

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Talking Through Graphics: An Empirical Study of the Sequential Integration of Modalities

Permalink

<https://escholarship.org/uc/item/0vn6m3cm>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors

Umata, Ichiro
Shimajima, Atsushi
Katagiri, Yasuhiro

Publication Date

2000

Peer reviewed

Talking Through Graphics: An Empirical Study of the Sequential Integration of Modalities

Ichiro Umata (umata@mic.atr.co.jp)¹

ATR Media Integration & Communications Research Laboratories;
Seika Soraku Kyoto, 619-0288 Japan

Atsushi Shimojima (ashimoji@jaist.ac.jp)

Japan Advanced Institute of Science and Technology;
1-1 Asahi Tatsunokuchi Nomi Ishikawa 923-1292 Japan

Yasuhiro Katagiri (katagiri@mic.atr.co.jp)

ATR Media Integration & Communications Research Laboratories;
Seika Soraku Kyoto, 619-0288 Japan

Abstract

An empirical investigation was conducted on the characteristics of language use in graphics communication settings. Graphics communications, such as dialogues using maps, drawings, or pictures, provide people with two independent systems of representation, spoken language and graphics. Drawing on our dialogue data, we will show that the presence of a graphical representation significantly changes the way the spoken language is used, extending its expressive capacity in most cases. As two remarkable uses of language affected in this way, we will report the phenomena of *mediated reference* and *dual description*, illustrating them with actual examples from our data. Finally, a quantitative analysis of our data will show that these special uses of language are indeed as common as conventional uses of language in the presence of graphical representations.

Introduction

Conversational exchanges that involve external graphical representations are quite common in our daily lives. People often give and ask directions by referring to maps, or they might draw a floor plan in discussing where to place furniture in a living room. Now, linguistic expressions denote objects and relations in the world. This denotation relation is governed by conventions inside the language. An utterance of a linguistic expression carries unique information about the world through these conventions. This is what the standard view of the semantics of language tells us. However, when we look at speech in conversational exchanges involving graphics, regardless of the language used, we will immediately notice utterances that do not conform to this standard picture.

This paper is a detailed examination of the impact of graphics on the use of language. Our data consist of actual two-party dialogues where participants draw or consult a map during verbal exchanges. We will focus on two remarkable uses of language, called “mediated reference” and “dual description,” that we found through an examination of our data. Both phenomena are clearly specific to dialogues involving some graphical representation, or at least, some external representation other than speech.

Briefly, *mediated reference* is a case where a linguistic expression reaches its “final” referent due to the fact that its “immediate” referent has a referential connection to this final one in the system of graphics. For example, our subjects often use the indexical “kore” (this) to refer to a building or some other landmark, although its immediate referent is clearly an icon on the map; the icon refers to the landmark in the system

of map, and this fact somehow enables the indexical expression to do so too. We will discuss more examples of mediated reference later, and introduce three more varieties of the phenomenon.

Dual description is a case where a declarative sentence is used to describe a fact that holds in the graphic as well as the corresponding fact in the situation represented by the graphic. Suppose, when asked about the number of stations between two particular stations, one counts the number of icons on a railroad map and says, “There are three stations in-between them.” Is this report concerned with the map itself, or with the mapped railroad? Is it reporting that the railroad map has three station icons between two particular station icons, or that the railroad system has three stations between the two stations? Whichever the answer may be, it seems clear that the speaker has managed to describe both facts with this sentence. Note that, on the semantics associated with the railroad map, the first fact means the second fact, and this semantic relation somehow underwrites the duplicative use of the sentence.

Both uses of language are so natural and common in a dialogue involving a graphical representation that people may not even be aware of the phenomena. In fact, their frequent occurrence in such settings suggests that they are not a deviant but rather a perfectly legitimate use of language. Yet the empirical research on the integration of linguistic and graphical representations has focused on the issue of how speech is used to disambiguate a graphic (Neilson and Lee 1994) or how a graphic is used to disambiguate speech (Lee and Zeevat 1990). The linguistic-graphic integration has been also studied from a logical point of view, but the focus has been on how a graphic expresses what cannot be easily expressed by a linguistic representation (Barwise and Etchemendy 1996, Shimojima 1999). For both views, the fundamental form of linguistic-graphic integration is a *parallel* one, where each mode of representation expresses information in its own way, but since one mode of representation expresses what the other form does not, they may work complementarily to each other. In contrast, the two phenomena that we are highlighting in this paper point to a rather different form of integration, where the presence of one mode of representation extends the expressive capacity of the other by affecting the way it is used. Our goal is to draw due attention to this *sequential* form of graphic-linguistic integration by demonstrating that the instances of that type of integration are commonly observed in actual human dialogues, as opposed to mere logical possibilities.

In the next section, we will describe the methods through

¹Also with Kobe University.

which we collected our dialogue data. The two subsequent sections are, respectively, qualitative descriptions of the phenomena of mediated reference and dual description, where we illustrate each phenomenon with examples of language use drawn from our dialogue data. In these sections, the phenomenon of mediated reference is classified into four different types, and the mechanism underlying dual description is analyzed. The final section is devoted to a quantitative description of the two phenomena, where we use “content phrasal unit” to quantify the frequency of mediated reference and dual description in the dialogue process. As it turns out, these graphic-oriented uses of language occur as frequently as standard uses in our dialogue data, indicating that the speakers are quite ready to exploit the graphical representations at hand to extend the expressive capacity of their language.

Data

The conversational data analyzed in this paper were gathered from a series of graphical communication experiments, which were conducted for a larger research project investigating the interactions between cognitive/communicative factors and graphical representations². Our data consist of 19 task-oriented dialogues, with a total length of 116 minutes. Of these dialogues, 14 involve “a map sketching task” while the other 5 involve “a GRE task.”

Map Sketching Task In this task, two subjects were asked to work together as partners to draw a map showing four landmarks in Nara (a local town) as accurately as possible. The subjects were seated in separate, soundproof rooms and worked together using a shared virtual whiteboard and a full duplex audio connection. All inputs to the screen were by stylus, and any writing or erasing by one participant would appear simultaneously on their partner’s screen. The subjects were video-taped during the task.

GRE Task In this task, two subjects were asked to solve a logical reasoning problem from the Graduate Record Examination (GRE). The problem was on possible route selections in a hypothetical truck delivery area. The subjects were again required to work on the problem together and it was suggested that drawing a diagram on the screen might help them to answer the question. All pairs drew diagrams and eight pairs among the nine drew map-like ones. The communication environment of the subjects was the same as that in the map sketching task. The GRE task differed from the Map Drawing task in that the subjects had to not only draw an accurate map, but also use it for problem solving.

Mediated Reference

Studies on the use of multimodal information in reasoning and communication have mostly focused on the complementary or parallel form of integration and have not closely examined the sequential form of integration of multiple modalities. Particularly, little is known about the directionality and the systematic nature of such integration. In our data, two directions were observed in referring to entities through the

sequential integration: mediated references via a representation system (a sketch map, a diagram, etc.) to the world and mediated references via the world to a representation system. We will call the former a *forward* mediated reference, and the latter a *backward* mediated reference. Further, mediated references can be observed between not only individuals but also relations. We will examine, in this section, the phenomena of mediated references observed in our graphical communication experiments. Figure 1 shows four possibilities of mediated references.

Individual Mediated Reference

Consider the following utterances:

- (1) (From the map data, pointing to a part of the map with the stylus)

de, koko-ni-ne, tasika Deiri-Sutoa-ga-ne,
and, here-DAT, probably, Daily Store-NOM,
kono kado-ni atta.
this corner-DAT was.
“And I think there’s Daily Store on this corner.”

- (2) (From the GRE data, pointing to a path on the map with the stylus)

kore-ga 100 desuka?
this-NOM 100 is
“This is 100km, isn’t it?”

In (1), the speaker was pointing to a part of the map, and the linguistic expression “*koko* (here)” and “*kono kado* (this corner) literally denoted a part of the map. However, there was just a blank space on this part of the map and there were no symbols that could be regarded as an icon of a store. If we assume that the speaker was talking about the map, this utterance would be regarded as simply meaningless or at most false. In this task setting, it is unlikely that the speaker was talking nonsense or lying. Therefore, this utterance was a statement not based on the map but on the real world. Here, the linguistic expressions “*koko*” and “*kono kado*” referred to some place in the world via the place on the map. Similarly, no signs showing the distance could be seen on the diagram in the case of utterance (2), and no suitable properties for the referent of the expression “100” could be found on the diagram. Consequently, this utterance was also a description of some situation regarding the delivery route, not a part on the diagram. In these cases, the reliable correspondence between the spatial configuration of a map and a place in the world enabled *forward* mediated references: references to places in the world through places on the map.

We can also find examples of *backward* mediated references in the data. Some of them are as follows:

- (3) (From the map data, pointing to the icon of Nara Park on the map)

ja, kore, moo-tyotto koen okkiku suru?
So this a-little-more park big make
“So, shall we make this park a little bigger?”

- (4) (From the map data, after realizing that they made a mistake)

²These experiments were designed by Patrick G.T. Healey, Nik Swoboda, Ichiro Umata and Yasuhiro Katagiri.

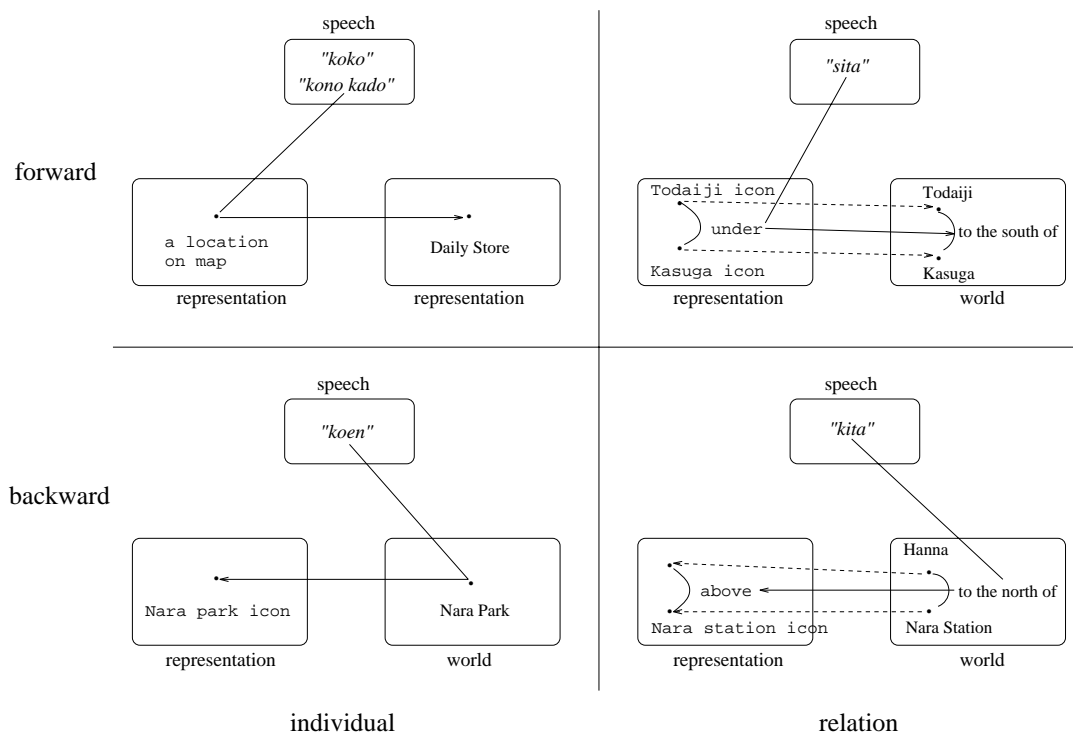


Figure 1: Four categories of mediated reference

a jaa kesimasyoo Teramati Kita
 well so erase Teramati-town Kita-town
 “Well, so, let’s erase the Teramati-Kita route.”

The linguistic expression “*koen* (park)” in utterance (3) literally denoted a real park. However, one cannot usually make a real park bigger, so this utterance cannot be regarded as a statement about the world. Rather, utterance (3) was a statement about some operation on the map, and “*koen*” referred not to the real Nara Park but rather to the icon on the map. Similarly, utterance (4) was a suggestion to erase the path on the diagram. They could not erase the delivery route itself because it was not assumed to be in the situation described by the GRE problem. In both statements, the objects in the world play an intermediate role, and the linguistic expressions refer indirectly to the icons on the map.

Relation Mediated Reference

So far, we have concentrated on mediated references between individuals. However, there are also examples of mediated references between relations in our data. Consider the following examples:

- (5) (From the map data, pointing to a part of the map with the stylus)

kokorahen-ni Toodaiji-ga aru
 around-here-DAT Todaiji-temple-NOM is
kara, kono sita-no hoo-kana?
 because this below-GEN direction-I wonder
 “Because Todaiji is around here, it (Kasuga-shrine) is probably below this, isn’t it?”

- (6) (From the map data, looking at his partner’s drawing on the map)

aa, soko zutto ue nobotteiku-to,
 Yes there straight up go
nyugakusiki-no toko dayo.
 entrance ceremony-GEN place is
 “Yeah, if you go straight up there, you can find the place where we had the entrance ceremony.”

In dialogue (5), the speaker intended to show the listener where Kasuga-shrine was. However, since there were no Kasuga-shrine icons on the map, the subjects were talking about a real-world situation via the map. However, the linguistic expression “*sita*,” which roughly means *under* or *below*, cannot be regarded as referring directly to real world relations; in the real world, Kasuga shrine is not *under* Todaiji temple but *south* of it. Here, “*sita*” referred to the real world relation (i.e. *to-the-south-of*) indirectly via the relation on the map (*under*). This was based on the semantic correspondence established between the map and the world. Similarly, in the case of utterance (6), there were no icons for “*nyugakusiki-no toko*” (the place where we had the entrance ceremony) on the map. Therefore, this utterance was describing a state in the real world, and the expression “*ue*” referred to the spatial relation in the real world (i.e. *to the north*) via the relation on the map (i.e. *up*) in (6).

Excerpt (7) includes an example of a *backward* mediated reference to a relation, as well as examples of individual backward references.

- (7) (From the map data, revising the position of the Nara Station icon)

Hanna Way-no yori kita-ni ittya
 Hanna Way-GEN than north-DAT go
akan?
 no good-Q

“So, we can’t draw it north of Hanna Way?”

Since one cannot change the place of the real train station, utterances in (7) were about the situation on the map. Thus, the linguistic expression “*Hanna Way*” in (7) referred backwardly to the icons on the map, not to a real world object. Similarly, the linguistic expression “*kita-ni*” (to the north of) makes a backward mediated reference to the relation of the places on the map (i.e. above) via the real world relation to the north of in this utterance³. Due to the preservation of the configurational constraints between the map system and the world, such mediated references to spatial relations are quite natural and are commonly found in conversations in which maps are used.

Informational Duality

Thus, when a map accompanies a dialogue, speakers can make a wide variety of indirect references, either individual or relational, forward or backward, through the systematic semantic relations established between the map and the mapped area. From the speaker’s point of view, this means increased freedom of reference with a limited vocabulary, but from the listener’s point of view, this might mean an increased likelihood that an utterance will become ambiguous in regards to the map itself or the mapped region. For we cannot settle this issue simply by assessing the literal referents of the expressions used, due to the possibility of indirect reference. Purely semantic disambiguation of this sort is generally not applicable.

Fortunately, listeners can often rely on pragmatic cues to resolve such ambiguity, as we have seen in the cases of (1)–(7). Generally, listeners can reject an interpretation of a statement if, on that interpretation, an utterance is to perform a speech act that is not felicitous in that context, such as (i) describing or checking the current position of an object not on the map, or (ii) requesting or otherwise discussing an operation on the mapped region that is impossible to address. The examples (1), (2), (5), and (6) correspond to (i), and thus they were considered not to be about the map, while (3), (4), and (7) correspond to (ii), and they were considered not to be about the mapped region.

However, our data also contain a number of utterances *not* subject to even such disambiguations. In the following, we will consider some of those examples and analyze the informational and functional structures of such utterances.

An Example

The following dialogue from our GRE data was conducted soon after the partners drew a graph-like map showing the routes connecting various towns, including Kawabata, Kitamati, and Hasimoto. The speakers are concerned with how many towns a truck driver can pass through in one day.

³The subjects erased the old icon and were just starting to draw a new one, so “*iku (itty)*” in this utterance expresses the movement of the icon on the map.

- (8) A: *kazoemasukanee?*
(Shall we count?)
B: *soosuruto.*
(If we do so, then....)
A: *kazoeruto 3-tu kanaa. 4-tu-wa tyotto muridesuyonee.*
(On my counting, it is three, I suppose. Four is not feasible, is it?)
B: *uun.*
(Hmm)
A: *Kawa toka dattara, kore moo sudeni 300 toka dakara, moo, Kita, kotti, Kita-ni*
(If this is Kawa or something, and if this is already 300 or so, well, Kita, here, to Kita [Mumbling indistinctively.]
Kawa-kara Kita-ni itte Hasi-de, kore 3-tu desuyone.
(Going from Kawa to Kita and then to Hasi, that’s three, isn’t it?)

The case in point is the last utterance of speaker A, which is underlined. On the one hand, one may well regard the names “*Kawa*,” “*Kita*,” and “*Hasi*” to denote the icons for Kawabata town, Kitamati town, and Hasimoto town. In this interpretation, the entire utterance concerns the map, and the speaker is reporting the following information:

- (9) There are three town icons on the path: the Kawa icon, the Kita icon and the Hasi icon.

The icons for Kawabata town, Kitamati town, and Hasimoto town already exist on the map, connected by a particular path on the map. Thus, reporting the number of town icons on this path, such as reporting (9), is a speech act that the speaker may well perform at this point. In fact, the above excerpt shows that prior to this utterance, the subjects have explicitly agreed to do such counting. Thus, one cannot reject the interpretation of the utterance as concerned with the map for any obvious reason.

On the other hand, it is also natural to regard “*Kawa*,” “*Kita*,” and “*Hasi*” as indirectly denoting the real towns, and if so, the utterance conveys the following information about a traffic route in the mapped region:

- (10) There are three towns on the route: Kawabata town, Kitamati town and Hasimoto town.

Recall that the present problem for the subjects is the maximum number of towns that a truck driver can pass through within one day. The reporting in (10) is directly relevant to the solution to this problem, and hence is a speech act quite likely to be performed at this point. Thus, the interpretation of this utterance as being about the mapped region cannot be rejected, either.

As we will see shortly, our data contain a number of examples of this type, where an utterance is as likely to be about the map as to be about the mapped region. Upon reflection, this type of utterance seems frequent in everyday dialogue involving some graphical representation—we do not always clearly explain this to listeners, or even to ourselves, whether our assertion or report is about the picture at hand or about the situation the picture depicts. So the question is: how can we understand such utterances? Are they *ambiguous* in the sense that: although such utterances are “really” concerned with either the map or the mapped region, they fail to provide sufficient cues to distinguish between them? If we frequently use such ambiguous utterances, how can we ever be successful in communication?

Analysis

On these questions, we propose that such utterances are *not* concerned with either the picture or the depicted situation *exclusively*. Rather, they are concerned with *both*, and thus handle information about the picture as well as information about the pictured situation. Such utterances are therefore not ambiguous in the above sense. The problem of ambiguity occurs only when we assume that such an utterance is concerned with only one subject matter. Here we explicitly discard this assumption for the kind of utterances in question. For example, the subject matter of the underlined utterance in (8) is not single, but dual, and the utterance reports the number of the town icons on a particular path on the map, *as well as* the number of the towns on the corresponding route on the mapped region.

But how is it ever possible for a single utterance to have such dual informational contents? Briefly, this is possible because representation is a transitive relation. Due to this principle, whenever an utterance represents a picture having a property α , and this property α on the picture in turn represents the depicted object having property β , the original utterance will also represent this object having property β . In the present case, the underlined utterance in (8) represents the map having the structural property of (9), and due to the semantic convention associated with the maps, a map with the property (9) represents the mapped region as having the structural property (10). Thus, by transitivity, the utterance also represents the mapped region having the property of (10). This is how a single utterance carries two pieces of information: one about the map and the other about the mapped region.

This mechanism may be made clearer by using the analogy of a copy machine. Suppose you make a copy d' of a document d , and then make a copy d'' of the copy d' that you just made. The copy d'' being a copy of the copy d' accurately represents d' more or less, and carries information about d' . Notice that this copy d'' also carries information about the original document d —we can look at the second copy d'' and learn what the original document d is like. (In fact, this is usually the main use of the second copy: we look at it in order to get information about the original document, often forgetting that it also carries information about the first copy.) Thus, the second copy d'' carries two pieces of information, one about the first copy, and the other about the original document. The second copy d'' carries the latter *via* the first copy, thanks to the transitivity of representation.⁴

Our claim is that the same thing happens in the case of the utterance in (8), where the mapped region in the truck delivery area is the original document d , the map of it is the first copy d' of d , and the utterance is the copy d'' of d' . The utterance in (8) carries information about the mapped region in the truck delivery area via the map, just as d'' carries information about d via d' . The utterance carries depicted information (9) and (10) about the map and the mapped region, just as d'' carries duplicated information about d' and d .

There are two major advantages to this claim. First, it avoids attributing ambiguity to utterances of this type that are

⁴The idea that the carrying of information is a transitive relation is called “the Xerox Principle” by Dretske (1981); this idea has been a focus of interest in situation theory (Barwise and Perry 1983) and was subsequently developed in qualitative information theory (Barwise and Seligman 1997).

found in spoken dialogues employing graphical representations. Therefore, such utterances are not particularly prone to misinterpretation, which is why these types of utterances can occur frequently without hindering smooth communications among speakers.

The second advantage of our analysis is that it gives us a way to capture a set of mechanisms provided by those “dual” utterances to facilitate problem solving processes involving the use of graphical representations. Recall, from our discussion on example (8), that a dual utterance occurs in a context where two different communicative acts are likely: one concerned with the picture and the other concerned with the depicted situation. In our reckoning, the speaker is considered to perform both acts with the utterance, without skipping or suppressing either act. A dual utterance effectively works as a bridge, and both the speaker and the hearer can engage in joint problem solving by matching and transferring information between the graphics domain and the problem domain. To clarify this point, let us suppose that utterance (8) was ambiguous and actually carried only one piece of information. If it was on the situation in the diagram, then the utterance itself would not convey the information about the world and would not directly lead to the answer to the question. If it was on the world situation, then it would show the answer but have no grounds for it. Under the assumption of dual information, the utterance provides both the answer and the basis for it at the same time: the information on the world based on the information on the graphics. Thus, in general, our proposal offers a more natural explanation of the use of such dual utterances in graphics communication, compared to theories that attribute a single informational content to it.

Quantitative Analysis

We have demonstrated that a combination of graphical representation and linguistic representation in a graphical communication setting provides us with a novel sequential method for integrating of the linguistic and graphical modalities in the form of mediated and dual references. Our analysis so far has been concerned with classifications and functions of instances of these new types of references.

In order to further establish that the sequential integration actually provides us with a viable and effective mechanism for communication, we conducted a quantitative analysis on the relative frequencies of the “new” forms of references, both mediated and dual references; we performed comparison with “conventional” direct references within our data obtained in our Map and GRE experiment. Furthermore, the different characteristics of each task were expected to result in a different distribution of the final referents of linguistic phrases. The Map corpus was expected to have more instances referring solely to the object in the graphics domain, because the aim of the task was to complete a map. On the other hand, the GRE corpus was expected to include fewer of such instances, because the aim was to solve the problems of the world domain and the graphics simply assist in that purpose.

Our corpus consists of 14,011 words (9,179 for the Map and 4,832 for the GRE), and the number of content phrasal units⁵ was 5,325 (3,394 for the Map and 1,931 for the GRE).

⁵A content phrasal unit is a minimum phrasal unit that has a con-

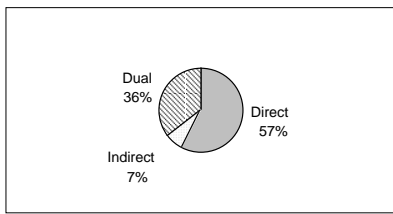


Figure 2: Relative usage frequencies of the direct, mediated and dual references)

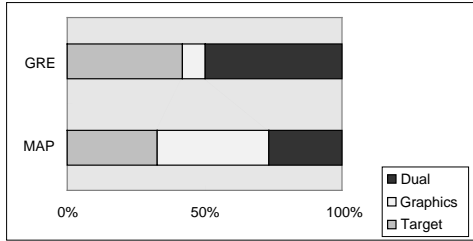


Figure 3: Distributions of final target domains of references for the Map and GRE data

Of them, 4,667 units were the ones describing the situations of the graphics and/or the world domain (2,875 for the Map and 1,792 for the GRE). We classified these units into the three categories shown above: direct, mediated and dual references.

Figure 2 shows the relative usage frequencies of the three types of references: direct, mediated and dual. Of all the reference occurrences, 57% were instances of direct references and 43% were instances of either mediated or dual references. This clearly shows that mediated and dual references are not mere theoretical possibilities or exceptional phenomena, but rather are mundane mechanisms routinely employed in actual communication.

Task characteristics of the Map sketching task and the GRE task can also be captured in quantitative terms. Figure 3 shows the distribution of final target domains of reference for the two tasks. A direct reference to the world and a forward mediated reference through the graphics to the world share the world as their final target domain of reference. Similarly, a direct reference to the graphics and a backward mediated reference through the world to the graphics eventually refer to the graphics as their final target domain. A dual reference is indeterminate as to its final target domain. The final target domains exhibit significantly different distributions between the two tasks ($\chi^2_{(4)} = 595.60, p < .001$). More concretely, (1) the Map data had more instances of graphic-only references, (adjusted residual: Map = 23.75, GRE = -23.75); (2) the GRE data had more instances of world-only references, (adjusted residual: Map = -6.27, GRE = 6.27); (3) the GRE data had more instances of dual references (adjusted residual: Map = -16.02, GRE = 16.02). Thus the assumption that the GRE data would have more world referents and fewer graphic referents than the Map data was supported. Furthermore, it is likely that dual references are strongly related to inferences on graphics, given the abundant instances of dual references in the GRE data.

tent word as its head.

Conclusion

Based on the data of spontaneous spoken dialogues involving graphic representations, we have analyzed the impacts of the presence of a graphic on the use of spoken language. We found (1) a pre-established semantic relation between a graphic and the situation depicted by it provides the speaker with rich possibilities of mediated references, including forward individual, backward individual, forward relational, and backward relational references; (2) the same semantic relation also lets the speaker use a declarative sentence to express dual pieces of information; (3) mediated reference and dual description are not exceptional but rather mundane mechanisms routinely employed in actual communication; (4) the characteristics of communicative contexts affect the distributions of the final referents of linguistic phrases. We further suggested that the use of dual descriptions is strongly related to inferences on graphics.

These findings indicate that in spontaneous human communications, spoken language and a graphic representation may be used in the *sequential* composition, where the latter affects the usage of the former to extend its expressive capacity. This is in stark contrast to the common view of the interaction between linguistic and graphic modalities, where the integration is made only at the level of multiple pieces of information expressed by the two modalities in individual manners. A *parallel* composition of this type is not the only form of the graphic-linguistic integration, and probably, not even a dominant form.

References

- Barwise, J., and E. Etchemendy (1995). Heterogeneous Logic. In Glasgow, J. I., Narayanan, N. H., and B. Chandrasekaran (Eds.) *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, 211–234 Cambridge, Mass.: MIT Press.
- Barwise, J., and J. Perry (1983). *Situations and Attitudes*. Cambridge, Mass.: MIT Press.
- Barwise, J., and J. Seligman (1997). *Information Flow: The Logic of Distributed Systems*. Cambridge, U.K.: Cambridge University Press.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, Mass.: MIT Press
- Lee, J., and H. Zeevat (1990). Integrating Natural Language and Graphics in Dialogue. In Diaper, D., Gilmore, D., Cockton, G. and B. Schackel. (Eds.) *Human Computer Interaction—INTERACT'90*, 211–234 Amsterdam:
- Neilson, I., and J. Lee (1994). Conversations with Graphics: Implications for the Design of Natural Language/Graphics Interfaces *International Journal of Human-Computer Studies* 40, 509–541.
- Shimajima, A. (1999). The Graphic-Linguistic Distinction: Exploring Alternatives *Artificial Intelligence Review* 13, 313–335.