3D CAD modeling using gestures and speech: Investigating CAD legacy and non-legacy procedures

Sumbul Khan¹, Bige Tunçer², Ramanathan Subramanian³, Lucienne Blessing⁴

¹Singapore University of Technology and Design
sumbul_khan@sutd.edu.sg

²Singapore University of Technology and Design
bige_tuncer@sutd.edu.sg

³Singapore University of Technology and Design
ramanathan@mymail.sutd.edu.sg

⁴Singapore University of Technology and Design
lucienne_blessing@sutd.edu.sg

Abstract. 3D CAD modeling using natural interaction techniques necessitates greater research into the modeling procedures employed by users. In a previously conducted experiment, we elicited speech and gestures input for 3D CAD modeling tasks for conceptual design. In this paper, we examine the 3D modeling procedures articulated by the participants, using gestures and speech, for creating basic 3D models of increasing complexity. We identified 3D modeling procedures and characterized them as CAD legacy and non-legacy procedures. Results show that (1) non-legacy procedures were employed by a considerable number of participants who had fair and high proficiency in CAD and (2) Non-legacy procedures with fewer steps were rated favorably by participants. Based on the results, we provide recommendations on key aspects of non-legacy procedures that need to be incorporated in CAD modeling programs to facilitate speech and gestural input.

Keywords: Gestures, 3D CAD modeling, Human Computer Interaction, computer aided design, natural interaction

1 Introduction

Speech and gestural interaction offers designers a powerful technique for creating conceptual 3D Computer Aided Design (CAD) models. Hand gestures are seen to play a key role in perceiving and describing spatial information [1]. Gestural interaction is especially relevant for designers due to its analogy to creative arts such as sketching, clay modelling and sculpting wherein the hands are used for form exploration. Experts opine that multi-modal interfaces using gestures and speech represent a research-level paradigm shift away from conventional Windows-Icons-Menus-Pointers (WIMP) interfaces, toward providing users with greater expressive
power, naturalness, and flexibility [2]. However, the success of multimodal systems using gestures and speech depends on the integration of knowledge of patterns that typify how people interact [2].

Elicitation studies are a popular technique to identify relevant gestures for human computer interaction. Research has pointed out the existence of a ‘legacy’ bias in gesture elicitation studies, which results from users’ previous experience with technology. Consequently, a number of studies have proposed measures to reduce the effect of legacy bias in gesture elicitation [3]. This issue is especially relevant for gesture elicitation studies for 3D CAD modeling, since current generation of designers are well versed with WIMP based CAD modeling programs.

While there is a small body of literature that investigates the speech and gestures used for 3D CAD modeling [4–6], there is virtually no investigation into the sequence of steps, or modeling procedures that designers would use to assemble a CAD model, with gestures and speech. Existing 3D CAD modeling tools are designed for a WIMP based point and click input, that require users to select tool widgets and enter commands in a procedural fashion.

Would designers, when creating 3D models with gestures and speech, follow the same procedures as when using WIMP input? One view is that when using a novel input technique, the conceptual knowledge of existing CAD users will be transferred to the new environment, and they will model using predictable routines from CAD modeling. We refer to these as CAD legacy modeling procedures.

An opposing view is that since gestures are ubiquitous, designers have an innate way of using gestures and speech for spatial thinking and reasoning, irrespective of the way they make models in CAD programs. Previous studies in mental imagery suggest that designers have intrinsic ways of using their ‘mind’s eye’ for visualizing spatial configurations, which are expressed using gestures and speech [7, 8]. Blindfolding exercises in think-aloud problem-solving sessions reveal how designers use gestures to sculpt and interact with shapes in space [8, 9]. The modeling sequences reported in these studies show how designers naturally used their hands and words to articulate the creation of objects and bear no resemblance to the modeling sequences as used in CAD programs. We refer to the modeling procedures that deviate from conventional CAD logic as non-legacy procedures.

For instance, let us consider the modeling of a twisted cylinder. In a CAD legacy procedure, twisting a cylinder would require the creation of a cylinder, and then specifying the two points of the axis of the twist and at least two reference points or the angle of the twist. On the other hand, in a non-legacy procedure perhaps a person would first articulate the concept of a cylinder using speech and gestures, and then then use a ‘twisting’ motion, possibly using both hands turning in opposite directions to articulate the twist, as though modeling with clay.

So far there is little investigation in the procedures designers employ in the articulation of 3D modeling tasks with gesture and speech in the early stages of design. For the development of a 3D CAD modeling interface based on natural human interaction, it is important not only to know the gestures and speech strings that are natural for designers, but also what modeling procedures they use. Especially since most current day designers are experienced in using CAD tools, it is crucial to know
if designers show a preference for legacy modeling procedures and consider these as easy.

To investigate these issues, we analyze data from a previously conducted experiment with 41 participants from architecture and engineering backgrounds to elicit gestures and speech for 3D CAD modelling tasks. In related studies, we presented a compilation of gestures from the experiment and an implementation of a prototype [4, 5, 10]. In this paper we examine the 3D modeling procedures articulated by the participants using gestures and speech, for creating basic 3D models of increasing complexity. While previous studies in gesture elicitation are positioned against legacy bias or to benefit from it, our approach was exploratory. The main contributions of this paper are:

- We present the modeling procedures articulated by designers using gestures and speech, for creating 3D CAD models of increasing complexity, and categorize them as legacy or non-legacy
- We present insights on users’ preferences of modeling procedures using measures such as frequency distribution, articulation time and participants ratings.
- Based on the results, we provide recommendations on the key aspects of non-legacy procedures that need to be incorporated in CAD modeling programs to facilitate speech and gestural input.

Our paper, in its investigation of the speech and gestures used by designers, may be seen as an effort to understand designers’ culture through their linguistic behavior. We attempt to define aspects of designers’ use of gestures and speech that had so far been implicit.

2 Background

This research addresses the issue of using gestures and speech for 3D CAD modeling in the conceptual design stage, wherein designers conduct massing studies by creating and manipulating three dimensional forms. CAD modeling involves the development of a virtual object, using step-by-step commands of instantiating primitives and using operations such as move, copy, rotate, union and subtract for manipulation and assembly. We use the term ‘modeling procedure’ or ‘procedure’ to imply the sequence of steps taken or operations performed to create a 3D CAD model, as used by Hamade et al. [11]. In the conceptual design stage, CAD modeling is often used to ideate the form of design artefacts, such as the initial ideas for buildings, products or automobiles. Hence, CAD modeling is relevant for both architects and engineers. Since design is an integral component of the design and practice of both groups, we use the term ‘designer’ to refer to both these professional groups.

Speech and gestures together are central to human communication, wherein they play complementary roles. Speech is based on conventions and conveys meaning discretely, relying on codified words and grammatical devices. On the other hand, gestures are idiosyncratic and imagistic [12]. The focus of this study is mid-air, touch-free bimanual gestural interaction (also known as unencumbered gestures in 3D
space), without any controller or external device. We define gestures as hand and arm movements in free space, which convey meaningful information.

Interfaces using gestures and speech offer an improved user experience and better control than conventional mouse and keyboard input [2]. A number of studies have investigated gestural interaction in the context of 3D CAD modeling environments [13, 14]. Gestural interaction is seen to be especially relevant for the conceptual design stage, when ideas are fluid [15]. Previous studies have categorized gestures for the creation of 3D environments, based on their origins [16]. Related approaches also include studies on modeling and analyzing tangible gestures [17]. Pioneering studies in using speech as input for CAD modeling include Bolt's 'Put-that-there' system [18], 'Talk and Draw' [19], Weimer and Ganapathy's [20] speech and glove-based gesture input. More recently, a number of studies report the development of multimodal interfaces for 3D CAD modeling [21, 22].

Research has established the potential of elicitation studies for the design of natural interfaces. An elicitation study is a participatory design method to elicit interactions from users. In an elicitation study, users are shown referents (an action's effects) and are asked to provide the corresponding signs (interactions that result in the given referent) [23–25]. The aim of an elicitation study is to derive interactions that end-users find preferable to those designed by Human Computer Interaction (HCI) professionals [23]. Elicitation studies have been conducted for a variety of applications such as surface computing [24] and web-based displays [23]. User-proposed gestures are easier to perform and more suitable than designer-created gestures [26]. User elicited gesture sets are considered user friendly and intuitive [23, 24].

Research has addressed the issue of legacy bias in interaction elicitation, which arises when participants in an elicitation study are well experienced with a previous technology. Largely, there are two approaches to address legacy bias in elicitation studies: (a) Studies that seek to reduce legacy bias [3, 27] and (b) Studies that seek to benefit from legacy bias [28]. The former approach argues that legacy bias limits the potential of user-elicitation methodologies for producing interactions that take full advantage of emerging application domains. Furthermore, they argue that since legacy interactions are based on an abstract relationship with previous experience with technology, they can become outdated with time. On the other hand, the latter approach reasons that legacy knowledge would make it easier to design new interactions, as familiar gestures are easier to recall, produce confidence, and are especially useful for groups of users [28].

Existing natural interfaces for 3D CAD modeling tools mostly use procedures that are researcher designed and have their basis in technical convenience [29, 30]. While there is some research in user elicited gestures for 3D CAD modeling [4–6], so far there has been no research investigating the modeling procedures that are natural for users. This study addresses this gap in literature by examining the 3D modeling procedures articulated by the participants, using gestures and speech in a previously conducted gesture elicitation experiment. In the following sections we provide a brief description of the experiment and the methods used for the analysis of the modeling procedures.
3 Method

3.1 Participants

We refer to the experiment described previously in [4, 5, 10]. We conducted a gesture elicitation experiment individually with 41 participants (21 female, 20 male) from architecture (49%) and engineering backgrounds (51%). Seventy percent of the subjects were from the age group 22-30 years, followed by the age group 31-40 years (15%), 18-21 years (10%) and 41-60 years (5%). All participants were fluent in English, which was a prerequisite for participation in the experiment.

This study employed a convenience sample, recruiting the participants who were readily available for the one-hour experiment. Hence, in our sample 93% participants were familiar with one or more CAD programs and only 7% were not familiar with any CAD programs. However, this should not be a source of bias, as this study is not focused on the comparison of the two groups. Rather it is interested in the behavior of the group that was proficient in CAD. Furthermore, given that CAD software is in widespread use, both in education [31] as well as professional practice in Architecture, Engineering and Construction (AEC) in the developed world [32], it is safe to assume that our sample size of over 90% participants being proficient in CAD is not very deviant from the representative sample of architects and engineers from this region. A previous conducted survey from this country has reported that 84.5% of architects and engineers reported using CAD in 2001 [33], and 88.2% firms from the construction industry from a neighboring country reported using CAD in 2011 [34]. The group that is under-represented in our sample would be the designers who are over 50 years of age, who have had little or no exposure to CAD. We believe this group would have not have a significant bearing on the insights from this research, as it would only constitute a minority in the projected user group for multi-modal 3D CAD modeling.

3.2 Experiment design

The aim of the experiment was to elicit mid-air hand gestures and speech commands for 3D CAD modeling for conceptual design. The experiment was conducted individually with each participant, in two sessions, A (gestures only) and B (gestures and speech). In the experiment, participants were seated in front of a screen on which images and video clips of the CAD modeling referents were shown, one by one.

The participants were shown fifteen referents in four categories: (1) Four tasks for the creation of CAD primitives, (2) Four tasks for manipulation of primitives, (3) Four navigation tasks (4) Three 3D CAD modeling tasks. Categories 1 to 3 were reported earlier in our previous studies [4, 5, 35]. In this paper, we report findings from the category of the CAD modeling tasks.

The three 3D modeling task referents shown to the participants had varying degree of complexity (Fig 1). Modeling task-I comprised a short cylindrical primitive with a square inset. Modeling task-II comprised three Frobel blocks, placed at 90 degrees rotation from one another. Modeling task-III comprised a set of three wedges, with
successive wedges scaled down, rotated and positioned at the edge of the previous. Each referent model was shown from 360° in a 15-second video clip.

![Cylinder with square inset](image1) ![Triangular blocks](image2) ![Triangular wedges](image3)

**Fig. 1.** The three modeling tasks shown to participants in the experiment.

The participants were given a scenario that a fictitious designer, Laura, in an adjacent space needed help in 3D CAD modeling. The participants’ task was to give instructions to Laura to create the object they saw on their screen. In this paper, we report findings from session B, wherein participants were free to use gestures or speech or both, as they wished. The order of the referents was randomized. After the completion of each task, the participant rated how easy they found the task, on a 5-point scale where 5 was ‘Very easy’. The experiment sessions were video recorded with two high speed cameras.

After the experiment, participants filled out a questionnaire and rated their skills in twenty 3D CAD modeling programs on a 5-point scale. The list of CAD modeling programs included all popular CAD modeling programs such as AutoCAD, Rhino, 3dsMax, SketchUP and Solidworks. Participants were also allowed to add entries to the list.

### 3.3 Coding and analysis

Since the focus of this study is natural articulation using both speech and gestures, in this paper we report findings from the three modeling tasks from session B. We coded and analyzed 41 x 3 = 123 video clips of 3D modeling tasks. Speech segments from the video clips were transcribed.

First, the authors jointly reviewed all samples to define the typical modeling procedures employed by participants. CAD legacy procedures were identified as those which closely follow the conventional CAD modeling routines of creating primitive objects and modifying them in a step-by-step procedure using classic CAD commands such as extrude, scale and rotate to create components which are then assembled...
Non-legacy procedures were identified as the procedures that deviated from the conventional CAD routines for 3D modeling. A coding manual was prepared based on the review of the samples.

Operation units in the modeling procedures were delineated based on the function performed by gestures, speech (verbal expressions) and their combination. Synonymic expressions were coded under one operation unit. For instance, expressions such as “Make a block” or “Create a cuboid” or “Make a box” were all coded as “Create cuboid”. Outliers were 5% or lesser in each category and were ignored for the purpose of this research.

Two coders, co-authors of the paper, independently coded the sample. Both coders were from architecture background and well versed with CAD software. Ten percent of the sample from each category was coded by both coders and was used to calculate the inter-coder reliability. Disagreements were resolved with discussion and the resulting overlapping sample was included in the final sample. Inter-rater reliability was good for each category (Cohen’s kappa>0.8).

CAD proficiency of participants was calculated as follows. Participants who reported themselves as “Excellent” in one or more CAD programs; or “Good” in two or more CAD programs, were marked as “High proficiency”. Participants who reported themselves as “poor” or “no CAD skills” in all CAD programs were marked as “Low or no CAD proficiency”. The rest of the participants who reported themselves as “Good” in less than two CAD programs, or “Fair” in a number of them, were marked as “Fair proficiency” in CAD. Hence, 66 % participants were found to have high proficiency, 27% participants had fair proficiency, and 7% participants had low proficiency.

4 Results

Over 90% of participants employed a combination of gestures and speech as a mode of interaction in modeling tasks I and II, and 100% in the case of modeling task-III.

We identified the modeling procedures employed by the participants for each modeling task (Fig 2). In modeling task-I, cylinder with square inset, we identified three modeling procedures (Fig 2a). Procedure I-1 followed a typical CAD routine. First a circle was created and extruded up. A rectangle was created on its top surface and extruded down. In procedure I-2, a different decomposition was followed, as it was visualized as the removal (or subtraction) of a block from the cylinder. In procedure I-3, participants started with a cylinder primitive. Then they instantiated a cuboid primitive and subtracted the latter from the former. The last operation was also articulated as embedding a cuboid into a cylinder. We categorized the embed operation as non-legacy, as it was different from an extrude down or a subtraction typically used in CAD programs. We categorized the procedures I-1 and I-II as legacy procedures, as these are based on the typical routines used in CAD modeling.

We delineated four procedures in modeling task-II, Frobel blocks (Fig 2b). In procedure II-1, participants first created a rectangle and extruded it. Subsequent blocks were created by copying, rotating and positioning the block. In procedure II-2,
participants instantiated three rectangles at once (using speech), extruded all and placed them one by one, by gesturing the rotation and positioning of each using their hands. Similarly, in II-3, participants started with the concept of three cuboids (using speech) and articulated the placement of each one by one. In procedure II-4, participants started with the concept of three cuboids, and then described details of the object such as its dimensions and proportions, before articulating the placement of each. In this category, we categorized only procedure I-1 as legacy. Procedures II-3 and II-4 were categorized as non-legacy. Procedure II-2 had both legacy and non-legacy operations.

In the third referent category, Wedges, five procedures were delineated (Fig 2c). In procedure III-1, participants first created a triangle and extruded it. The triangular block was then copied, scaled and positioned. The last three operations were then repeated. In procedure III-2, participants first created three triangles, using gestures to position them in space. Subsequently, participants extruded the three triangles one by one. In procedure III-3, participants created the triangle in its correct position, and then extruded it. The two operations were then repeated twice. In procedures III-4 and III-5, participants articulated the creation of three triangular prisms (using speech), then articulated the placement of each, specifying both rotation and position, using gestures and speech. In procedure III-5, additionally, participants describe the hypotenuse using speech and gestures, before articulating the placement of the next wedge. In this category, we characterized procedures III-1, III-2 and III-3 as legacy whereas III-4 and III-5 were characterized as non-legacy. Examples of legacy and non-legacy modeling procedures are given in Fig. 3.

We investigated the number of participants who employed legacy and non-legacy procedures, based on their CAD proficiency (Fig 4a). Participants with low or no CAD proficiency employed non-legacy modeling procedures. Sixty percent participants with high CAD proficiency employed non-legacy modeling procedures for modeling task III. Non-legacy procedures in tasks II and III were also had high medians of participants’ ratings of ease (Mdn_{II-3}= 4, Mdn_{II-4}= 4, Mdn_{III-4}=3, Mdn_{III-5}=3) (Fig 4b).

We analyzed the average time participants spent on articulating the modeling tasks using gestures and speech. Overall, the average time spent on articulating modeling task-I was the least ($M=23.28$ s). The average time participants spent on modeling tasks II was 29.17 seconds and on modeling task III was 42.37 seconds. Spearman’s rank order correlation between articulation time (in seconds) and participants’ ratings of ease revealed a statistically significant correlation only for the modeling task-I (Table 1).

We also analyzed the average time participants spent articulating the modeling tasks based on modeling procedures (Fig 4c). In modeling task-I, cylinder with square inset, average time spent on procedure I-2 was the least with 17.75 seconds (SD=5.76), and the average time spent articulating procedure I-3 was the most ($M=27.14$ s, SD=13.19). In modeling task II, Froebel blocks, average time to articulate procedure II-3 was the least ($M=28.25$, SD=16.58), with procedures II-1($M=32.8$, SD=9.35) and II-4 ($M=33.75$, SD=7.22) taking more time. For modeling task III,
Wedges, average articulation time for procedures III-2 ($M=29$ secs, $SD=8.29$) and III-4 ($M=34$ secs, $SD=22.47$) was much lesser than the legacy procedure III-1 ($M=51.4$ secs, $SD=15.53$) or the non-legacy procedure III-5, in which participants articulated details ($M=56.12$ secs, $SD=14.33$).
Fig. 2. Modeling procedures identified for the three modeling tasks (a) Modeling task I- (Cylinder with inset) (b) Modeling task-II (Froebel blocks) (c) Modeling task-III (Wedges) (*Non-legacy operations shown with grey infill and legacy operations with no infill).
Modeling task-I (Cylinder with square inset): Procedure I-1 (CAD Legacy)

CREATE CIRCLE

S: Draw a circle

EXTRUDE

S: Extrude

CREATE RECTANGLE

S: And draw a square

EXTRUDE DOWN

S: And extrude it down

(a)

Modeling task-II (Proebel blocks): Procedure II-1 (CAD Legacy)

CREATE RECTANGLE

S: Make a rectangular box

EXTRUDE

S: 

COPY

S: Copy it

ROTATE

S: Rotate it 90 degrees

POSITION

S: and put it on the corner of the 1st box.

COPY

S: Copy the 2nd box

ROTATE

S: Rotate it 90 degrees

POSITION

S: and put it on the corner of the 2nd box

(b)

Modeling task-II (Proebel blocks): Procedure II-3 (Non-legacy)

CREATE THREE CUBOIDS

S: Three identical extruded rectangles

PLACE

S: like that

PLACE

S: and at 90 degrees

PLACE

S: and another one on top

(c)
Fig. 3. Examples of legacy and non-legacy modeling procedures using gestures and speech (a) Procedure I-1 (CAD legacy), (b) Procedure II-1 (CAD legacy), (c) Procedure II-3 (Non-legacy), (d) Procedure III-3 (CAD legacy), (e) Procedure III-4 (Non-legacy) (*S= accompanying speech)
Fig. 4. (a) Frequency distribution of modeling procedures based on CAD proficiency of participants (b) Articulation times (in seconds) of modeling procedures (c) Medians for participants' ratings of ease of modeling procedures
Table 1. Result from Spearman’s rank order correlation between articulation time (in seconds) and participants’ ratings of ease.

<table>
<thead>
<tr>
<th>Modeling task</th>
<th>R</th>
<th>P (two-tailed)</th>
<th>Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-Cylinder with square inset</td>
<td>-0.38555</td>
<td>0.02218</td>
<td>Yes</td>
</tr>
<tr>
<td>II-Froebel blocks</td>
<td>-0.05317</td>
<td>0.76161</td>
<td>No</td>
</tr>
<tr>
<td>III-Wedges</td>
<td>-0.11663</td>
<td>0.50464</td>
<td>No</td>
</tr>
</tbody>
</table>

5 Discussion and conclusion

The objective of this exploratory study was (1) to investigate the modeling procedures used by designers in a previously conducted speech and gesture elicitation experiment for CAD modeling for conceptual design (2) to determine participants’ preferences of CAD modeling procedures using gestures and speech.

5.1 Identification of CAD modeling procedures

We identified 3D modeling procedures used by the participants in three modeling tasks and characterized them as CAD legacy and non-legacy procedures. As pointed out in previous studies that complex referents elicited lesser gestural agreement [24], we found a greater number of modeling procedures for the high complexity modeling task, as there are several ways to decompose complex referents. Hence, we identified three procedures in modeling task-I (two of which were CAD legacy), four procedures in modeling task-II (one of which was CAD legacy) and five procedures in modeling task-III (three of which were CAD legacy). Modeling task-I, being a simple model, had fewer procedures and distinguishing features to delineate any procedure as clearly non-legacy. Some procedures were identified to have both legacy and non-legacy operations, such as I-3 and II-2.

5.2 Preferences of modeling procedures

Our chief finding from this study was the significant use of non-legacy modeling procedures by participants who had fair and high proficiency in CAD (56% in modeling task-II, and 51% in modeling task-III). Since most participants had fair and high proficiency in CAD, it could have been assumed that the conceptual knowledge of CAD would be transferred to the speech and gesture based input, and that most participants would employ CAD legacy procedures. However, the distribution of fair and high CAD proficiency participants across legacy and non-legacy procedures suggests that although CAD legacy knowledge has some impact on participants’
modeling techniques, participants also relied on their intrinsic ways of using gestures when articulating modeling tasks.

Furthermore, our findings show that some CAD legacy procedures such as the ones that involve step-by-step copying and manipulating (II-1 and III-1) could be cumbersome for users when using gestures and speech, as these procedures got lower ratings than other procedures for the same modeling task (Mdn\(_{II-1}\) = 3, Mdn\(_{III-1}\) = 2.5). Participants rated more favorably when they employed non-legacy procedures that involved fewer steps. For instance, non-legacy procedures II-3 and III-4 had low articulation time and corresponded to more favorable ease ratings (Mdn\(_{II-3}\) = 4, Mdn\(_{III-4}\) = 3) than other procedures for the same modeling task. Similarly, CAD legacy procedures that involved direct placement, such as procedures III-2 and III-3 were rated better than the legacy procedure III-1, which had more steps. While previous studies assert that legacy interactions, due to their familiarity, are easier for users [28], our results show that in the case of CAD modeling, legacy procedures corresponding to long sequences could be tedious.

We did not find a statistically significant correlation between articulation time and ease ratings in modeling tasks II and III. It is interesting to take note of modeling procedures II-4 and III-5, wherein participants described details of the modeling object, such as the dimension and proportions, with greater articulation times, were also rated favorably by participants. This could be attributed to participants being more confident with their descriptions when they provided details. Wobbrock et al. [24] report a similar finding that gestures that took longer to perform were rated as easier, possibly because they were 'smoother or less hasty'.

### 5.3 Recommendations for CAD modeling programs

We thus conclude that both legacy knowledge and non-legacy knowledge are relevant for different purposes in 3D CAD modeling. The fact that a large number of participants employed non-legacy procedures and rated them favorably suggests that non-legacy procedures, along with legacy procedures, need to be incorporated as valid forms of input in 3D modeling programs. Legacy procedures could be useful for modeling simple objects, which require fewer steps. Their inclusion would also provide familiar modeling techniques for users who are engrained in conventional CAD thinking. However, for the assembly of complex objects, it would be useful to employ non-legacy procedures which could reduce the number of steps taken to model the object. This approach is substantiated with HCI literature that states that, users do not want to perform complex sequences of gestures to achieve basic goals [28].

Figure 5 shows the differences between legacy and non-legacy procedures that we observed in the three modeling tasks. Based on the analysis of the data elicited from the participants, we provide recommendations on the key aspects of non-legacy procedures that need to be incorporated in CAD modeling programs to facilitate speech and gestural input:

- Speech and gesture based CAD modeling should allow a higher level of abstraction in input. We observed that in non-legacy procedures, participants
employed less atomic concepts, for example using “cylinder” instead of “circle” and “extrude”.

- We recommend the development of “compound” commands that combine spatial input such as rotation and positioning into a single operation. Speech and gestural input gives greater freedom in expression of spatial concepts. For example, in our data set, we identified participants employing the operation “place” that incorporated both ‘rotation’ and ‘position’ (specifying the location of the object).

- We recommend the use of simultaneous creation and manipulation of many objects using speech, as demonstrated by the operation “Extrude the three triangles”

- Speech and gestural modeling programs should also facilitate the instantiation and manipulation by specifying objects in relation to other objects. Examples include, “Extrude the 2nd triangle” or “Create a bigger triangle”. This recommendation is based on the observation of how participants referred to objects in our experiment and is supported by literature [18].

- We recommend the use of routines that allow direct input, easier position finding that are more amenable to speech and gestural input.

<table>
<thead>
<tr>
<th>CAD legacy input</th>
<th>Non-legacy input</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D Shape</td>
<td>Less atomic primitives</td>
</tr>
<tr>
<td>Extrude</td>
<td></td>
</tr>
<tr>
<td>Rotate</td>
<td>Place</td>
</tr>
<tr>
<td>Position</td>
<td></td>
</tr>
<tr>
<td>Instantiate or manipulate one</td>
<td>Instantiate/Manipulate many</td>
</tr>
<tr>
<td>Create</td>
<td>Instantiate/Manipulate by specifying objects in relation to others</td>
</tr>
<tr>
<td>Copy</td>
<td></td>
</tr>
<tr>
<td>Manipulate</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. Key differences observed between CAD legacy and non-legacy modeling procedures
5.4 Limitations and future work

This study has several limitations. Firstly, in our gesture elicitation experiment designers already had the end-product in front of them, which is not the case in real-world design scenarios where the end-product is a vague mental model. Therefore, experiment participants may have had greater clarity of the modeling procedures while developing the models. Secondly, people articulate spatial information in different ways, based on their education, experience, socio-cultural and linguistic background [38]. This study is tilted towards English speaking populace of south Asia. Thirdly, we computed CAD proficiency based on participants’ self-reported proficiency in various CAD programs. While this could be source of bias, we do not expect a major discrepancy due to this issue as most participants from our experiment were computer savvy. Finally, this paper only reported an initial investigation on the interaction of gestures and speech for CAD modeling. Due to time constraints, we could investigate only three basic modeling tasks.

Analyses of a greater number of modeling tasks involving complex geometry, as well in-depth investigation of the interaction of gestures and speech in CAD modeling procedures are necessary for further development of gesture and speech based interfaces for CAD modeling for conceptual design.

5.5 Concluding remarks

This study highlights the relevance of the modeling process in CAD and its inter-relationship with the interaction technique used. The internal logic of 3D modeling programs must reflect the natural modeling techniques of designers.

We presented an analysis of how designers used gestures and speech to create virtual models in space. This blend of gestures and speech reflects a linguistic culture of designers which must be defined and analyzed for the development of successful human computer interfaces. While most research on natural interfaces for CAD modeling focus largely on gestures, to our knowledge, this is the only study that examines users’ CAD modeling procedures articulated with gestures and speech. Investigations in this area are crucial for the development of user friendly, natural interfaces for 3D CAD modeling. We conclude that multi-modal interaction techniques using gestures and speech are so paradigm shifting that they will dictate changes in the very modeling structures of extant CAD programs.

Acknowledgements. This research was supported by SUTD-MIT International Design Centre (IDC) grant number IDG21500109, under the Sustainable Built Environment Grand Challenge, and Visualization and Prototyping Design Research Thrust.

References


15. Pratini, E.: Modeling with gestures: sketching 3D virtual surfaces and objects


