

Towards Many Gestures to One Command: A User Study for Tabletops

Yosra Rekik¹, Laurent Grisoni^{1,2}, Nicolas Roussel¹

¹ INRIA Lille Nord Europe, France,

² LIFL, University of Lille 1, CNRS, France

yosra.rekik@inria.fr, laurent.grisoni@univ-lille1.fr,
nicolas.roussel@inria.fr

Abstract. Multi-touch gestures are often thought by application designers for a one-to-one mapping between gestures and commands, which does not take into account the high variability of user gestures for actions in the physical world; it can also be a limitation that leads to very simplistic interaction choices. Our motivation is to make a step toward many-to-one mappings between user gestures and commands, by understanding user gestures variability for multi-touch systems; for doing so, we set up a user study in which we target symbolic gestures on tabletops. From a first phase study we provide qualitative analysis of user gesture variability; we derive this analysis into a taxonomy of user gestures, that is discussed and compared to other existing taxonomies. We introduce the notion of *atomic* movement; such elementary atomic movements may be combined throughout time (either sequentially or in parallel), to structure user gesture. A second phase study is then performed with specific class of gesture-drawn symbols; from this phase, and according to the provided taxonomy, we evaluate user gesture variability with a fine grain quantitative analysis. Our findings indicate that users equally use one or two hands, also that more than half of gestures are achieved using parallel or sequential combination of atomic movements. We also show how user gestures distribute over different movement categories, and correlate to the number of fingers and hands engaged in interaction. Finally, we discuss implications of this work to interaction design, practical consequences on gesture recognition, and potential applications.

Keywords: Tabletop; multi-touch gesture; gesture recognition; interaction design.

1 Introduction

Tabletops have become very good candidate systems for interactive setups. Users in such systems tend to use different class of gestures (e.g. [12]), which implies that any application shall potentially integrate ways to handle such variety. In the meantime, multi-touch gestures are often thought by application designers for a one-to-one

mapping between gestures and commands, which does not take into account the high variability of user gestures; it can also be a design choice that leads to simplistic interaction choices. There is, to our knowledge, no solution yet to this issue, nor even any path of research drawn yet to solve it. Our motivation is to make a step toward many-to-one mappings between user gestures and commands, by understanding user gestures variability for multi-touch systems. Most studies are task oriented, and allow to exhibit best matches between gesture type, and elementary tasks or commands. Some results show little agreement among users in mapping between gestures and their effect [40]. In order to provide application designers with knowledge that will help designing good many-gestures-to-one-command mappings, we need another experimental approach, which is to exhibit, for a specific type of gesture, all the possible gestural representations that users may achieve. Instead of studying relation between gestures and tasks, we propose to study relation between gesture and underlying symbolic pattern. We advocate, following [1], for the need to construct designed interaction languages, and we provide, in this article, elements that describes how gesture variability may be integrated in such interaction design. We think that a good question, as a start, is: *how do people draw symbolic gestures using their hands on tabletops, and how do they express variability for the same symbol?*

For answering this, we set up a user study in which we target symbolic gestures on tabletops. We have designed our experiment as a two-fold work (see Section 3). From a first phase study, we provide a taxonomy of user gestures; this taxonomy integrates three different views of user gesture (semantic, physicality, *movement structure*); we discuss this in regard to other existing taxonomies, and exhibit its quality. We introduce, through this taxonomy, the notion of *atomic* movement, that may be combined throughout time (either sequentially or in parallel), and can be used to structure user gesture. In a second phase, a user study is conducted with specific class of gesture-drawn symbols; from this study, and according to the provided taxonomy, we evaluate user gesture variability with a fine grain quantitative analysis. Our findings indicate that user equally use one or two hands, also that more than half of gestures are achieved using parallel or sequential combination of atomic movements. We also show how user gestures distribute over different movement categories, and correlate to the number of fingers and hands engaged in interaction. Finally, we discuss implications of this work to interaction design, practical consequences on gesture recognition, and potential applications.

2 Related Work

In the process of developing surface computing technologies, several studies can be found in an attempt to grasp and unify the rich vocabulary of gestural multi-touch interaction, e.g., [28, 22, 16]. We structure the description of related work into three parts: user-centered studies of interaction gestures for tabletops, taxonomy propositions, and formalisms for multi-touch gestures.

A lot of work has been done these past years on user-defined interaction for multi-touch. Given the versatility of free-hand multi-touch gestures and the high variety of users behaviors in producing them [26], user-centric approaches have been at the

heart of many research studies on gestures for multi-touch systems. Rather than bounding users to an arbitrary set of gestures defined by system designers (e.g., [25, 31, 37]), Wobbrock et al [40], followed by others (e.g., [29, 5]), adopt a guessability methodology [39] to build up a user-defined gestures set for classical control actions and object manipulations. Finding its fundamentals in [30, 33, 6], this approach consists in presenting the effect of an action to users and then asking them to invent the corresponding gesture. The gesture which was consistently performed by the largest number of users is then retained to be representative of the corresponding action. Within these studies, valuable discussions are reported about the different characteristics of user-defined gestures, e.g., number of fingers, hand pose, etc. However, the focus is on the design and analysis of a one-to-one mapping between gestures and their actions. In [11], Henze et al suggest to derive and compare multiple gesture sets rather than a single one. Their findings indicate that this is a beneficial approach to reduce the risk to exclude promising candidates for gestures. In a field study investigating the variety of gesture performed by people, Hinrichs et al [12] found that users choice of gestures was influenced by the interaction context in which the current action occurred and not only based on preferences for a given gesture for a particular action. They suggest that a many-to-one mapping is also desirable to strengthen the design of gestural interaction techniques. While being specific to pen gestures, Long et al [22,23] studied perceived gesture similarities. Since evaluation of similarities is a complementary problem to the one addressed in this article, such a work can help explaining differences between gestures classes.

Several works also target taxonomy of multi-touch gestures. Wobbrock et al. [40] are among the first to establish a unified taxonomy for surface gestures. They provide a coarse-grain classification along four categories: form, nature, binding, and flow. While these categories are all important, the main focus of our work falls within the *form* category, which captures *how* gestures are performed by users, and its relationship to the *nature* category, which captures the users *semantic* interpretation of gestures. In Wobbrock's taxonomy, the nature category distinguishes symbolic, physical, metaphorical and abstract gestures. The form category distinguishes static or dynamic pose and path for each hand. In the same spirit, Wu et al. [41] describe the process of gesture performance as a finite state machine, with start position (registration), a dynamic phase (continuation), and end position (termination), similar in concept to that described in Charade [3]. Freeman et al [7], in the context of the design of a gesture-learning tool, expanded the form category along three dimensions: registration pose, continuation pose, and movement. Although that taxonomy of Freeman et al presents a sound picture of the large variety of multi-touch gestures, it is not adapted to model variability of association between gestures and symbols.

Remarkable recent researches are being conducted on the formal specification and the reliable recognition of multi-touch gestures. GeForMT [15] provides a formal abstraction of multitouch gestures using a context-free grammar. A discussion of that formalism is given in respect to Wobbrock et al. [40] taxonomy in an attempt to show how it can capture users gestures. Gesture Coder [24] recognizes multitouch gestures via state machines. Proton [19, 18] describes multi-touch gestures as regular expressions modeling a whole sequence of touch events. GestIT [35] is a proof of concept

library implementing a meta-model based on compositional operators and Petri Nets to describe multi-touch gestures. All these software-oriented frameworks and languages provide system sound specifications allowing to express complex multi-touch gestures. Nevertheless, it is not obvious how they can apply to capture in a comprehensive and faithful manner the behavior and variability of non technical users in producing gestures. The implications for multi-touch interaction systems to support the variety of users choices in a transparent manner open in fact new opportunities but raises many challenges. For example, recent studies have tackled the difficult issue of designing robust multi-touch recognizers abstracting away the use of multiple fingers [14], or multiple differently ordered strokes [2, 20, 38] for the same symbol. Interestingly, many of those recognizers were designed without prior in-depth analysis of users behaviors thus only taking into account a limited designer-oriented vision of users choices. In this context, a better understanding of users choices is relevant to reduce the dualism of how and who must be adapted to the other: the user or the system [8]. As discussed later in this paper, our findings enlighten the features that deserve deeper modeling efforts for future system-oriented gesture formalizations.

3 User Study

3.1 Overview and Rationale

Eliciting user behaviors have been proved extremely useful to help the design of new, strong and flexible interaction tools (e.g., [33, 40, 5, 9]). Our goal is to broaden the range of possible responses we can get from users and to gain insights into users' variability when issuing a multi-touch command. More generally, we want to elicit the different ways users perceive and issue a multi-touch gesture with the ultimate goal of determining the rules leading to a better definition of what shall be a gestural language for multi-touch interaction. For that purpose, we ask participants to appeal to their imagination to perform different gestures, at the aim of grasping and analyzing the variability and dynamic of user behavior.

As a result, our proposed study is divided in two phases: a first phase, that has two goals: familiarize participants with the interactive surface, and more importantly, observe and analyze their 'intuitive' interaction styles, within an uncontrolled experimental procedure where user can both choose symbol and gesture that represents it. This phase is intended to construct a taxonomy that is used as a basis for the remainder of the study. In the second phase, we achieve quantitative analysis of how users draw symbols, using explicit instructions (name of the symbol, number of variations) and asking participants to explore the different ways to achieve specific symbols. The detailed experimental procedure and context is described in the following.

3.2 Participants

A call for participation has been made using mailing lists and the advertising lobby screens available at our lab and its institutional partners. The call targeted people who were not user-interface designers. In final, we collected results from 30 volunteers, among them 14 were female. Age of participants ranged from 20 to 57 years (average

age was 28.4 years). All participants were right-handed. Participant occupations included secretary, chemists, biologists, electronic and mechanics experts, researcher in networks and telecommunications and graduate students. Participant nationalities include different European, African and Asian countries. Self-reported expertise of participants with touchscreen devices were found to significantly differ in the type of interactive surface ($\chi^2 = 62.27$, $p < 10^{-11}$, $\phi = 0.72$). Fig. 1 summarizes this; we can see that none of our participants previously used a Tablet Pc nor Tabletops. As a result they are completely novice to the interactive surface used in our experiment.

Table. 1. Distribution of usage of touchscreen devices among our participants.

	Smart Phone	Tablet	Tablet PC	Tabletops and surfaces
Never	8	18	30	0
Occasional	9	9	0	0
Regular	13	3	0	0

3.3 Apparatus

The study took place in our lab, where we had set up a Microsoft Surface 1 measuring 24"x18". Only the experimenter (one author) and the subject were present during the study. During each experiment session, participants' hands were videotaped and the experimenter observed user behavior and took detailed notes. Author notes and videos from each participant were then used in our analysis.

3.4 Procedure

As sketched previously, our study consisted of two phases:

•**First phase:** The interactive surface is presented to the participant and he/she is explained that the surface accepts multiple fingers. The participant is then asked to perform any multi-touch gesture that comes to his mind and meaningful to him/her. No further comment or request is made that may suggest degrees of freedom, such as number of hands to use, number of fingers, etc. For each performed gesture, users are asked to describe it in a think-aloud protocol. Users have 3 minutes to represent all the gestures they can think of; they are also free to stop before 3 minutes in case they consider the task to be over.

•**Second phase:** we explicitly provide the participants with the following sequence of 8 symbolic forms (corresponding to a subset of Microsoft Application gestures [27]): circle, square, triangle, vertical line, horizontal line, corner, V and Caret. For each symbol, the participant is asked to perform the symbol using four different manners. Participants are only told the name of the symbols to perform, by oral instruction, and we do not show them any image of the required symbols.

Both phases are conducted consecutively for each participant with a small pause in between. Participants were not constrained by any timing issues when performing their gestures. To prevent any screen content from influencing the gestures participants were performing, we provided no visual feedback of gesture input, e.g., [40,14].

4 Results From the First Phase

4.1 General observations

In the first phase of our study, we collected 618 user-made gestures. The number of gestures per participant ranged from 8 to a maximum of 46 gestures. Although collected gestures were having broad properties, similar features were observed among different participants. Without surprise participants produced different forms using interchangeably one or more fingers and one or two hands, to draw different kinds of symbols (e.g., line, circle, square, triangle, etc), alphanumeric characters (i.e, letters and numbers), shaped (e.g., tree, heart, flower, star, bird). 6 participants, being regular users of iPads or Smart Phones, additionally performed gestures mimicking standard control actions such as double tap, or rotational, translational and scaling patterns. From our collected data, we were also able to extract several observations about the physical engagement of participants. In particular, 26 (resp. 27) over the 30 participants have used at least once a single (reps. two) hand(s). 24 participants moved simultaneously both hands in symmetric poses. 4 participants alternated from one hand to another in a sequential style. 6 participants used one hand to perform a gesture while their second hand was hold in a stationary pose as to draw a static reference guiding the other hand. This observation holds for gestures performed with a single hand using the thumb and the index. 1 participant used to move hands in the air and touching the surface with her fingers from time to time. 1 participant used exclusively static hand posture on the surface. Except for these two cases, the relative movement of participants fingers was the rule guiding the achievement of gestures. Neither the number of fingers nor their type seemed to us as a conscious parameter that participants were intentionally thinking about. We did also notice no particular preference on the start and the end positions of performed movements. Participants mostly used their right hands when moving from left to right, and inversely they used their left hands when moving from right to left. However, we did not notice other apparent rules applying to the direction of movements nor to the size of their trajectories.

4.2 Qualitative analysis: user gesture as atomic movements combined over parallelism and sequentiality

From the gestures collected during the first phase, we extracted several observations that provide elements about how users perform gestures.

- Atomic movements: The very recurrent observation in participants' behavior is that they grouped their fingers into unitary blocks moving in a consistent manner, while being completely free from the microscopic timeless notion of touch as may be handled by the system. We found that number of contact fingers does not impact the accomplishment of their movements, as long as involved fingers are close to each others. The interesting observation is that the notion of proximity is relative to user-proper referential and seems to be hardly definable in absolute and universal manner from a system point-of-view. Users referential can in fact be substantially scaled up or down from the performance of one gesture to another one. However, it

tends to stay constant and consistent over time and through possibly multiple movements composing the same single gesture. From this observation, we introduce the notion of *atomic movement* which reflects users' perception of the *undividable* role that a group of fingers is playing when performing a gesture. From our observations, users atomic movements are mostly in reference with the imaginary trail of a group of fingers which position is evolving in closely related movements. An atomic movement can have an internal state that can change depending on hands shape, fingers arity, velocity, direction, etc. However, state changes do not alter the role an atomic movement is playing in users' mind and its primary intention. From our observations, we distinguish between two categories in participants movements depending on whether (i) the trail corresponding to fingers is stationary or (ii) it implies an embodied motion. As a practical examples, variable number of fingers, from one or two hands, moving together following the same path or being held stationary to delimit or point a region in the interactive surface, are among the most frequently observed atomic movements.

- Parallelism:** For some gestures, participants combined the movements of their fingers simultaneously, in either a symmetric or an asymmetric style. From a geometrical perspective, symmetry occurs mostly between two atomic movements performed in parallel on the surface such as the trajectory of the one was the mirrored image of the other. For instance, users are observed to produce circled pattern by moving fingers from both hands in parallel such as the trajectory of each hand forms a semi-circle. On the other side, asymmetry in gestures occurs when participants were holding some fingers stationary upon the surface and simultaneously moving some others. For instance, users are observed to produce circled patterns by touching a region of the surface with one hand and simultaneously moving fingers from the other hand all around. From a physicality perspective, bi-handed parallel movements are mostly attended with the use of the same fingers combination on each hand, while the use of one hand mostly engages the use of the index and the thumb. From these observations, we found that symmetry in users gestures can be described by the parallelism expressed by atomic movements.

- Sequentiality:** We observed that some participants often operate in a sequential manner by iteratively posing and moving fingers on the surface, then releasing and posing fingers again at a new location binding the set of already performed movements. Users sequential movements imply more than a time pause or direction change. They are performed using one hand, as well as alternating different hands or fingers, and mixing parallel atomic movements with single atomic movements. In this class of interaction style, movements are mixed and matched both in time and in space according to users specific referential. This referential does not map perfectly with the system. For example, the boundaries of strokes induced by the atomic multi-finger movements are never perfectly matching with one another, though we think that participants intended to do so in their minds.

From this analysis, user gestures can be modeled using atomic movements, possibly combined along with parallelism and sequentiality. Fig. 1 provides a simple situation illustrating this with three gestures produced by different participants.

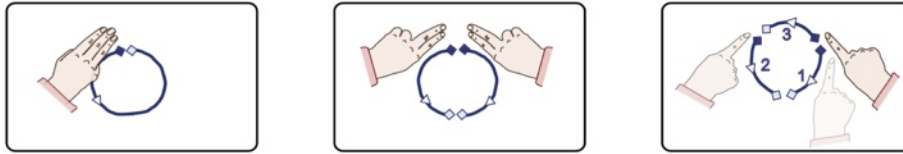


Fig. 1. Example of Atomic movements combination through time. From left to right: Elementary atomic movement, parallel movement, sequential movement.

4.3 An embodied taxonomy of multi-touch gesture

To capture the space in which our participants were conceiving and producing gestures, we propose the multi-level layered taxonomy summarized in Fig. 2. It is worth noticing that the levels of our taxonomy do not model separable attributes to be characterized individually. Instead, they represent the different aspects of a single unified dynamic mechanism ruling users in the achievement of a multi-touch gesture.

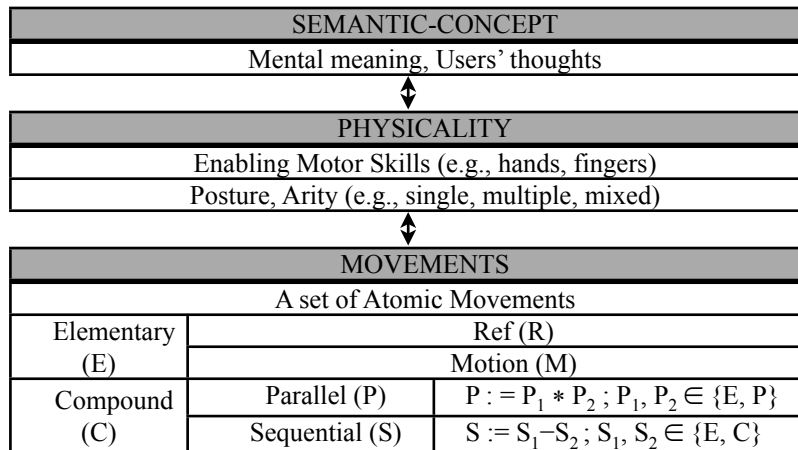


Fig. 2: A multi-level model for users' gestures. The '*' (resp. '-') refers to movements performed simultaneously (resp. sequentially) in time.

At the high level of our taxonomy, we capture the fact that a multi-touch gesture emerges from what user's mind is modeling before even touching the surface. In this respect, an external observer can only try to guess the *semantic concept* hidden in user's gesture, since it might be the case that the gesture it-self is not sufficient to fully reveal user's thought — which is in accordance with previous studies [36,13,40]. From a neurological perspective, hands and fingers are controlled and coordinated by human motor system at the aim of achieving a desired task. The *physicality* level thus captures the motor control allowing users to project the semantic level into the inter-

active surface. The *movement* level is the consequence of the motor goal expressed by hands and fingers motions in order to infer unitary blocks building the whole gesture.

The movement level is at the core of our taxonomy since it constitutes the interface between the user and the interactive surface/system. Consistent with our observations, we propose to structure this level according to two generic classes built in a recursive manner. At the low level of the recursion, we find the class of gestures formed with an elementary atomic movement. An elementary atomic movement can be either of type stationary (Ref) or Motion as discussed previously in our qualitative observations. The *Compound* class refers to the recursive composition of a set of atomic movements. It is expanded in two categories depending on the time combination of composing atomic movements. The *Parallel* category refers to users making two or more different but simultaneous movements. The *Sequential* category refers to users performing a set of atomic movements, being possibly parallel or elementary, holding and releasing hands or fingers, on and from the surface, in a discrete iterative manner.

4.4 Users variability

Our taxonomy is the result of a qualitative empirical synthesis of a wide range of collected gestures. We found that the three levels of our taxonomy contribute leveraging and unifying the high variety of users gestures. In fact, users gestural variations can be elicited as the result of the mental picture and the time-space composition of atomic movements, as well as their physical mapping into users fingers and hands.

At the semantic level, the global pattern induced by movements is the most apparent attribute that users where instantiating in several different manners. However, gestures with similar global patterns can have different properties, e.g., their composing atomic movements can be in different classes. At the physical level, variations in the number of fingers and hands are a natural outcome for most participants. Finally, users variations can be captured at the movement level by eliciting the different possible time combinations of atomic movements (Motion and Ref) as well as their number which can vary from a gesture to another and from a user to another.

4.5 Comparison with existing taxonomies

Comparing to previous taxonomies, the semantic concept of our taxonomy relates to the nature category defined by Wobbrock et al [40]. In that study, users were shown the effect of a gesture, then they was asked to issue the gesture. Hence, the nature category is tightly related to the action of the gesture. In our first-phase study, we did not ask participant to perform any precise action. Thus, the semantic concept level only reveals the meaning of the gesture without mapping it to the type of a particular action. On the other hand, physicality in our taxonomy relates to the form category sketched in [40] and expanded by Freeman et al [7]. The registration, continuation and movement dimensions described within the form category there-in did not directly result from a specific user-centric study, since the intention of Freeman et al work was primary focusing on teaching the user how to perform a gesture. Although, those dimensions provide a sound picture of *how* users may perform a gesture, we find that the physicality and movement levels of our user-centric taxonomy comple-

ments and refines in many aspects the empirical work of Freeman et al. For example, Freeman’s distinguishes between two types of movements: *path* and *no path*, depending on whether the hand moves along a surface path or not. In our work, we explicitly distinguish between how users perform gestures (Physicality) and the notion of movements. In this respect, the Movements level introduces a new dimension in users gestures and consistently renders the embodiment of gestural multi-touch interaction. The semantic concept behind users gestures can then be captured within an embodied and coherent flow engaging the cooperation of users fingers and hands to materialize the inter-relation between a set of unitary atomic blocks composing the gesture.

5 Results from the Second Phase

In our second-phase experiment, participants were asked to produce symbols in four different manners. We were able to adequately classify all the gestures from this experiment phase using our taxonomy. In addition to that, no pair of gestures associated to the same symbol, and identified as different by one user, falls into the same category according to our taxonomy. As an example, appendix shows the set of gestures that we noticed in the experiment in the case of the circle symbol.

In this section, we discuss taxonomic breakdowns of the variety of gestures proposed by participants. 2 participants were excluded from the provided statistics since they made less than the four gestures required per symbol in our experiment. 5 participants produced more than four different gestures for some symbols. We constraint our analysis to only the four first ones. Overall, we have retained $28 \times 8 \times 4 = 896$ gestures that are analyzed in the following.

5.1 Physicality/Movement inter-dependency

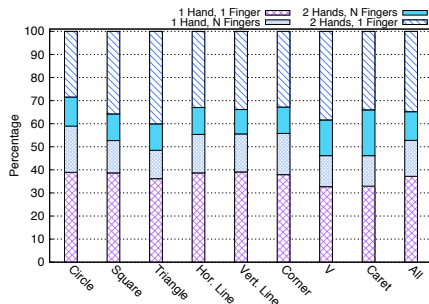


Fig. 3: Hands and Fingers (per hand) ratio.

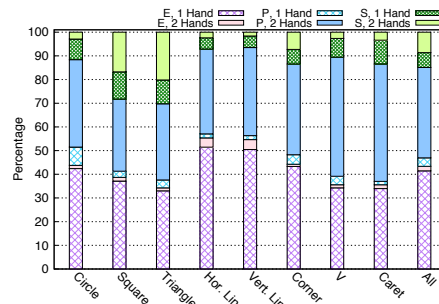


Fig 4: Movement and Hands ratio (E: Elementary, P: Parallel, S: Sequential).

Fig. 3 shows the ratio (averaged over all users) of gestures performed with one and two hands for each symbol and overall. We also incorporate the amount of fingers (single or multiple) engaged per each single hand. If the gesture is movement-compound, we count it multi-finger if at least one hand was engaged with more than one finger. We can see that users-gestures are fairly distributed over one hand (52.77%) and two hands (47.22%). Although, participants used more often a single finger per hand (78.17%), a significant ratio of gestures where multiple fingers are

used per hand can still be reported (21.83%). A Friedman test revealed that Symbols' type does not have a significant effect on the ratio of two handed gestures performed by users.

Fig. 4 shows gestures ratios classified by movement categories, where we further distinguish between one-handed and two-handed gestures. A Friedman test revealed a significant effect of Symbols' type on the ratio of movement categories (Elementary: $\chi^2 = 32.55$, $p < 3.10^{-5}$; Parallel: $\chi^2 = 21.98$, $p < 2.10^{-3}$; Sequential: $\chi^2 = 50.12$, $p < 1.10^{-8}$). In the elementary atomic category, a post-hoc test using Wilcoxon test showed the significant differences of the couple of symbols (vertical line, horizontal line) and the other symbols. We attribute this to the fact that this couple of symbols does not imply direction change in fingers movements so that elementary atomic movements are the more natural to conceive for users. In the Parallel category, significant differences were found between the couple of symbols (V, Caret) and the other symbols. Actually, the ratio of parallel two-handed gestures performed for these two symbols is higher compared to the other symbols. This can be explained by the fact that these symbols can be more easily mapped into users two hands. In the Sequential category, significant differences were found between the couple of symbols (square, triangle) and the other symbols. These two symbols are in fact clearly different from the others by the number of stroke combinations that can be used to perform them. Overall, we can see that users produced atomic and compound parallel gestures in approximately the same proportion (resp. 43.34% and 41.76%), while the compound sequential category is represented in a relatively non-negligible ratio of 14.89%. A Chi-square test with Yates' continuity correction revealed that the percentage of two-hand and one-hand gestures significantly differed by movement category ($\chi^2 = 523.34$, $p < 2.10^{-16}$, $\phi = 0.78$). We can in fact remark the high correlation between two-handed gestures (resp. one-handed) and the parallel movement category (resp. elementary).

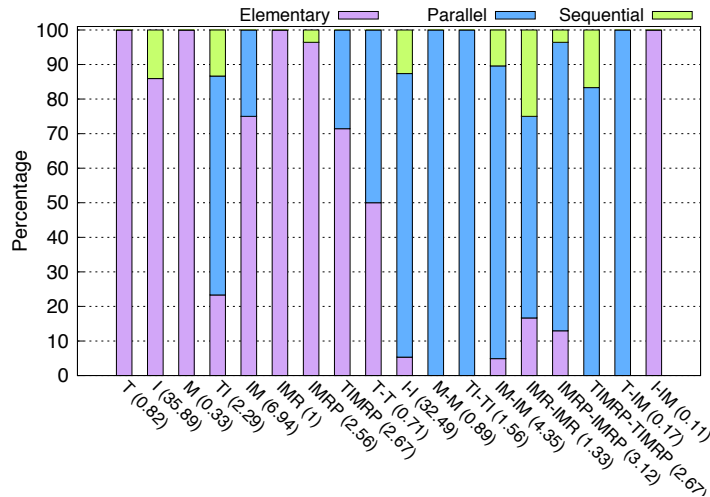


Fig. 5: Movement category ratio according to Hands-Fingers Combinations. T, I, M, R and P, denote resp. the thumb, index, middle, ring, and pink. The sign '-' distinguishes between left and right hand. Numbers in braces refer to the ratio of the finger combination over all users.

Finally, Fig. 5 shows the different combinations of hands fingers and their mapping to the movement categories. Overall we observed 18 possible one-handed and two-handed finger combinations. The parallel movement category is represented in 12 combinations among them only the index-thumb and index-middle is one handed. The other parallel two-handed combinations show a high similarity in the type of fingers used per hand. We can notice the absence of gestures engaging the pink or the ring in an elementary atomic movement. Whenever these two fingers are used, they appear in combination with the middle, index, and/or thumb fingers, by inducing the same multi-finger atomic movement. These combinations reflect the natural (comfortable) motor capabilities of users as well as the affordance of hand movements and their dependencies – which is consistent with previous studies on the mechanical/neurological relationship between fingers and their kinematics, e.g., [15, 27, 41].

5.2 Users' Transition-Frequency Automaton

Gestures' properties were not random over the four trials allowed per symbol. In order to capture users thoughts and priorities in conceiving the different manners of producing a gesture, we study in this section the evolution of gestures properties over time using probabilistic automaton [34], mapping gestures properties into states and users variations into transitions. Fig. 6 shows four such automata in a comprehensive informal manner. The first three automata are user-centric while the fourth one provides a more system-centric perspective as it will be discussed in the following.

Every initial state of the automata depicted in Fig. 6 refers to participants starting the experiment. Columns refer to subsequent gestures produced by participants. The rows of the first automaton classify gestures according to whether they are one-handed or two-handed. Those in the second automaton classify gestures according to the movement category. The third automaton distinguishes between gestures where every composing atomic movement is single-finger and those where at least one atomic movement is multi-finger. Finally, the fourth automaton classifies gestures depending on whether exactly one touch is involved throughout the whole gesture, or multiple touches are involved. The main difference with the third automaton is that touches are viewed relative to the system and not to users. The numbers in each cell is then computed as the average ratio over all users of gestures found in the corresponding state. This provides gestures distribution over time and can be interpreted as the empirical probability of user's gesture property being mapped the corresponding state. Similarly, transitions depicted by labeled rows show the average ratio of participants moving from a state to another, which can be interpreted as the empirical conditional probability of falling in the subsequent gesture type knowing the type of the present gesture. For example, the initial state of the first automaton reads as 0.74 of participants perform the first gesture with one hand, or alternatively as users perform a two-handed gesture first with probability 0.26. Given that a user performs the first gesture with one hand, there is a probability of 0.45 that the outcome of the second gesture is two-handed. Notice that the cells in each column sum to one, which provides the empirical probability distribution (and thus the average ratio) of corresponding gesture types. These automata are actually averaged over users and symbols so that they only represent the behavior of an 'average' user over all symbols. We did make a more fine

symbol-dependent analysis, not shown here, which revealed that different states for the four gesture trials are observed within every symbol; however the probability transition is different from a symbol to another and from a user to another.

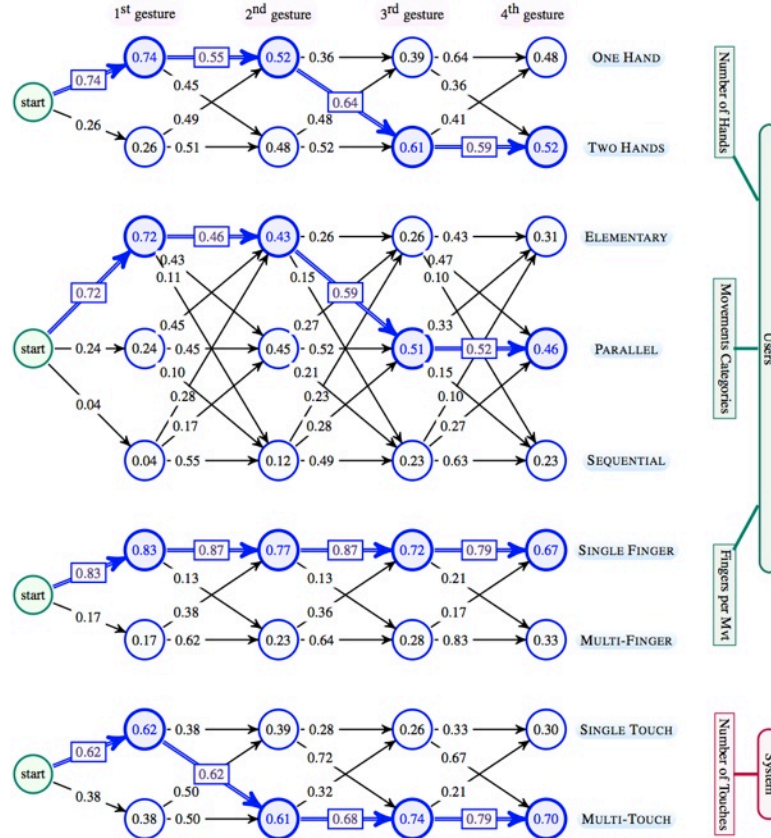


Fig. 6. Gestures Transition-Frequency Automaton (average over all symbols and users) according to (from top to bottom): Number of Hands, Movement categories, Number of fingers per atomic movement, and Number of contacts.

Fig. 6 provides a time-dependent information about gesture type frequencies and variations. For example, as one can see in bold lines referring to the most likely gesture transitions, users start more likely with one-handed, elementary-atomic-movement, single-fingered, single-touch gestures. Then-after, they are more balanced in their choices consistently switching to two-handed and parallel movements. The empirical probability that users start with two-handed or parallel-movement gestures is relatively significant (0.26 and 0.24) and grows sharply as users advanced in the experiment. This is to contrast with fingers usage since a single finger per movement is most often used all along the produced gestures. We interpret this as bi-manual usage and movement variations being the most significant features ruling users' mind in performing the different set of gestures. Although the sequential-movement strategy is unlikely as a starting strategy (0.04), it is interesting to remark that users falling in this state are more likely to produce the same type of movements in subsequent ges-

tures. We can interpret this as the sequential mode offering more degree of freedom in producing different gestures by consistently playing with hands movement combination. Finally, we remark that from the system perspective, gestures involving multiple touches on the surface are significantly represented all along the experiment mostly because users are either engaging their two hands or performing parallel movements.

6 Implications for gestures-based application design

In this section, we discuss the implications of our results for gesture design, surface technology, and user interfaces. In particular, we address points that seem important in order to design application that authorize several gestures for one command.

6.1 Movement matters more than posture

Our user study demonstrates that the movement induced by fingers motion matters for participants more than hand posture. Over all our participants, only two participants performed static gestures where fingers or hands were maintained stationary (Ref). Only in this case, the motor skills (blob type, posture, arity) used to structure the gesture are important, while their movements are not. Static gestures (set of Ref atomic movements), where hands/fingers posture is crucial, go beyond available classical multi-touch surfaces and need further sensing and input processing technology. In contrast, gestures where movements on the surface are crucial, are more accurate to the available knowledge and expertise on processing multi-touch input. Therefore it is our opinion that movement-based gestures provide more space to fully take advantage of nowadays multi-touch technology, so that their study and understanding should be the priority in the short term. However, current trends in augmenting surface computing technologies with new sensing facilities are also compatible with our user study. Advances in these directions would allow to enrich multi-touch surface input vocabulary so that gestures embodiment and versatility can be better encapsulated and exploited within gesture interfaces.

6.2 Interaction gesture: a multi-level, multi-view phenomenon

Our study reveals that the variations of users gestures for the same command can be structured and classified by the specific properties of a set of atomic movements. Although the notion of atomic movement and the role it plays in our gesture taxonomy constitutes a low level abstraction of what users them-selves are modeling, it should as well serve for designers as a basic tool in the process of thinking, formalizing and setting up multi-touch gestural interaction techniques. Designing for multi-touch gestures as multi-movement entities embodied in users thoughts would then push a step towards shortening the gap between designers vision and the way multi-touch gestures are perceived and produced by end-users. In particular, an atomic movement is by definition not sensitive to the number of fingers or the number of hands being used, so that it enables to unify and to leverage previous studies recommending to not distinguish gestures by number of fingers, e.g. [40]. Thinking about multi-movement multi-touch gestures, one have to keep in mind that the interdependency between the set of atomic movements forming a multi-touch gesture highly

depends on users motor control over time and over space. Two main alternatives can be elicited depending on whether one hand or two hands are considered. In both cases, it is more likely that the movements occur in parallel that is simultaneously in time. In addition, users are more likely to engage a single finger in the performance of one elementary atomic movement, though this should not serve as *the* rule.

Besides allowing to provide guidelines for the design of multi-touch gestures, the atomic movement perspective allows to expand in a comprehensive, yet precise, manner the space of possible mappings between a command and users gestures. By investigating the different possible combinations at the movement level, a variety of single and multi-finger, single and two-hand gestures can be supported, which can: (i) improve flexibility, (ii) not penalize users by offering adequate response, and (iii) make sure that the variety of users choices leads to a gratifying interactive experience.

6.3 Gesture recognition needs to be deeply rethought

From the aspect of system feasibility, our study raises new challenges for the generic encoding and the reliable recognition of multi-movement multi-touch gestures. In fact, a formal and rigorous system-computable definition to what is an atomic movement is first needed. From the quality of such a definition depends the design of system embedded programs that determines a faithful representation of users' atomic movements and enable a consistent processing and interpretation of users gestures. One promising research path is to augment existing multi-touch frameworks based on formal grammars (such as *proton++* [18] and others [24, 15, 35]) with both (i) declarative language elements that capture the notion of touch closeness in an elementary atomic movement, as well as with (ii) new compositional operators that render the time and space relation of atomic movements. The goal would be to automatically encompass users mental model features like: the independence of movements from the number of fingers, the possible variations in the combination of movements etc, within such formal frameworks. For patterned shape gestures, the challenge is more on the extraction of the different strokes implied by users movements. For example, it is not clear how recognizers in the \mathcal{S} -family [38] can handle the fact that a stroke could be constructed by users using a variable number of fingers. Recently, Jiang and al. [14] proposed an algorithm to extract a single stroke from the different trajectories of multiple fingers on the surface. However, this reduction is incompatible with the fact that multiple strokes can interleave in time, e.g., drawing a circle or a square or triangle or V or caret using two symmetric parallel atomic movements will be recognized as a line. We argue that the state-of-the-art recognizers for multi-touch gestures have to be rethought to support usability and consistently take into account the variety of users gestures in issuing a command. One path can be to take advantage from the consistency of the notion of touch closeness with respect to every user global time-space referential when performing atomic movements.

7 Conclusion and Future Works

In this paper, we investigated the different gestures that users adopt to issue the same symbol. We provided direct implications of our findings for the design of table-

tops gesture-based applications. For reusability purposes, it is important to elicit the 'natural' alternatives available for users to perform a symbolic command. In future work, it would be interesting to investigate all the potential uses of such a variability integration, and its practical impact on gestural interaction techniques. Besides, it can be interesting to explicitly ask users to perform gestures in different classes and to evaluate their preferences and ranking of each class. Taking variability into account within gesture recognition, and integrating it into available gesture recognizer for tabletops and multi-touch systems is also a challenging issue which still has to be addressed in the future. A more general issue related to our user-study is to investigate to what extent our findings on user's variability can be applied to other type of gesture detection devices which do not require a contact surface, e.g., Kinect.

Acknowledgments

This work is partially funded by Interreg IV-A 2-seas SHIVA project.

References

1. Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S., Tzu Kwan Valino Koh, J.: Understanding Mid-Air Hand Gestures: A study of Human Preferences in Usage of Gesture Types for HCI, Microsoft Research TechReport MSR-TR-2012-111
2. Anthony, L., Wobbrock, J. O.: \$n\$-protractor: a fast and accurate multistroke recognizer. *GI*, pp. 117--120. Canadian Information Processing Society, (2012)
3. Baudel, T., Beaudouin-Lafon, M.: Charade: remote control of objects using free-hand gestures. *Commun*, 36(7):28--35. ACM (1993)
4. Bevans, A.: Investigating the effects of bimanual multitouch interaction on creativity. *Proc. of Creativity & Cognition*, pp. 451--452. ACM (2011)
5. Cohe, A., Hachet, M.: Understanding user gestures for manipulating 3d objects from touchscreen inputs. *GI*, pp. 157--164. Canadian Information Processing Society (2012)
6. Epps, J., Lichman, S., Wu, M.: A study of hand shape use in tabletop gesture interaction. *CHI EA*, pp. 748--753. ACM (2006)
7. Freeman, D., Benko, H., Morris, M. R., Wigdor, D.: Shadowguides: visualizations for in-situ learning of multi-touch and whole-hand gestures. *ITS*, pp. 165--172. ACM (2009)
8. Good, M. D., Whiteside, J. A., Wixon, D. R., Jones, S. J.: Building a user-derived interface. *Comm. of the ACM*, pp.1032--1043. ACM (1984)
9. Guilford, S. J.: *The nature of human intelligence*. McGraw-Hill series in psychology.
10. Hager-Ross, C., Schieber, M. H.: Quantifying the independence of human finger movements: Comparisons of digits, hands, and movement frequencies. *The Journal of neuroscience*'00, pp. 8542--8550 (1967)
11. Henze, N., Löcken, A., Boll, S., Hesselmann, T., Pielot, M.: Free-hand gestures for music playback: deriving gestures with a user-centred process. *MUM '10*, pp. 16:1--16:10, ACM.
12. Hinrichs, U., Carpendale, S.: Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. *CHI*, pp. 3023--3032. ACM (2011)
13. Huang, T. S., Pavlovic, V. I.: Hand gesture modeling, analysis, and synthesis. *IEEE International Workshop on Automatic Face and Gesture Recognition*, pp. 73--79. (1995)
14. Jiang, Y., Tian, F., Zhang, X., Liu, W., Dai, G., Wang, H.: Unistroke gestures on multi-touch interaction: supporting flexible touches with key stroke extraction. *IUI*, pp.85--88. ACM (2012)
15. Kammer, D., Wojdziak, J., Keck, M., Groh, R., Taranko, S.: Towards a formalization of multi-touch gestures. *ITS*, pp. 49--58. ACM (2010)

16. Karam, M., Schraefel, M.C.: A taxonomy of gestures in human computer interactions. Technical report, Univ. of Southampton (2005)
17. Kin, K., Hartmann, B., Agrawala, M.: Two-handed marking menus for multitouch devices. *ACM Trans. Comput.-Hum. Interact.*, pp.16:1--16:23 (2011)
18. Kin, K., Hartmann, B., DeRose, T., Agrawala, M.: Proton++: acustomizable declarative-multitouch framework. *UIST*, pp. 477--486. ACM (2012)
19. Kin, K., Hartmann, B., DeRose, T., Agrawala, M.: Proton: multitouch gestures as regular expressions. *CHI*, pp. 2885--2894. ACM (2012)
20. Li, Y. Protractor: a fast and accurate gesture recognizer. *CHI '10*, pp. 2169--2172, ACM.
21. Lin, J., Wu, Y., Huang, T. S.: Modeling the constraints of human hand motion. *HUMO*, pp. 121--126. IEEE Computer Society (2000)
22. Long, A. C., J. A. Landay, Jr., Rowe, L. A.: Implications for a gesture design tool. *CHI*, pp. 40--47. ACM (1999)
23. Long, A. C., J. A. Landay, Jr., Rowe, L. A., Michiels, J.: Visual similarity of pen gestures. *CHI*, pp. 360--367. ACM (2000)
24. Lü, H., Li, Y.: Gesture coder: a tool for programming multi-touch gestures by demonstration. *CHI '12*, pp. 2875--2884. ACM.
25. Malik, S., Ranjan, A., Balakrishnan, R.: Interacting with large displays from a distance with vision-tracked multi-finger gestural input. *UIST*, pp. 43--52. ACM (2005)
26. D. McNeill. *Hand and Mind: What Gestures Reveal about Thought*. Univ. of Chicago Press (1992)
27. Microsoft. Application gestures and semantic behavior (windows) Website. Available at: [http://msdn.microsoft.com/en-us/library/windows/desktop/ms704830\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/ms704830(v=vs.85).aspx)
28. Morris, M. R., Huang, A., Paepcke, A., Winograd, T.: Cooperative gestures: Multi-user gestural interactions for co-located groupware. *CHI*, pp. 1201--1210. ACM (2006)
29. Morris, M. R., Wobbrock, J. O., Wilson, A. D.: Understanding users' preferences for surface gestures. *GI*, pp. 261--268. Canadian Information Processing Society (2010)
30. M. Nielsen, M. String, T.Moeslund, and E. Granum. A procedure for developing intuitive and ergonomic gesture interfaces for hci. pp. 409--420. Springer Berlin Heidelberg (2004)
31. Rekimoto, J.: Smart skin: an infrastructure for free hand manipulation on interactive surfaces. *CHI*, pp. 113--120. ACM (2002)
32. Schieber, M. H., Santello, M.: Hand function: peripheral and central constraints on performance. *J Appl Physiol*, pp. 2293--2300 (2004)
33. Schuler, D., Namioka, A.: *Participatory Design: Principles and Practices*. L. Erlbaum Associates Inc. (1993)
34. Segala, R., Lynch, N.: Probabilistic simulations for probabilistic processes. *Nordic J. of Computing*, pp. 250--273 (1995)
35. Spano, L. D., Cisternino, A., Paterno, F.: A compositional model for gesture definition. *HCSE*, pp. 34--52, Springer-Verlag (2012)
36. Thieffry, S.: Hand gestures. In *The Hand* (R. Tubiana, ed.), pp. 488,492, Univ. of Chicago.
37. Tse, E., Shen, C., Greenberg, S., Forlines, C.: Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. *AVI*, pp. 336--343. ACM (2006)
38. Vatavu, R.-D., Anthony, L., Wobbrock, J. O.: Gestures as point clouds: a \$P recognizer for user interface prototypes. *ICMI*, pp. 273--280, ACM (2012)
39. Wobbrock, J. O., Aung, H. H., Rothrock, B., Myers, B. A.: Maximizing the guessability of symbolic input. *CHI EA*, pp.1869--1872. ACM (2005)
40. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. *CHI*, pp. 1083--1092. ACM (2009)
41. Wu, M., Shen, C., Ryall, K., Forlines, C., Balakrishnan, R.: Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. *ITS*, pp. 185--192. IEEE CS (2006)

Appendix:

In Fig. 7, we summarized a representative set of gestures produced by users for the circle symbol by adequately classifying them according to our taxonomy. We show different movement categories and their mapping into fingers and hands. Although Fig. 7 shows gestures relative to the circle symbol, the depicted fingers and hand poses, as well as the induced atomic movements fairly holds for the other symbols.

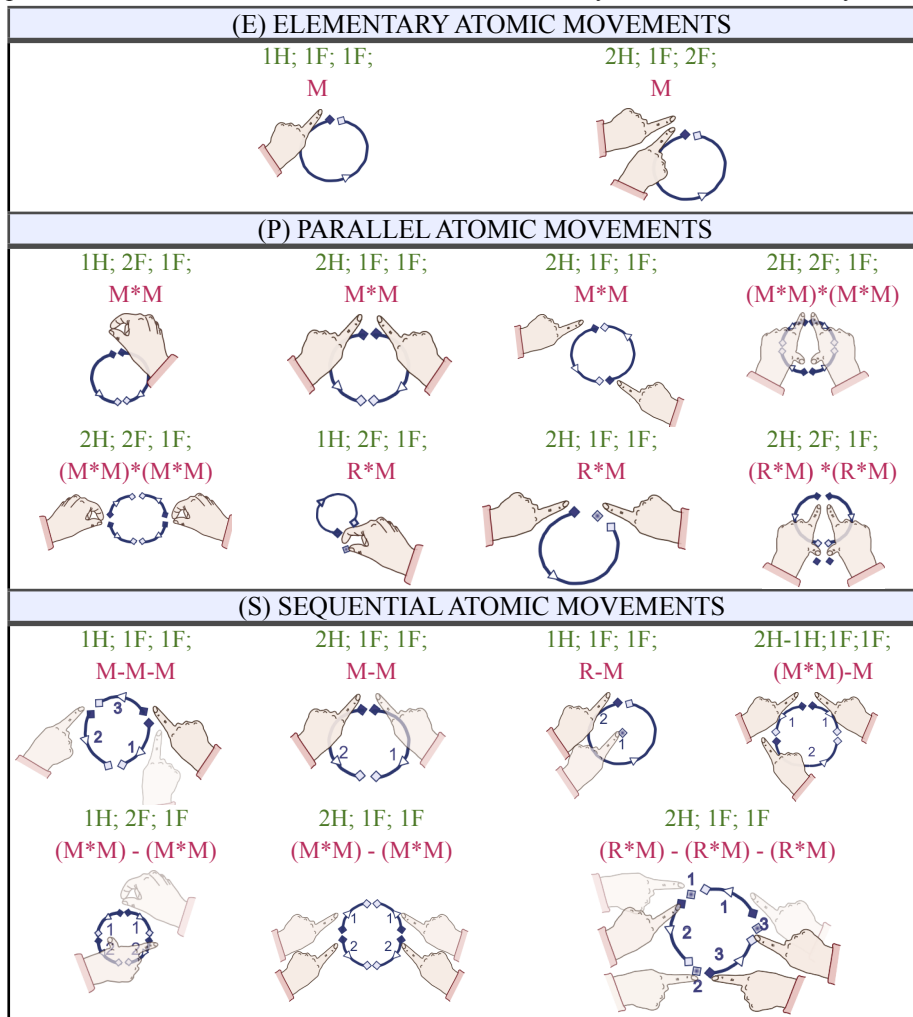


Fig. 7. A representative set of gestures for the circle symbol. We show respectively the number of hands used; the number of fingers per hand; and the number of fingers per movement (e.g., 2H; 1F; 1F; reads as: two hands, one finger per hand, one finger per movement). The atomic movements (R: Ref or M: Motion) and their time composition is also explicated. Gestures depicted with the Index could also be performed by more fingers, e.g. middle+index+ring.