	
Система движения	последовательность	+	sequentia	motusratio
	одновременность	*	simultaneus	
	цикличность	<>	repetitio	
Численный показатель	временной отрезок (мс)		tempus intervallum	secundum numerum indicator
Природа и функции	Название	Символ	Латинское название	
Части тела	рука (в целом)	M	manus	pars corporis humani
	ладонь	P	palma	
	кисть	MO	manus	
	предплечье	AB	antibrachium	
Направление (траектория) движения	вверх	su	sursum	directio (trajectoria) motus
	вниз	de	deorsum	
	середина	med	medium	
	вправо	dx	dextra	
	влево	si	sinistra	
	по часовой стрелке	shm	secundum horologii motum	
Положение руки	вытянутый	ex	extensus	manus positio
	согнутый	an	ancus	
Форма движения	круг	c	circulus	forma motus
	полукруг	sc	semicirculus	

Нотация позволяет описывать как движение целиком, так и фрагменты движения. Пример формулы движения с использованием нашей нотации:

:1000:{*si Ma*} (1)

Описание движения с использованием конкретной системы нотации можно проиллюстрировать на примере поднятия руки, как показано на рисунке 1. Для точного описания состояния части тела используется следующая модель (демонстрируется для левой руки):

- М (manus) – обозначает руку/кисть
- Левый нижний индекс (si):
Указывает выбор руки, где “si” означает левую руку (аналогично, “dx” означает правую руку).
- Правый верхний индекс (ex):
Указывает положение руки, где “ex” обозначает вытянутую руку.
- Правый нижний индекс (de):
Обозначает направление, где “de” означает вниз.

Использование фигурных скобок {} вокруг нотации указывает на то, что эта часть движения является состоянием, а :1000: указывает на продолжительность этого состояния в миллисекундах. Таким образом формула (1) передает начальное состояние выполнения движения.

Это можно перевести в более привычную для машины форму, что приведет к:

:1000: {_si_M^ex_de} (2)

Некоторые положения и состояния руки предполагаются по умолчанию, в частности, состояние полностью вытянутой руки считается по умолчанию. Следовательно, такие состояния могут явно не отмечаться в нотации. Поэтому (2) упрощается до:

:1000: {_si_M_de} (3)

Используя эти правила нотации, мы можем описать движение, показанное на рисунке 1, следующим образом:

:1000:{si_M_de}:1000:[si_M_su]:1000:{RE} (4)

Из этой формулы очевидно, что она не описывает промежуточные состояния: движение левой руки переходит непосредственно из положения руки, вытянутой вниз, в положение руки, вытянутой вверх. Может показаться, что в формуле отдельно описывается движение (например, тут: 1000:[si_M_su]). Однако это не так: здесь фиксируется конечное состояние результата движения, а размещение его в части кода, обозначающего движение нужно для того, чтобы а) обозначить само движение, б) время его выполнения и в) возможности одновременного исполнения с другим действием (пример дан ниже). В то же время поскольку конечное состояние уже представлено в формуле, нет смысла дублировать его запись, поэтому формула завершается фрагментом 1000:{RE}, указывающим на то, что это состояние, длящееся 1000 мс, обозначено в блоке движения, т.е. является {si_M_su}.

В этом случае движение выполняется наиболее эффективным образом. Однако, если мы хотим выполнить это движение иначе, нам нужно определить **промежуточные состояния** для положения руки. Это приведет к тому, что рука будет двигаться по другой траектории. Например, поднятие руки через промежуточное положение, где рука вытянута в сторону, и затем пауза на 1 миллисекунду (1:{RE} обозначает состояние покоя на 1 миллисекунду), можно описать следующим образом:

:200:{si_M_de}:1000:[si_M_si]:1:{RE}:1000:[si_M_su]:200:{RE}

Рисунок 1. Поднятие руки в приложении Blender (три среза движения)
Figure 1. Raising the Arm in Blender (three motion slices)

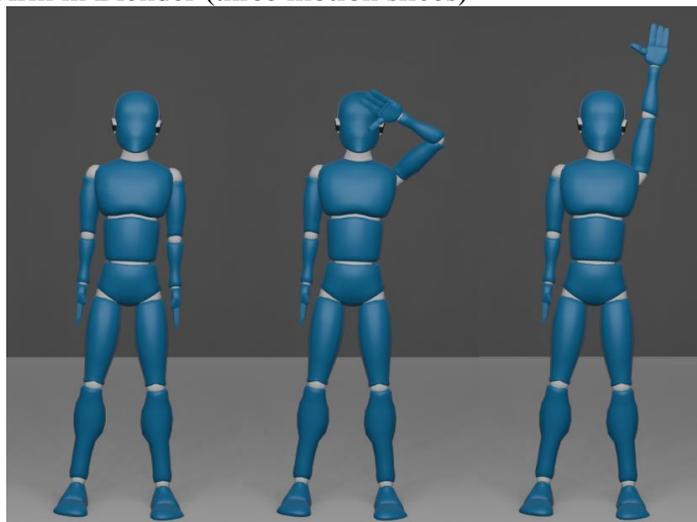
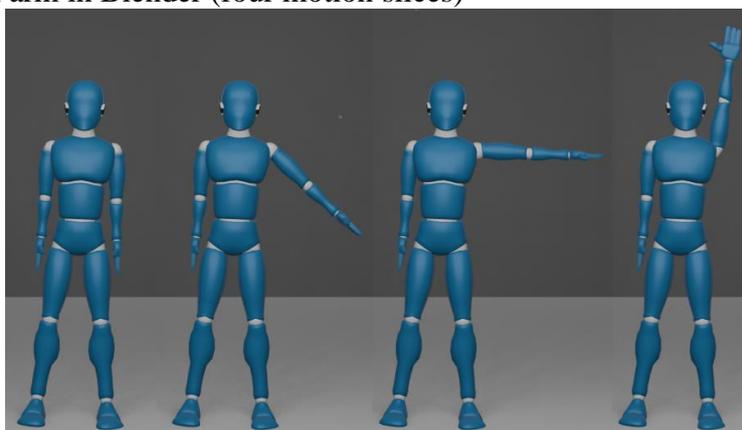


Рисунок 2. Поднятие руки в приложении Blender (четыре среза движения)
Figure 2. Raising an arm in Blender (four motion slices)



На рисунке 2 видно, что конечное положение руки совпадает с положением, показанным на рисунке 1: рука оказывается вытянутой с ладонью, обращенной вперед. Чтобы указать другое конечное положение, нам нужно использовать дополнительную нотацию для уточнения точной ориентации. Например:

```
:200:{si_M_de}:1000:[si_M_si]:1:{RE  
}:1000:[si_M_su*MO_dx_]:200:{RE}
```

Разберем конструкцию [si_M_su*MO_dx_]:

- Указывается конечное положение левой руки (вверх).
- Знак '*' обозначает одновременность, то есть другое действие происходит одновременно с положением руки вверх.

- "MO" представляет ладонь, а "dx" указывает, что ладонь повернута вправо. Для левой руки это означает, что ладонь повернута к телу.

Этот пример иллюстрирует, что направления и положения частей тела, описанные в таблице 1, могут быть применены к различным частям тела в зависимости от их анатомических возможностей. Например, для ладони "MO_at" (вперед) означает поворот руки так, чтобы ладонь была обращена вперед. Таким образом, положение руки в ее конечном состоянии, как показано на рисунках 1 и 2, может быть точно описано с использованием этой системы нотации.

Так как нотация предназначена для моделирования движений компьютеризированными системами, в них

есть предустановленные настройки (конфигурации), по которым осуществляется действие частей тела антропоморфной 3D-модели. Например, поднятие руки вверх производится через промежуточное состояние сгиба руки в локте, поэтому такое движение не требует специального маркирования. В то же время такой подход ставит проблему наиболее адекватных человеческим движениям настроек, которые могут быть получены с помощью статистического анализа множества моделируемых движений людей с разными психо-биологическими характеристиками (например, возраст, пол, темперамент и др.), что является перспективой исследования.

До сих пор описание движений проводилось на примере одной части тела; но как нотация работает с одновременным движением нескольких частей тела, например, обеих рук?

Для моделирования симультанно-сукцессивных движений (одновременного исполнения действий несколькими частями тела в их временной динамике) вводятся следующие понятия:

1. Независимость (автономность) частей тела 3D-модели, ограниченная только анатомическими и физиологическими возможностями человека, зафиксированными в настройках 3D-модели.

2. «Дорожки записи», с помощью которых осуществляется независимое описание движения каждой части тела (для каждой части тела вводится своя «дорожка»).

3. Событие движения, под которым понимаем а) любое изменение состояния/движения части тела, вызванное внешними причинами (другой частью тела или предметом – например, столкновением руки с другой частью тела или предметом) или б) комплексное движение частей тела, при котором задаются обязательные состояния этих частей относительно друг друга во времени. Отдельно следует отметить, что, несмотря на автономность движения части тела, его движение может в силу связанности и инерционности

оказывать влияние на другую часть тела (например, поворот руки связан с поворотом корпуса). Такие инерционные изменения естественны и не являются событиями.

Событие привязано ко времени, но происходит в пределах некоторого выделенного промежутка Δt . Параметры Δt можно использовать для расчета и оценки точности моделируемых движений. Событие обозначаем символом “#” в верхнем левом индексе части тела, например, описание события (#1) контакта двух рук перед грудью: для левой руки $[_{si}^{\#1}M_{dx}^{*at}]$ и для правой $[_{dx}^{\#1}M_{si}^{*at}]$. События имеют свои индексы, что позволяет отличать одно событие от другого и сопоставлять фрагменты формулы, описывающие разные части тела.

Отдельно отметим, что комплексные движения (как например, «руки вместе» в СМА), в нашем случае могут быть определены отдельно и храниться в виде вариантов формулы. Иными словами, те ли иные паттерны состояний или движений могут быть реализованы в виде единиц некоего словаря со своими возможными формальными представлениями.

Реализация системы управления движением

Разработка системы управления движением, основанной на лингвистической нотации проходит несколько этапов: нотацию жестов, построение внутреннего представления движения и его перевод в исполняемые движения, демонстрируемые с использованием программ для 3D-моделирования и анимации.

Первый этап – разбор нотации (парсинг формулы). Парсинг включает разбор строки нотации, выделение термов и создание абстрактного синтаксического дерева (AST). Покажем процесс парсинга на примере разбора формулы: $:400:\{_{dx}M_{2la-de-2at}\}:500:[_{dx}M_{la-su-2at}]:400:\{RE\}$

1. Выделение отдельных термов:

1. Первый терм: $:400:\{_{dx}M_{2la-de-2at}\}$

- Время: 400 миллисекунд.
- Тип: состояние ({}).
- Детали положения:
 - dx_M: правая рука.
 - 2la: движение влево с коэффициентом 2.
 - de: движение вниз.
 - 2at: движение вперед с коэффициентом 2.
- Объект термина: Создается State Term с атрибутами:
 - time = 400.
 - node = dx_M.
 - details = {'2la', 'de', '2at'}.
- 2. Второй терм: :500:[_dx_M_la-su-2at_]
 - Время: 500 миллисекунд.
 - Тип: движение ([]).
 - Детали направления (конечное состояние):
 - dx_M: правая рука.
 - la: движение влево.
 - su: движение вверх.
 - 2at: движение вперед с коэффициентом 2.
 - Объект термина: Создается Move Term:
 - time = 500.
 - node = dx_M.
 - details = {'la', 'su', '2at'}.
- 3. Третий терм: :400:{RE}
 - Время: 400 миллисекунд.
 - Тип: состояние ({}).
 - Детали: RE (фиксация состояния).
 - Объект термина: Создается State Term:
 - time = 400.
 - details = {'RE'}.

Второй этап – построение внутреннего представления движения. На основе AST строится внутреннее представление движения, состоящее из серии кадров. Эти кадры содержат информацию о том, какая кость (часть тела) исполнителя относится к ним, как они влияют на его положение и конкретный момент времени с начала движения, когда этот кадр актуален. В этой серии записываются только ключевые кадры – крайние положения примитивных движений. Например, если исполнителю

нужно обвести букву “L” рукой, в серии будут включены ключевые кадры для положений руки внизу, вверху и вверху справа. Интерполяция движения является обязанностью исполнителя.

Третий этап – преобразование внутреннего представления в конечное движение. Внутреннее представление движения передается исполнителю, который преобразует его в конечный результат, видимый пользователю. Исполнителем может быть робот или любая программа, работающая со скелетной анимацией, включая игровые движки. В этом исследовании в качестве исполнителя используется Blender 3D, программа для 3D-моделирования и анимации. Blender 3D позволяет создавать пользовательские дополнения с использованием Python и предоставляет обширный набор готовых графических интерфейсов и программных утилит для работы с 3D-графикой.

Результаты и обсуждение

В данной статье в качестве результатов мы рассматриваем демонстрацию возможностей нотации описывать естественные движения человека, взятые из фрагментов видеозаписей мультимодального корпуса описания людьми различных эмоциональных состояний. В данном случае реализуется основное назначение разрабатываемой жестовой нотации – служить целям создания бесшовных инструментов перехода от одной формы человеческой активности к другой в различных интерфейсно-опосредованных коммуникациях с помощью универсальной и точной системы для кодирования и описания движений человека, включая жесты, позы и сложные взаимодействия частей тела.

На данном этапе мы не ставим задачи полного и точного темпо-ритмического и пространственного воспроизведения машиной движений человека в силу имеющихся на сегодняшний день технических проблем. Поэтому сразу обозначим рамки формального описания:

нотация будет осуществляться только для движений рук.

Комплекс движений описывается как совокупность состояний для каждой руки, представленных в двух записях (т.н. дорожках). В формальной записи переход от одной дорожки к другой отделяется запятой (данный знак используется только в этой функции).

Производимая формальная запись движения, в конечном счете, сформировало и наше понимание жеста, который с точки зрения семантики коммуникативно направлен и семантически целостен, а с точки зрения формы представляет собой движение от одного состояния до другого (включая промежуточные состояния прохождения движения, если они были).

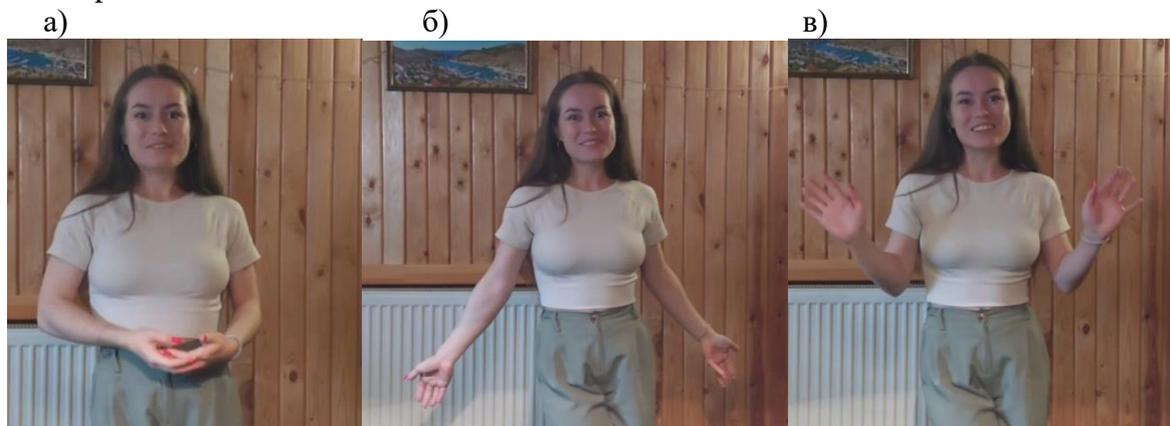
Рассмотрим фрагменты движений из специализированного мультимодального корпуса эмоций и жестов. Эмоции в данном корпусе условно можно поделить на положительные и отрицательные. Количественные и качественные

характеристики движений, сопровождающих эти эмоции, могут варьироваться. В качестве примера рассмотрим два комплекса движений: один, используемый в контексте положительной эмоции «Восторг» (рисунок 3), и другой, сопровождающий отрицательную эмоцию «Злость» (рисунок 4), а также соответствующие языковые контексты их применения (см. пример 1 и пример 2). В качестве иллюстрации будут использоваться скриншоты состояний движения рук информантов, которые служат для модельного представления в виде формул движения.

Каждый из представленных комплексов движений, выбранных нами для иллюстрации, отличается рядом параметров. В отношении первого комплекса, который встречается в контексте эмоции «Восторг», следует отметить его упрощенную структуру. Кроме того, данное движение является примером синхронного движения обеих рук.

Рисунок 3. Иллюстрация комплекса движений в рамках эмоции «Восторг» и его составляющие.

Figure 3. Illustration of the complex of movements within the framework of the emotion “Delight” and its components.



Данный комплекс движений наблюдается в следующем языковом контексте:

Пример 1. ...И вот (**Согнутые руки сложены одна ладонь на тыльную сторону ладони другой руки**) полет на парашуте (**Разведение рук в разные стороны вниз**), я была в восторге

(**Поднятые по разным сторонам согнутые руки**)...

При описании эмоции «Восторг» информант использует этот комплекс движений (Рисунок 3), который можно разделить на три основных состояния: а) руки согнуты и сложены ладонь на тыльную сторону ладони; б) разведение

рук в стороны вниз; в) поднятые по разные стороны согнутые руки.

При детализированном описании этих состояний можно выделить следующие характеристики: начальное состояние (а) представляет собой исходное положение, в котором руки согнуты в локтях под углом 90 градусов, левое предплечье направлено вперед и вправо, правое – вперед и влево, ладони сложены одна на другую на уровне живота (ладони вверх). Данный компонент можно представить в виде термов ($\{_{dx_B_de_+dx_AB_ad-at-si_}\}$ и $\{_{si_B_de_+si_AB_ad-at-dx_}\}$). В состоянии (б) правая рука вытянута вниз и вправо, левая – вниз и влево ($[_{dx_AB_de-la_}]$ и $[_{si_AB_de-la_}]$). Наконец, последнее состояние (в)), являющееся завершающим этапом комплекса, характеризуется тем, что правая рука согнута, с предплечьем, направленным вверх и вправо, и левая рука также согнута, с предплечьем, направленным вверх и влево ($[_{dx_AB_su-la_}]$ и $[_{si_AB_su-la_}]$). За каждой частью компонента, обозначающего движение,

следует фиксация с конкретным измерением времени, например, (:190:{RE})

Этот комплекс движений может быть закодирован четко в виде следующей формулы:

:415:{ $_{dx_B_de_+dx_AB_ad-at-si_}$ }:375:[$_{dx_AB_de-la_}$]:765:{RE}:690:[$_{dx_AB_su-la_}$]:190{RE},:415:{ $_{si_B_de_+si_AB_ad-at-dx_}$ }:375:[$_{si_AB_de-la_}$]:765:{RE}:690:[$_{si_AB_su-la_}$]:190:{RE}

Структура данной формулы состоит из шести термов, по три для каждой руки. Эти термы распределены по двум дорожкам, по одной для правой и левой руки соответственно. При сравнении термов и их последовательности в обеих дорожках можно отметить их полную идентичность, что указывает на синхронность движений обеих рук. Кроме того, каждый терм можно соотнести с определенным языковым контекстом, что наглядно представлено в таблице 2.

Таблица 2. Соотношение термов и контекстов эмоции «Восторг»

Table 2. The ratio of terms and contexts of the emotion “Delight”

Термы	Дорожка 1	:415:{ $_{dx_B_de_+dx_AB_ad-at-si_}$ }	:375:[$_{dx_AB_de-la_}$]:765:{RE}	:690:[$_{dx_AB_su-la_}$]:190:{RE}
	Дорожка 2	:415:{ $_{si_B_de_+si_AB_ad-at-dx_}$ }	:375:[$_{si_AB_de-la_}$]:765:{RE}	:690:[$_{si_AB_su-la_}$]:190:{RE}
Языковой контекст		И вот	полет на парашюте	я была в восторге

Далее излагается иллюстрация комплекса движений, сопровождающих отрицательную эмоцию «Злость».

Что касается комплекс движений в рамках эмоции «Злость», то он сопровождается следующей языковой контекст:

Пример 2. (Обе руки опущены вниз) Злость – это когда постоянно хочется прямо подраться (Руки согнуты, предплечья правой и левой рук поднимаются вверх исходятся – правое влево, левое вправо) или ты хочешь

(Правая рука согнута, предплечье вверх, левая рука вытянута вниз и немного в бок) стукнуть по столу (Согнутые руки сложены одна ладошка в кулак на другую), сказать (Правая рука согнута, предплечье правой руки влево и немного вверх, левая рука вытягивается вперед, вправо, опускается вниз с пятерной амплитудой), что нет (Руки вытянуты вниз и расходятся – правая влево, левая вправо и пересекаются), ты не прав...

Рисунок 4. Иллюстрация комплекса движений в рамках эмоции «Злость» и его составляющие.

Figure 4. Illustration of the complex of movements within the framework of the emotion “Anger” and its components.



Интерпретация движений в данной последовательности в отличие от предыдущей, – сложней, поскольку включает несинхронизированные и дополнительные действия для правой (dx) и левой (si) руки, структурированные по термам, в двух основных дорожках.

Для правой руки движение начинается с исходного положения, в котором рука опущена вниз ($[_dx_M_de_]$). Далее следуют несколько компонентов. Второй компонент (б)) ($[_dx_AB_at-ad-si_]$) проявляется, когда предплечье правой руки вытягивается вперед, сгибается к плечу и движется влево, после чего происходит фиксация конечного положения ($\{RE\}$). Далее следует третий компонент (в)) –

предплечье поднимается вверх и движется влево ($[_dx_AB_su-la_]$), за чем также следует фиксация ($\{RE\}$).

Далее происходит первое событие (г)) ($[_dx^{\#1}AB_at-ad-2si_]$), когда предплечье вытягивается вперед, сгибается к плечу и движется влево с удвоенной амплитудой. После фиксации положения ($:200:\{RE\}$) сталкиваемся со следующим компонентом (д)) ($[_dx_AB_at-ad-2si-su_]$), где предплечье снова вытягивается вперед, сгибается к плечу, движется влево с удвоенной амплитудой и поднимается вверх. Этап завершается фиксацией ($\{RE\}$). Последнее событие (е)) включает опускание руки вниз и движение влево с утроенной амплитудой ($[_dx^{\#2}M_de-3si_]$), с последующей фиксацией ($\{RE\}$).

Левая рука повторяет схожую последовательность движений. В исходном положении (а) она также опущена вниз ($\{_{si_M_de_}\}$). Следующий компонент (б)) включает вытягивание предплечья вперед, сгибание к плечу, движение вправо и опускание вниз на ($[_{si_AB_at-ad-dx-de_}]$), после чего фиксируется положение ($\{RE\}$). Далее (в)) предплечье опускается вниз с удвоенной амплитудой и движется влево ($[_{si_M_2de-la_}]$), с фиксацией ($\{RE\}$).

Первое событие (г)), в течение которого предплечье вытягивается вперед, сгибается к плечу и движется вправо, можно отметить как ($[_{si^{\#}1^{\wedge}AB_at-ad-dx_}]$), с последующей фиксацией ($\{RE\}$). Затем следует компонент движения (д)), где предплечье вытягивается вперед, движется вправо и опускается вниз с пятерной амплитудой ($[_{si_AB_at-dx-5de_}]$), после чего положение фиксируется ($\{RE\}$). Последнее событие (е)) включает опускание руки вниз и движение вправо с утроенной амплитудой ($[_{si^{\#}2^{\wedge}M_de-3dx_}]$), завершающееся фиксацией ($\{RE\}$).

Общее описание данного комплекса движений с учётом их временного диапазона можно кодировать следующим образом:

:240: $\{_{dx_M_de_}\}$:1535: $[_{dx_AB_at-ad-si_}]$:3365: $\{RE\}$:335: $[_{dx_AB_su-}$

$la_]$:175: $\{RE\}$:405: $[_{dx^{\#}1^{\wedge}AB_at-ad-2si_}]$:685: $\{RE\}$:240: $[_{dx_AB_at-ad-2si-su_}]$:130: $\{RE\}$:100: $[_{dx^{\#}2^{\wedge}M_de-3si_}]$:1170: $\{RE\}$,:240: $\{_{si_M_de_}\}$:1535: $[_{si_AB_at-ad-dx-de_}]$:3365: $\{RE\}$:335: $[_{si_M_2de-la_}]$:175: $\{RE\}$:405: $[_{si^{\#}1^{\wedge}AB_at-ad-dx_}]$:685: $\{RE\}$:240: $[_{si_AB_at-dx-5de_}]$:130: $\{RE\}$:100: $[_{si^{\#}2^{\wedge}M_de-3dx_}]$:1170: $\{RE\}$

Исходя из представленной формулы, можно отметить, что рассматриваемая последовательность представляет собой сложный комплекс движений рук, который включает в себя не только элементарные манипуляции, но и более сложные формы взаимодействия между руками, что свидетельствует о высоком уровне сложности координации и синхронизации движений в данной последовательности. Таким образом, можно подчеркнуть, что предлагаемая жестовая нотация представляет собой эффективный инструмент для систематизации и отображения не только примитивных, но и таких сложных движений. Данная нотация позволяет адекватно и детально фиксировать любые жестовые взаимодействия, что обеспечивает точное представление их структуры и последовательности.

Таблица 3. Соотношение термов и контекстов эмоции «Злость»

Table 3. The ratio of terms and contexts of the emotion “Anger”

	№ термина	1	2	3
Дорожки	Дорожка 1	:240: $\{_{dx_M_de_}\}$:1535: $[_{dx_AB_at-ad-si_}]$:3365: $\{RE\}$:335: $[_{dx_AB_su-la_}]$:175: $\{RE\}$
	Дорожка 2	:240: $\{_{si_M_de_}\}$:1535: $[_{si_AB_at-ad-dx-de_}]$:3365: $\{RE\}$:335: $[_{si_M_2de-la_}]$:175: $\{RE\}$
Языковой контекст			Злость – это когда постоянно хочется прямо подраться	или ты хочешь
	№ термина	4	5	6
Дорожки	Дорожка 1	:405: $[_{dx^{\#}1^{\wedge}AB_at-ad-2si_}]$:685: $\{RE\}$:240: $[_{dx_AB_at-ad-2si-su_}]$:130: $\{RE\}$:100: $[_{dx^{\#}2^{\wedge}M_de-3si_}]$:1170: $\{RE\}$
	Дорожка 2	:405: $[_{si^{\#}1^{\wedge}AB_at-ad-dx_}]$:685: $\{RE\}$:240: $[_{si_AB_at-dx-5de_}]$:130: $\{RE\}$:100: $[_{si^{\#}2^{\wedge}M_de-3dx_}]$:1170: $\{RE\}$
Языковой контекст		стукнуть по столу	сказать	что нет ты не прав

Представленные в таблицах 2 и 3 нотации фрагментов речежестового поведения можно обогатить дополнительными параметрами, первая часть из которых представляет собой динамически изменчивые параметры поведения, основанные, во-первых, на уже обозначенных параметрах (последовательности термов формулы, речевой материал, время протекания жестов), во-вторых, включающие метрики, выводимые из этих параметров, и в-третьих, факультативные физиологические характеристики, которые можно измерять в экспериментальной ситуации (пульс, объем дыхания, сатурация и др.), а вторая часть состоит из факультативных постоянных (для протекающей коммуникации) характеристик, описывающих социально-демографические, психологические и физиологические параметры человека (пол, возраст, образование, рост, размах рук и др.).

Построение иерархии метрик для анализа жестов и их ограничений опирается на обширные исследовательские материалы. В ряде работ последних лет (Liu et al., 2022; Qi et al., 2024; Tonoli et al., 2024; Yang et al., 2023; Yoon et al., 2020; Zhi et al., 2023) обсуждаются такие метрики, как Mean Squared Error (MSE), Mean Average Error (MAE), Fréchet Gesture Distance (FGD), Gesture Area Coverage (GAC) и др., которые сравнивают движения с эталоном, оценивают естественность и схожесть движений, анализируют динамические характеристики (скорость, ускорение и др.) и учитывают пространственное покрытие жестов.

Полагаем, что некоторые метрики могут быть использованы и в нашем случае для оценки речежестового поведения. При этом мы исходим из того, что поскольку интегральной поведенческой единицей в нашем случае является терм (как конечное состояние движения, которое в нотации может быть представлено переходом от предыдущего состояния к последующему посредством движения, т.е. как $\{termK\} \rightarrow [termK+1]$), то все измерения должны быть привязаны к термам. Из такого подхода следуют некоторые метрики:

1. L – длина траектории движения руки. Здесь и в дальнейшем вычисления

можно проводить как для фрагмента ($termK \rightarrow termK+1$), так и для всей цепочки движения одной руки или двух рук: $term1 \rightarrow term2 \rightarrow \dots \rightarrow termN$.

2. S – площадь под кривой траектории движения руки или рук в рамках одного жеста, т.е. выделенной поведенческой единицы $termK \rightarrow termK+1$ и всего движения.

3. V – объем захваченного жестом ($termK \rightarrow termK+1$) пространства. Здесь могут быть разные методики измерения: от пространства жеста одной руки или двух рук до пространства, рассматриваемого в жесте между двумя руками.

Параметры L , S , V при описании реальной коммуникации (без использования 3D-моделей) нуждаются в нормировании; мы предлагаем ввести нормирующий параметр $= \frac{L}{I1} \cdot \frac{S}{I2} \cdot \frac{V}{I3}$, где $I1$ – рост, $I2$ – размах рук для L , N^2 для S и N^3 для V .

4. $speech$ – объем речевого материала, привязанного ко времени произведения жеста, т.е. к $termK \rightarrow termK+1$.

5. $temp$ – метрика, характеризующая динамику изменения захваченного жестами пространства: $temp = \frac{L}{I1} \cdot \frac{S}{I2} \cdot \frac{V}{I3} \cdot \frac{speech}{I4} \cdot \frac{temp}{I5}$. Метрика $temp$ может вычисляться иначе, например, как отношение количества термов или количества речевого материала к единице времени.

В целом, запись аргументов функции Behavior можно представить следующим образом:

Behavior = $\{(term_1, speech_1, t_1, L_1, S_1, V_1, \dots), \dots, (term_N, speech_N, t_N, L_N, S_N, V_N, \dots); (pulse_1, tidal\ volume_1, \dots)^*, \dots, (pulse_N, tidal\ volume_N, \dots)^*; (gender, age, N, temp...)^*\}$, где знаком $*$ отмечены группы факультативных параметров.

Заключение

Создание формализованной системы кодирования движений человека для интерфейсно-опосредованной коммуникации в различных машинно-генерируемых средах, как мы попытались показать, является сложной, но решаемой научной проблемой. В ходе проведенного исследования была представлена система Пермской жестовой нотации и выделена единица описания поведения человека – терм, представляющий запись перехода от предшествующего фиксированного

состояния части тела в рамках его движения к последующему фиксированному состоянию. Терм как единица дискретизации движения выравнивает относительно себя и другие формы активности (т.е. выступает как единица когнитивно-поведенческая), в качестве которых можно рассматривать речь (мы говорили об объеме речевого материала, хотя могут быть и иные измеряемые параметры, например, акустические), а также мимику или данные носимых устройств, передающих сведения о пульсе, объеме дыхания и др.

Исходя из того, что вероятно, каждое движение человека имеет семантику, т.е. не является случайным и/или хаотическим в последовательности актов человеческого поведения, мы должны в пределе требовать от машины такой же семантизации двигательной активности и синхронизации ее с речевой семантикой сообщения. Полное и точное темпо-ритмическое и пространственное воспроизведение машиной движений человека (т.е. по сути копирование хронотопа человека его виртуальным аватаром) возможно с помощью обращения к комбинированию методов и технологий компьютерного зрения с современными нейросетевыми моделями-трансформерами. Конечной целью такого подхода является генерация формулы движения в разработанной формально-языковой нотации на основе анализа поступающего на вход модели видеоряда движений человека. Сгенерированная формула может затем передаваться исполнителю (виртуальной 3D-модели или физическому роботизируемому устройству). Воспроизведение закодированных в формуле движений исполнителем нужно в том числе и для анализа точности и реалистичности сгенерированных анимаций путём сравнения с реальными движениями с последующими корректировкой формулы и дообучения модели.

Очевидно, что качество воспроизведения модели речежестового поведения человека машиной в конечном счете должно оцениваться самим человеком. Важно, чтобы машины не просто повторяли физические движения, но и адекватно передавали семантические и эмоциональные

значения этих движений, соответствующие человеческим ожиданиям и культурным нормам. В этом контексте появляется запрос к созданию отдельной ветви общей семантики – техносемантики как семантики, генерируемой антропоморфными машинами в процессе человеко-машинного взаимодействия и адресованной коммуницирующему с машиной человеку (о части проблем техносемантики см. Belousov et al., 2024).

Синтез движений, который ориентирован не только на механическое воспроизведение, но и в перспективе – на смысловое содержание, является важной составляющей в создании более эффективных и человекоцентричных систем коммуникации в различных машинно-генерируемых средах. При этом важно учитывать не только визуальные и кинематические параметры, но и скрытые смыслы, которые несут жесты и движения в контексте общения. Эта задача становится все более актуальной в условиях развития антропоморфных машин и больших языковых моделей, которые стремятся не просто подражать человеку, но и коммуницировать с ним на уровне, близком к человеческому.

В заключение отметим, что дальнейшее развитие систем кодирования движений, в том числе и представленной в данной работе, и применение их в различных областях, таких как киноиндустрия, компьютерные игры, образование, медицина, спорт и робототехника, могут стать основой для новых форм коммуникации, в которых машины не просто выполняют команды, но и активно участвуют в семантически значимых взаимодействиях, поддерживая полноценный и интуитивный диалог с человеком.

Список литературы

- Abbie M. Movement notation // The Australian journal of physiotherapy. 1974. Vol. 20 (2). Pp. 61–69. [https://doi.org/10.1016/S0004-9514\(14\)61177-6](https://doi.org/10.1016/S0004-9514(14)61177-6)
- Bashan M., Einbinder H., Harries J., Shosani M., Shoval D. Movement Notation: Eshkol and Abraham Wachmann. Köln: Verlag der Buchhandlung Walther König, 2024.

- Belousov K. I., Sazina D. A., Ryabinin K. V., Brokhin, L. Yu. Sensory Technolinguistics: On Mechanisms of Transmitting Multimodal Messages in Perceptual-Cognitive Interfaces // *Automatic Documentation and Mathematical Linguistics*. 2024. Vol. 58, Issue 2. Pp. 108–116. <https://doi.org/10.3103/s0005105524700079>
- Benesh R., Benesh J. *An Introduction to Benesh Dance Notation*. London: A & C Black, 1956.
- Bernardet U., Fdili Alaoui S., Studd K., Bradley K., Pasquier P., Schiphorst, T. Assessing the reliability of the Laban Movement Analysis system // *PLoS ONE*. 2019. Vol. 14 (6): e0218179. <https://doi.org/10.1371/journal.pone.0218179>
- Birdwhistell R. L. *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture*. Washington, DC.: Foreign Service Institute, 1952.
- Bull P., Doody J. P. *8 Gesture and body movement*. De Gruyter eBooks, 2013. Pp. 205–228. <https://doi.org/10.1515/9783110238150.205>
- Calvert T. *Approaches to the Representation of Human Movement: Notation, Animation and Motion Capture* // *Dance Notations and Robot Motion*, Springer Tracts in Advanced Robotics. 2015. Pp. 49–68. https://doi.org/10.1007/978-3-319-25739-6_3
- Dael N., Mortillaro M., Scherer K. R. The Body Action and Posture Coding System (BAP): Development and Reliability // *Journal of Nonverbal Behavior*. 2012. Vol. 36 (2). Pp. 97–121. <https://doi.org/10.1007/s10919-012-0130-0>
- Dell C. *A Primer for Movement Description: Using Effort-shape and Supplementary Concepts*. New York: Dance Notation Bureau Press, 1977.
- Duprey S., Naaim A., Moissenet F., Begon M., Chèze, L. Kinematic models of the upper limb joints for multibody kinematics optimisation: An overview // *Journal of Biomechanics*. 2017. Vol. 62. Pp. 87–94. DOI: 10.1016/j.jbiomech.2016.12.005
- Ekman P., Friesen W. V. *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists, 1978.
- El Raheb K., Ioannidis Y. From dance notation to conceptual models: a multilayer approach // *Proceedings of the 2014 International Workshop on Movement and Computing, MOCO, ACM*. New York, 2014. Pp. 25–30.
- El Raheb K., Buccoli M., Zanoni M., Katifori A., Kasomoulis A., Sarti A., Ioannidis Y. Towards a general framework for the annotation of dance motion sequences // *Multimed Tools Appl*. 2023. Vol. 82. Issue 3. Pp. 3363–3395. <https://doi.org/10.1007/s11042-022-12602-y>
- Eshkol N., Wachmann A. *Movement Notation*. London: Weidenfeld and Nicolson, 1958.
- Farnell B. M. *Movement Notation Systems // The World's Writing Systems* / ed. Peter T. Daniels. Oxford University Press, 1996. Pp. 855–879.
- Frishberg N. Writing systems and problems for sign language notation // *Journal for the Anthropological Study of Human Movement*. 1983. Vol. 2 (4). Pp. 169–195.
- Frey S., Hirsbrunner H-P., Jorns U. *Time-Series Notation: A Coding Principle for the Unified Assessment of Speech and Movement in Communication Research*. Tübingen: Gunter NarrVerlag, 1982.
- Grushkin D. A. *Writing Signed Languages: What For? What Form?* // *American Annals of the Deaf*. 2017. Vol. 161 (5). Pp. 509–527. <https://doi.org/10.1353/aad.2017.0001>
- Guest A. H. *Dance Notation: The Process of Recording Movement on Paper*. New York: Dance Horizons, 1984.
- Guest A. H. *Labanotation: The System of Analyzing and Recording Movement* (4th ed.). New York: Routledge, 2005. <https://doi.org/10.4324/9780203823866>
- Harrigan J. A. *Proxemics, Kinesics, and Gaze // The New Handbook of Methods in Nonverbal Behavior Research*. 2008. Pp. 136–198. <https://doi.org/10.1093/acprof:oso/9780198529620.003.0004>
- Izquierdo C., Anguera M. T. Movement notation revisited: syntax of the common morphokinetic alphabet (CMA) system // *Front. Psychol*. 2018. Vol. 9:1416. <https://doi.org/10.3389/fpsyg.2018.01416>
- Karg M., Samadani A.-A., Gorbet R., Kuhnlenz K., Hoey J., Kulic D. Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation // *IEEE Transactions on Affective Computing*. 2013. Vol. 4. Issue 4. Pp. 341–359. <https://doi.org/10.1109/t-affc.2013.29>
- Kendon A. *Gesture // Annual Review of Anthropology*. 1997. Vol. 26 (1). Pp. 109–128. <https://doi.org/10.1146/annurev.anthro.26.1.109>
- Key M. R. *Nonverbal communication: a research guide and bibliography*. Metuchen, N.J.: The Scarecrow Press, 1977.
- Kilpatrick C. E. *Movement, Gesture, and Singing: A Review of Literature // Update: Applications of Research in Music Education*. 2020. Vol. 38. Issue 3. Pp. 29–37. DOI: 10.1177/8755123320908612

Laban R. von, Lawrence F. C. Effort: Economy of Human Movement / 2nd ed. London: Macdonald & Evans, 1974.

Laumond J., Abe N. Dance Notations and Robot Motion. Cham (ZG): Springer International Publishing AG, 2016. <https://doi.org/10.1007/978-3-319-25739-6>

Liu H., Zhu Z., Iwamoto N., Peng Y., Li Zh., Zhou Y., Bozkurt E., Zheng, B. BEAT: A Large-Scale Semantic and Emotional Multi-modal Dataset for Conversational Gestures Synthesis // Computer Vision – ECCV 2022. 2022. Pp. 612–630. <https://doi.org/10.48550/arXiv.2203.05297>

Murillo E., Montero I., Casla M. On the multimodal path to language: The relationship between rhythmic movements and deictic gestures at the end of the first year // Frontiers in Psychology. 2021. Vol. 12. Pp. 1–8. <https://doi.org/10.3389/fpsyg.2021.616812>

Novack A. M., Wakefield E. M. Goldin-Meadow S. What makes a movement a gesture? // Cognition. 2016. Vol. 146. Pp. 339–348. <https://doi.org/10.1016/j.cognition.2015.10.014>

Qi X., Liu C., Li L., Hou J., Xin H., Yu, X. Emotion Gesture: Audio-Driven Diverse Emotional Co-Speech 3D Gesture Generation, IEEE Transactions on Multimedia. 2024. Pp. 1–11. <https://doi.org/10.1109/TMM.2024.3407692>

Streeck J. The Significance of Gesture: How it is Established // Papers in Pragmatics. 2010. Vol. 2. Issue 1-2. <https://doi.org/2.10.1075/iprapip.2.1-2.03str>

Shafir T., Tsachor R., Welch K. B. Emotion Regulation through Movement: Unique Sets of Movement Characteristics are Associated with and Enhance Basic Emotions // Frontiers in Psychology. 2016. Vol. 6. <https://doi.org/10.3389/fpsyg.2015.02030>

Stults-Kolehmainen M. A. Humans have a basic physical and psychological need to move the body: Physical activity as a primary drive // Frontiers in Psychology. 2023. Vol. 14. <https://doi.org/10.3389/fpsyg.2023.1134049>

Tonoli R. L., Costa P. D. P., Marques L. B. d. M. M., Ueda L. H. Gesture Area Coverage to Assess Gesture Expressiveness and Human-Likeness // International Conference on Multimodal Interaction (ICMI Companion '24), 4–8 November 2024, San Jose, Costa Rica. ACM, New York, NY, USA. <https://doi.org/10.1145/3686215.3688822>

Trujillo J. P., Vaitonyte J., Simanova I., Özyürek A. Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research // Behavior Research Methods. 2018. Vol. 51 (2). Pp. 769–777. <https://doi.org/10.3758/s13428-018-1086-8>

Van Elk M., van Schie H. T., Bekkering H. Short-term action intentions overrule long-term semantic knowledge // Cognition. 2009. Vol. 111. Issue 1. Pp. 72–83.

<https://doi.org/10.1016/j.cognition.2008.12.002>

Yang S., Wu Z., Li M., Zhang Z., Hao L., Bao W., Zhuang H. QPGesture: Quantization-Based and Phase-Guided Motion Matching for Natural Speech-Driven Gesture Generation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. Pp. 2321–2330. <https://doi.org/10.48550/arXiv.2305.11094>

Yoon Y., Cha B., Lee J.-H., Jang M., Lee J., Kim J., Lee G. Speech Gesture Generation from the Trimodal Context of Text, Audio, and Speaker Identity // ACM Transactions on Graphics. 2020. Vol. 39. Issue 6.

<https://doi.org/10.1145/3414685.3417838>

Zhi Y., Cun X., Chen X., Shen X., Guo W., Huang S., Gao S. LivelySpeaker: Towards Semantic-Aware Co-Speech Gesture Generation // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023. Pp. 20807–20817.

<https://doi.org/10.1109/ICCV51070.2023.01902>

Авторы прочитали и одобрили окончательный вариант рукописи.

The authors have read and approved the final manuscript.

Конфликты интересов: у авторов нет конфликтов интересов для декларации.

Conflicts of interests: the authors have no conflicts of interest to declare.

Белоусов Константин Игоревич, доктор филологических наук, профессор кафедры теоретического и прикладного языкознания, Пермский государственный национальный исследовательский университет, Пермь, Россия.

Konstantin I. Belousov, Doctor of Philology, Professor, Department of Theoretical and Applied Linguistics, Perm State University, Perm, Russia.

Талески Александр, кандидат филологических наук, доцент кафедры теоретического и прикладного языкознания, Пермский государственный национальный исследовательский университет, Пермь, Россия.

Aleksandar Taleski, Candidate of Philological Sciences, Associate Professor, Department of Theoretical and Applied Linguistics, Perm State University, Perm, Russia.

Агаев Артем Русланович, разработчик программного обеспечения, ООО «Хьюмен Семантикс», Пермь, Россия.

Artem R. Agaev, software developer, Human Semantics LLC, Perm, Russia.