Coding gestural behavior with the NEUROGES–ELAN system

HEDDA LAUSBERG
German Sport University Cologne, Cologne, Germany

AND

HANSLOETJES
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

We present a coding system combined with an annotation tool for the analysis of gestural behavior. The NEUROGES coding system consists of three modules that progress from gesture kinetics to gesture function. Grounded on empirical neuropsychological and psychological studies, the theoretical assumption behind NEUROGES is that its main kinetic and functional movement categories are differentially associated with specific cognitive, emotional, and interactive functions. ELAN is a free, multimodal annotation tool for digital audio and video media. It supports multileveled transcription and complies with such standards as XML and Unicode. ELAN allows gesture categories to be stored with associated vocabularies that are reusable by means of template files. The combination of the NEUROGES coding system and the annotation tool ELAN creates an effective tool for empirical research on gestural behavior.

Methods in Gesture Research: A Short Introduction to the State of the Art

Knowledge about gestural behavior is theoretically relevant for understanding cognitive, emotional, and interactive processes, but it also has far-reaching practical implications for diagnostics and therapy in the clinical context, for learning processes, and for obtaining communicative competencies. Thus, research on gestural behavior is conducted in several academic disciplines, including neuroscience, medicine, psychology, linguistics, anthropology, and social sciences. Especially because of the shift in recent decades toward transferring information through visual media, a thorough knowledge of how gesture reflects and affects the gesturer’s thinking and feeling, as well as of how it influences the recipient’s cognitive and emotional processes, is becoming increasingly important.

Despite many examinations carried through in various disciplines, the state of knowledge about the neuropsychology of gestural behavior has not developed far beyond the popular level (i.e., “body language”). This situation is partly due to methodological deficiencies in the analysis of gestural behavior. In many studies, gesture units and gesture types are not clearly defined, and interrater agreement is rarely examined. Furthermore, the assumptions behind the diverse gesture classifications applied in the studies are not made transparent, and sometimes there is an implicit claim that the particular coding system suits all research questions. In addition, different forms of presenting data (e.g., in terms of total numbers \( n \) of gestures—\( n/\text{min}, n/5\ \text{min}, n/100\ \text{words}, n/\text{time speaking}, n/\text{time of interview}, \) or logarithms—or of handedness—right vs. left hands only; right, left, or both hands; no differentiation of hand laterality; or handedness indices) are an obstacle for comparing the results of different studies and for gradually building up a corpus of knowledge on gestural behavior. By way of example, the apparent contradictions between study results on the relationship between aphasia (loss of the ability to produce and/or comprehend language) and gesture production/comprehension—that is, between an equally distributed deficit and a compensatory use of gestures (see, e.g., Lott, 1999)—are mainly due to a lack of scrutiny during development of the study methods.

A similar situation exists in the area of media annotation, even when only the digital era is considered. A variety of tools are available to the public, each designed for a certain group of users with a specific task or a particular platform (operating system) in mind. The data produced by these tools are stored in various ways and in different formats. This is another obstacle to interoperability and comparison, in addition to those mentioned above.

The XML (Extensible Markup Language) data produced by the majority of contemporary tools are a significant improvement on the old-time proprietary formats. XML data are relatively easy to transform, but nevertheless the conversion is rarely lossless. In an attempt to improve interoperability between tools from multiple modalities, a group of developers in collaboration with selected tool users launched an initiative that resulted in a proposal for
Module I refers to the kinetic features of a hand movement (movement vs. rest; trajectory and dynamics; location of acting) and therefore represents a type of analysis that could potentially be submitted to a 3-D video analysis. Just like the categories of Modules II and III, the categories of Module I have been chosen because they are related to neuropsychological states or have neurobiological correlates (see the Theoretical and Empirical Background of the NEUROGES Categories section below).

Each Module I category is precisely defined—for example, repetitive movements are defined by a specific trajectory and dynamics.

Trajectory. The hand moves forth–back–forth or, more often, on the same trajectory—for example, up–down–up–down—in the air or on the body. The trajectory can be varied in one dimension—for example, up–down–up–down movement while simultaneously moving forward. It is a characteristic feature of repetitive movements that the repetition results in a metrical or rhythmical movement. The repetitive gesture unit can include a preparation phase in which the hand is moved from a resting position or—if another gesture unit (e.g., a phasic gesture) preceded—from an active position directly, with a straight trajectory, to the position where the actual repetitive hand movements start. In the same vein, the repetitive gesture unit can also include a retraction phase, in which the hand is moved back directly, with a straight trajectory, to a resting position or—if another gesture unit follows—to an active position.

Module I: Kinetic gesture coding. Following the algorithm given in Figure 1, first the right-hand movements are tagged (Step 1) and coded (Steps 2 and 3), and then so are the left-hand movements.

![Figure 1. Algorithm for kinetic right-/left-hand coding (Module I).](image-url)
neuroges–elan 843

Module III: Functional gesture coding. In Module III, the gesture units, which have been defined in Modules I and II, are coded with respect to their function and are further classified as types (Figure 3).

Although good interrater agreement can be easily achieved for kinetic (or formal) gesture parameters (such as in Modules I and II), it is more difficult to operationalize gesture function (Wallbott, 1989). However, the observer deduces the meaning of a gesture from a specific combination of kinetic features (hand shape, hand orientation, path, effort, gaze, gesture space, etc.). Therefore, the gesture types in Module III are defined precisely by a specific set of kinetic variables (Table 1).

**Theoretical and Empirical Background of the NEUROGES Categories**

The NEUROGES coding system is grounded in empirical psychological and neuropsychological research. The theoretical assumption behind NEUROGES is that its main kinetic and functional gesture categories are differentially associated with specific cognitive (spatial cognition, language, praxis), emotional, and interactive functions. This implies that the different gesture categories may be generated in different brain areas. In the following sections, two examples are given to illustrate the link between NEUROGES categories and either neuropsychological function or a neurobiological correlate, respectively.

**Example: The Module I “on body” category and the gesturer’s state of internal regulation.** For a long time, researchers have proposed that gestures in which...
the hand acts on the body (also termed *autistic gestures* or *body-focused gestures*) are related to different forms of self-regulatory processes. Furthermore, the substantial empirical evidence that these “on body” gestures are displayed more often with the left hand than with the right (even in right-handers) provides valuable information with respect to their neurobiology. Because spontaneous hand preferences reflect the activation of the contralateral hemisphere (Hampson & Kimura, 1984; Verfaellie, Bowers, & Heilman, 1988), the relative left-hand preference for “on body” gestures indicates right-hemispheric activity. Given that the right hemisphere is critically involved in the generation of the vegetative components of emotional responses and of the concomitant subjective experiences, as well as in emotional communication (Blonder, Burns, Bowers, Moore, & Heilman, 1995; Fernández-Carriba, Loeches, Morcillo, & Hopkins, 2002; Gainotti, 1999; Moscovitch & Olds, 1982), an association between psychological states of stress, right hemispheric activation, and the execution of left-handed “on body” gestures for the purpose of self-regulation seems plausible.

In the clinical domain, as early as 1935, Krout reported that in psychoanalytic interviews, certain “autistic” movements (equivalent to the Module I category “on body”) were reliably elicited through association experiments. Sainsbury (1954) confirmed that “on body” hand movements especially occur in relation to emotionally loaded issues. Other clinicians (for a review, see Joraschky, 1983) suggested that among schizophrenic and autistic patients, the increased body stimulation these gestures represent is an attempt to reinforce bodily boundaries, and thus that these “on body” gestures have a self-regulatory function.

---

**Figure 3.1. Algorithm for functional gesture coding (Module III) (continued on next pages).**
In a field-dependent mode of perceiving, perception is strongly dominated by the overall organization of the field, and parts of the field are experienced as “fused.” In a field-independent mode of perceiving, parts of the field are experienced as discrete from their organized background. The tendency toward one or the other style of perceiving is a consistent pervasive characteristic of an individual’s perception that is related to intelligence, self-image, and reaction to conflicts. Furthermore, the continuous body-focused activity was more prominent in a number of contexts: in interviews with cold interviewers versus those with warm, empathetic interviewers (Freedman et al., 1972); when discussing personal topics versus

The most substantial work on body-focused gestures was initiated by Norbert Freedman, who postulated that body-focused gestures—as compared with object-focused communicative gestures—indicate a withdrawal from communication for the purpose of self-stabilization (Freedman, 1972; Freedman & Bucci, 1981). During free association or in semistructured interviews, field-dependent subjects (Witkin et al., 1954) showed significantly more continuous body-focused activity (cf. the Module I category “continuous”), especially hand-to-hand (cf. the Module II category “act on each other”), than did field-independent subjects (Freedman & Bucci, 1981; Freedman, O’Hanlon, Oltman, & Witkin, 1972; Sousa-Poza & Rohrberg, 1977).

In a field-dependent mode of perceiving, perception is strongly dominated by the overall organization of the field, and parts of the field are experienced as “fused.” In a field-independent mode of perceiving, parts of the field are experienced as discrete from their organized background. The tendency toward one or the other style of perceiving is a consistent pervasive characteristic of an individual’s perception that is related to intelligence, self-image, and reaction to conflicts. Furthermore, the continuous body-focused activity was more prominent in a number of contexts: in interviews with cold interviewers versus those with warm, empathetic interviewers (Freedman et al., 1972); when discussing personal topics versus

Figure 3.2. Algorithm for functional gesture coding (Module III) (continued).
The left-hand preference for "on body" gestures is not only associated with psychopathological states, but also seems to be a general feature of this gesture category. Several neuropsychological studies have reported a left-hand preference for "on body" gestures ("self-touch"), as compared with a right-hand preference for communicative gestures ("free movements"), which are most often executed distant from the body (Dalby, Gibson, & Grossi, 1980; Kimura, 1973a, 1973b; Lavergne & Kimura, 1987). Furthermore, Blonder et al. (1995) reported a left-hand preference for grooming and fidgeting, which contrasted with a right-hand preference for "symbolic" gestures. This pattern of hand preferences seems to be quite stable over the lifespan. Already in babies from 1 to 6 months of age, a left-hand preference for self-touch and right-hand preference for communicative gestures was observed (Trevarthen, 1996). In line with the research findings in adults, Trevarthen suggested that the babies’ self-touch served the purpose of self-regulatory withdrawal, whereas other gestures demonstrated externally oriented activity.

As stated above, the relative left-hand preference for "on body" gestures indicates activity in the right hemisphere (Hampson & Kimura, 1984; Verfaellie et al., 1988) that is well compatible with psychological states of emotional stress and a need for internal regulation.

**Example:** The Module III “panto-image” category and the role of the left hemisphere. In a system-

---

<table>
<thead>
<tr>
<th>Objective Purpose</th>
<th>What is the gesture type?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving an obvious purpose</td>
<td>Self-touch functional</td>
</tr>
<tr>
<td>Body-adjacent manipulation functional</td>
<td></td>
</tr>
<tr>
<td>Object manipulation functional</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position Shift</th>
<th>What is the gesture type?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changing position</td>
<td>Hand shift</td>
</tr>
<tr>
<td>Arm shift</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incomplete Gestures</th>
<th>What is the form?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abortive (gesture or shift)</td>
<td></td>
</tr>
<tr>
<td>Preparation</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td>Retraction</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Several Gestures in a Unit</th>
<th>What is the form?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td></td>
</tr>
</tbody>
</table>

---

These empirical studies all confirmed Freedman’s original hypothesis that “on body” gestures, especially when they occur in a continuous manner, are associated with psychological states of stress, and thus might serve self-regulatory functions. Similar concepts concerning the self-regulating function of “on body” gestures have been put forward in social psychology (“self-adaptors”; Ekman & Friesen, 1969), ethological research (Morris, 1978), and evolutionary anthropology (Wallis, 2004).

The left-hand preference for “on body” gestures is not only associated with psychopathological states, but also seems to be a general feature of this gesture category. Several neuropsychological studies have reported a left-hand preference for “on body” gestures (“self-touch”), as compared with a right-hand preference for communicative gestures (“free movements”), which are most often executed distant from the body (Dalby, Gibson, & Grossi, 1980; Kimura, 1973a, 1973b; Lavergne & Kimura, 1987). Furthermore, Blonder et al. (1995) reported a left-hand preference for grooming and fidgeting, which contrasted with a right-hand preference for “symbolic” gestures. This pattern of hand preferences seems to be quite stable over the lifespan. Already in babies from 1 to 6 months of age, a left-hand preference for self-touch and right-hand preference for communicative gestures was observed (Trevarthen, 1996). In line with the research findings in adults, Trevarthen suggested that the babies’ self-touch served the purpose of self-regulatory withdrawal, whereas other gestures demonstrated externally oriented activity.

As stated above, the relative left-hand preference for “on body” gestures indicates activity in the right hemisphere (Hampson & Kimura, 1984; Verfaellie et al., 1988) that is well compatible with psychological states of emotional stress and a need for internal regulation.

**Example:** The Module III “panto-image” category and the role of the left hemisphere. In a system-

---

<table>
<thead>
<tr>
<th>Objective Purpose</th>
<th>What is the gesture type?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving an obvious purpose</td>
<td>Self-touch functional</td>
</tr>
<tr>
<td>Body-adjacent manipulation functional</td>
<td></td>
</tr>
<tr>
<td>Object manipulation functional</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position Shift</th>
<th>What is the gesture type?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changing position</td>
<td>Hand shift</td>
</tr>
<tr>
<td>Arm shift</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incomplete Gestures</th>
<th>What is the form?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abortive (gesture or shift)</td>
<td></td>
</tr>
<tr>
<td>Preparation</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td>Retraction</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Several Gestures in a Unit</th>
<th>What is the form?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td></td>
</tr>
</tbody>
</table>
atic study on split-brain patients, Lausberg, Cruz, Kita, Zaidel, and Pito (2003) demonstrated that the production of pantomime gestures—the “panto-image” category in Module III—relies on a specific left hemispheric competence. Three patients with complete callosotomy and, as control groups, 5 patients with partial callosotomy and 9 healthy subjects were examined for their ability to pantomime the use of visually presented objects (Module II category “panto-image”) and to manipulate objects with their hands (Module III category “object manipulation functional”). In each condition, 11 objects were presented to the subjects, who either pantomimed or demonstrated the object use with either hand. Two independent raters evaluated their videotaped movements. Although the functional object manipulations were perfect with either hand in all three groups, the split-brain patients displayed conceptual apraxic errors in the panto-image condition with only their left hands. Since functional object manipulation was unimpaired for either hand, both hemispheres must be competent to execute this gesture type, but the isolated right hemisphere was impaired in executing the panto-image gesture type. Thus, the production of panto-image gestures relies on a specific left hemispheric competence.

A recent study on healthy subjects (Lausberg, Heekeren, Kazzer, & Wartenburger, 2009) confirmed the critical role of the left hemisphere in the production of the panto-image gesture type. The cerebral activation patterns during functional object manipulations and panto-image gestures were examined by functional magnetic resonance imaging. In the panto-image condition, there was activation in the left superior/middle temporal gyrus relative to the functional object manipulation condition; that is, this area was activated when either hand executed the panto-image gestures. In contrast, functional object manipulations revealed large bitemporally distributed homologous areas of activations for either hand relative to the panto-image condition. These results suggest that panto-image gestures and functional object manipulations partially differ in their cerebral representations. Specifically, the ability to perform panto-image gestures relies on specialized left hemispheric functions.

The NEUROGES Coding System and ELAN

NEUROGES aims at combining findings from different disciplines into one system of well-defined gesture categories. Widespread application of this system will improve the transparency and comparability of research results. Given the important role of visual digital media in contemporary (neuro)psychological research practice, there is a need for an annotation tool capable of capturing the idea of gesture-type modularity and allowing reuse of the gesture-type definitions. ELAN is such a tool; its controlled vocabularies are apt to store the gesture-type codes, and its template feature allows for the creation of consistent and comparable corpuses. These and other features of ELAN are discussed in more detail in the next section.

Description of ELAN

ELAN (www.lat-mpi.eu/tools/tools/elan) is one of the annotation tools that have come into existence with the maturation of digital media. Originating in the early 21st century, it was designed to support audio as well as video annotation and to be versatile enough to accommodate different fields of research. Whereas some tools are specialized—for instance, for field linguistics or audio-only speech analysis—ELAN has always been a multipurpose, multimodal annotation tool (Brugman & Russel, 2004).

One of ELAN’s predecessor tools, developed at the Max Planck Institute for Psycholinguistics (MPI), was Media Tagger. Two major disadvantages of Media Tagger were that it was a Mac-only application and that its data were stored in a proprietary QuickTime file. By the time ELAN was being developed, both Java, a platform-independent programming language, and XML, a format for structured textual data, had gained ground. These techniques were adopted for the MPI tools. ELAN is therefore available for Windows, Mac OS X, and Linux, and its .eaf file format is XML-schema based. All of the annotation data are in Unicode.

Media handling in ELAN. Technically, the most challenging part in developing the application was the handling of video playback. The Java Media Framework, designed to provide a multiplatform media solution, never lived up to expectations. Simply stated, there is no high-performance, pure-Java media solution. For the Windows version of ELAN, therefore, software components have been developed to integrate a native, Direct Show–based player in ELAN, resulting in improved accuracy and performance. For Mac OS X, ELAN uses QuickTime for Java, which is provided by Apple as a Java bridge to their QuickTime Player. For the Linux version of ELAN, the media files are handled by JMF.

ELAN can display up to four video files that can either be integrated in the main window or detached in a separate, resizable window. It is even possible to associate up to six video files that can each be hidden or shown through a selection menu. A rich set of media controls is available, allowing for stepping through the media with different step sizes. With milliseconds as the time units, ELAN provides for optimal spatial as well as temporal inspection of the media. When only video without sound is used during the annotation task—which is the primary procedure in NEUROGES—segmentation with video-frame precision is sufficient.

Annotations and tiers, the basic building blocks. An annotation in its most elementary form is a textual label or tag associated with a segment of the media, which is defined by a begin time and an end time. Most annotation tasks start with identifying the segments and applying a value to each one. Annotations are grouped in tiers, which are a form of layers; a tier acts as a container for annotations. Annotations on the same tier cannot overlap, and typically annotations on the same tier refer to the same kind of event (speech, gestures, posture, etc.). The user can define and create as many tiers as needed, and the tiers can be grouped hierarchically. When a tier is part of a hierarchy, additional constraints apply to its annotations; for example, an annotation on a child tier has to be contained within the time boundaries of an annotation on its parent tier. In general, tier dependencies are used to add
annotations to annotations—for example, to add glosses to words or morphemes, or to create subdivisions of a larger unit. The advantages of hierarchical tiers are that relations are made explicit and that time alignment is inherited. The advantage of independent tiers is maximum freedom; the only constraint is that annotations on the same tier cannot overlap. The choice depends on the focus of the research and the user’s preference.

For any kind of tier, a user can fill in a number of attributes. In the case of NEUROGES, the tiers are not part of hierarchies; instead, separate, independent tiers are created in which all gesture categories can be coded. Two of the tier attributes are used to create tier groups; the “participant” attribute specifies the subject to which annotations refer, and the “annotator” attribute specifies who created the annotations on a certain tier—that is, the rater.

These attributes are reflected in the naming conventions of the NEUROGES tiers. For example, tiers have names such as A_rh_unit_R1, A_rh_unit_R2, and so forth. The prefix—for instance, “A” or “B”—refers to the participant, and the suffix—“R1” or “R2,” in our examples—to the rater. With tiers for each hand separately and for both hands combined, this can easily lead to dozens of tiers per transcription. When visualizing the data, the tiers can be hidden, and they can be sorted in several ways. This level of detail, expressed by the sheer number of tiers, accommodates assessment of interrater agreement per gesture type, which is considered an important step in validation of the system. Experience with the NEUROGES system so far has led to new requirements for ELAN with respect to such calculations.

Controlled vocabularies, template files, and statistical analysis. A few features of ELAN render it a particularly useful tool for research based on a solid theoretical foundation and a well-defined classification system, such as the NEUROGES coding system with its defined gesture categories. The first of these features is ELAN’s controlled vocabulary facility. A controlled vocabulary is a user-definable list of values that logically belong to the same category or group. By creating a number of controlled vocabularies, each of which holds the values of specific aspects of the gesture categories, an instantiation of the theoretical model can be constructed. In the annotation editing process, the vocabularies are used to present the proper list of values, thus accelerating the transcription task and making it less prone to errors. The NEUROGES coding system contains three controlled vocabularies, some of which contain values that can be applied in a cascading manner, from a more general value to a more specific one—for example, from “phasic” (Module I, Step 2) to “phasic distant” (Module I, Step 3).

Another important feature of ELAN for use with NEUROGES is its template files. An ELAN template file holds the skeleton of a transcription file: the definitions of the tiers and the controlled vocabularies. A template document can be used as the basis for new transcription documents. Documents based on the same template constitute a consistent and comparable set of resources. Sharing templates between teams highly improves the interoperability of research resources. In the same vein, the NEUROGES–ELAN template improves interrater training and agreement. Either two raters can code the same file and then make their codings visible or, alternatively, they can code in separate files and then join their work by using ELAN’s Merge Transcriptions functionality. The templates for the NEUROGES system are available for download through the ELAN Third-Party Resources page.

The search system of ELAN provides the means to construct complex queries on the basis of the structural and temporal relationships between annotations. A query can be executed in either a single file or multiple files, selected by the user. This allows for the researcher to, for instance, quickly find all instances of certain co-occurring events, conveniently open the respective files, and jump to the right position in each video. The search results can be exported to a tab-delimited text file for further processing by an application for statistical analysis—that is, to Excel and then to SPSS.

ELAN promotes the calculation of interrater agreement not only for the NEUROGES values (as described above) but also for the temporal overlaps between gesture units (e.g., for Module I, Step 1) by using the function Create Overlap (which performs a logical AND function) in conjunction with Merge Tiers (which performs logical OR). Thus far, determining interrater agreement in their demarcations of natural units of behavior (i.e., when does a behavior unit start and when does it end?) has been an unsolved problem.

Conclusion

To summarize, substantial empirical evidence indicates that the main NEUROGES kinetic and functional movement categories are differentially associated with specific cognitive, emotional, and interactive functions. The combination of the NEUROGES coding system and the annotation tool ELAN results in an effective tool for research on gestural behavior. On the one hand, ELAN provides the perfect technical basis for applying the NEUROGES coding system, and on the other, experience with the NEUROGES system has provided useful insights for the refinement of ELAN, resulting in new or improved tool functionality, such as fully integrated calculation of contingency tables and determination of interrater agreement.

In a recent study by an independent research group, the NEUROGES–ELAN system proved to be an objective and effective instrument for research on gestural behavior (Sassenberg, Foth, Wartenburger, & van der Meer, 2009).

Author Note

The development of the NEUROGES coding system was supported by German Research Foundation (DFG) Grant LA 1249/1-3. Correspondence related to this article may be sent to H. Lausberg, Dept. of Neurology, Psychosomatic Medicine, and Psychiatry, German Sport University Cologne, Am Sportpark M"{u}ngersdorf 6, 50933 Cologne, Germany (e-mail: h.lausberg@dshs-koeln.de).

References


(Manuscript received November 1, 2008; revision accepted for publication March 13, 2009.)