# Exploring User-Defined Gestures for Lingual and Palatal Interaction

Santiago Villarreal-Narvaez[1,2], Jorge Luis Perez-Medina[2†] and Jean Vanderdonckt[2†]

[1]Namur Digital Institute, University of Namur, 61 Rue de Bruxelles, Namur, 5000, Belgium.
[2]Louvain Research Institute in Management and Organizations, Université catholique de Louvain, 1 Place des Doyens, Louvain-la-Neuve, 1348, Belgium.

Contributing authors: santiago.villarreal@unamur.be; santiago.villarreal@uclouvain.be; jorge.perezmedina@uclouvain.be; jean.vanderdonckt@uclouvain.be;
[†]These authors contributed equally to this work.

**Abstract**

Individuals with motor disabilities can benefit from an alternative means of interacting with the world: using their tongue. The tongue possesses precise movement capabilities within the mouth, allowing individuals to designate targets on the palate. This form of interaction, known as lingual interaction, enables users to perform basic functions by utilizing their tongues to indicate positions. The purpose of this work is to identify the lingual and palatal gestures proposed by end-users. In order to achieve this goal, our initial step was to examine relevant literature on the subject, including clinical studies on the motor capacity of the tongue, devices detecting the movement of the tongue, and current lingual interfaces (e.g., using a wheelchair). Then, we conducted a Gesture Elicitation Study (GES) involving twenty-four ($\boldsymbol{N = 24}$) participants, who proposed lingual and palatal gestures to perform nineteen (**19**) Internet of Things (IoT) referents, thus obtaining a corpus of **456** gestures. These gestures were clustered into similarity classes (**80** unique gestures) and analyzed by dimension, nature, complexity, thinking time, and goodness-of-fit. Using the Agreement Rate ($\boldsymbol{Ar}$) methodology, we present a set of sixteen (16) gestures for a lingual and palatal interface, which serve as a basis for further comparison with gestures suggested by disabled people.

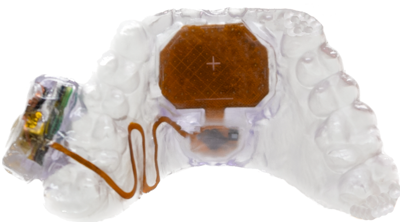
**Keywords:** Gesture interaction, tongue interaction, internet-of-things, Gesture elicitation study

# 1 Introduction

According to the analysis conducted by Villarreal-Narvaez et al. [1] on a corpus of 216 articles, it was found that none of the proposed gestures by Gesture Elicitation Studies (GES) involve the use of the tongue as a body part.

The tongue is a multifaceted organ that can execute different movements to create speech sounds, particularly those that require the tongue to touch, sweep, or come close to the palate, teeth, or alveolar ridge. These movements include rolling, folding, twisting, cloverleaf, and pointing [2]. The tongue's tip boasts impressive motor skills and mobility [3]. Unfortunately, anyone can become a victim of an accident or illness that results in significant mobility loss or quadriplegia. In such cases, communication may be limited to eye movements, blinks, or, in some cases, the tongue [4]. It would be relevant to explore how individuals with paralysis can remotely control regular household items like televisions or lamps. The device may consist of a simple infrared camera placed at the mouth [5], an intra-oral form with a touch-sensitive device provided with electrodes [6] (see Figure 1a) or a conventional touchpad located on the roof of the mouth can be operated using the tip of the tongue, as shown in Figure 1b [7].

(a) An illustration of a mechanism that can detect tongue movements through electrodes [6].



(b) *MouthPad^* [7] by Augmental.

**Fig. 1**: Examples of devices that allow interaction through the lingual and palate.

The main advantage of using the tongue to interact with computers or intelligent environments is that this little member is a muscle that maintains its activity in people who have suffered an accident and have been seriously injured. These people only have active cranial nerves to move their tongue and eyes, allowing them to interact with systems [8].

This article aims to report the results of a study of users' favorite gestures [9, 10] by the tongue and palate, considered a tactile surface like a trackpad [11]. We present: (i) a set of the most proposed gestures for 19 Internet of Things (IoT) referents based on the Agreement Rate ($Ar$) [12] and (ii) Classifying Lingual and Palatal Gestures.

Twenty-four individuals with varying ages and degrees of familiarity with computer science were interviewed to determine that a gesture must be associated with an action in particular (referent or command). Before proceeding, the participant undergoes a creativity test. Afterward, a sequence of actions (referents), such as turning on a light, is presented to them one after the other. The participants were instructed to suggest a gesture that best corresponds with the given action. There is no information regarding the reliability and precision of the device utilized, allowing for a great deal of creativity from the participants. In total, 19 referents are presented to the participants. A post-processing analysis of the participants' gestures is carried out to determine the concordance rates between these and the possible implementation of the gestures retained in the apparatus. For each gesture made, each participant gives a score on the quality of the proposed gesture (Agreement rate). Moreover, the time of reflection between the presentation of the referent to the participant and the realization of the gesture is measured. Finally, each participant evaluates the experiment using a questionnaire and gives their opinion on the experience. This procedure is standard for GES [1, 9, 10], where users show their preferences on an interaction surface [13, 14], on a mobile device [15], in front of a public display [16], for radar-based gestures [17], for nose gestures [18] or with the abdomen [19].

Gemicioglu et al. [20] analyzed tongue-based interaction and selected eight gestures from the literature according to seven factors critical to designing. The Usability factors were: Accessibility, Privacy, Feedback, and Simplicity while the selected Recognition factors were: Signal strength, Distinguishability, and Rarity. Chen et al. [21] investigated the design space and usability of mouth-based microgestures for wearable devices moving toward the face. Their research aims to facilitate intuitive interactions with these devices by exploring user preferences and developing a practical gesture set. While both studies examine interaction methods involving the mouth, our study stands out for its emphasis on lingual and palatal gestures proposed by end users to execute referents. We also focused on the interaction of people with motor disabilities with systems. In section 2, we analyze current studies on this topic. Our research methodology is detailed in section 3, while the results and discussion are presented in section 4. Section 5 discusses our findings and future work.

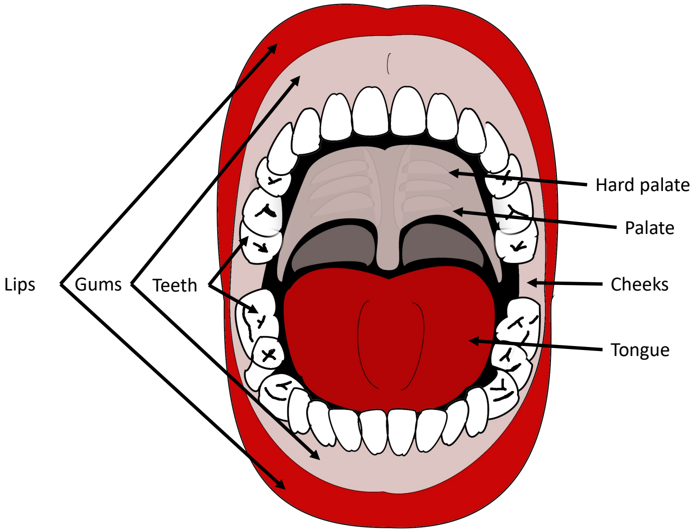



**Fig. 2**: Glossary - Buccal cavity parts

# 2 Related work

## 2.1 Glossary

In this section, we list alphabetically parts of the buccal cavity (mouth) that interact in Lingual and palatal gestures; see Figure 2.

1. *Cheeks* are the lateral walls of the mouth; they are elastic and are composed of skin.
2. *Gums* are the ones that cover the roots of the teeth and support the teeth.
3. *Hard Palate* is the upper part of the mouth; it is located behind the upper teeth, and its surface is irregular and hard.
4. *Lips* are fleshy and are the outer part of the mouth. The articulation of the lips allows speech and blocks the mouth.
5. *Palate* is located in the upper part of the back of the mouth; its surface is soft and flexible.
6. *Teeth* are the calcified parts of the mouth. Their primary function is to cut and grind food.
7. *Tongue* is the primary muscle; its mobility is wide and allows speech, and the Tongue is the sense of taste.

## 2.2 Lingual and Palatal Gestures interactions

The field of Tongue-based Human-Computer Interaction (THCI) is emerging. The Tongue can indeed be exploited for a wide range of tasks and actions; most several clinical studies focus on evaluating tongue movement [5, 22] and disabled people, but not only on them [23–27].

Yousefi et al. [23] assessed the Tongue Drive System (TDS), a wearable wireless technology enabling individuals with motor impairments to control computers and wheelchairs using tongue motion. Participants with tongue piercings completed four tasks over five sessions, investigating the TDS learning process and limitations. Tongue rings with sealed magnetic tracers were used. Comparing the tongue-TDS to an index finger keypad revealed insights into tongue-related factors, enhancing TDS usability and similar assistive technologies. Performance measures like throughput, error rate, and reaction time improved significantly, with some reaching a plateau.

Lukaszewicz [24] introduced a novel computer user interface for individuals with disabilities. The interface utilizes ultrasound imaging of specific regions on the Tongue's surface for recognition. Multiple ultrasound images were captured and analyzed to assess usability. Representative areas were identified, and the boundaries of tongue surface movement were determined. A medical ultrasound sensor was used as the input device, and a sequence of digital image frames was stored on a computer. The position of the tongue surface was determined based on the distance between the main echo and the ultrasound sensor. The signals indicating tongue position will be utilized for controlling speech synthesis.

Sahni et al. [26] addressed silent speech recognition challenges where vocalized audio is unavailable or noisy. The researchers developed a wearable system comprising the Tongue Magnet Interface (TMI) and the Outer Ear Interface (OEI). TMI tracks tongue movement using a magnetometer on Google Glass, while OEI measures jaw-induced ear canal deformation using earmold proximity sensors. A dataset of 1901 utterances of 11 phrases silently mouthed by six participants was collected. Hidden Markov model-based techniques were employed for phrase recognition. The study shows promising results for user-dependent recognition.

Ito et al. [25] aimed to introduce facial expression training as a treatment for managing facial expression in clinical practice. It focuses on teaching patients, including those with cleft lips and/or Palate, to move their facial muscles voluntarily. The experiment had two phases to assess the feasibility of using a support interface for training. Phase 1 results showed potential for introducing facial expression training with the proposed method. Phase 2 aims to evaluate the support interface's effectiveness further, providing deeper insights into its utility.

The research [6] explored various methods and techniques for detecting movement, position, and language abilities to create THCI (Tangible Human-Computer Interaction) systems. Two minimally invasive infrared systems were presented, including an electroplatographic system illustrated in Figure 1a. To improve the usability and acceptance of THCI systems, a set of usability criteria based on characteristics such as non-intrusiveness, efficient signal input conditioning, system reliability, effectiveness of control strategy, and ease of learning and use were suggested.

Augmental [7] has developed *MouthPadˆ* (Figure 1b), a tongue-driven interface that controls a computer, smartphone, or tablet via Bluetooth. It is created explicitly for the user using a 3D scan of their mouth. The apparatus is made of water-resistant dental resin, similar to the material used in dental appliances. This device utilizes a machine-learning algorithm in its processor to analyze signals from the user's tongue position and pressure, enabling them to control the cursor. Users can control the cursor using their tongue in a centralized area. Right-clicking is done with a sipping gesture, while left-clicking is achieved by pressing the tongue against the palate. It is inclusive for all individuals, regardless of disability or lack thereof.

Jain and Joshi [27] introduced the "Tongue Drive System," an assistive technology enabling individuals with severe disabilities to control their environment using their tongue. It utilizes a mouthpiece with magnetic Palaters to measure the tongue's magnetic field, wirelessly transmitting and processing the signals to control a wheelchair. The system overcomes the limitations of previous technologies, offering a cost-effective solution. The design process is analyzed, and the system is based on a PIC microcontroller and RF module. Successful testing confirms the validity of the approach.

Type by teeth (TYTH) allows users to interact with a computing system by tapping on their teeth. This interaction is analogous to using a finger to type on a keypad, except that the tongue substitutes for the finger and the teeth for the keyboard [28].

A tongue-driven system may rely on the Hall Effect to drive a wheelchair [29]: a magnet permanently attached to the tongue using adhesive tissues and tongue piercing generates a magnetic field that will vary around the mouth. These variations are sensed by an arranged and ay of Hall effect productive or mounted on the headset outside the mouth. These sensors wirelessly transmit raw data to a microcontroller and its associated microprocessor to process signals for controlling a wheelchair. This technology is reported to provide faster, smoother, and more convenient control than traditional devices, such as a joystick, especially for people with minimal or no movement ability in their upper limbs.

Abbas et al. [30] have developed a wireless, non-contact assistive technology called the Tongue Controlled Device (TCS). This device allows individuals with paralysis or limited limb motion to access computers and control their environment using tongues. A small magnet is implanted or attached to the tongue, and its movement is detected by magnetic sensors placed on a headset or orthodontic brace. These sensors wirelessly transmit signals to a microcontroller. The microcontroller then sends the signals to a portable digital assistant carried by the user. Users can use these commands to access a computer, control a wheelchair or prosthetics and interact with their surroundings. The potential applications are vast, With multiple possible orders based on teeth or tongue gestures.

Regarding brain-computer interfaces, Kæseler et al. [8] used single-trial electroencephalography (EEG) to detect and classify tongue movements. Their results determined that at least four tongue movements (left, right, down, and up) can be reliably used to interact with a system.

Mat Zin et al. [31] reviewed of the different Electropalatography (EPG) in speech therapy. These devices are used for contact detection of the tongue and hard palate. The result, a list of EPGs with their modifications and specific characteristics, describes the evolution of the devices and the incorporation of new technology to provide patients with greater comfort. Finally, the researchers suggest that the study's advanced EPG system can be expanded for other uses and not just limited to monitoring. This reflection makes us think that we could use them for our interest in gesture detection and associate it with an input mechanism.

Two studies [32] examine tongue movements and the activity of suprahyoid muscles through electromyography (EMA). The researchers tested six palatal tongue movements, namely extrusive, right-and-left lateral, upward, and downward, touching the Palate with the tongue tip, and handling the palatal aspects of the maxillary second molars. By utilizing a wireless electrode-type system in the suprahyoid, they discovered that the EMA of the suprahyoid muscles precisely reflects the tongue's movements. This breakthrough enables a device to detect the tongue's activity as input.

Li et al. [33] introduced a non-contact sensor called Tongue-n-Cheek, which uses an array of micro-radars to capture tongue gestures. These micro-radars are proximity sensors that detect muscle movements during tongue gestures. By employing a unique signal processing algorithm, it translates these movements into gestures. Tongue-n-Cheek demonstrates high effectiveness with 95% gesture recognition accuracy and minimal delay.

Chen et al. [21] conducted three studies. The first study involved participants brainstorming and suggesting microgestures using their mouths to facilitate communication. The second study explored the physical and mental workload of the proposed gestures, which were then classified based on the interaction. For the third survey, the researchers provided participants with a list of the top 20 preferred gestures from study 2, along with descriptions, and asked them to choose the most appropriate gesture for a task. While our research is similar to Chen et al.'s study, we focused on using the tongue as the primary reference point, resulting in 80 unique gestures compared to the 20 preferred gestures. Additionally, our contribution is based on a more comprehensive approach that considers the tongue's movement and interaction with other parts of the mouth rather than solely focusing on the mouth's mechanisms.

Gemicioglu et al. [20] introduced a system utilizing multiple modes and devices to evaluate different sensors' effectiveness in accurately distinguishing tongue gestures. The research emphasizes the practicality and potential advantages of utilizing tongue gestures as a non-invasive and convenient method of interacting with head-mounted displays while ensuring user confidentiality. The study focused on identifying factors critical to designing (seven factors)

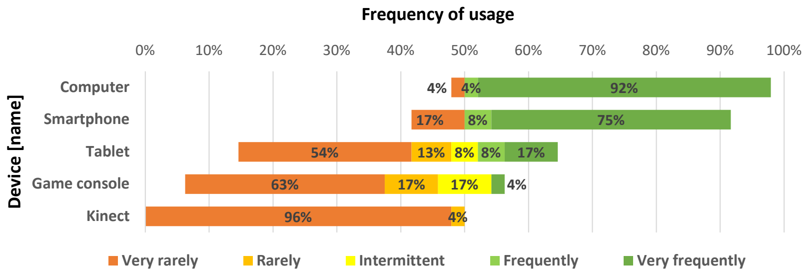



**Fig. 3**: Frequency of usage of devices per participant.

tongue gestures that are usable (Accessibility, Privacy, Feedback, and Simplicity) in various settings and can be effectively recognized (Signal strength, Distinguishability, and Rarity) by classification algorithms. Several factors affect its ergonomics. These include the form and placement of sensors that detect tongue gestures, the types and number of distinguishable tongue gestures available, the feedback and guidance provided to the user during input, and the cognitive and emotional aspects of using these gestures, such as mental workload, privacy, social acceptance, and user satisfaction. As a result, they presented eight gestures with the tongue and mouth closed to interact (plus two control gestures). These gestures were later used in another study [34] to combine with gaze tracking as a point-and-click interface in three hands-free games and a musical instrument. Unlike our study, the gestures were intended for users, while end-users proposed our group of gestures without limitations (privacy).

# 3 Experiment

All participants adhered to the same experimental procedure to ensure successful data collection.

## 3.1 Participants

To ensure that the results accurately reflect the general population. For this experiment, we selected twenty-four no-handicap participants (12 women and 12 men, with a Male/Female ratio of 1) ranging from 15 to 66 years old (with a mean age of 30.70 and Standard Deviation of 15.05). Of the participants, 63% were students; only one was left-handed, while 23 were right-handed. The average score on the creativity tests was 60.92, slightly below the average of 63.45. None of the participants had prior knowledge of the tongue apparatus, and only one had no computer experience. Four participants reported rarely using a smartphone. Most participants had low experience with video games and Microsoft Kinect, while their experience with tablets was average, as shown in Figure 3.

**Table 1**: List of referents.

| ID | Referent | ID | Referent |
|---|---|---|---|
| 1 | Turn TV on | 11 | Turn light off |
| 2 | Turn TV off | 12 | Brighten light |
| 3 | Start player | 13 | Dim light |
| 4 | Volume up | 14 | Turn heating on |
| 5 | Volume down | 15 | Turn heating off |
| 6 | Next | 16 | Turn alarm on |
| 7 | Previous | 17 | Turn alarm off |
| 8 | Turn AC on | 18 | Answer phone |
| 9 | Turn AC off | 19 | Hang up phone |
| 10 | Turn light on | | |

## 3.2 Referents

As this is a GES, we will refer to a user interface feature that can be controlled independently with a command as a "referent," which is equivalent to a "function" [1]. The 19 referents of the experiment are presented in Table 1.

## 3.3 Procedure

Two members of the group welcomed the participants at the experimental site. The environment consisted only of a desk and a computer. Each participant must have no link with the members of the group questioning him to ensure the neutrality of his answers. We divided the experiment into three phases.

**Pre-test Phase**. The participants were first asked to fill out the consent form and a demographic questionnaire, including questions about their knowledge of connected devices. Then, the participants scored between 1 and 7 according to their expertise and frequency of device use. Then, the participant's level of creativity was assessed before the elicitation procedure using an online generic creativity test (testmycreativity). The test yields a score ranging from 0 to 100, with higher scores indicating greater creativity. The score is calculated based on participants' responses to 40 questions (See Appendix A). In GES, the participants must propose a gesture for a referent, action, or task in an environment [1]. Therefore, it requires a level of creativity on the part of the participants; we consider that the performance in this test should be equal to or greater than 50/100.

**Test Phase**. After successful pre-tests, the experience begins. The participant is introduced to the tongue detection device, and each mouth part (rough palate, hard and soft, floor of mouth, teeth, etc.) is briefly explained. This ensures that the participant and our group use the same vocabulary, which is crucial for an unbiased experience. Without a common understanding, the conclusions drawn from the experience could be considered obsolete. To prevent this, necessary arrangements have been made. Next, the 19 referents are presented randomly to the participant as studies [9, 12, 35]. Participants perform and rate the tongue gesture for each referent on a scale of 0 to 10. We also record the reflection time. It's important to note the variability in the order of referents, as having a fixed order could significantly impact the results. This highlights the importance of the scientific research methodology and the need for clarity and transparency to avoid ambiguity.

**Post-test Phase**. Once the participant finishes using the device, they are asked to complete a post-test questionnaire (see Section 4.6). The questionnaire includes closed questions with predefined answers and open-ended questions for participants to express their opinions freely. While the closed questions are quickly processed, the open-ended questions require more elaborate manual work. This step is crucial to the experimental process and is often overlooked. Our observations have shown that gathering participant feedback after the fact is beneficial for two reasons. Firstly, personal feelings can provide a better interpretation and understanding of the results. Secondly, it can assist future researchers in adapting the experimental setup to be more relevant and effective for those with quadriplegia.

## 3.4 Results

During the study, 456 lingual and palatal gestures were observed and analyzed. Out of these, it was found that the participants used 80 different gestures. These gestures were categorized based on how the tongue interacted with the oral cavity, such as touching, tapping, shooting, turning, blocking, and sweeping. Some of these gestures were single-stroke, while others required multiple strokes [36]. To categorize and name the gestures (as shown in the last two columns of Table 2), we followed the guidelines of Braner and Hort's work [37] (four main categories and nine subcategories), and also, our inspiration for this approach comes from the original GES by Wobbrock et al. [10], which used a touch surface. In our case, we view the tongue as the equivalent of the hand and the remaining mouthparts as the touch surface.

Our categorization of the “Groups” and “Subgroups” was informed by previous research on GES. The term “Group” is utilized to differentiate whether the participant solely used their tongue or if it interacted with another part of their mouth, akin to the ”locale” dimension proposed by Piumsomboon et al. [38], which encompasses on-the-surface, in-the-air, and mixed locale. Regarding the “Subgroup”, we analyzed whether the gestures were static (To perform the tongue pose, one must hold their tongue in a fixed position) or dynamic (The tongue position is altered in a specific area) [10] and pinpointed the exact part of the mouth involved.

It's essential to consider both the oral components and the location of a gesture to differentiate between them. For example, touching the right cheek versus the left cheek with the tongue are two distinct gestures. To assess the overall coherence between gestures, we utilize the AGATe software [12], which calculates the coherence rate between tongue movements. The program uses the formula developed by Wobbrock et al. [9] to determine the Agreement

| Group | Subgroup | Category | Gesture |
|---|---|---|---|
| In-the-air | Static | Swallow | 1: tongue |
| | | Bend | 2: tongue side<br>3: tongue via hard palate<br>4: tongue via high teeth<br>5: tongue between teeth |
| | | Tongue out | 6: to up<br>7: straight<br>8: to left<br>9: to right<br>10: to down<br>11: U-shaped |
| | Dynamic | Tongue out | 12: to up twice or more times<br>13: to down twice or more times<br>14: to right twice or more times<br>15: to left and right<br>16: to right and left<br>17: to down and up<br>18: and bite<br>19: and in<br>20: to up and down<br>21: to down and right<br>22: to up-left and up-right<br>23: to down-left and down-right |
| | | Rotate | 24: tongue clockwise<br>25: tongue anti-clockwise |
| On-the-surface | Palate | Touch | 26: palate<br>27: palate twice<br>28: high-low palate<br>29: hard palate 3 times<br>30: soft palate 3 times<br>31: soft palate<br>32: lower palate<br>33: hard palate<br>34: Hard-low palate<br>35: palatal papilla |
| | | Sweep | 36: front-back palate<br>37: back-front palate<br>38: left-right upper palate<br>39: clockwise palace circle<br>40: anti-clockwise palace circle |
| | Teeth | Touch | 41: top in teeth twice<br>42: two times high then low teeth<br>43: a tooth<br>44: front upper teeth<br>45: inner upper teeth<br>46: inner lower teeth<br>47: top lower teeth<br>48: right lower teeth<br>49: left upper teeth<br>50: front lower teeth |
| | | Sweep | 51: between left-right teeth<br>52: right-left teeth<br>53: left-right teeth<br>54: right-left top bottom teeth<br>55: back-forth teeth |
| | | Block | 56: pinch<br>57: tongue between center-teeth<br>58: two or more times |
| | Lips | Touch | 59: left corner<br>60: closed mouth<br>61: upper lip<br>62: lower lip in the middle |
| | | Sweep | 63: left-right upper lip<br>64: right-left bottom lip<br>65: right-left upper lip<br>66: back-forth upper lip<br>67: up-down lip |
| | Gums | Touch | 68: upper gum<br>69: lower gum |
| | | Sweep | 70: up-down gum |
| | Cheek | Touch | 71: left cheek then right cheek<br>72: right cheek then left cheek<br>73: cheek twice<br>74: cheek<br>75: left cheek in the middle<br>76: left cheek in the middle 3 times<br>77: left cheek up<br>78: right cheek |
| | | Sweep | 79: clockwise right-cheek circle<br>80: anti-Clockwise left-cheek circle |

**Table 2**: Classification of lingual and palatal gestures.

rate, which is calculated using formula (1) as cited in the study by Vatavu et al. [35].

$$A(r) = \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|}\right)^2 \geq AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|}\right)^2 - \frac{1}{|P|-1} \tag{1}$$

To characterize the degree to which participants simultaneously agree for referents $\{r_1, r_2, ..., r_k\}, 2 \leq k \leq |P|$. The co-agreement rate for $k > 2$ referents is determined by using the formula (2). Where $P$ corresponds to the set of gesture proposals for referent $r$ and P corresponds to the size of this set. $P_i$ represents the subset of propositions of identical gestures of the set $P$.

$$CR(r_1, r_2, ..., r_k) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{k} \delta_{i,j}, n = \frac{1}{2}|P|(|P| \; - 1) \tag{2}$$

As part of this study, an Agreement rate of 0.029 was measured. This one is relatively weak and is explained by the great diversity of the gestures retained.

In addition, specific gestures are often performed several times by the same person for different referents. We also often found that the gesture of a participant associated with a particular referent does not coincide with the same referent for another individual. In the series of referents presented, most of the actions are opposed. Usually, the power button on a TV remote is used to turn the TV on and off. It might seem logical for people to use the same gesture to turn other devices on and off, like sticking out their tongues. However, there is often a lack of agreement about these gestures, mostly because people forget what was previously suggested.

A complete classification of the gestures is carried out to claim the majority of gestures. For example, the gestures “touch the second canine on the right” or “touch the second molar on the left” are grouped in a new category called “touching a tooth”. This second sort considers the more global aspect of the gestures made. However, it is still essential to distinguish gestures that make sense from the action taken, such as pulling one’s tongue up or down to select the next or previous item. From this new sort, it turns out that 80 gestures are retained with an average Agreement rate of 0.035, as shown in Figure 4.

The Agreement Rate, which is still relatively low, reflects the mismatch of participants as to the use of tongue only. This shows, in particular, an existing link with a time of generally critical reflection, as presented in the rest of this document. No participant is used to performing such an exercise. However, it is interesting to highlight tongue movements with the highest Agreement rates as gesture No. 20 “Next” (0.091 Agreement rate) with the gesture “tongue out up and down”; its opposite, the gesture No. 17 “Previous” (0.062 Agreement rate) with the gesture “tongue out down and up” and finally the gesture No. 7 “Turn light on” (0.083 Agreement rate) with the gesture “tongue out (straight)”. The gestures performed for “Next” and “Previous” referents are naturally linked

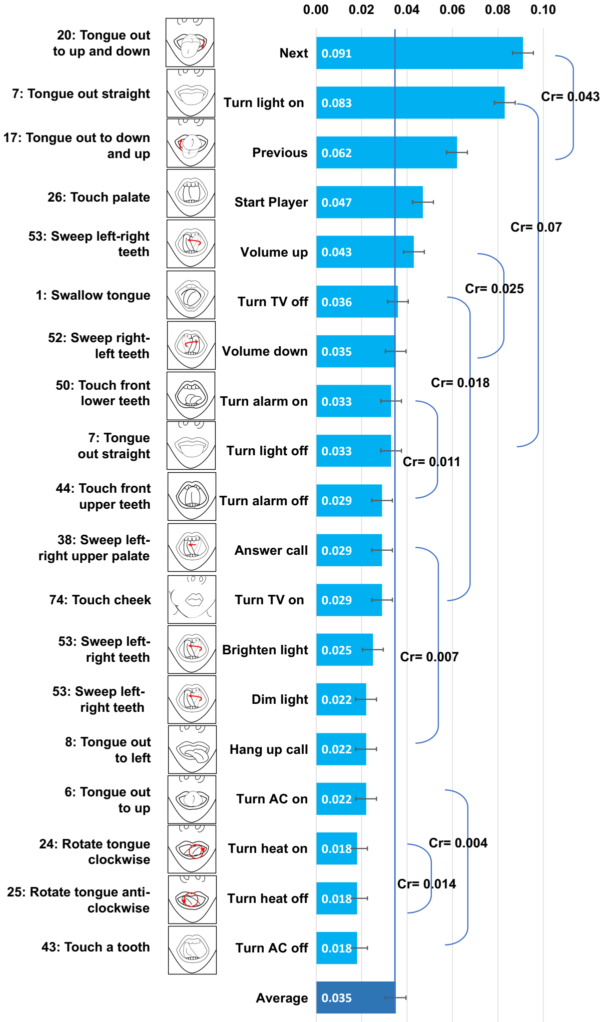



**Fig. 4**: Most elicited gesture by referent and Agreement rate of gestures by referent. Error bars indicate the standard deviation of the agreement rate. Co-agreement rate between opposite referents (Cr).

to the gesture that one would make on a smartphone to select the next or previous item of a list with the help of a finger. The gesture of tongue out (straight) instead of "Turn light on" may refer to the pressure exerted on the switch to "Turn light on". The latter can be relatively performed by known physical gestures.

The Frequency of reproduction of gestures is shown in Figure 5. As stated above, several gestures can be associated with the same motion. The most frequent gestures are shown in Table 3.

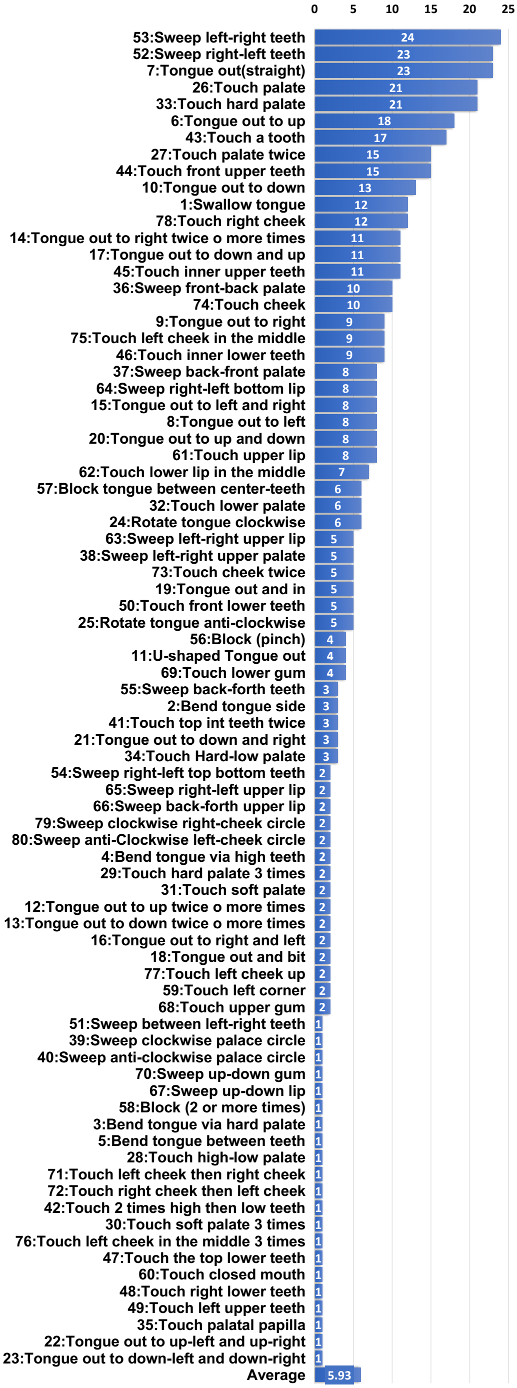



**Fig. 5**: Frequency of occurrence of gestures

**Table 3**: The most frequently elicited gestures.

| Times | Gesture ID | Gestures |
|---|---|---|
| 20 | 53 | Sweep left-right tooth |
| | 57 | Tongue out straight |
| | 26 | Touch palate |
| | 33 | Touch hard palate |
| | 52 | Sweep right-left tooth sweep |
| >10 | 6 | Tongue out straight up |
| | 43 | Touch tooth |
| | 27 | Touch palate two times |
| | 44 | Touch front lower teeth in front |
| | 14 | Tongue out two times |
| | 10 | Tongue out down |
| | 78 | Touch cheek right |

Based on the analysis of gesture frequency, most elicited gestures appear relatively simple, except for repetitive actions, almost palate and pulling the tongue. Interestingly, none of the most commonly executed gestures were associated with the referents that received the highest Agreement ratings. This suggests a lack of spontaneity and natural flow in the experiment design. Additionally, examining the relationship between opposing referents, such as turning a television on and off, may be worthwhile. There are nine opposite referents listed, as illustrated in Figure 4.

Surprisingly, most gestures have a low Agreement rate except for the third co-referent, "Next/Previous", which has an impressive Agreement rate of up to 0.043. In the case of other referents, a low Agreement rate implies that the associated gestures must not be "Inverted" between opposite referents if they are to be retained. For instance, it would be logical to use "Tongue out to the right to answer call" and "Tongue out to the left to hang up call".

# 4 Analysis and Discussion

After gathering data through experience, analyzing and interpreting the results is essential. This involves extensive Classification of gestures and overall assessments by participants. To do this, we classify gestures based on dimensions, nature, complexity, participant score, and duration of reflection time. This helps identify the gestures that were successful in achieving their intended purpose.

## 4.1 Classification of gestures by number of dimensions

The Classification of gestures is based on the number of dimensions of tongue movements, as depicted in Figure 6.

0D. For example, a gesture can be classified as 0D at the touch of an oral part or simply holding the right tongue out of the mouth. Given the device (infrared

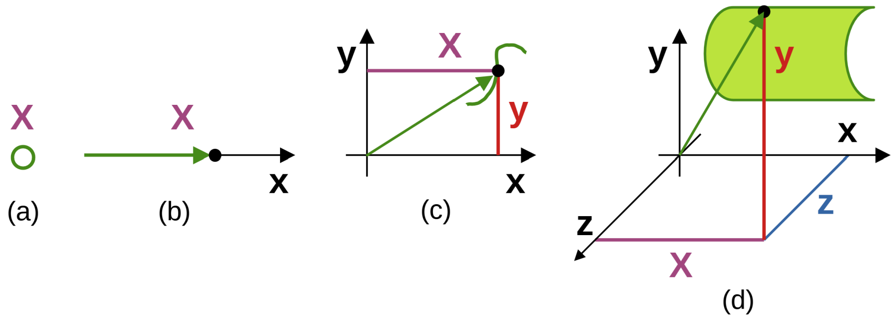



**Fig. 6**: Illustration of the gestures by dimension: a) 0D, b) 1D, c) 2D and d) 3D.

camera), the movements concerned are mainly those of the type "To type", "To touch" and "To shoot" (in the direction to leave). The gestures "Touch" and "Pull" take into account the duration during which the tongue remains fixed (zero acceleration on contact or when out of the mouth for a while). The "Typing" gestures take into account the tricky aspect of the contact (rarely acceleration). It is also decided to use the gestures "To type" and "To touch" with repetitions (for example "Type 2 times"). In more detail, the tongue can reach the internal and external parts of the mouth. Internally, we find: teeth (front and back), gums (up and down), cheeks (left and right), palates (hard, soft, and rough), and the floor of the mouth. Externally, the tongue can reach the lips (up and down) or pull the tongue. For these parts, it is also desirable, in case useful, to differentiate the direction of the tongue (for example: "Touch the front upper teeth"). It is also possible to touch several parts of the mouth simultaneously. Thus, a huge number of gestures are possible.

1D. 1D gestures are performed when the tongue moves in a single direction with usually a more or less constant speed. In this case, the tongue can reach the same parts of the mouth as in the case of 0D gestures. In this category are mainly the gestures of type: "Sweep part of the mouth" or "Tongue out and sweep". For sweeps, depending on whether the gestures have meaning about the referent, it is necessary to consider the direction of the movement (for example: from left to right, from right to left, from top to bottom, from bottom to top, etc.).

2D. The tongue describes the gestures in 2D according to a surface belonging to a plane (2 directions). It is possible to describe 2D gestures in contact with the palate (hard, soft, and rough), the floor of the mouth, or the cheeks.

3D. 3D gestures describe a surface in space (3 directions). Examples of gestures can be "Turning" the tongue inside or outside the mouth. In this category are also taken more original gestures such as "Bend" the tongue or "Deform in the shape of a U".

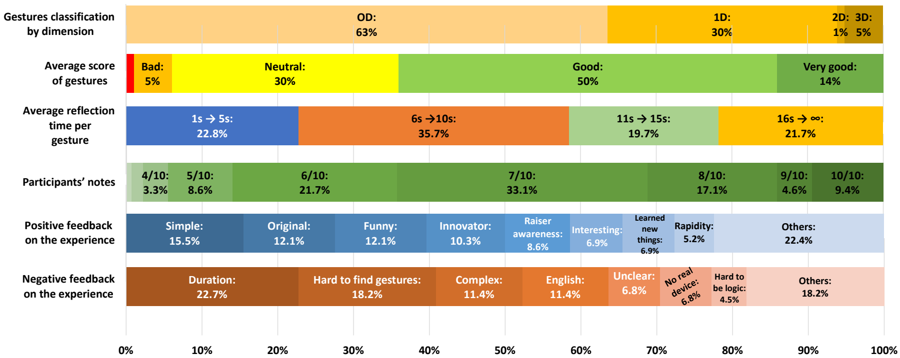



**Fig. 7**: Frequency of Classification of gestures by dimension, Average score of gestures, Average reflection time per gesture, Participants' notes, Positive feedback on the experience, and Negative feedback on the experience .

Consequently, 63% of the gestures, which amounts to 289, are of 0D type. Out of these, 135 are categorized as "touch", 75 as "pull", and 68 as "type". Category 1D comprises roughly 30% of the gestures, which is 139 in number and includes 94 "sweep inside the mouth" gestures and 45 "pull the tongue and sweep" gestures. The remaining gestures, only 1% of which are 2D and 5% of which are 3D, are shown in Figure 7.

It's worth noting that around 75% of the observed gestures involved some form of contact with the mouth, and a similar proportion of gestures were made inside the mouth. This suggests that participants found it easier to imagine relatively simple gestures.

## 4.2 Classification of gestures by nature

The gestures are classified according to their nature, similar to the approach taken by McNeill [36, 39] and Wobbrock. et al. [10]:

1. Physical: the gesture described by the participant as if he were acting on the object. e.g., work as if you were pushing the button to turn on the television by pulling the tongue.
2. Metaphorical: the gesture refers to a metaphor; in other words, the relationship with the referent is indirect. For example: perform a palace scan to answer a call, as is the case with your finger on most smartphones.
3. Symbolic: The gesture makes a symbolic reference linked to the action to be performed at a distance. For example: Draw a "T" on the palace to turn on/off the television concerning the first letter of the object in question.
4. Abstract: the gesture described is not directly related to the action to be performed remotely. For example: approaching both sides of the U-shaped tongue to turn on the heat, so the two activities are unrelated.

This Classification appeals to the subjectivity of the experimenters. The analysis results are shown in Figure 8. First, no symbolic gesture has been

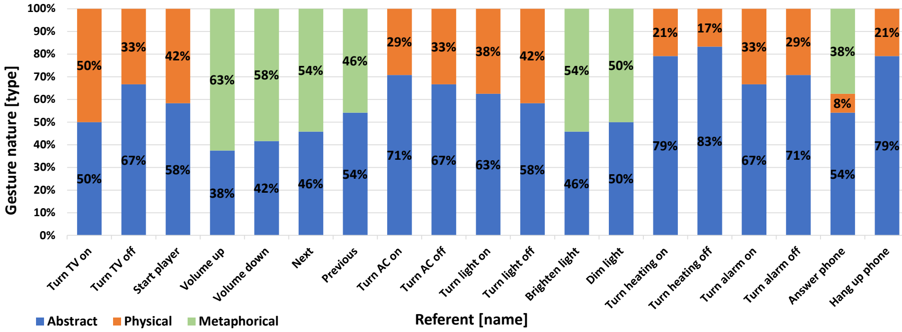



**Fig. 8**: Classification of gestures by nature.

made, given the apparatus used and the actions to be performed. Next, most of the gestures elicited are of the order of the abstract (60.2% on average). Secondly, physical gestures (20.8% on average) and metaphorical gestures (19% on average). Given the apparatus and the actions to be performed, the large number of abstract gestures coincides with earlier analyzes in which the participants already had difficulty finding a generally logical meaning with the referent.

We observed metaphorical gestures for actions such as “Volume up/down”, “Next”, “Previous” or “Brighten/Dim Light” through sweep-type gestures with the tongue, like what it could be doing on a smartphone device. Finally, we find gestures of the physical type for referents of the kind “Turning on/off + object” with a gesture like exerting pressure on a buccal part or pulling the tongue, referring to the button of sort On/Off of the objects.

After analyzing the relationship between the referents and the gestures proposed by participants, we found that referents with context that can vary their value or position, such as volume up/down, brighten/dim light, and next/previous, are associated with dynamic gestures. These gestures involve the tongue moving to create single or multiple strokes, such as a left-to-right sweep of the teeth to increase volume. Interestingly, we also observed that gestures in the opposite direction corresponded with opposite references, such as a right-to-left sweep of the teeth. For the referents linked to activating or deactivating (On/Off), the participants proposed static gestures, where the tongue maintained a position or only made contact with another part of the mouth, e.g., sticking out the tongue, touching a tooth, or retracting the tongue. Except for heating, the participants may have associated this with regulating the temperature and rotated the tongue clockwise to turn on and anti-clockwise to turn off.

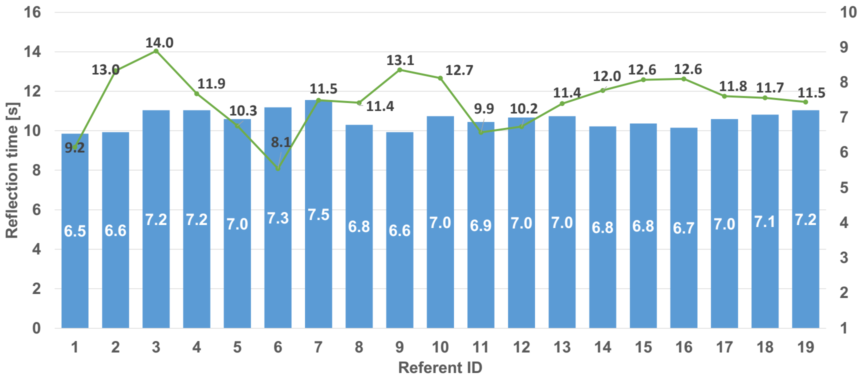



**Fig. 9**: Correlation reflection time - Score.

## 4.3 Classification of gestures by complexity

A study was conducted to analyze the complexity of gestures. The gestures were categorized as simple (done in one stroke) or composed (made by combining multiple traits). The results showed that 61% of the gestures were uni-stroke, while 39% were multi-stroke.

The analysis of complexity presented here shows that participants preferred simple gestures over complicated ones. The more straightforward gestures involved touching or slapping the tongue, while more complex ones included actions like "pulling the tongue and swiping left to right", "tapping the left cheek then the right" or "sticking the tongue out twice upward". One possible explanation for this preference is that, when faced with a new situation and time pressure, participants naturally opted for simplicity, prioritizing ease and speed over complexity, even more than they might have in a typical, everyday scenario.

## 4.4 Classification of gestures by score

Following each gesture performance, the participant was asked to rate their gesture on a scale of 1 to 10 to evaluate its quality. Different categories of scores were defined to classify the gestures. Scores 1 and 2 are considered very bad, 3 to 4 are bad, 5 to 6 are average, 7 to 8 are good, and 9 to 10 are excellent. The frequency of scores is illustrated on average in Figure 7.

Most participants showed great enthusiasm for the proposed gestures, with around 64% of gestures receiving good and very good scores, while only 6% were rated bad or very bad. A majority of participants expressed satisfaction with their actions, despite low Agreement rates. Notably, referent No. 13, which involves decreasing luminous intensity, mainly received good grades, despite a low Agreement rate. However, due to participants' lack of familiarity with using only tongue movements to communicate, agreeing on the gestures to perform was challenging. Nevertheless, 50% of the notes attributed to gestures

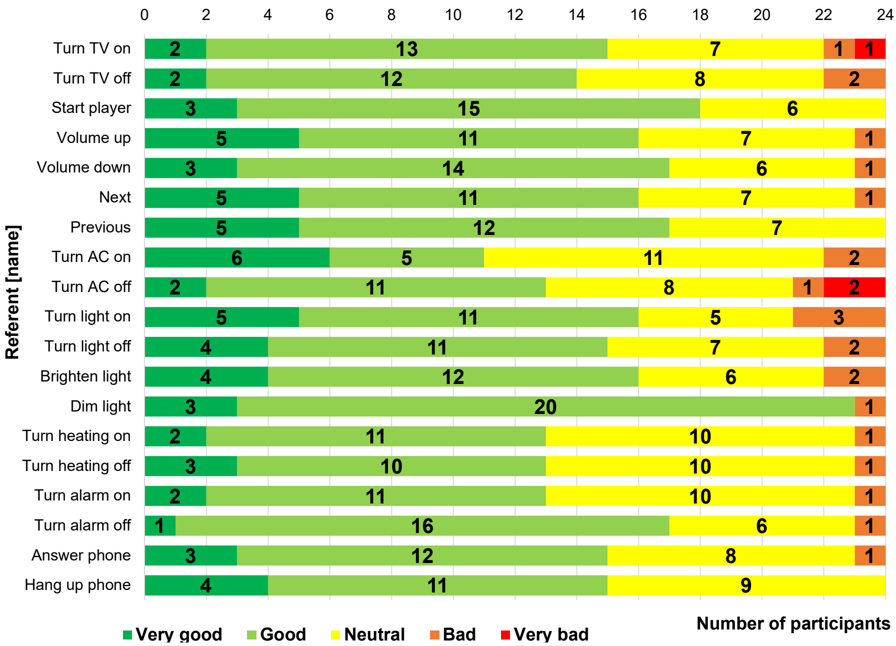



**Fig. 10**: Scores of gestures by referent.

were 7/10 and 8/10, indicating a high level of success. Figure 9 provides further insight into the distribution of scores.

## 4.5 Classification of gestures by the time of reflection

The time participants take to reflect on the gesture they want to achieve can provide valuable insights [40]. A shorter reflection time usually indicates a more natural and coherent gesture. However, when participants take too little time to reflect, they may make a gesture that does not align with the desired action. To simplify things, the reflection times are grouped into five-second categories (always rounded to the nearest integer), as illustrated in Figure 10 and in greater detail in Figure 11.

Only 22.8% of gestures are performed with a relatively short reflection time (1 to 5[$s$]). This is relatively minor, thus reporting the difficulty of some participants to find a multitude of gestures only with the tongue, especially after some actions have already been done. 35.7% of the gestures are in the category of gestures of average duration (from 6 to 10[$s$]). In contrast, about 40% of the gestures are performed after a considerable time for reflection (more than 10[$s$]). To deepen the analysis, examining the correlation between reflection times and mean scores by referent would be helpful, as shown in Figure 9.

We can notably note that referent No. 6 (“Next”) has the best Agreement rate, the average reflection time is the lowest, and the given score is the second best. The referent No. 7 (“Previous”) scored the highest score. It is also

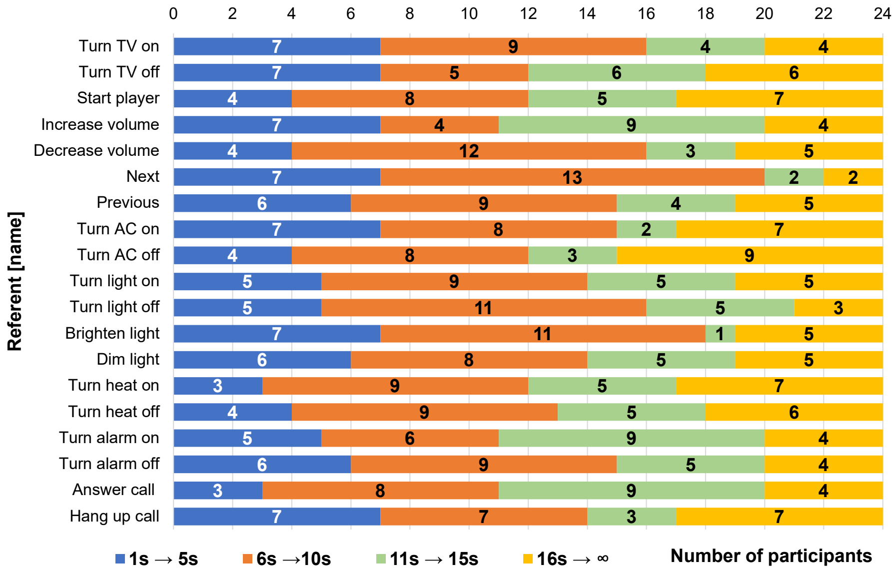



**Fig. 11**: Reflection time of gestures by referent.

interesting to note that referents such as No. 8, 9, 4, and 15 ("Turn AC/heat on/off") somewhat surprised or confused the participants, who needed more reflection time and gave their choices lower scores. For referent No. 3, the reflection time is the most important, and the assigned score is high, thus going against intuitive reasoning. The referents are generally achievable on a smartphone-type device, have frequently needed less time for reflection, and have better ratings because they seem more logical for the participant.

## 4.6 Experiment appreciation

Following each study, participants were asked to complete *the IBM Post-Study System Usability Questionnaire (IBM PSSUQ questionnaire)* [41]. It is a tool used for evaluating the usability of systems. The questionnaire consists of 19 items measuring user satisfaction, such as system usefulness, ease of use, and interface quality. The responses are collected on a 7-point Likert scale, ranging from strongly agree to disagree strongly, or the participant can opt to abstain from answering questions irrelevant to the study. We utilized a proposed questionnaire version by UIUXTrend consisting of 16 questions and maintained a scale of 1 to 7 (See Figure 13).

The average score for most questions ranges from 4 to 6, as demonstrated in Figures 12 and 13. However, participants deemed questions 7, 8, 9, and 10 off-topic, yet some still provided scores for these questions. Participants rated question 12 highly regarding clarity of information and question 13 for the acceptability of the proposed interface. However, the ratings for questions 1 and 2, which are related to the ease of use of the system, were the lowest.

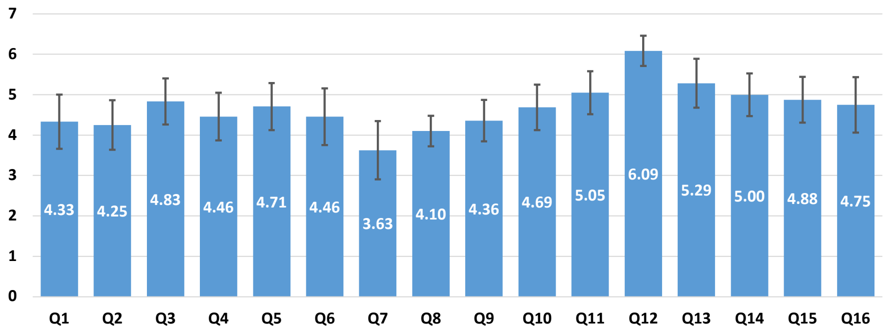



**Fig. 12**: Score by question.

Of the participants, 15 were satisfied with the design, while seven were not. Furthermore, nearly half of the participants were uncertain about the device's learning ability and felt uneasy using it.

Along with the IBM PSSUQ questionnaire, participants were asked to provide three positive and negative aspects of their experience. Only relevant and notable occurrences were considered to draw accurate conclusions. Participants faced various challenges, such as navigating between opposite gestures (e.g., turning the TV on/off), finding gestures, and understanding the creativity test. However, they also highlighted positive aspects of the experience, including the device's originality, innovation, practicality, and entertainment value. Refer to Figure 7 for more information on participants' opinions.

## 4.7 Implications for Designing Palatal and Lingual Gestures

Throughout the elicitation process, participants could perform their gestures without any restrictions. However, we did not gather information on whether the participants would feel comfortable performing these gestures in private or public settings, such as work meetings. We assumed that the participants were in their personal or family space, so their proposed gestures should not be influenced by their surroundings or any criticism from others. It's important to note that some gestures, like No. 7, where the user sticks their tongue out to turn off the light, may elicit laughter from onlookers. Additionally, many suggested gestures require the user to open their mouth, which can be awkward in public settings. If others don't understand these gestures, it could lead to confusion or mockery, harming the user's social standing.

We relied on the frequency of open-mouth gestures proposed by end-users (see Figure 5) as they felt at ease suggesting these gestures. In contrast to Gemicioglu et al. [20], who proposed eight gestures with closed mouths for privacy purposes.

This work and the one carried out by Chen et al. [21] are independent studies. They focus on gestures made by the mouth; for example, there is the

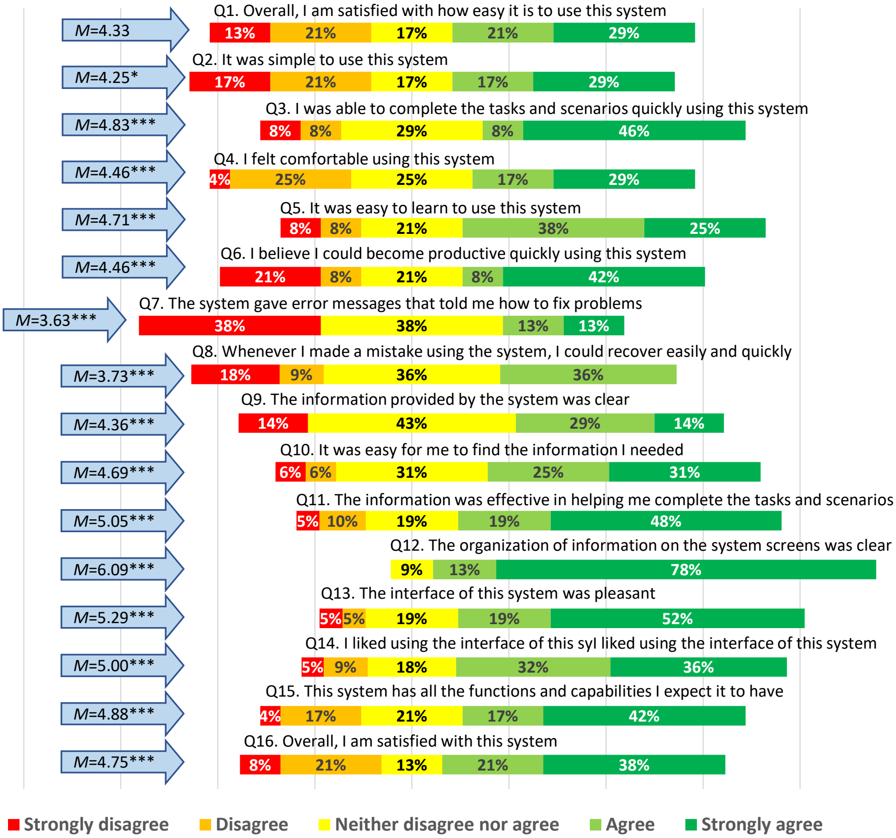



**Fig. 13**: Distribution of scores by question. $M$ represents the mean score of each question and the level of significance results from a Wilcoxon sign-rank test to determine whether the scores depart from the median $Mdn$=4.

gesture to close the mouth (G6); this gesture would not be taken into account in our study since we focus on gestures that can be produced with the tongue and can interact with parts of the mouth. In study 3, it is illustrated that 12 of the 20 (from G1 to G20 gestures) gestures were with the tongue. These 12 gestures are found in our classification (see Table 2). G2, G3, G11, G2, and G15 (in total, five gestures) match the gestures associated with references from our study (see Figure 4). Users prefer to propose gestures where the tongue interacts with the teeth and the palate.

We suggest the following:

- ▲ Consider that the gestures are made with the mouth open.
- ▲ For opposite referents or commands, such as turning the volume up or down, consider gestures that move in opposite directions.
- ▲ Users prefer to interact with their teeth and palate.

The gestures that received agreement for executing a referent are displayed in Figure 4. The participants tended to favor gestures where the tongue moved independently and interacted with the teeth as a secondary option.

# 5 Conclusion and future work

For many people, controlling remote objects is no big deal, thanks to the increasing use of smart gadgets. However, for those with disabilities, it can be a complex task. This study focuses on using a device to control objects through tongue movements. To determine appropriate gestures to remember, 24 participants were interviewed, and their elicited gestures were sorted. Surprisingly, there was little correlation between the participants' gestures.

Nonetheless, of the 80 gestures retained, many related to actions that can be performed on devices like smartphones. We matched 16 gestures with 19 IoT referents based on the Agreement Rate calculation. However, the inexperience of users in using their tongue as a remote control tool led to relatively long reflection times to make gestures, resulting in a lack of spontaneity and concordance among participants. Despite this, participants were somewhat satisfied with their proposals. Finally, after analyzing the results and as a conclusion of our discussion, we establish implications for the design of a lingual and palatal interaction.

The participants perceived the lingual and palatal interaction as very useful in raising awareness about the cause of the disability. There is great enthusiasm, and interest in the development and standardization of use in the future. However, this will necessitate new studies on acceptance. In Section 2, we provide examples of this interaction.

Although our work in part of Section 2 focuses on the study of mobility and the ability to move the tongue of people with disabilities or paraplegia, the participants in our experience were non-disabled. This means that the proposed gestures may not consider the challenges of people with disabilities. As future work, we consider carrying out a GES with disabled people, comparing the final gestures, and evaluating the difference between the two studies. Waiting for a lingual and palatal interaction to become popular, that is, systems that interact with gestures of the language are developed, we could compare these with the set of gestures of the system and gestures proposed by end-users. This study could be replicated to assess the influence on users when an interface becomes popular and commercial systems are available.

# Appendix A Creativity test

TestMyCreativity is an online tool that comprises 40 questions. We specifically selected this test due to its convenience in administration and its ability to assess a wide range of factors. The test evaluates various dimensions of creativity, including *abstraction*, which involves extracting fundamental concepts from ideas. It also measures *connection*, which focuses on identifying relationships between seemingly unrelated elements. The test assesses *perspective*, which entails the ability to shift viewpoints in terms of space, time, and other individuals. Additionally, it examines *curiosity*, which reflects a willingness to challenge and improve established norms. The test also measures *boldness*, which encompasses pushing boundaries beyond conventional limits. *Paradox* is another dimension evaluated, which involves embracing and effectively working with contradictory concepts. Furthermore, the test assesses *complexity*, which evaluates an individual's ability to handle and process a large amount of information. Finally, *persistence* is examined, focusing on the determination to continue seeking stronger solutions even after generating satisfactory ones. Figure A1 shows an example of the questions asked in this questionnaire.

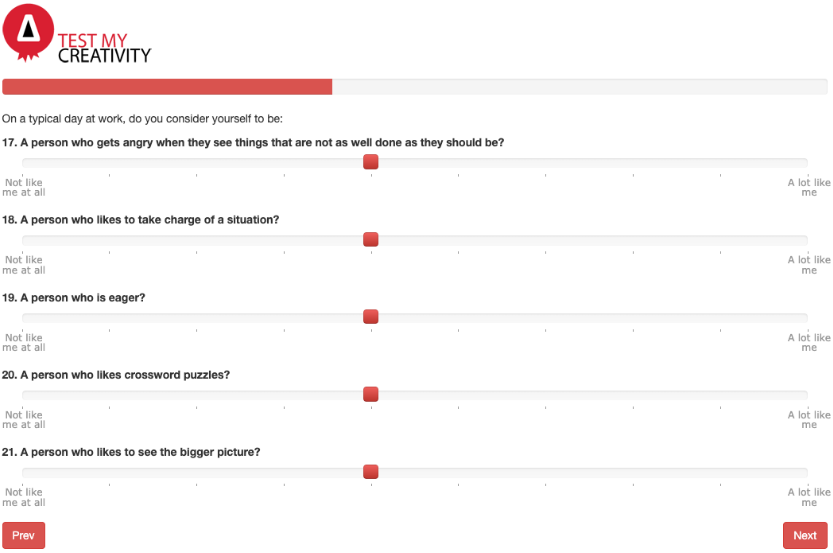



**Fig. A1**: Example of creativity test questions (TestMyCreativity)

# References

[1] Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.-D., Wobbrock, J.O.: A systematic review of gesture elicitation studies: What can we learn from 216 studies? In: Proceedings of the 2020 ACM Designing Interactive Systems Conference, pp. 855–872. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3357236.3395511

[2] Katz, W.F., Mehta, S.: Visual feedback of tongue movement for novel speech sound learning. Frontiers in Human Neuroscience **9** (2015). https://doi.org/10.3389/fnhum.2015.00612

[3] Caltenco, H.A., Lontis, E.R., Boudreau, S.A., Bentsen, B., Struijk, J., Andreasen Struijk, L.N.S.: Tip of the tongue selectivity and motor learning in the palatal area. IEEE Transactions on Biomedical Engineering **59**(1), 174–182 (2012). https://doi.org/10.1109/TBME.2011.2169672

[4] Rupp, R., Biering-Sørensen, F., Burns, S.P., Graves, D.E., Guest, J., Jones, L., Read, M.S., Rodriguez, G.M., Schuld, C., Tansey-MD, K.E., Walden, K., Kirshblum, S.: International Standards for Neurological Classification of Spinal Cord Injury: Revised 2019. Topics in Spinal Cord Injury Rehabilitation **27**(2), 1–22 (2021) https://meridian.allenpress.com/tscir/article-pdf/27/2/1/3069530/i1082-0744-27-2-1.pdf. https://doi.org/10.46292/sci2702-1

[5] Colak, A., Yilmaz, A., Sut, N., Memis, D., Ulucam, E., Cigali, B.S., Bozer, C.: Investigation of the availability of tongue movements in mallampati classification. Saudi Med J **32**(6), 607–611 (2011)

[6] Khan, M.M., Sherazi, H.I., Quain, R.: Tongue-supported human-computer interaction systems: a review. In: Proceedings of 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1410–1415 (2014). IEEE

[7] Augmental: Mouthpadˆ : A tongue-controlled interface that lets people explore the digital world hands-free. In: Augmental Press Release (2023). Augmental. https://www.augmental.tech/

[8] Kæseler, R.L., Andreasen Struijk, L.N.S., Jochumsen, M.: Detection and classification of tongue movements from single-trial EEG. In: 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 376–379. https://doi.org/10.1109/BIBE50027.2020.00068. ISSN: 2471-7819

[9] Wobbrock, J.O., Aung, H.H., Rothrock, B., Myers, B.A.: Maximizing the guessability of symbolic input. In: CHI'05 Extended Abstracts on Human

Factors in Computing Systems, pp. 1869–1872 (2005). ACM

[10] Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09, pp. 1083–1092. Association for Computing Machinery, New York, NY, USA (2009). https://doi.org/10.1145/1518701.1518866. https://doi.org/10.1145/1518701.1518866

[11] Berthellemy, M., Cayez, E., Ajem, M., Bailly, G., Malacria, S., Lecolinet, E.: Spotpad, locipad, chordpad and inoutpad: investigating gesture-based input on touchpad. In: Proceedings of the 27th Conference on l'Interaction Homme-Machine, p. 4 (2015). ACM

[12] Vatavu, R.-D., Wobbrock, J.O.: Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15, pp. 1325–1334. ACM, New York, NY, USA (2015). https://doi.org/10.1145/2702123.2702223. http://doi.acm.org/10.1145/2702123.2702223

[13] Liang, H.-N., Williams, C., Semegen, M., Stuerzlinger, W., Irani, P.: User-defined surface+ motion gestures for 3d manipulation of objects at a distance through a mobile device. In: Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction, pp. 299–308 (2012). ACM

[14] Morris, M.R., Wobbrock, J.O., Wilson, A.D.: Understanding users' preferences for surface gestures. In: Proceedings of Graphics Interface 2010, pp. 261–268 (2010). Canadian Information Processing Society

[15] Ruiz, J., Li, Y., Lank, E.: User-defined motion gestures for mobile interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 197–206 (2011). ACM

[16] Rodriguez, I.B., Marquardt, N.: Gesture elicitation study on how to opt-in & opt-out from interactions with public displays. In: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, pp. 32–41 (2017). ACM

[17] Magrofuoco, N., Pérez-Medina, J.L., Roselli, P., Vanderdonckt, J., Villarreal, S.: Eliciting contact-based and contactless gestures with radar-based sensors. IEEE Access **7**, 176982–176997 (2019). https://doi.org/10.1109/ACCESS.2019.2951349

[18] Pérez-Medina, J.L., Villarreal, S., Vanderdonckt, J.: A gesture elicitation study of nose-based gestures. Sensors **20**(24), 7118 (2020). https://doi.org/10.3390/s20247118

[19] Vo, D.-B., Lecolinet, E., Guiard, Y.: Belly gestures: body centric gestures on the abdomen. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, pp. 687–696 (2014). ACM

[20] Gemicioglu, T., Winters, R.M., Wang, Y.-T., Tashev, I.J.: Tongue gestures for hands-free interaction in head worn displays. In: Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers. UbiComp/ISWC '22 Adjunct, pp. 38–40. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3544793.3560363. https://doi.org/10.1145/3544793.3560363

[21] Chen, V., Xu, X., Li, R., Shi, Y., Patel, S., Wang, Y.: Understanding the design space of mouth microgestures. In: Designing Interactive Systems Conference 2021. DIS '21, pp. 1068–1081. Association for Computing Machinery. https://doi.org/10.1145/3461778.3462004. https://doi.org/10.1145/3461778.3462004

[22] Hiiemae, K.M., Palmer, J.B.: Tongue movements in feeding and speech. Critical Reviews in Oral Biology & Medicine **14**(6), 413–429 (2003) https://doi.org/10.1177/154411130301400604. https://doi.org/10.1177/154411130301400604. PMID: 14656897

[23] Yousefi, B., Huo, X., Kim, J., Veledar, E., Ghovanloo, M.: Quantitative and comparative assessment of learning in a tongue-operated computer input device—part ii: Navigation tasks. IEEE Transactions on Information Technology in Biomedicine **16**(4), 633–643 (2012). https://doi.org/10.1109/TITB.2012.2191793

[24] Lukaszewicz, K.: The ultrasound image of the tongue surface as input for man/machine interface. In: Rauterberg, M., Menozzi, M., Wesson, J. (eds.) IFIP TC13 International Conference on Human-Computer Interaction, pp. 825–828 (2003). IOS Press

[25] Ito, K., Takami, A., Hanibuchi, S., Nishida, S., Yagi, M., Uematsu, S., Sigenaga, N., Takada, K.: Clinical usefulness of human-computer interface for training targeted facial expression: Application to patients with cleft lip and/or palate. In: Salvendy, G., Smith, M.J. (eds.) Human Interface and the Management of Information. Information and Interaction, pp. 513–521. Springer, Berlin, Heidelberg (2009)

[26] Sahni, H., Bedri, A., Reyes, G., Thukral, P., Guo, Z., Starner, T., Ghovanloo, M.: The tongue and ear interface: A wearable system for silent speech recognition. In: Proceedings of the 2014 ACM International Symposium on Wearable Computers. ISWC '14, pp. 47–54. ACM, New York, NY, USA (2014). https://doi.org/10.1145/2634317.2634322.

http://doi.acm.org/10.1145/2634317.2634322

[27] Jain, M., Joshi, H.: Tongue operated wheelchair for physically disabled people. International Journal of Latest Trends in Engineering and Technology **4**(1), 587–595 (2016). https://doi.org/10.21172

[28] Nguyen, P., Bui, N., Nguyen, A., Truong, H., Suresh, A., Whitlock, M., Pham, D., Dinh, T., Vu, T.: Tyth-typing on your teeth: Tongue-teeth localization for human-computer interface. In: Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services. MobiSys '18, pp. 269–282. ACM, New York, NY, USA (2018). https://doi.org/10.1145/3210240.3210322. http://doi.acm.org/10.1145/3210240.3210322

[29] Chavan, A., Patil, B., Jadhav, S.: Wheelchair operated by tongue motion. International Journal for Research and Development **3**(8), 86–90 (2015)

[30] Abbas, U., Khan Turk, S., Zakir Shaikh, M., Kumar, H., Panhwer, A., Kumar, N.: Design and implementation of advanced wireless tongue drive/operated system for paralyzed, disabled & quadriplegic patients. International Journal of Advanced Research **4**(4), 587–595 (2016). https://doi.org/10.21474/IJAR01/258

[31] Mat Zin, S., Md Rasib, S.Z., Suhaimi, F.M., Mariatti, M.: The technology of tongue and hard palate contact detection: a review. BioMedical Engineering OnLine **20**(1), 17 (2021). https://doi.org/10.1186/s12938-021-00854-y

[32] Oguchi, H., Watanabe, T., Nakamura, N., Watanabe, S.: Influence of tongue movements on masticatory efficiency. Dent Oral Craniofac Res **2**(6) (2016). https://doi.org/10.15761/DOCR.1000187

[33] Li, Z., Robucci, R., Banerjee, N., Patel, C.: Tongue-n-cheek: non-contact tongue gesture recognition. In: Proceedings of the 14th International Conference on Information Processing in Sensor Networks. IPSN '15, pp. 95–105. Association for Computing Machinery. https://doi.org/10.1145/2737095.2737109. https://doi.org/10.1145/2737095.2737109

[34] Gemicioglu, T., Winters, R.M., Wang, Y.-T., Gable, T.M., Paradiso, A., Tashev, I.J.: Gaze & tongue: A subtle, hands-free interaction for head-worn devices. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. CHI EA '23. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3544549.3583930. https://doi.org/10.1145/3544549.3583930

[35] Vatavu, R.-D., Wobbrock, J.O.: Between-subjects elicitation studies: Formalization and tool support. In: Proceedings of the 2016 CHI Conference

on Human Factors in Computing Systems, pp. 3390–3402 (2016). ACM

[36] McNeill, D.: Gesture and Thought. Phoenix Poets Series. University of Chicago Press, Chicago, USA (2008). https://books.google.be/books?id=N0SmyU4TKRwC

[37] Brauer, J., Holt, T.V.: Tongue thrust classification. The Angle orthodontist **35**, 106–12 (1965)

[38] Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) Human-Computer Interaction – INTERACT 2013, pp. 282–299. Springer, Berlin, Heidelberg (2013)

[39] McNeill, D.: Hand and Mind: What Gestures Reveal About Thought. University of Chicago Press, Chicago, USA (1992). https://books.google.be/books?id=-C0KVfpUqKgC

[40] Jégo, J.-F., Paljic, A., Fuchs, P.: User-defined gestural interaction: A study on gesture memorization. In: 2013 IEEE Symposium on 3D User Interfaces (3DUI), pp. 7–10 (2013). https://doi.org/10.1109/3DUI.2013.6550189

[41] Lewis, J.R.: Psychometric evaluation of the pssuq using data from five years of usability studies. International Journal of Human–Computer Interaction **14**, 463–488 (2002)