Is there a way to make a meaningful shape from a collection of the individual words of the English language? What would that shape look like? Why would this approach be preferable to a traditional representation – a printed dictionary, for example? These were among the questions we asked when we set out to visualize the English language as a single entity.

The first assumption we made was that our daily interaction with three-dimensional objects would give us leverage in establishing meaning through familiarity. We set out to create an object that would, as a result of regular observation, begin to afford the same level of information about its state as the plant you observe on your breakfast table each morning.

The initial step in the process of developing the formal attributes of the object was the identification of salient ways to categorize our data (individual words). Because the intent of the project was to propose a system that would function primarily as a reference tool, the dimensions we chose were tailored both to enable the retrieval of a particular word and to facilitate the recognition of broad trends throughout the base of data. Presumably, we are all used to locating words in a dictionary by wading through an alphabetic ordering. Based largely on this familiarity and the efficiency this familiarity affords, we chose alphabetic organization as our first axis of description.

The remaining axes of description, or "descriptors," were chosen for their perceived ability to contribute to the overall expressiveness of the resultant form: word familiarity and temporal etymology. Determined on a word-by-word basis, the most likely methodology to be used in the measurement of the familiarity of a particular word would be based on the number of times that word appeared in a statistically significant corpus. Temporal etymology would be determined by the first recorded use of a word. Once plotted along the axes, the polygonal representations of the words would be packed along their vectors to the origin. Additional expressiveness would be added to the resultant form through the assignment of color to a particular word based on the national origin of that word: for instance, words with a Germanic root would share a common hue.

The interface to this reference tool would allow the user to extract information specific to a particular word such as its definition, etymology, or illustration (if appropriate). One could also use Gradus to browse synonyms and antonyms, replicating the essential functionality of a thesaurus. The emergent form would be tornado-like – relatively few words at the tail of the object (the beginning of the language) and a burgeoning of words following the industrial revolution through the 20th century. Herein lies one of the major advantages of this approach over traditional means: unlike printed dictionaries, Gradus offers information about its content through its form.

Gradus, as an expressive, sculptural representation of the English language offers an efficient way to develop an understanding of the nature and the history of specific words and the language as a whole.

Pages from the oldest extant (1689) Gradus ad Parnassum. Printed on vellum, there were only fifty impressions made of this edition. Noted

Matt Grenby
MIT Media Lab
20 Ames Street
Cambridge, Massachusetts 02139 USA
grenby@media.mit.edu