

---

# Unravelling the Language of Eating

**Nicole J.J.P. Koenderink**

Wageningen UR FBR  
Bornse Weilanden 9  
6708 WG Wageningen, NL  
[Nicole.Koenderink@wur.nl](mailto:Nicole.Koenderink@wur.nl)

**Jan L. Top**

VU University Amsterdam  
Wageningen UR FBR  
De Boelelaan 1081  
1081 HV Amsterdam, NL  
[Jan.Top@wur.nl](mailto:Jan.Top@wur.nl)

**Abstract**

Understanding eating behaviour requires observations of how consumers choose what to eat and of what they actually eat. Recognizing eating gestures is difficult and labour intensive. We propose an ontology of elementary eating gestures to make visual observations of eating behaviour objective and quantifiable. Moreover, this ontology facilitates automating this task. The eating gesture vocabulary is published as an ontology and is

*Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honoured. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).*

UbiComp '14, September 13 - 17 2014, Seattle, WA, USA  
Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-3047-3/14/09...\$15.00.  
<http://dx.doi.org/10.1145/2638728.2641340>

**Andrea van Doorn**

Industrial Design,  
Delft University of Technology  
Landbergstraat 15, 2628 CE Delft  
The Netherlands  
[a.i.vandoorn@tudelft.nl](mailto:a.i.vandoorn@tudelft.nl)

available for use.

**Author Keywords**

Eating gesture; ontology; consumer science; computer vision; social interaction; eating

**ACM Classification Keywords**

I.2.4: Semantic networks, J.4 Social and behavioural sciences

**Introduction**

Measuring food choice and food intake is relevant for the following consumer oriented disciplines. First, we mention **consumer science**. An important area of consumer research aims at determining how eating behaviour is influenced by ambient factors, such as the presence of other people [13]. Literature shows that social facilitation – the presence of others during consumption – influences e.g. food intake [7], food appreciation [4], and meal duration [6]. The presence of a social model – someone who behaves in a pre-instructed way – is reported to influence the behaviour of the other meal companions in terms of actual consumption [10]. One research objective is to determine how people are influenced by others when eating their meal in a natural eating setting. Second, we think of **health care**. The nutritional condition of people in a medical setting, such as in hospitals, rehabilitation centres or at home, largely determines their well-being. Monitoring the eating behaviour of these people helps determining their eating behaviour and detecting undesirable situations. Specific eating

disorders can be monitored, e.g. on people with Parkinson's disease. Third, **marketing** is a field of interest. In order to understand food selection by consumers, the effect of ambient factors and factors that determine preference and liking it is useful to measure consumer behaviour objectively.

Observing food intake requires objective and unobtrusive observation techniques. Visual recordings using video are well suited for this purpose. However, analysing these recordings may lead to results that are difficult to compare and to quantify, due to the lack of a demarcation and disambiguation between different gestures and movements.

In this paper, we propose a formal framework that defines elementary eating gestures, the **Eating Gesture Ontology**. The purpose of this ontology is to enable **objective annotation of recordings of eating people**, for example with computer vision software. As a result, observations from different studies can be uniformly interpreted, compared and merged to yield new insight. The framework also helps automating the analysis of the recorded data.

This paper is organised as follows. First, we discuss the literature that inspired us. Next, we describe the method used to determine the formal framework of eating gestures. In the results section, we present the Eating Gesture Ontology. We end with conclusions and recommendations for future work.

### **Related work**

To measure the influence of social factors on eating behaviour, a number of methods is frequently used. Meals can be weighed before and after a meal when

one is interested in the amount of intake [4,6,9], diaries can be kept by the consumers [4], questionnaires can be filled in [3,5]. These methods are indirect, since the eating persons are not observed during their meal, only the effect of eating is measured or a report of the activity is made afterwards. Sometimes researchers manually annotate the actions of consumers [1,5] during an intervention study. To the best of our knowledge, no studies have been reported in which the gestures of people during their meal are systematically charted.

In this paper we create a theoretical framework to represent the gestures of people while eating. This framework helps to systematically study the behaviour of eating consumers. In the domain of knowledge engineering and knowledge representation, ontologies are frequently used to capture such knowledge [14]. An ontology is an explicit specification of a model of part of the world that is represented for some purpose [9]. An ontology contains a set of concepts and relations between these concepts. It can be used as a domain model [8]. Ontologies capture static knowledge in a problem-solving independent way [14]. They are used in many areas where different groups of people need to communicate about a common field, such as bioinformatics [2], the biomedical domain [12] and the agrifood domain [11]. We propose to model the framework of eating gestures as an ontology in order to have a shared and commonly understood basis for future research in this area of consumer studies.

### **Method**

In our study we have limited our scope to modelling eating gestures for healthy people at lunch during work time.



Figure 1. A screenshot of one of the two recorded lunches. The faces of the consumers have been obscured because of privacy reasons.

#### *Gathering of data*

We have used video recordings of eating people to identify gestures that occur in an eating setting. To gather such data, we have used ceiling-mounted cameras in the Restaurant of the Future facility<sup>1</sup> of Wageningen UR. In this restaurant customers regularly have their lunch. All customers are informed that their selection and eating behaviour can be monitored at all times. The ceiling of the Restaurant of the Future is equipped with Panasonic WV-CS950 pan-zoom-tilt cameras that we have used to observe consumers in two different circumstances. Initially, we have observed two lunches that were chosen at random from the lunches that took place in the Restaurant of the Future on two separate days. The first lunch consisted of two people and took 9:38 minutes. The second lunch consisted of six people and took 9:15 minutes. In

<sup>1</sup> <http://www.restaurantofthefuture.wur.nl>

Figure 1, a screenshot of one of these sessions is displayed. During these lunches, the consumers were aware that they could be recorded, but no specific intervention has been done.

Secondly, we have prepared an experiment in which three groups of two people were each invited twice for a lunch. Each lunch was paid for by this study. During the experiment, the following items were present at the table: a jar of milk, a jar of juice, 6 sandwiches (3 x white, 3 x whole wheat), 1 plate with cheese slices, meats and eggs (savory sandwich fillings), 1 plate with salad, 1 bowl with cups of butter and various sweet sandwich fillings (marmalade, chocolate sprinkles), and a pile of paper napkins. Moreover, each participant had a bowl of soup, a plate, a set of cutlery (spoon, knife, fork), and 2 glasses. These lunches took 6:34, 8:16, 19:16, 21:03, 21:37, resp. 22:31 minutes. Each of these lunches was recorded with two cameras. Each camera recorded the lunch with one of the participants in view (see Figure 2 for a frontal view of one of the two participants of one of the lunches).

#### *Annotation of videos*

After recording both types of lunch sessions, the second step of the study was to identify gestures in the videos and annotate the videos accordingly. To do this, we needed to decide on the granularity level of gestures. As an example, we can consider "taking a bite" as an elementary step. Does this gesture start with putting butter on the bread, by folding the sandwich, by lifting the sandwich to the mouth or by putting one's teeth in the sandwich? We have decided to divide the gestures in relatively small building blocks, which can be composed to make more complex movements. This choice has the consequence that for example "taking a

bite” is restricted to putting food into the mouth. The gesture starts close to the mouth; bringing the food to the position close to the mouth is considered to be a different building block (that can be followed by both “taking a bite” or by “putting down the bread”).

With this decision in mind, we analysed the recorded videos from the first session and determined the observed gestures for one of the consumers per lunch sequentially. We assigned names to the gestures and organised them in categories. Next, we used the gestures as defined in the first session to annotate the videos from the second session using The Observer XT<sup>2</sup> software. When gestures were detected that did not occur in the first session, these were added to the collected set of eating gestures.

## Results

The annotation of the videos has resulted in a collection of 26 *eating gesture* classes, 10 *actuator* classes, 5 *object* classes and 5 *gesture* location classes. In the next sections, we describe the concepts in more detail.

### *Eating gestures*

The eating gesture ontology consists of six categories, each consisting of a number of eating gesture classes (see Table 1). Each eating gesture class is further specified by the *instances* of the class. For example, the gesture “move food to mouth” has instances “move soup to mouth with a spoon”, “move bread to mouth with one hand”, “move bread to mouth with two hands” et cetera. These instantiations inherit the properties, as

described in the next section, of the eating gesture class to which they belong.

Category	Description	Corresponding classes of eating gestures
Prepare food gestures	Actions aimed at preparing food for consumption.	Cut, Open package, Pick up package, Put down package, Place spread on bread, Pour, Stir
Prepare consumption gestures	Those actions that precede consumption.	Pick up food or drink, Pick up cutlery, Move to mouth
Consume gestures	Those actions that deal with the actual consumption of food.	Take a bite (or sip), Chew food, Swallow food
Finalise consumption gestures	Those actions that follow directly after consumption.	Move away from mouth, Put down food, Put down cutlery
Clean gestures	Those actions aimed at cleaning oneself or ones environment.	Lick, Pick teeth, Pick up cleaning aid, Put down cleaning aid, Stack, Wipe
Non-eating gestures	Gestures that are not primarily aimed at the consumption of food, but that deal with the interaction with others at the table or with fidgeting.	Express with hands, Express with head, Express with shoulders, Play

Table 1. An overview of the identified gesture categories with their descriptions and the corresponding subclasses of eating gestures, 26 in total.

We have identified a total of 196 gesture instances in the observed lunch setting.

### *Locations, actuators and objects*

Besides concepts, we have also identified properties of these concepts. Each eating gesture instance has (i) a *start location*, (ii) an *end location* and (iii) an *actuator* that moves (iv) an *object* from start to end location.

<sup>2</sup> <http://www.noldus.com/human-behavior-research/products/2/the-observer-xt>

The gesture “move food to mouth”, e.g., starts in the “personal space” of the eater and ends at the “mouth”. The “hand”, sometimes with a piece of cutlery in it, brings a specific “food” item to the mouth.

The **start and end location** of an eating gesture may be used to facilitate automated gesture identification. For this reason, we have observed which locations occur in the eating context. These are: mouth, head, torso (body locations); table; personal space; object locations (e.g. the location of a food item, glass or bowl).

The actuators of a gesture consist of either just the hands or of objects that are held in one or both hands. We have seen the following **actuators**: head or head part; hand or hand part; one hand with utensil, cleaning aid or package; two hands; two hands with utensil or cleaning aid. Finally, the following **objects** have been used in our eating experiments: table ware, cutlery (utensils); food, waste, package (perishables); clothes, body part (person bound object); napkin, toothpick (cleaning aids); table.

The locations and actuators may be used to filter out probable gestures at some point in time. Examples of annotated eating gestures can be found in Figure 2.

### **Conclusion**

We have shown that it is possible to create an eating gesture ontology by observing video images of eating behaviour. This ontology can be used to annotate recordings of eating behaviour.

From the analysis of the annotations, we make the following observations. First, to study the behaviour of

consumers in an eating context, gestures that are not directly related to the intake of food should be included as well. Second, some gestures seem to appear in frequently-occurring patterns. These observations are relevant for the automated study of eating behaviour.

Future work can now focus on (i) identifying which patterns of eating gestures occur, and (ii) which interaction patterns between eating companions occur. If one is able to detect interaction patterns, one can for example study the hypothesis that mirroring behaviour between eating companions is correlated with feelings of well-being.

### **Acknowledgements**

This publication was supported by the Dutch national program COMMIT. We thank the participants in the experiments.

### **References**

- [1] Areni, C.S., and Kim, D. The influence of background music on shopping behavior: classical versus Top-40 music in a wine store. *Adv Consum Res* 20 (1993), 336-340.
- [2] Bard, J.B.L., and Rhee, S.Y. Ontologies in biology: design, applications and future challenges, *Nature Reviews Genetics* 5 (2004), 213-222.
- [3] Bell, R., Meiselman, H.L., Pierson, B.J., and Reeve, W.G. Effects of adding an Italian theme to a restaurant on the perceived ethnicity, acceptability, and selection of foods. *Appetite* 22, 1 (1994), 11-24.
- [4] Bellisle, F., and Dalix, A.M. Cognitive restraint can be offset by distraction, leading to increased meal intake in women. *Am J Clin Nutr* 74 (2001), 197-200.
- [5] Caldwell, C., and Hibbert, S.A. The influence of music tempo and musical preference on restaurant patrons' behavior. *Psychol Market* 19 (2002), 895-917.

[6] Clendenen, V.I., Herman, C.P., and Polivy, J. Social facilitation of eating among friends and strangers, *Appetite* 23, 1994, 1-13.

[7] De Castro, J.M., and De Castro, E.S. Spontaneous meal patterns of humans: influence of the presence of other people. *Am J Clin Nutr* 50, 1989, 237-247.

[8] Falbo, R.A., Guizzard, G., and Duarte, K.C., An Ontological Approach to Domain Engineering, In *Proc. SEKE'2002*, 351- 358.

[9] Gruber, T.R. A translation approach to portable ontology specifications. *Knowledge Acquisition* 5, 1993, 199-220.

[10] Hirsch, E.S., and Kramer, F.M. Situational influences on food intake. In: *Marriott BM, ed. Nutritional needs in hot environments*. Washington, DC: National Academy Press (1993), 215-243.

[11] D. Lukose, World-Wide Semantic Web of Agriculture Knowledge, *Journal of Integrative Agriculture*, 11 (5), 2012, 769-774.

[12] Smith, B., Ashburner, M., Rosse, C., Bard, J., et al, The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nature Biotechnology*, 25 (11), 2007, 1251-1255

[13] Stroebele, N., and De Castro, J.M., Effect of ambience on food intake and food choice, *Nutrition* 20 (9), 2004, 821-838.

[14] Studer, R., Benjamins, V.R., and Fensel, D., Knowledge Engineering: Principles and methods, *Data & Knowledge Engineering* 25, 1998, 161-197.

[15] Van Dale, *Basiswoordenboek Nederlandse Gebarentaal*, ed: T. Schermer, C. Koolhof, 2009, Van Dale Uitgevers, Utrecht.

		
<p><u>Put soup on spoon</u> ∈ Pick up food  loc: location of bowl -&gt;location of bowl  object: soup  actuator: one hand with spoon</p>	<p><u>Move soup to mouth with spoon</u> ∈ Move to mouth  loc: location of bowl -&gt; mouth  object: soup  actuator: one hand with spoon</p>	<p><u>Take a sip of soup with spoon</u> ∈ Take bite  loc: mouth -&gt; mouth  object: soup  actuator: one hand with spoon</p>

Figure 2: Frames from a video sequence in which the person to the right is eating soup. In the table, for each eating gesture the name, its class, its start and end location, the object and actuator involved are listed. The persons have given permission to show their faces.