

Why are Analogue Graphics and Natural Language both Needed in HCI?

Niels Ole Bernsen

Centre for Cognitive Science (CCS), Roskilde University
PO Box 260, DK-4000 Roskilde, Denmark
phone: +45 46 75 77 11; fax: +45 46 75 45 02
email: nob@cog.ruc.dk

The combined use of language and analogue graphics for the expression of information probably is as old as language itself. The paper addresses the question why we need both the expressions of natural language and analogue graphics for the representation of information. It is argued that analogue graphics and natural language have the complementary expressive virtues of specificity and focus, respectively. Their corresponding lack of focus and specificity, respectively, explain why (a) both have developed a number of mechanisms for coping with these deficiencies and (b) why their combination may often have superior expressive power. Since specificity follows from the analogue character of analogue graphics rather than from their graphic character, analogue sound and touch representations are analysed to explore whether results from the analysis of analogue graphics and their complementarity with natural language can be transferred to other analogue modalities of expression. The paper exemplifies the comparatively new field of Modality Theory.

1. INTRODUCTION

Natural language can be used to represent virtually anything and it may therefore seem enigmatic why analogue graphical expressions are sometimes preferred to natural language expressions for certain representational purposes. On the other hand, once the expressive power of analogue graphics has been realised, it may become less evident why natural language representations would ever be needed if it were not for the fact that speaking or writing is often more practical than drawing or creating animations and videos. The answer to these two questions seems to reside in two complementary features of natural language and analogue graphical expression. The features are that natural language expressions are *focused but lack specificity* while analogue graphical representations are *specific but lack focus*. This paper attempts to clarify the issues involved and to explore some of the consequences of the basic distinction between specificity and focus.

The work described forms part of the European ESPRIT Basic Research project GRACE which ultimately aims at providing a sound theoretical basis for usability engineering in the domain of multimodal representations. Whereas the enabling technologies for multimodal (including virtual reality) representation are growing rapidly, there is a lack of theoretical understanding of the principles which should be observed in mapping information from some task domain into presentations at the human-computer interface in a way which optimises the usability of the interface, given

the specific purposes of the computer artifact being designed. Part of the research agenda of GRACE is to analyse in depth the differences in expressive power between different generic representational modalities such as (spoken or written) natural language and analogue graphics (Bernsen 1993b, c).

The plan of the paper is as follows. Sect. 2 provides the concepts needed in the analysis to follow. Sect. 3 presents the distinction between specificity and generality. Sect. 4 presents the distinction between focused and unfocused representation. Both specificity (*cum* lack of focus) and focus (*cum* lack of specificity) are representational virtues, and their respective representational implications are described in Sects. 5 and 6. Since both representational virtues have their corresponding weaknesses, it is not surprising that the widespread use of natural language and analogue graphics has led to the invention of mechanisms which to some extent serve to remedy those weaknesses (Sect. 7). On the other hand, given those weaknesses, one obvious way of trying to eliminate them is to combine the representational modalities of natural language and analogue graphics into multimodal representations (Sect. 8). The representational virtue of specificity in analogue graphics turns out to be due not to their graphical character but to their analogue character. Analogueness, however, is a property not only of graphics but of other representational modalities as well. The implications for sound and touch are explored in Sect. 9. Finally, Sect. 10 is a review of results.

2. SOME RELEVANT CONCEPTS

Some of the central concepts we shall need are explained in this section.

2.1 External and Internal Representations

The representations or representational modalities we shall be dealing with are primarily *external* representations, that is, they are embodied in some medium of expression such as graphics, acoustics or haptics, and are hence external to the human cognitive system and intersubjectively accessible. This is true of written or spoken words and sentences and of analogue graphics on computer screens or on paper. External representations are considered as representations by the human cognitive system and are primarily, as far as we are concerned, produced by data structures in computers and other items of information technology. It is important not to confuse external representations with the representations which are *internal* to the human cognitive system. Spoken or written natural language, when considered as external representations, are generally non-analogue. This does not preclude that the internal representations evoked by natural language are to some extent and in some sense analogue representations. External representations are interpreted as representations by an observer, and interpretation is an internal cognitive process. The properties of specificity and focus central to this paper derive from the fact that natural language and analogue graphics provide very different means of supporting the interpretation of external representations. For this reason, we cannot avoid the issue of internal representations entirely in what follows.

2.2 Analogue and Non-Analogue Representations

The distinction between *analogue* and *non-analogue* (external) representations designates the difference between representations, in whatever modality, which represent through sharing at least one dimension of information with what they represent and representations which represent through conventional pairing between representation and what is represented. Most analogue representations, such as photographs or diagrams, share many dimensions of information with what they represent, whereas

others, such as graphs, share only one or a few dimensions of information with what they represent. As long as we focus only on external representations, the analogue/non-analogue distinction is clear in most cases. In practice, however, the distinction sometimes can be difficult to draw primarily because of the existence of *levels of abstraction* in analogue representation, whether the representation be a sound, a piece of graphics such as a diagram or a tactile/kinaesthetic one. A highly abstract diagrammatic representation, say, of a computer network showing servers, terminals, wiring, etc., may have so few recognisable topological similarities with what it represents that it may just as well, arguably, be considered a non-analogue representation of what it represents. The less recognisable similarity there is between what is represented and its representation, the more we may have to rely on additional knowledge of the representational conventions used in order to decode particular representations. In the limit, where we find, i.a., natural language, we have to rely exclusively on representational conventions.

Another problem in applying the analogue/non-analogue distinction is that it is sometimes unclear how *real* are the states of affairs which appear to be represented in analogue representations. The equator, for instance, is nearly always represented on maps, but what does this representation correspond to? An arbitrary triangular icon, on the other hand, perhaps resembles many triangular shapes to be found in nature or culture, so is it really arbitrary after all or is it rather a highly abstract analogue representation? These two examples may be distinguished using the criterion that the equator on the map does represent a fixed topological property of the globe whereas the triangular icon really is intended as being arbitrary - one might just as well have used a circle or something else again. What matters are exclusively the representational conventions imposed on it. In any case, the 'reality' represented in analogue representations is certainly more comprehensive than the tangible world of spatio-temporal objects, situations, processes and events. In another example, a conceptual graph does have a topology but in this case it appears justified to assume that the topology is not an analogue representation of conceptual relations because such relations do not themselves appear to be topological. Conceptual graphs, therefore, are non-analogue diagrams. However, it is not evident at this point that the topology criterion just described will be able to resolve all problems about the analogue versus non-analogue character of particular external representations. We may have to accept the existence of an undecidable 'grey' area between analogue graphical diagrams and non-analogue graphical diagrams which are often alternatively called 'abstract' or 'conceptual' diagrams. The sound and touch domains may pose similar decidability problems.

2.3 Arbitrary and Non-Arbitrary Representations

The distinction between *non-arbitrary* and *arbitrary* representational modalities marks the difference between external representations which, in order to perform their representational function, rely on an already existing system of meaning and representations which do not do so. The reason why this distinction tends to be overlooked is that, in most cases, it coincides with the distinction between analogue and non-analogue representation. For the purpose of this paper, however, it is important to note that the external representations of spoken and written language constitute exceptions to this rule. They are non-analogue *and* non-arbitrary.

The separation between the analogue/non-analogue distinction, on the one hand, and the arbitrary/non-arbitrary distinction, on the other, does seem quite important. It provides a broad and intuitive justification of why natural language can compete successfully with graphics for many representational purposes in human-computer interfaces and elsewhere. Despite being non-analogue considered as a form of external representation, natural language builds on an already existing system of meaning. If one

does not understand the particular natural language in question, one does not have access to its corresponding system of meaning, but the system of meaning 'is' there nevertheless. And the separation between the analogue/non-analogue and arbitrary/non-arbitrary distinctions demonstrates that explanations of why, e.g., natural language modalities are in some cases inferior, and in others superior, to analogue graphical modalities cannot simply be provided by appealing to the analogue/non-analogue distinction.

2.4 Representational Modalities

We need not go deeply into the question of what is 'really' an external representational *modality*. The problem is not that the question is particularly difficult to answer but, rather, that the term 'modality' is being used in widely different ways in the literature. Explicating one's favoured sense of 'modality', therefore, is both an exercise in contrastive semantical decision-making and an effort in conceptual analysis. Elsewhere (Bernsen 1994), a 'pure' (or unimodal) modality has been characterised as consisting of a specific *medium* and a *profile* constituted by its properties as selected from the following list of binary opposites: analogue/non-analogue, arbitrary/non-arbitrary, static/dynamic, linguistic/non-linguistic. A 'medium' is a physical substrate having a set of perceptual qualities accessible to humans such as a set of visual properties. 'Pure' modalities can be combined into multimodal representations. Given this conceptual apparatus, e.g., spoken language, written language and analogue static graphics come out as different pure representational modalities. It is possible that the current confusion surrounding the notion of 'modality' in the literature is due to the assumption that modalities are entities characterisable through one single property, if we could only identify that property. By contrast, the medium/profile notion of modalities assumes that modalities are complex-property entities.

2.5 The AG Domain

We know that natural language is capable of representing virtually everything, including 1-D, 2-D and 3-D spatial domains, the temporal domain and both concrete non-spatial and so-called abstract (non-spatial) domains. Analogue graphics can represent that to which they have an analogue relationship, i.e., the spatio-temporal domain, temporal events and processes being of course best represented in dynamic analogue graphics. The discussion below deals with natural language representations of the representational domain of analogue graphics which for the sake of brevity may be called *the AG domain*. It is important to note that the AG domain is significantly broader than the domain of access of human vision. Scientific visualisation, for instance, enables the visualisation of many spatio-temporal domains to which human vision has no access, such as intonation patterns in spoken language.

2.6 Limitations to Analogueness

Analogue graphics are not analogue, purely and simply. There are always limitations to the analogue mapping between analogue graphics and what they represent, even in the cases of photographs and videos. These limitations derive from aspects such as the *degree of selective abstraction* of the graphics, their *degree of resolution* and their *spatial dimensionality*. Yet other sources of lack of analogue mapping between analogue graphics and what they represent should be disregarded here since they deal with different phenomena. One example is one and the same object being viewed from one *perspective* and analogously represented from a different perspective. Another, one and the same object being viewed from one *distance* and analogously represented from a different distance.

3. SPECIFICITY VERSUS GENERALITY

One of the two related, main differences between the respective representational powers of natural language and analogue graphics seems to be the specificity of analogue graphics vs. the generality of natural language expressions.

Natural language represents the AG domain through its individual expressions drawing upon an arsenal of more or less shared, general and stereotypical internal representations (or concepts) based for the most part on common visual experience. Being general and stereotypical, these representations always leave open and undetermined a certain *interpretational scope* (cf. Bernsen & Svane 1994). A description in natural language normally leaves out a wealth of individual features of the entities in the AG domain which it describes. Recipients may or may not mentally or otherwise fill in by themselves the details omitted in the description and thereby exploit or avoid to exploit the interpretational scope of the description. The term 'interpretational scope' should in this context be interpreted in a rather strong sense. Our general concepts may be structured in many ways as frames, scenarios, scripts, image schemata, etc., but, strictly speaking, even conceptual features such as defaults belong to the interpretational scope of concepts rather than to their core meaning.

The interpretational scope of a particular description in natural language can be incrementally narrowed and determined through the addition of further linguistic expressions. In the AG domain, however, this process tends to be lengthy and complex whenever the aim is to render all the properties of the entities being described. Arguably, the expression-addition process will in principle never succeed in providing a representation which is informationally equivalent to an analogue graphical representation of the same entities. One way of ensuring informational equivalence would be to require that the natural language description allows an exact and intersubjective, analogue graphical reconstruction of the AG domain described. One may attempt to devise exceptions to this principle of non-equivalence, but even if such exceptions do exist they will be unimportant by comparison to the domain where the principle holds. For instance, it might, perhaps, be possible to reconstruct the informationally equivalent analogue graphics corresponding to the following description: "A perfect circle with a diameter of 2 centimeters drawn in completely black ink and in a perfect 1 millimeter wide brush stroke". Even in this case more needs to be said on, e.g., the nature and structure of the surface on which the analogue graphics were drawn.

Analogue graphics, on the other hand, represent the AG domain through representing details of individual entities. Analogue graphics represent *that over which* the corresponding, abstract and general natural language expressions are abstractions and generalisations. To be sure, the extent to which this is the case depends on the degree of abstraction of the analogue graphics used, on their degree of resolution and their spatial dimensionality. However, to the extent to which analogue graphics represent individual detail, no interpretational scope is left open. In this sense, analogue graphics are *specific* as compared to the corresponding natural language expressions, and independently of the degree of abstraction of the graphics and their degree of resolution. Remember that we are always comparing a piece of analogue graphics with its corresponding natural language description or descriptions. This having been said, it is of course the case that, to the extent that analogue graphics embody some degree of abstraction and lack of resolution, they themselves leave open an interpretational scope. So the fact that both natural language and analogue graphics may leave open an interpretational scope should not be misconstrued as stating that, given a certain level of abstraction of a piece of

analogue graphics, its meaning may be identical or informationally equivalent to that of the corresponding natural language expression. This is virtually never the case. However abstract a piece of analogue graphics is, the meaning it expresses is always more specific than that of the corresponding linguistic expression. In a simple example, there are infinitely many specifically different graphic ways of representing an angle of 60 degrees. These all fall within the interpretational scope of the otherwise exact natural language expression 'an angle of 60 degrees'.

The example just provided helps clarify what is meant here by 'the linguistic expression corresponding to a piece of analogue graphics'. In using natural language we hardly ever attempt to go to the length of trying to provide descriptions which are informationally equivalent to some analogue graphical representation in the AG domain. Instead, we use expressions such as 'an angle of 60 degrees' and such expressions are sufficient for the communicative purpose at hand. However, such expressions leave open large interpretational scopes. Had we been using analogue graphical representations for the same communicative purpose instead, parts of the interpretational scope left open would have been closed.

Speaking now of *internal* representations, it would seem to follow directly that the internal representations to be posited by cognitive science as constituting the general meaning or sense of natural language expressions are not 'analogue' in the sense in which analogue graphics is analogue. These meanings or senses are generally like variables rather than constants. This is how they succeed in subsuming indefinite numbers of specifically different instances. And since the contents of our perceptual experience are never like variables but always consist of specific instances, the general meanings of natural language expressions cannot be analogue. A *specific mental model* created by some individual of a state of affairs in the AG domain which has been expressed through a general expression in natural language, on the other hand, might well be analogue in more or less the sense of analogue graphics. However, such a mental model would be one which exploited the interpretational scope of the natural language expression in question. To avoid any misunderstanding it may be pointed out here that, just like analogue graphics, mental models of entities in the AG domain may be quite abstract and low-resolution and do not have to incorporate more specificity than done by the most selectively abstract piece of analogue graphics.

The distinction between specificity and generality may be said to reflect a difference between 'direct' and 'indirect' external representation. Natural language represents the AG domain *indirectly* in the sense of representing *via* the general concepts of natural language. Analogue graphics, by contrast, represents the AG domain *directly* in the sense of not having to represent this domain *via* general concepts.

4. UNFOCUSED VERSUS FOCUSED REPRESENTATION

The second main difference between the respective representational powers of natural language and analogue graphics seems to be the focused nature of the representations provided by natural language vs. the unfocused nature of analogue graphics.

Natural language expressions and descriptions are focused as compared with the corresponding analogue graphics. This contrast is closely related to that of the generality of natural language expressions vs. the specificity of analogue graphics. Natural language expressions focus on a particular aspect of what is being represented and leave open an interpretational scope. Analogue graphics close the representational scope and, for that very reason, do not focus. In the example of Sect. 3 above, the purely graphic representation of an angle of 60 degrees does not tell whether it is the 60 degrees which

matter (and they would have to be measured first) to the representer, whether what matters is the fact that an angle is being represented or whether what is being represented is something third. When, on the other hand, natural language is being used to state the fact that an angle is 60 degrees, no irrelevant detail is involved and the statement is focused. It is important to note that 'focused' does not imply 'picking out a particular detail'. 'Focused' does imply picking out something which is then being expressed, but it need not be a detail of a larger whole and might just as well be the larger whole itself. Focusing, in other words, may operate at any level of detail.

Specificity implies that many different representational purposes may be satisfied by one and the same analogue graphic representation which, therefore, remains unfocused until further information has been provided. The viewer may happen to focus on particular aspects of the analogue representation but, barring contextual implications, is in no position to know if this is the focus intended by the representer. Focusedness implies that only one representational purpose is at least, and normally, being intended which, therefore, remains otherwise unspecific and leaves open an interpretational scope.

5. IMPLICATIONS OF SPECIFICITY

Representational specificity is a powerful property of analogue graphics. This section explores some of its implications in terms of useful properties of analogue graphics. So far, no principle has been found which might help establish an exhaustive list of such implications. Implications are stated in a somewhat coarse-grained format leaving out more or less obvious qualifications which would need to be made in an exhaustive presentation. In this and the following section (Sect. 6), the reader should bear in mind that we are only speaking about implications of specificity and focusedness, respectively. That is, we are only dealing with the strengths of representation deriving from these properties in analogue graphics and natural language, respectively. In Sect. 7 we shall take a full view of analogue graphic and natural language expressions.

5.1 Representational Exhaustiveness

The potential of analogue graphics for achieving representational exhaustiveness, or one-to-one mapping with what is represented, follows from their specificity and is limited by their dimensionality as compared with the dimensionality of what is represented as well as by their degrees of abstraction and resolution. A 2-D map, for instance, cannot provide 3-D specifics; or a piece static graphics such as a process diagram (Bernsen 1993a), while somehow capable of representing movement, cannot provide its specifics. Process diagrams seem to represent movement and processes through the way they are being read (or interpreted) by people who use their domain knowledge to exploit the interpretational scope of the diagrams. Given their lack of specificity, the internal representations evoked by natural language expressions lack the potential for representational exhaustiveness.

5.2 Smooth Mapping

The specificity of analogue graphics allows them to smoothly map what is to be represented into the representation, their smooth mapping potential only being limited by their dimensionality and degrees of abstractness and resolution. Smooth mapping preserves whatever continuous transitions between properties are needed for the representational purpose at hand. The internal representations evoked by natural language expressions, being general and having an interpretational scope, lack the property of smooth mapping.

5.3 Direct Measurement

The specificity of analogue graphics allows direct measurements to be performed on the representation, which reflect the properties of what is represented. The potential for direct measurement is bounded by dimensionality and by the degrees of abstractness and resolution of the graphics. The internal representations evoked by natural language expressions, being general and having an interpretational scope, lack the property of direct measurement.

5.4 Approximate Inference

The specificity of analogue graphics allows approximate inferences to be performed on the representation, which reflect the properties and qualities of what is represented. The potential for approximate inference is bounded by dimensionality and by the degrees of abstractness and resolution of the graphics. Natural language expressions, being general and having an interpretational scope, lack the property of allowing approximate inference. However, natural language expressions do allow a form of approximate inference via the stereotypical concepts they evoke. Such inferences can be performed as well on the corresponding analogue graphics.

5.5 Direct Entity Identification

The specificity of analogue graphics provides the informational basis for subsequent direct identification of the particular entities represented. The generality and stereotypical character of the internal representations evoked by natural language expressions makes difficult the subsequent identification of the particular entities represented. This is why the police prefers photographs of robbers to linguistic descriptions. It is true that we manage pretty well in everyday life with natural language expressions for entity identification. The reason why we do so seems to be the widespread use of (linguistic) indexical reference, definite description and proper names (see below). The police would normally prefer to know the full name of a robber rather than his or her linguistic description.

5.6 Easy Update Connectivity

The introduction of a new entity into a piece of analogue graphics immediately allows an updating of its spatial (or spatio-temporal) relationships to all other entities represented. The introduction of a new piece of information into a series of natural language expressions describing something in the AG domain enforces the use of more or less complex inferences in order to update the internal representation of what is being described.

5.7 Substitution for Direct Experience

The specificity of analogue graphics means that they can be used as substitutions for direct perceptual experience, for instance in enhanced reality or virtual reality technologies. The internal representations evoked by natural language expressions, being general and having an interpretational scope, lack this property.

6. IMPLICATIONS OF GENERALITY AND FOCUSEDNESS

Just as representational specificity is a powerful property of analogue graphics, generality and focusedness are powerful representational properties of linguistic expressions. This section explores some of their implications in terms of useful properties of natural language. So far, no principle has been found for establishing an exhaustive list of such implications.

6.1 Abstraction

Abstraction allows natural language to 'directly' represent abstractions over experience in the AG domain. Such abstractions cannot be represented in analogue graphics. A simple example is that it is impossible to graphically represent colour in general. Because of their inherent specificity, analogue graphics have a limited potential for representing abstractions as compared to the corresponding natural language expressions. This is a profound advantage of linguistic expression which seems to reflect the fact that our repertoire of internal representations includes a large number of general and stereotypical concepts in the AG domain in addition to specific mental models. Such concepts can be conceived of as organised into abstraction hierarchies. The colour green, for instance, is already an abstraction which cannot be represented as such in analogue graphics. At a higher level of abstraction, the concept of colour subsumes all our abstract concepts of individual colours. At a still higher level, the concept of visual properties of entities (almost) subsumes the abstract concept of colour together with other concepts. Natural language allows us to freely focus on the appropriate level of abstraction.

Whereas colour in general cannot be represented in analogue graphics, the full colour spectrum can be represented in analogue graphics to an arbitrary degree of resolution. Generalising this observation, it would seem that any part of the AG domain can be represented in analogue graphics, to an arbitrary degree of exhaustiveness, namely as collections of specific instances. After all, the AG domain concepts of natural language are built from specific observed instances by the neural circuitry of the brain. However, communication in natural language would be impossible if we always had to include information on such specifics. Instead, natural language makes it possible to navigate freely at the abstraction levels above the specifics in the AG domain to realise particular communicative purposes at the constant price of operating within an interpretational scope.

6.2 Relevance Decidability

Given their non-focused character, it can be difficult to decide with respect to a piece of analogue graphics what is and what is not relevant to a specific representational or communicative purpose. It can therefore be difficult or impossible to identify the representational purpose behind a piece of analogue graphics in the first place. Given their focused character, the corresponding natural language expressions do not raise this problem. This is not to deny, of course, the existence of irrelevant discourse. But natural language is '*made for relevance*', i.e., for making relevant descriptions at appropriate levels of abstraction. Relevance does not pose a problem for linguistic expression in the sense in which specificity poses a problem for natural language. There is reason to believe that far more cases of communication error arising through the use of natural language arise from lack of specificity than from lack of relevance (cf. Bernsen & Svane 1994). "Be (sufficiently) specific!" is a much more important injunction to include in a practically oriented set of conversational postulates than is the injunction "Be relevant!".

6.3 Beyond the Analogue Media

Natural language expressions can represent many types of entity which lie outside not only of the AG domain but outside of the representational potential of external

analogue media as a whole, including highly abstract concepts such as 'truth' or 'justice'. Given the notion of abstraction hierarchies of Sect. 6.1 above, even such concepts would seem to have some basis in specific occurrences. However, the properties of those specific occurrences that make them suitable for creating abstractions such as 'truth' or 'justice' cannot be captured in analogue media of representation.

6.4 Reasoning

Natural language expresses a number of important logical and epistemic operators which have no obvious equivalents in the domains of external analogue media, such as 'not', 'or' or 'if-then', and which can be essential to the realisation of specific purposes of information representation or communication. Again, the properties of specific occurrences or situations that make them suitable for creating such abstractions cannot be captured in analogue media of representation. To some extent, the importance of logical and epistemic operators for the representation of information has been taken into account in the graphic medium. A common solution is to add standardised abstract iconic representations to analogue graphics. For instance, a cigarette with a big X across it means that smoking is not allowed. Because of their standardised character, such abstract icons act as non-analogue and *non*-arbitrary external representations just like those of natural language (cf. Sect. 2 above).

7. DEFICIENCY-HANDLING MECHANISMS

The complementarity between natural language and analogue graphics representations has two main implications which will be discussed in this section and Sect. 8, respectively. The first is that each type of representation includes a number of mechanisms which are internal to that mechanism and whose function is to 'patch up' their respective deficiencies of expression. Thus, a number of *focusing mechanisms* have evolved in analogue graphics and natural language makes use of various *specificity mechanisms* for achieving increased specificity. These types of mechanism enable analogue graphics and natural language to overcome, to some extent, their respective, inherent expressive deficiencies and hence to realise a broader scope of information representation. And both types of mechanism can be seen to inherit their respective deficiency-handling capabilities from the complementary representational type. The second implication is that *the multimodal combination* of the two types of external representation offers many opportunities for benefiting from the strengths of each.

Viewing the mechanisms to be presented below as deficiency-handling devices is of course to adopt one perspective on those mechanisms among other, equally possible perspectives. Undoubtedly, some or most of these mechanisms have been present during the entire life-time of the representational types we are considering. The purpose of presenting those mechanisms as deficiency-handling devices is to emphasise the complementarity between specificity and focus. From another, equally valid and compatible, perspective, the isolated use, for a large variety of purposes, of each of the natural language or analogue graphics representational modalities can be seen as an effort to achieve as much specificity *and* focus as possible within the basic representational constraints on a particular modality. From this latter perspective, one is likely to emphasise the extent to which specificity and focus in particular instances of natural language or graphical representations are matters of degree.

7.1 Focusing Mechanisms in Analogue Graphics

Perhaps not surprisingly, given what has been said above, but worth pointing out anyway is the fact that the focusing mechanisms of analogue graphics all appear to trade analogueness for focus. Each focusing mechanism achieves its results by *decreasing* the analogue relationship between representation and what is represented. This happens at a price, of course, namely that focused analogue graphics, in various ways and to varying degrees, lose many of the virtues of specificity pointed out in Sect. 5 above. The primary advantage obtained by focusing, on the other hand, is an increase in relevance decidability which thus lets analogue graphics share one of the important advantages of natural language. A second advantage obtained through some types of focusing is that analogue graphics succeed in approaching the abstract representational qualities of natural language (cf. Sect. 6.1 above).

7.1.1 Selective Removal of Specificity. Selective removal of specificity is a useful mechanism for increasing the focus and hence the communicative relevance of analogue graphics. For instance, if one cannot clearly see from a piece of analogue graphics which kind of dog or tree is being represented, it may be contextually likely that what is represented is simply a dog or a tree. A step upwards in the abstraction hierarchy has thus been achieved. Removing background and other communicatively irrelevant entities from a piece of analogue graphics equally serves to enhance their focus. The filtering of information in order to make important features or structures appear more prominently represents a combination of selective removal of specificity and aspect enhancement (see below). What has been called here selective removal of specificity is often termed 'selective abstraction'. Note that this is *not* abstraction in the sense in which natural language expressions are abstract because specificity is being preserved in the process. However, selective removal of specificity shares with linguistic abstraction the effect of opening an interpretational scope which may be why the term 'abstraction' is often being used (ambiguously) in both cases.

7.1.2 Dimensionality Reduction. Dimensionality reduction is a form of selective removal of specificity. Many representational or communicative purposes can be achieved by using a lower spatio-temporal dimensionality than that characterising the entities being represented. For instance, many spatial layouts do not require representation in 3-D; many spatio-temporal processes and events can be expressed purely in the spatial domain and even 2-D representations are often sufficient for doing that.

7.1.3 Enhancing Aspects for Saliency. Analogue graphics have many different mechanisms for enhancing certain aspects of what is represented. Such mechanisms serve to increase the comparative saliency of certain aspects in the context of the graphics as a whole. In static graphics, relative enhancement of contours, differences in colouring, encircling, distortion of proportions, foregrounding, static simulation of dynamic zooming, and scaling and selective enlargement of entities all serve this purpose. These mechanisms can also be used in dynamic graphics which have an additional repertoire for aspect enhancement including dynamical change of colours, contours, shapes and sizes, zooming and scaling, blinking or oscillation, movement and so on. In addition, as indicated earlier, we have a small arsenal of standardised abstract (non-analogue, non-arbitrary) icons some of which can be used for saliency-enhancement. The most common example is the use of arrows for focusing purposes in analogue graphics. In books for bird-watchers, for instance, it is common to use arrows to point to discriminatory features among otherwise closely resembling species. Without these arrows, the analogue graphical bird representations provided would be less than

half as useful for the support of species identification tasks, which offers a powerful illustration of the non-focused character of analogue graphics.

7.1.4 Dwelling and Repetition. Dwelling and repetition are two other focusing mechanisms which are primarily used in dynamic graphics but which may be used in static graphics as well.

7.2 Specificity Mechanisms in Natural Language

When using natural language for representational purposes one nearly always faces the problem of how to sufficiently reduce representational scope. Obvious examples include the description of complex spatial layouts or faces, but the problem is much more general and failure to solve it often causes communication error. It seems likely that, e.g., underspecified instructions lead to much wasted effort in the workplace. We have seen that the use of focusing mechanisms in analogue graphics happens at the price of reducing analogueness. The use of specificity mechanisms in natural language, on the other hand, does not necessarily happen at a price such as reduced focus or generality. And when a price has to be paid, its nature depends on the particular specificity mechanism used.

7.2.1 Lengthy Description. Increasing the comprehensiveness of a description (instruction, etc.) is a key method for reducing the interpretational scope of linguistic expressions. As remarked earlier, one virtually never achieves complete specificity this way. However, specificity sufficient for a given communicative purpose can often be achieved. The widespread use of summaries, repetition, statements of 'key points' and so on, testifies to the fact that the longer a description (instruction, etc.) becomes, the easier it becomes for recipients to lose its overall focus. The architecture of the standard news article in newspapers is that of a staged increase in (length and) specificity of description.

Interestingly, even though natural language is, by itself, a focused type of representation it also has mechanisms for focus enhancement. In the written natural language modality which exploits the graphical medium of expression, this is done through the use of graphical mechanisms such as underlining, italics, different font sizes, relative positioning of text bits, etc. In principle, all the analogue graphics saliency-enhancement mechanisms might be used as we do to some extent when annotating text written by others. It is common knowledge that these mechanisms are often misused, i.e., used unnecessarily, which serves to re-emphasise the inherent focusing power of natural language. In spoken language, focus can be marked through acoustic mechanisms such as change of rhythm or loudness of expression.

7.2.2 Indexical Reference. Indexical reference designates the strongest family of specificity mechanisms in natural language. In the broad sense of indexical reference with which we are concerned, indexical reference ties linguistic expressions to something specific which is or can be known from past, present or future experience. In tying linguistic expression to experience, indexical reference makes use of proper names, definite descriptions (e.g., "it's on the table in the living room") and indexicals. Extra-linguistic acts of indexical reference such as pointing serve the same purpose of tying language to experience in order to achieve specific internal representation of the topics of discourse. In the AG domain, the use of linguistic indexicals can be viewed as analogous to combining linguistic expressions and analogue graphics (e.g., "this is my sister Charlotte"). Descriptions of entities in the AG domain are of course strongly supported by indexical reference. However, when indexical reference to specific entities

in the AG domain is used for illustration rather than identification (see Sect. 8 below), the specificity of the entities referred to may easily cause a decrease of generality and focus which then has to be remedied, for instance through the use of several different illustrations or by lengthening of the accompanying linguistic description.

7.2.3 Metrics. The inclusion of more or less exact metric properties of entities in linguistic descriptions is an important means of reducing interpretational scope. As suggested by the analysis of the example of an angle of 60 degrees in Sect. 4 above, use of metrical properties appears to be a necessary, but virtually never a sufficient condition for achieving fully specific natural language representations.

7.2.4 Use of Metaphor and Analogy. The difficulty of completely representing specific entities in the AG domain linguistically explains why the use of metaphor and analogy may dramatically increase the specificity of natural language descriptions. What a metaphor or an analogy contributes is to add, in a single word or phrase, an entire complex of features to the description which has already been provided. This can be much more efficient than one's having to painstakingly add literal expression upon literal expression to constrain interpretational scope. For instance, if a male person is correctly described as resembling a mole in face and posture it may be possible to pick out that person from a large crowd without the need of any further information.

8. THE INTEGRATION OF NATURAL LANGUAGE AND ANALOGUE GRAPHICS

The integration of natural language and analogue graphics for purposes of information representation in the AG domain offers the opportunity of combining the virtues of each generic form of representation. In such integrated multimodal representations, analogue graphics, whether static or dynamic, contributes specificity of representation and natural language contributes focus of representation, i.e., the virtues listed in Sects. 5 and 6 above are being combined. Adding to the power of combined representation, natural language can be used in expressing relevant information from outside the AG domain including abstract concepts and concepts facilitating reasoning in the AG domain. Moreover, many common types of multimodal linguistic/analogue graphics representation consist of combinations of linguistic expression and analogue graphics to which focusing mechanisms have been applied. In principle, of course, all the deficiency-handling mechanisms described in Sect. 7 above can be applied in multimodal representations. The result is a form of multimodal representation which can represent information in a way which is adequately focused and adequately specific at the same time. In the limit, such combinations can adequately substitute for direct experience in the AG domain and may hence serve as reality enhancements and substitutes for, e.g., training purposes. Let us review some well-known combination mechanisms.

8.1 Annotating

Naming of geographical locations on maps, entity parts in a diagram or persons in photographs, insertion of feature names or descriptions, defining the interpretation of graphical elements, or using written language 'bubbles' in cartoons are examples of linguistic annotation of analogue graphic representations. Other important forms of annotation of analogue graphics are the use of accompanying written or spoken natural language text. Annotation makes it possible to obtain the focus needed for a given representational purpose without loss of specificity.

8.2 Illustrating

Illustrating and annotating mirror one another. In annotation, it is the analogue graphics which are central to the representation and natural language serves to focus the graphical representation for a given communicative purpose. In illustrating something, it is the linguistic representation which is central to the representation and the analogue graphics are used for providing *specific models* of the subject-matter of linguistic representation. When this subject-matter is one of general concepts (in the AG domain) and/or reasoning with such concepts, illustration is all that analogue graphics can provide. Illustration supports the creation of appropriate mental models of the subject-matter at issue while linguistic expression more or less successfully prevents the loss of generality and focus. Books on mushrooms, for instance, can get people killed if this is not done properly.

Both annotation and illustration are widely used in what is commonly known as 'multimedia' representations of information.

8.3 Mutual Disambiguation and Redundancy

A third way of combining linguistic expression and analogue graphics is to make them disambiguate each other. A common example is the combination of word icons and analogue graphical icons in computer interfaces. Both are equally central to the multimodal representation they jointly constitute.

In this paper we have been mainly dealing with natural language and analogue graphics and their combinations as used for the 'output' representation of information on, e.g., computer screens. In all such scenarios, the user is a passive recipient of information and the representer's task is one of optimising focus and specificity for given communicative purposes. Researchers have begun to explore combinations of natural language and graphics as input modalities to computers (e.g., Lee & Zeevat 1990, Klein & Pineda 1990). This line of work does not seem likely to change the main points made in this paper.

8.4 Abstract or Conceptual Graphics

Strictly speaking, this topic lies outside of the scope of this paper. It may however be remarked that the combination of linguistic expressions and non-analogue, arbitrary graphical structures such as points, lines, boxes, etc. is a widely used method of combining general conceptual information expressed in written natural language with useful properties of non-linguistic graphical expression such as perspicuous ordering, segmentation, grouping and so on. Visual programming languages exploit this method.

9. SPECIFICITY AS A CONSEQUENCE OF ANALOGUENESS: IMPLICATIONS FOR SOUND AND TOUCH

The specificity of analogue graphics seems to derive from its analogue character rather than from its graphic character. Analogueness is a property of other representational modalities than those of static or dynamic, diagrammatic or non-diagrammatic analogue graphics. This observation opens the perspective that other analogue representational modalities, such as analogue sound and touch, share many of the 'virtuous' properties of analogue graphics which derive from their specificity. This, indeed, seems to be the case. If we reconsider the seven properties of analogue graphics identified in Sect. 5 above, we find that six of these are characteristic of analogue sound and touch as well. They are:

- *representational exhaustiveness;*
- *smooth mapping;*
- *direct measurement;*
- *approximate inference;*
- *direct entity identification;*
- *substitution for direct experience.*

Only the property of *easy update connectivity* seems to be questionable with respect to touch and sound.

Furthermore, analogue sound and touch share the limitations of analogue graphics noted in Sect. 6 above with respect to:

- *relevance decidability;*
- *abstraction;*
- *confinement to the analogue medium;* and
- *lacking capability for representing logical and epistemic operators.*

The non-focused character of analogue sound and touch representations raises the question to what extent we find the same focusing mechanisms in the AS (analogue sound) and AT (analogue touch) domains as we found in the AG domain. The mechanisms were:

- *selective removal of specificity;*
- *dimensionality reduction;*
- *enhancing aspects for saliency;*
- *dwelling and repetition.*

Selective removal of specificity, enhancement for saliency and dwelling and repetition can be used in the AS and AT domains just as in the AG domain. Selectively specific sound diagrams, for instance, constitute a useful addition to the representational repertoire of current computers. Sound, being one-dimensional, cannot be subjected to dimensionality reduction. The dimensionality of touch is a complicated issue which will not be pursued here. As in the AG domain, use of these mechanisms in the AS and AT domains imply reduction in analogueness of representation.

Lacking in specificity, natural language needs the same specificity-enhancement mechanisms in the AS and AT domains as were needed in the AG domain:

- *lengthy description;*
- *indexical reference;*
- *metrics;*
- *use of metaphor and analogy.*

In other words, the basic distinction between specificity and focus appears to generalise rather smoothly into a characterisation of the basic differences between the use of spoken and written (and touch, for that matter) natural language, on the one hand, and the use of analogue representation in the AG, AS and AT domains, on the other.

Several implications seem to follow. The first is that analogue sound and touch representations may individually profit from being combined with natural language representations in the same ways as can analogue graphics, i.e. through the mechanisms of:

- *annotating*;
- *illustrating*;
- *mutual disambiguation and redundancy*.

Abstract or conceptual sound or touch 'diagrams' may not currently be widely used, but they are certainly possible in principle.

The second implication is that the integration of analogue graphics, sound and touch can be used to increase the scope of external representation towards the achievement of true virtual reality representation. However virtually real such representations become, linguistic representations will preserve their complementary virtues. These can be used, therefore, for annotating combined AG, AS and AT domain representations just as the latter can be used for illustrating abstract and general linguistic representation.

10. CONCLUDING DISCUSSION

The distinction between specificity and focus seems to be quite fundamental to the understanding of the representational capabilities and limitations of natural language, on the one hand, and analogue graphics, sound and touch representations on the other. Mapping out some of the implications of this distinction, as has been attempted above, seems to provide a principled basis for addressing the representational strengths and weaknesses of a multitude of interface and other representational modality combinations some of which are only now becoming technologically feasible. While the number of pure generic interface modalities are relatively limited and can be analysed in a principled manner, their actual or possible multimodal combinations are many and diverse (Bernsen 1993b, c). There seems to be no way of coping with this complexity other than through departing from the analysis of a small number of basic properties such as those of specificity and focus. In this way, we may be able to arrive at principled answers to many questions in the comparatively new field of modality theory, among which the celebrated puzzle: "When is a picture worth a thousand words?" (cf. Hovy & Arens 1990). The answer to this one has, in fact, been indicated above.

REFERENCES

- Bernsen, N.O. & Svane, B. (1994). Communication failure and mental models. *Proceedings of the 13th Scandinavian Conference on Linguistics*. Roskilde University, 1992 (in press).
- Bernsen, N.O. (1993a). Matching Information and Interface Modalities. An Example Study. *Esprit Basic Research project GRACE Deliverable 2.1.1*.
- Bernsen, N.O. (1993b). A research agenda for modality theory. In Cox, R., Petre, M., Brna, P. and Lee, J., Eds. *Proceedings of the Workshop on Graphical Representations, Reasoning and Communication*. World Conference on Artificial Intelligence in Education, Edinburgh, August 1993, 43-46.
- Bernsen, N.O. (1993c). Modality Theory: Supporting multimodal interface design. To appear in *Proceedings from the ERCIM Workshop on Multimodal Human-Computer Interaction*, Nancy, November 1993 (in press).
- Bernsen, N.O. (1994). Foundations of multimodal representations. A taxonomy of representational modalities. To appear in *Interacting with Computers* 1994.

- Hovy, E. & Arens, Y. (1990). When is a picture worth a thousand words? Allocation of modalities in multimedia communication. Paper presented at the AAAI *Symposium on Human-Computer Interfaces*, Stanford.
- Klein, E. & Pineda, L.A. (1990). Semantics and graphical information. In D. Diaper (Ed.), *Human-computer interaction - INTERACT '90*. Amsterdam: Elsevier.
- Lee, J. & Zeevat, H. (1990). Integrating natural language and graphics in dialogue. In D. Diaper (Ed.), *Human-computer interaction - INTERACT '90*. Amsterdam: Elsevier.

Acknowledgements: The work described in this paper was carried out under Esprit Basic Research project 6296 GRACE whose support is gratefully acknowledged. Discussions on Michael May's GRACE work on graphical features and types led the author to the hypothesis of the complementarity of specificity and focus and to the derivation of the features mentioned in Sects. 5 and 6, most of which had been identified by Michael. Kenneth Holmquist has provided valuable comments on an earlier draft.