

Aspects of Digital Gesture Lexicography

Ulrike Lynn

Faculty of Humanities, Technische Universität Chemnitz, Germany

ulrike.lynn@phil.tu-chemnitz.de

Abstract

The multi-modality of speech requires an extension or revision of the applicable lexicography. This article will examine points to consider when designing or planning a gesture dictionary and will attempt to identify features that should be present regardless of the focus of the gesture collection. Where appropriate, these features are compared with similar structures in spoken language lexicography. We will discuss what advantages are to be gained by conceiving of the dictionary in digital from right from the outset and what problems may arise that are unique to a digital edition.

Index Terms: gesture, digital, dictionary, lexicography

1. Introduction

The computer technology that we have available to us today invites new and creative thinking about how a dictionary can be designed and realized. Ideas that were once unfeasible or unrealistic are suddenly possible. We will attempt to incorporate these possibilities into a logically structured conception of gesture lexicography. We will especially be interested in a non-linear presentation with flexible reference and search functions that could ideally be created on the fly to cater to a specific user's needs, and the inclusion of multimedia content to enhance the multi-modal aspects of gesture.

The aspects of lexicography we discuss here are based on the 'Dictionary of Contemporary Physical Contact Gestures in the Mid-Atlantic region of the United States' [1], currently the only dictionary of emblematic gestures, which we will use as a foundation for this article. The goal of that dictionary was to develop a structure for presenting, analyzing and organizing a specific type of gesture. We will take a closer look at the results and see if it is possible to extract and generalize features that could be used in a gesture dictionary regardless of focus.

In the course of compiling the dictionary, the author identified areas of future expansion that, due to constraints of time and space, did not find a place in that work. The possibility of a digital version sometime in the future was mentioned. It hardly seems relevant anymore, even just a few years later, to consider doing a print version of a gesture dictionary being newly planned. As we discuss the aspects of lexicography extracted from the aforementioned work, we will present ideas for how these elements can benefit from digitalization.

We intentionally use the term 'digitalization' as distinct from 'digitization'. This reflects the view that what we seek is more than the translation of a work designed for print into some digital form. This would be the digitization of an existing work into a pdf or even into html form. The work is not designed nor structured to take advantage of digital tools. We can think of digitization as a *static digital* form. Digitalization, on the other hand, would be planning the work from the outset using tools and techniques that the digital realm provides. Creating, in other words, a *dynamic digital*

version that will change, adapt and grow through continued input from both author(s) and users.

Zimmer [9] makes a similar distinction between e-texts and hypertexts. E-texts, in this conception, are published online but do not take advantage of any of the expanded presentation possibilities that the internet allows. They can as well be printed out without any loss of information or usability. Hypertexts, on the other hand, can not exist on paper. They incorporate other media that can be combined and linked in any number of ways. They allow communication on the visual and auditory levels simultaneously and can only be accessed through a digital device.

The contrast between digitization and digitalization takes this a step further. It is not only the incorporation of multimedia elements and the ability to connect information with single-click access that we hope to bring to digital lexicography. We would also hope to add a fluid non-hierarchical structure that can respond to a user's needs as well as the ability to collaborate or incorporate feedback to create a dynamic work. This also avoids the potentially loaded term 'hypertext' which still tends to conjure an image of simple text that is linked to something else.

We will proceed logically through the foundation work to see what may be adapted or generalized to suit a broad range of desired outcomes. Another opportunity for expanding the 'Dictionary of Contemporary Physical Contact Gestures' [1] mentioned by the author is to investigate the historical development of the included gestures. This we will refer to as *etymology* even if it is so only in a broadened sense. With this in mind, we will use this as an example when we need one to discuss applying the generalized principles. We will then examine how each of these structural elements might possibly be improved through the use of a dynamic digital format rather than printed or otherwise static form. In conclusion we will address some questions that arise solely within the context of digitalization and will attempt to answer them.

2. Inclusion Criteria

Any gesture dictionary must be in some sense limited in scope. There will always be a concentration or a specific goal. The gestures selected for inclusion must necessarily be a finite set taken from a much larger pool of potential gestures. This is readily apparent in the context of the physical contact gesture dictionary which deals with a specific type of gesture, those involving physical contact, practiced within a specific geographical region, the Mid-Atlantic United States.

Gestures can be selected for inclusion for any of a number of reasons. They may, for example, fit a certain research concentration or share a specific attribute, such as being solely speech-accompanying, or they may be drawn from a specific culture or region. The decision of what to include may seem too obvious to deserve mention, but it is essential when creating a new dictionary. The selection criteria must be explicitly and exactly defined to prevent the work from expanding to unmanageable proportions, both for the author and for the user. It also makes clear, again for both author and

user, what the main objective of the dictionary is and exactly what audience it is designed to serve. Inclusion criteria can significantly simplify the data collection process, though it is important to note that they can also be a form of preconception and should be treated as such. They have the potential to cause you to see the specific attribute everywhere or perhaps miss it in borderline cases that may subconsciously have already been decided.

A significant advantage to a dictionary in digital form is that a series of dictionaries, each with a different set of inclusion criteria, could be connected as modular units within a larger whole. Different groups of researchers could compile their dictionaries then use a centralized framework to publish their work. This allows a level of integration that would not be possible with individual books nor even with individual digital versions. An example helps illustrate this: research group A publishes a dictionary of gestures in cultural space x using a shared framework. Research group B then publishes a dictionary tracing the etymology of the gestures collected by group A. This information then becomes available as supplemental links within the original work. Should group C then publish a dictionary of gestures drawn from cultural space y within the framework, the possibility for comparative study is greatly expanded.

3. Data Collection

The particular details of this stage are so unique to each project that it is hardly possible to generalize an approach. That it is arguably the most important stage in a work should be apparent to anyone deciding to undertake a gesture dictionary. This is, however, an area that can be radically changed in the context of a digital edition. Gestures are primarily observed directly, whether through live interaction or through recorded visual media. It is rare to find written descriptions that are sufficiently detailed for lexicographic purposes. Modern techniques of data mining that may be suitable for the written word are not yet advanced enough to parse an equivalent amount of information that must be visually extracted. Primary source material, once collected, is then condensed into what becomes the dictionary entries but then usually disappears into the background. Once freed from traditional space constraints of a book format, however, some portion of this supporting material could be accessed, if needed, from the main dictionary entry. This has a parallel in the use of example sentences used to demonstrate use of a word in a spoken-language dictionary.

As we will see later, the context in which a gesture can be found plays a role in the organization of the dictionary. The explanation and presentation of this context information could be well established and presented by multimedia source material. Even something so simple as the frequency of an observed gesture as documented in source materials could help provide a general clue as to the pervasiveness of a particular gesture.

Data collection can also be greatly expanded once a dictionary is planned in digital form. The idea of crowd sourcing information is an interesting one and similar to wiktionary-type trends in spoken language dictionaries. In addition to the same questions of organization, quality control and authorship that arise in that context, a gesture dictionary also presents unique hurdles. The necessarily large amount of photo and video information makes distributed collaboration more difficult than simple editing of text within a browser. Using distributed collaboration could, however, greatly enhance the resolution and accuracy of collecting gestures for inclusion out of a given gesture space. We will come back to

this idea of collaboration later as we discuss potential pitfalls in the digitalization of gesture lexicography.

4. Presentation

Once the inclusion criteria have been defined and a set of gestures collected through suitable means, they must be presented in some logical form. Here we are not talking about the organization of all the entries, this we will discuss in terms of navigation, but rather the presentation of each specific gesture entry. Here we already digress from speech dictionaries in that photo/video material or illustration is more or less required. While a written description of the gesture may or may not accompany the entry, it is almost unthinkable to leave out a visual representation.

When using still photos it is important to capture the essence, Kendon's *stroke* [6], of the gesture. The 'Dictionary of Contemporary Physical Contact Gestures' [1] provides a sequence of still images where the preparation or post-gesture movement sequences are important to understanding the gesture. This problem of selecting the 'right' moment of a gesture goes away in the context of a dynamic digital edition. With only minimally more production overhead, videos can be produced and embedded that depict the entire gesture sequence; from pre-gesture state of rest to post-gesture state of rest. In this way, no decision has to be arbitrarily made as regards the essence of the gesture.

This brings us to the question of where such photo or video material should come from. In our source dictionary, one finds the gestures collected recreated for the purposes of the photos in order to provide a unity of presentation. This is a practical as well as an aesthetic choice. It provides a level of control so that extraneous elements can be eliminated from the scene. It also provides a uniform experience for the user when cross-referencing or comparing gestures. This centralized production of multimedia material, as we will discuss later, is one of the main obstacles to group collaboration. During the data collection phase, however, much additional video and photo material was collected and there is no reason this cannot provide background or additional information as needed. It is nearly unthinkable to plan an etymological gesture dictionary without multimedia depictions of various stages of the gesture development. The changes over time may be too subtle to capture in words or could be made dramatically clear when presented sequentially in the course of a single video.

Depending on the goal of the dictionary, it may also be desirable to shift focus away from or otherwise completely remove certain elements from the multimedia depiction. In the 'Dictionary of Contemporary Physical Contact Gestures' [1], a decision was made to blur the facial expressions of the subjects performing the gestures for the photographic material as this detracted from the analysis of the communicative intent of the gesture and in some cases collapsed the polysemy of the gesture meaning. In deciding to recreate gestures and produce original material it is important for extraneous elements to fade into the background as much as possible so that the foreground is reserved solely for the gesture.

Comprehensive video material may also render written descriptions of a gesture extraneous. However, it must be said that a written description can subtly focus the viewer on the key elements of the gesture that are then reinforced by the visual depiction.

Another possibility provided by digitalization is the simple use of sound in cases of accompanying speech or colloquial spoken renderings of gestures. These could easily be embedded as audio files to enhance the user experience and also to avoid confusion when dealing with the sometimes

creative spelling of interjections or sounds uttered while performing a gesture.

4.1. Naming

Naming the entries should not be taken lightly. We will discuss this in more detail in the context of navigation, but it deserves mention here. The name is the first point of contact for the user with a gesture and establishes it as a formal entity. It should be easy to understand what gesture is meant without going so far as to provide initial analysis. Names can be a brief description of the movement but it is important that they remain on the morphological level without being burdened with meaning. As an example; 'Placing a hand on someone's back' is preferable to 'Demonstrating Support', which would already provide information at the level of meaning.

5. Definition

5.1. Analytical Tools

For a collection of gestures to truly be considered a dictionary, it is important to provide some framework of analysis that allows a discussion of meaning. What exactly this framework is could be unique to each project. For the 'Dictionary of Physical Contact Gestures' [1], the main focus was on the communicative intent and content of the gestures. They were analyzed, therefore, in terms of Austin/Searle's [8] speech-act theory. This provided analytical tools that allowed classification and structured interpretation of the communication. It may be, however, that a specific dictionary is not concerned primarily with communication. By way of example, a dictionary concerned with tracing the development of use-movement into gesture would have a different focus and, therefore, need a different set of analytical tools, perhaps those of philology or semantic shift. In our view, it is desirable to see what can be borrowed from the world of spoken communications. These theories provide a generally more robust context for analysis as they will be better tested than anything developed uniquely for a gesture dictionary. They also strengthen the connection between spoken and non-verbal communication and can provide an element of familiarity. The key, of course, is that a system of analysis be defined and applied consistently. This is ultimately what supplies any dictionary with its usefulness and will play a deciding role in judging its quality.

5.2. Context

It may be important, again depending on the focus of the dictionary, to stipulate certain contexts in which the gesture may or may not be found. This, in our opinion, should be part of the analytical framework in order to have a clear taxonomy of contexts. One option would be the register labels of the Oxford Dictionary to provide predefined descriptions of the tonality of a gesture to help to contextualize them.

It is also advisable to indicate when a gesture belongs to the passive gesture vocabulary of a particular region or group. These would be gestures that are recognized but not actively used. This can be important for completeness when documenting the gestures of a given time and place.

6. Variations

6.1. Expression

Gestures are fluid and their use is constantly developing. It is a given that during the data collection, many gestures will be

observed with slight variations at the morphological level. It is best to identify a stereotype for the gesture. That can be decided through frequency of use or even through simplicity; the unadorned version of the gesture becomes the dictionary entry and any other versions are treated as a variation of that. The decision as to what constitutes a variation must be made by the lexicographer. Taking a handshake as an example; we would hardly consider turning the hand 45 degrees from the vertical plane a variation, though it was witnessed many times. A speech dictionary also does not attempt to document every possible regional difference in pronunciation. Adding a second hand to clasp the shaking hand, though, is different enough to deserve mention but is not, in and of itself, a unique gesture. The list of variations at the level of expression should be comprehensive enough to cover most use cases without being so detailed as to overburden the user. These variations, like the main entry, can benefit from embedding in video form

6.2. Meaning and Use

As seen above, once a gesture has been classified in some way using the analytical tools of choice, it is possible to compare or contrast variations at the level of meaning. In the case of a dictionary primarily concerned with communicative content, one gesture generally has several meaning variations; a range of communicative content intended by use of the gesture. In this case, a single entry will have a number of 'definitions'.

6.3. Combining

Another pertinent fact can be the combinability of gestures. It may be, rather than being an extension or variation of a given gesture, that a gesture with its own individual entry may be combined with another independent gesture in the course of one communicative act. We are speaking here of simultaneous performance of the gestures and not successive performance as would be found in gesture dialogue. The scope of possibilities of successive gesture chains is simply too broad to be a part of a gesture dictionary unless, of course, this is the primary focus of the work. One could imagine, however, that in the context of a digital version complex chains n gestures long might be generated algorithmically then verified through human input. Gestures that are frequently combined should be indicated as such to facilitate cross-reference and to underscore their combined use in a gesture vocabulary. Some gestures are not combinable but it would only be in a special case that it would make sense to indicate non-combinability.

7. Navigation

Non-linear organization differentiates dynamic digital editions from both linear audio and video documents as well as documents designed for print. Here we will see the true advantages of this mode of presentation once a dictionary is freed from the limitations of a print-based edition.

A crucial element of lexicography is the design of the user interaction. Creating a navigation structure demands understanding the target audience and thinking through a variety of use cases. Each project will have different demands, again based on the focus of the specific dictionary. It will likely prove beneficial to have gestures be locatable by using either the expression or the meaning levels. That means if a user has seen a gesture but doesn't know what it means, they can search based on the physical action. This can be through the titles which, we have seen, are a brief description of the movement involved.

Should a user, however, want to collect all gestures that communicate something specific, they could consult an index

which categorizes the gestures according to general communicative intent. The following categories are to be found in the 'Contemporary Physical Contact Gestures' [1]: Greeting Gestures, Attention Gestures, Confirmation Gestures, Institution Gestures, Consolation Gestures, Encouragement Gestures, Affection Gestures, Attraction Gestures, Sexual Relation Gestures, Assistance Gestures, Aggression Gestures, Playful Power Gestures and Indirect Contact Gestures.

Whether or not these categories are useful to a specific project is irrelevant. The main point is that there should be some sort of organization based on the level of meaning that allows gestures with similar communicative intent to be easily discovered and grouped.

7.1. Linking

Clearly one of the main advantages to digitalization is to be found in dynamic linking. Instead of a table of contents or an index full of page numbers, each entry is simply accessible as a link within the current view. This also obviates the need for a numbering system which can have the impression of giving the entries a hierarchy that is not intended. Additionally, a print edition can only have a limited number of navigation methods. Otherwise the problem becomes how to notate them all for a user and still retain some sense of order. If one gesture links to ten others because of their similarity of form and another twenty due to their similarity of meaning, there is hardly room for any other navigation criteria in the jungle of page numbers and cross-references. This is, naturally, not a problem in digital form. As each gesture is only one click away from any other, there is no longer reason to avoid bringing any two in relation when the user no longer needs to flip endlessly through a book. It is easy to imagine in addition to two main navigation schemes - at the expression and meaning levels - there being any number of navigation matrices that are customized to a user's needs. As they would only be treating the dictionary entries as a dataset, they could be added and removed after the fact without altering the work itself. Each type of navigation could remain in place and be dynamically populated with the content appropriate to the gesture currently being viewed.

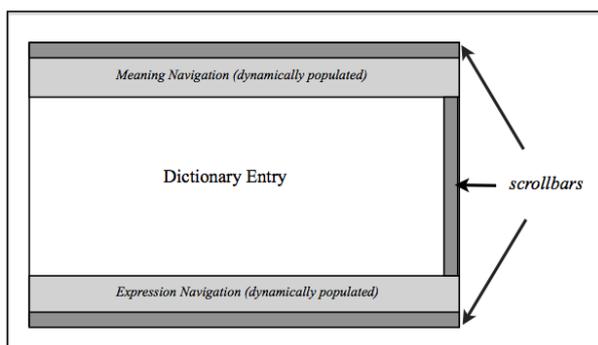


Figure 1: Example of a possible navigation scheme.

7.2. Performance Navigation

Yet another possibility in the digital realm, though perhaps not immediately realizable, is navigation of a gesture dictionary simply by performing the gesture in question. This is, naturally, very much dependent on the final delivery medium of a digital gesture dictionary. Assuming, however, that it is designed to be viewed on a camera-equipped device, it is not hard to imagine gesture performance-based navigation. With sinking costs of such technology and commercial products

such as the Kinect or the Myo easily available, it will not be long before this is feasible.

8. Discussion

We have seen along the way that there are many advantages to be gained by conceiving of gesture dictionaries in digital form. There are, however, some issues that need to be addressed. In many ways, these are the same problems that are currently the focus of so much debate in the digital humanities in general. Any new plan to create a gesture dictionary, however, should at least consider ways to mitigate these issues even if they can not be solved completely.

8.1. Medium

The first question is in what format the dictionary should be made available. It is one thing to say 'a digital version rather than print' but that implies that there is only one possible digital version. Generally, saying this means wanting to build a website around the content. But even this is not as straightforward as would be hoped. First of all it is important that the dictionary remain accessible and compatible for as long as possible. This may not be that difficult in the case of purely text as most speech dictionaries are, but even just talking about embedding video complicates the issue. A simple web version has to support the wide (and ever increasing) variety of devices with which users will access it. It should, in the interest of longevity and compatibility, adhere to web standards. All multimedia content needs to be unencumbered by intellectual property issues and should, ideally, be in a format that is also free to use, and that will remain so. An active site will need server space and bandwidth, and that for the entire planned life of the work. It will also require maintenance, both routine and proactive to avoid obsolescence.

Which brings us to:

8.2. Competence

In planning a publication team, it is important to have people with hard digital skills. In concrete terms, this means programmers, possibly interface designers and someone who can 'publish' and maintain the dictionary. If this means interdisciplinary collaboration, that is all the better. It does, however, need to be considered when budgeting and planning.

It should also be taken into account that if the gesture researcher(s) themselves are not doing the programming that there is a possibly detrimental time-lag between idea and realization that should also be taken into account.

8.3. Collaboration

As touched upon briefly above, digitalization provides robust and powerful tools for collaboration. These tools, however, can not simply be blindly implemented. There is the primary question 'who has access to these tools?' That is to say, with whom do we want to collaborate? Naturally if one is designing a dictionary focusing on a specific culture or region, it would be helpful to have input from its members as regards to the performance and interpretation of meaning of gestures. But when the net is cast too wide, the quality of the information suffers. Not only that, as we saw in section 2, a very specific definition of the target gestures is essential to the success of a project which will also potentially limit the group of potential collaborators.

Perhaps contribution to the dictionary is strictly limited to users; scholars, researchers or academics. Who then decides who has access and what are the restrictions in place to

prevent abuse or vandalism? In many ways, these questions have been addressed by many wiki-based projects. But as already noted, what happens when editing is not simply a matter of typing text into an input field? How would crowdsourced multimedia content be produced? Would it need to be centrally produced to the specifications of the 'crowd'?

This brings up the problem of authorship. For better or worse, the academic field is a meritocracy. If a group of authors sets a work loose for digital collaboration, how much credit do they get? Will an institution support a project with running maintenance costs that is so widely distributed as to bring no immediate prestige gain? A new proposal by the creator of the wiki concept for so-called *federated wikis* attempts to alleviate some of this confusion by preserving articles intact but allowing them to be changed when incorporated into another's work. (<https://medium.com/backchannel/the-failed-promise-of-deep-links-aa307b3abaa5> and <http://fed.wiki.org/>)

One can imagine a mechanism by which users can easily submit feedback regarding the content or usability of a digital dictionary. This would then need to be curated and, as needed, implemented by the authors. Or, as in section 7.1 - Linking, there could be an interface allowing custom interaction with the dictionary as dataset. Inventive methods of accessing or cross-referencing the data could then potentially be integrated into the dictionary itself. Another option would be to have the possibility to preserve a user's *link trail* - the path they chose on the fly to navigate the dictionary - as an option that can be shared with future visitors. This would be a way of sharing not just a link but access to a specific interaction with the information. Yet another variation would be to have the data collection phase open for collaboration with as few barriers to users as possible. Once the phase is deemed completed, the collected input is collated and published by the authors in a much more restricted way.

A dictionary can, of course, be authored by a core group and the collaboration limited in scope so as not to cause any confusion as to authorship. Our only point is that the level and mechanism of collaboration should be discussed right from the start.

9. Conclusions

We have proposed a series of modular characteristics that should be present in any type of gesture dictionary and discussed ideas for how to enrich a dictionary through the use of currently available digital tools. We made a distinction between dynamic and static digital editions in which the dynamic fully embraces the possibilities for interaction, multimedia content, collaboration and fluidity of presentation that technology allows. A static digital edition, on the other hand, is rooted in print-based thinking and is, in our opinion, no longer relevant in our increasingly interconnected lives. Some of the issues that arise in planning and implementing such a dictionary have been presented with the understanding that, as the field matures and more works are published in this manner, some of these problems will disappear while others, inconceivable at the moment, may take their place. We have the possibility now to realize ideas in lexicography that until now were impractical or even outlandish. The nascent field of gesture lexicography is, unburdened as it is with the residue of a long history, is the perfect place to stretch the limits of what is considered possible in creating a reference work.

10. References

- [1] Author, "Keep in Touch – A Dictionary of Contemporary Physical Contact Gestures in the Mid-Atlantic region of

the United States", Phd Dissertation, Technische Universität Berlin: Digital Repository of Technische Universität Berlin. Online: <https://opus4.kobv.de/opus4tuberlin/frontdoor/index/index/docId/3484>, 2011

- [2] Author, Gestures in dictionaries: Physical contact gestures. In: C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill and J. Bressem (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (HSK 38.2). De Gruyter Mouton, 2014. 1502-1511
- [3] Author, Levels of abstraction. In: C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill and J. Bressem (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (HSK 38.2). De Gruyter Mouton, 2014. 1702-1712
- [4] Fricke, E., J. Bressem and C. Müller, Gesture families and gestural fields. In: C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill and J. Bressem (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (HSK 38.2). De Gruyter Mouton, 2014. 1630–1640.
- [5] Kendon, A., Die wechselseitige Einbettung von Rede und Geste. In: C. Schmauser, T. Noll (Hrsg.), *Körperbewegungen und ihre Bedeutungen*. Arno Spitz, 1998.
- [6] Kendon, A., *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [7] Noll, T., Wie blättert man in einem Gestikon? In: C. Schmauser, T. Noll (Hrsg.), *Körperbewegungen und ihre Bedeutungen*. Arno Spitz, 1998.
- [8] Searle, J., *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.
- [9] Zimmer, E. D., *Die Bibliothek der Zukunft*. Hoffmann und Campe, 2000.