

FIGURE-GROUND RELATIONS IN HUMAN-COMPUTER DIALOGUE MODELING

Olena Begal

Lesya Ukrainka Eastern European National University, Lutsk,
Ukraine

Summary: This article is devoted to cognitive aspects of establishing figure-ground relations in view of developing virtual interlocutors capable of maintaining a dialogue with humans. The nature of figure-ground recognition is reinterpreted from the perspective of computer as an active participant of dialogue interaction with the aim of improvement of its communicative competence. The results of the research suggest the ways of optimizing the functioning of automatic dialogue systems. The components which are crucial to the establishment of figure-ground relations have been modeled for virtual interlocutors, implementation of which in the development of the linguistically competent chat-bot made it possible to achieve the natural and coherent dialogical communication between the human and the machine.

Keywords: figure-ground relations, cognitive aspects, dialogue, virtual interlocutor, modeling.

Statement of the problem and its significance. One of the key tasks of modern Artificial Intelligence studies is the successful performance by computers of the communicative and cognitive functions native to humans. With the purpose of conducting an in-depth analysis of the cognitive properties of communicative interaction, we should pay attention to the notion of figure-ground relations. Looking at the process of communication from this perspective will provide an opportunity for revealing the important parameters and principles which should be modeled as essential components of effective virtual interlocutors. Finding the ways of successful establishment of figure-ground relations in automatic dialogue systems will significantly facilitate their communicative competence.

The analysis of the recent research works and publications. E. Rubin [Rubin 1915], a Danish psychologist, was the first who suggested differentiating between the concepts of figure and ground. His theories became influential within Gestalt psychology. These notions were introduced into linguistics by L. Talmy [Talmy 1972, 1983, 2000]. He conducted a thorough investigation of figure and ground, which was summed up in the table of definitional and associated characteristics of these two notions (see Table 1).

B.L. Whorf [Whorf 1956] studied the ways of segmenting situations, events and experience. He pointed out that it is important to analyze and segment our experience irrespective of the influence of the language we speak, in the way similar for every person.

Although considerable amount of research has been devoted to the figure-ground relations, few attempts have been made to investigate peculiarities of their modeling from the perspective of human-computer interaction.

The subject matter of the study is the figure-ground relations in human-computer dialogue.

The objective of this article is to state the cognitive aspects of establishing figure-ground relations, which are essential for modeling the effective and competent virtual interlocutor.

The tasks of the study are the following:

1) to characterize the key cognitive aspects of establishing figure-ground relations;

2) to model the components of the virtual interlocutor, which are crucial to the establishment of figure-ground relations in the dialogue with human.

The presentation of the basic material and reasoning of the study results. The notions of Figure and Ground were first seriously considered within the field of Gestalt psychology and are closely related to the processes of orientation, preliminary acquaintance and focusing. The need of distinguishing these notions was pointed out by Danish psychologist E. Rubin [Rubin 1915].

In the field of Cognitive linguistics the notions of Figure and Ground became the key ones for the conceptualization of space. L. Talmy introduced these notions into linguistics to refer to the objects that serve as a focus against the ground of a certain environment [Talmy 1972]. L. Talmy gives the following definitions: «*The Figure* is a moving or conceptually movable entity whose site, path, or orientation is conceived as a variable the particular value of which is the relevant issue. [...] *The Ground* is a reference entity, one that has a stationary setting relative to a reference frame, with respect to which the Figure's site, path, or orientation is characterized» [Talmy 1983: 232].

L. Talmy also suggests some comparative characteristics of figure and ground (see Table 1) [Talmy 2000: 315]:

Tab. 1. Comparative characteristics of figure and ground [Talmy 2000]

	Figure	Ground
Definitional characteristics	Has unknown spatial (or temporal) properties to be determined.	Acts as a reference entity, having known properties that can characterize the Figure's unknowns.
Associated characteristics	more movable smaller geometrically simpler (often pointlike) in its treatment more recently on the scene/ in awareness of greater concern/ relevance less immediately perceivable more salient, once perceived more dependent	more permanently located larger geometrically more complex in its treatment more familiar/ expected of lesser concern/ relevance more immediately perceivable more backgrounded, once Figure is perceived more independent

The foregoing comparative characteristics of figure and ground, suggested by L. Talmy, distinctly systematize their peculiarities and provide an ample opportunity to single out the distinctive features. Having transferred the process of figure-ground recognition to the sphere of communication, we get a new interpretation of the principles of reaching the discourse coherence. The current topic of conversation acts as a figure, on which the attention of speaker is focused. The ground can be represented by context and all of the extralinguistic parameters of communication, including time, place, situation and so on. Figure-ground relations help to establish connections necessary for understanding and to ensure the realization of successful communication. The sense center of the message, just as the figure, is not always clearly specified and explicitly expressed, but, being supported by a certain context (ground), it becomes the definite focus of attention. Identification of the main theme of expression takes place on the conceptual level, when more relevant unit comes to the foreground from among the general context. Thus the key topic, just as the figure, directly depends on its background context, because it gets interpretation and acquires its semantic weight within this ground.

To analyze the realization of figure-ground relations in dialogues with existing automatic virtual interlocutors, we have chosen chat-bot A.L.I.C.E. [A.L.I.C.E. www], as it is one of the strongest programs of this type and has won the Loebner Prize (an annual competition in Artificial Intelligence that awards prizes to the chat-bot considered by the judges to be the most human-like) three times. Let's pay our attention to the following example:

1.

Elena: What is better to dress for cocktail party?

ALICE: I always prefer the more expensive one.

Elena: So it is better to dress the most expensive outfit for cocktail party.

ALICE: What's the best one?

Elena: Probably a little black dress is the best one.

ALICE: When is little black dress not the best one?

Elena: I think that little black dress looks always perfect.

ALICE: Can you think of one instance where it never happens?

Elena: Well, when you climb the mountain a little black dress will be not appropriate.

ALICE: Do you think I should do it differently?

In the dialogue № 1 the situation of going to the «*cocktail party*» serves as a ground. In this context «*a little black dress*» is viewed as a sense center of communicative interaction and receives the status of a figure. The validity of the singled out subject of conversation is conditioned by the background context. Besides, the interlocutors – the human and the chat-bot A.L.I.C.E. – consider the possibility of background situation change («*when you climb the mountain*»), in which the central figure loses its relevance. In these circumstances, the key subject of conversation directly depends on its background context.

In the next example the function of a ground is performed by the human's question «*Do you know any jokes?*», which forms a communicative situation, in

which it is expected that the chat-bot should tell a joke. The jokes of the virtual interlocutor A.L.I.C.E. are perceived as a figure in this case.

2.

Elena: Do you know any jokes?

ALICE: (1) What do you get when you cross a crazy cow and a banned parrot?

Elena: I don't know.

ALICE: canned carrot.

Elena: Is it funny?

ALICE: It could be.

Elena: Any other jokes?

ALICE: (2) What do you get when you cross a poppy and electricity?

Elena: What?

ALICE: flower power.

Elena: Cool! That is a good joke!

ALICE: Wow! Thanks. Next question?

It becomes clear from the context of conversation that the automatic dialogue system is joking and not asking serious questions. Therefore, the existence of the background situation helps to set the interlocutors on the subject of the talk. Consequently, it leads to proper reaction to the cues.

The mechanisms of singling out the figure from the general ground on the cognitive level are common to all people in nature. Similar thoughts are met in works of B.L. Whorf. He wrote that if we compared the ways, in which different languages in a different manner «segment» the same situation or certain experience, we would feel the need of ability to analyze and «segment» this experience irrespective of the influence of the language we speak, in the way similar for every person. B.L. Whorf proves it in the following example: «There is one thing on which all observers of the appearance of a running boy will agree, at least after questioning or experimental testing – that it can be divided into parts – and they will all make the division in the same way. They will all divide it into (1) a figure or outline having more or less of motion (the boy) and (2) some kind of background or field against which, or in which, the figure is seen» [Whorf 1956: 162-163]. Analyzing some event or situation through the prism of figure-ground relations, we receive an opportunity to achieve proper interpretation of the event and establish necessary links between its components. To represent the events of the objective reality in a universal and clear way for everyone is an important task, which can be solved through the conceptual extraction of the key sense centers, which correlate with the global context as figure and ground.

Differentiation of these notions allows us to understand how we focus on the key elements during interaction with the objective reality. By singling out the main figure of conversation, the participant of communicative interaction activates in his mind necessary mental models and receives an opportunity to maintain the dialogue without breaking its coherence.

Within the framework of our investigation, we have developed the linguistically competent virtual interlocutor which received the name of «Joseph Lin-

guiston». It is able to maintain a dialogue with users about scholars of various periods and directions of linguistics, their works and notable ideas. To provide the chat-bot with the ability to single out the figure from the general ground properly, we have developed an effective approach to structuring knowledge concerning linguists and their achievements. All of the data are given in the form of ontological knowledge representation which is based on the principle of tree hierarchy. It is not just a stock of separate information, but forms an integrated whole, among the constituent parts of which associative links are established. In programming environment the structural elements of the knowledge database are arranged in the following way (see figure 1):

Fig. 1. The fragment of knowledge representation of the linguistically competent virtual interlocutor «Joseph Linguiston» in its programming realization

```
{
  "id": "1",
  "name": "Noam Chomsky",
  "who": {
    "name": "Noam Chomsky",
    "prp": "he,his,him"
  },
  "kids": [
    {
      "id": "1.1",
      "name": "Life",
      "kids": [
        {
          "id": "1.1.1",
          "name": "Facts",
          "kids": [
            {
              "id": "1.1.1.1",
              "Q1": "What is Noam Chomsky's full name?",
              "K1": "name",
              "A": "Chomsky's full name is Avram Noam Chomsky."
            },
            {
              "id": "1.1.1.2",
              "Q1": "When was Noam Chomsky born?",
              "K1": "born",
              "Q2": "What is the date of Noam Chomsky's birthday?",
              "K2": "birthday",
              "Q3": "When is Noam Chomsky's birthday?",
              "K3": "birthday",
              "A": "Noam Chomsky was born on December 7th, 1928."
            }
          ],
        }
      ],
    }
  ],
}
```

```

    "id": "1.1.1.3",
    "Q1": "Where was Noam Chomsky born?",
    "K1": "born",
    "A": "Noam Chomsky was born in the East Oak Lane neighborhood
of Philadelphia, Pennsylvania [Barsky 1997: 9].",
  },
  {
    "id": "1.1.1.4",
    "Q1": "Is Noam Chomsky alive?",
    "K1": "alive",
    "A": "Yes, Noam Chomsky is still alive. He resides in Lexington,
Massachusetts [http://en.wikipedia.org/wiki/Noam\_Chomsky].",
  }
]
},
{
  "id": "1.1.2",
  "name": "Family",
  "kids": [
    {
      "id": "1.1.2.1",
      "Q1": "Tell me about Noam Chomsky's family.",
      "K1": "family",
      "Q2": "What do you know about Noam Chomsky's family?",
      "K2": "family",
      "Q3": "Who were Noam Chomsky's parents?",
      "K3": "parents",
      "Q4": "Tell me about Noam Chomsky's father.",
      "K4": "father",
      "Q5": "Tell me about Noam Chomsky's mother.",
      "K5": "mother",
      "Q6": "Where was Noam Chomsky's father born?",
      "K6": "father,born",
      "Q7": "Where was Noam Chomsky's mother born?",
      "K7": "mother,born",
      "A": "Chomsky's father, Dr. William \"Zev\" Chomsky was born in
Ukraine, but had fled to the United States in 1913 to avoid conscription into
the army. His mother Elsie Simonofsky was a native of what is present-day
Belarus who grew up in the United States. They both were teachers at the
religious school with William eventually rising to the position of school
principal [Barsky 1997: 9-10].",
    },
    {
      "id": "1.1.2.2",
      "Q1": "Tell me about Noam Chomsky's family.",
      "K1": "family",
    }
  ]
}

```

```

"Q2": "Who were Noam Chomsky's parents?",
"K2": "parents",
"Q3": "What do you know about Noam Chomsky's family?",
"K3": "family",
"A": "Chomsky was born in the family of school teachers."
},
{
  "id": "1.1.2.3",
  "Q1": "Is Noam Chomsky married?",
  "K1": "married",
  "Q2": "Tell me about Noam Chomsky's wife.",
  "K2": "wife",
  "Q3": "Does Noam Chomsky have a wife?",
  "K3": "wife",
  "Q4": "Noam Chomsky's wife",
  "K4": "wife",
  "A": "He was married to Carol Doris Schatz until her death from
cancer in December 2008 [Barsky 1997: 13].",
},
{
  "id": "1.1.2.4",
  "Q1": "Does Noam Chomsky have children?",
  "K1": "children",
  "Q2": "Tell me about Chomsky's children.",
  "K2": "children",
  "A": "Chomsky has two daughters, Aviva (b. 1957) and Diane (b.
1960), and a son, Harry (b. 1967)
[http://en.wikipedia.org/wiki/Noam_Chomsky].",
}

```

In this fragment it is shown how the data are structured within this automatic dialogue program. Let's explain the meaning of structural components:

- *"id"* – a unique identifier of the theme which indicates the structural level of hierarchical organization (for example, *id* which has the value of 1 will be one level higher than *id* with the value of 1.1, and *id* which has the value of 1.1 will be on the same level as *id* with the value of 1.2). Introduction of the *id* into the system allows to arrange the data in their relationship to each other.
- *"name"* – a name of the structural level which identifies its topicality.
- *"who"*: {
"name": "Noam Chomsky",
"prp": "he,his,him"
} – the purpose of this unit is to put in correspondence to the word, that is given under the name *"name"*, pronouns that can denote it. When the user activates certain theme by his inquiry, the pronoun, which corresponds to the key notion, becomes activated too. If the interlocutor has activated the global theme

«Noam Chomsky», his next request «Tell me about his family» will be analyzed by the chat-bot as «Tell me about Noam Chomsky's family». This method of arranging nouns and pronouns in their direct correspondence provides virtual interlocutor with the ability to establish reference in the dialogue with human correctly.

- "kids" – an array of child elements which indicates the subordination of the subsequent themes to the previous one in hierarchical terms.
- "Q1" – "Qn" – a set of prototypical questions which can activate the answer-reaction attached to them, if the input inquiry of the user corresponds to these prototypical questions.
- "K1" – "Kn" – the basic words of prototypical questions, the presence of which is checked by the system for additional control over the level of correspondence of the user's inquiry to prototypical questions.
- "A" – an answer-reaction assigned to a certain set of prototypical questions.

Having structured the data about linguists, their notable ideas and scientific works in accordance with the foregoing principle, we provided the linguistically competent virtual interlocutor with accurately systematized knowledge base. Relying on this knowledge representation, the chat-bot can relevantly establish relations between the key notions of the user's input message and identify the main theme of conversation (the figure), taking into account the general context (the ground). When the automatic dialogue system analyzes the human's input sentence, it compares this input with prototypical questions that are stored in its knowledge base. The prototype is deemed appropriate, if the structure of the input sentence as well as the number of the key words correspond to it. When prototype is chosen, it activates the answer-reaction assigned to it and the theme, to which this block of information belongs. The activated theme serves as the figure and will have this status until the user activates another theme by his input request. While the human and the chat-bot are communicating within a certain theme, the program will give the information, stored under the name of this theme, on the basis of associative links.

During the analysis of user's input stimulus, the weight is assigned to its words, depending on the part of speech, determined by the tagger. The weight of the word reveals its status in the semantic load of the input sentence. Obviously, the focus on nouns is essential for the establishment of the input sentence topicality, as they bear the name of the subject of conversation. Therefore, their weight should be the biggest. The verbs are the next in the weight scale. It is explained by the importance of determining the action, which takes the subject of conversation or which is performed on it. The adjectives have the average weight significance etc.

Let's analyze the example of a dialogue with the linguistically competent virtual interlocutor «Joseph Linguiston»:

3.

Bot: Bonjour! I'm Joseph Linguiston! Enter your name, please.

I: My friends call me Lena.

Bot: Nice to meet you, Lena. What linguist are you interested in?

I: Let's talk about Confucius.

Bot: Feel free to ask me about Confucius's: life, notable ideas, main concepts, theories and works.

Due to the orientation section of the dialogue, the virtual interlocutor directs the user into the appropriate thematic channel.

I: Who was he?

Bot: Confucius was a Chinese teacher, editor, politician, and philosopher of the Spring and Autumn period of Chinese history [<http://en.wikipedia.org/wiki/Confucius>].

I: What is he famous for?

Bot: Confucius famously emphasized the moral commitment implicit in a name: «Good government consists in the ruler being a ruler, the minister being a minister, the father being a father, and the son being a son... If names be not correct, language is not in accordance with the truth of things» [Chinese Text Project. The Analects www (Analects 12.11,13.3)].

Would you like to know more about his main concepts?

I: Sure.

Bot: Confucius believed that social disorder often stemmed from failure to perceive, understand, and deal with reality. Fundamentally, then, social disorder can stem from the failure to call things by their proper names, and his solution to this was the rectification of names [http://en.wikipedia.org/wiki/Rectification_of_names].

I: Tell me more about it.

Bot: The rectification of names is the Confucian doctrine that to know and use the proper designations of things in the web of relationships that creates meaning, a community, and then behaving accordingly so as to ensure social harmony is The Good [http://en.wikipedia.org/wiki/Rectification_of_names].

Should I tell you more about it?

Time after time, Joseph Linguiston takes the initiative in the dialogue by offering to tell more about the subject of conversation («*Would you like to know more about his main concepts?*»), («*Should I tell you more about it?*») and gives its interlocutor the right to make a choice.

I: No. Tell me better about Lev Shcherba.

Bot: What do you choose?

I: Tell me about Shcherba's family.

Bot: As far as I know, he was born into the family of an engineer [Российская Академия Наук. Щерба Лев Владимирович www].

I: What do you think about him?

Bot: I consider him to be a great Russian linguist and lexicographer [http://en.wikipedia.org/wiki/Lev_Shcherba].

I: What are his main achievements?

Bot: Basing himself on the ideas of Baudouin de Courtenay, Lev Shcherba elaborated the concept of the phoneme. He became the founder of the so-called «Leningrad school» of phonology, which included M.I.Matusevich and L.R.Zinder among others and carried on a polemic with the «Moscow school» [http://en.wikipedia.org/wiki/Lev_Shcherba].

Would you like to know more about his notable ideas?

I: Yes, I would like.

Bot: In contrast to Ferdinand de Saussure, Lev Shcherba recognized three rather than two objects of study: speech activity, language systems, and language material. He placed emphasis on the question of the capacity of the speaker to produce sentences never previously heard, a question which would become important to the linguistics of the later twentieth century. He also emphasized the importance of experiments in linguistics, particularly that of negative results, developing methods which became important for field study [http://en.wikipedia.org/wiki/Lev_Shcherba].

I: Sorry, I can't talk longer.

Bot: Goodbye, Lena! Have a nice day!

The dialogue № 3 is an example of successful communication of the linguistically competent virtual interlocutor «Joseph Linguiston» with the human. The chat-bot appears to be an active interlocutor that starts conversation and acquaints the user with the extent of its knowledge about linguist (e.g. «*Feel free to ask me about Confucius's: life, notable ideas, main concepts, theories and works*»). It allows the human to form an idea of the topicality range of the subsequent dialogue. In the situation, when the user requests to tell more («*Tell me more about it*»), the automatic dialogue system successfully establishes referential link between the pronoun «*it*» and «*the rectification of names*» and continues to disclose the initiated topic.

Joseph Linguiston correctly singles out the figure (the main theme) from the general ground (context of the dialogue interaction). At the beginning of this dialogue, the figure is represented by «*Confucius*». The chat-bot gives proper information concerning this scholar. Then attention is focused on «*the rectification of names*». Being supported by the context, it becomes a figure, as it is one of *Confucius'* notable ideas.

At the next stage of the dialogue, a switch to another topic is realized: «*Tell me better about Lev Shcherba*». The user's request is too general. In order to specify what exactly the interlocutor wants to know, the chat-bot says: «*What do you choose?*». It helps to reach an agreement on subject of conversation. Having received more detailed request, the automatic dialogue system gives the required answer.

Overall, the communicative behavior of the linguistically competent virtual interlocutor can be characterized as an interaction, focused on cooperation and reaching mutual understanding with the human, what ensures the efficiency of dialogical communication.

Conclusion. On the basis of the analysis which was conducted, we can draw the conclusion that it is possible to improve the communicative competence of automatic dialogue systems by taking into account the cognitive aspects of establishing figure-ground relations. To construct an effective virtual interlocutor, able to single out the figure from the background context, it is necessary to provide it with:

- 1) the knowledge representation in the form of hierarchical tree ontology with associative links;
- 2) the algorithm of assigning weight to notional words;

- 3) the mechanism of analyzing the user's input inquiry;
- 4) the orientation section in the dialogue which helps to direct the user into the appropriate thematic channel.

The perspectives of further research consist in the possibility to improve the dialogical communication between humans and computers by providing the virtual interlocutors with reliable components that imitate the communicative and cognitive functions, performed by humans. The results of the analysis will favor the development of the general theory of human-machine communication.

REFERENCES

- Российская Академия Наук. Щерба Лев Владимирович [www – Российская Академия Наук. Щерба Лев Владимирович](http://www.ras.ru/nappelbaum/df920c3d-0174-46ca-bfe2-a32cc513eabf.aspx) [The Russian Academy of Sciences. Shcherba Lev Vladimirovich]. Available at: <https://www.ras.ru/nappelbaum/df920c3d-0174-46ca-bfe2-a32cc513eabf.aspx> (retrieved 17.08.2014)
- A.L.I.C.E. [www – A.L.I.C.E. Artificial Intelligence Foundation](http://www.pandorabots.com/pandora/talk?botid=f5d922d97e345aa1). Available at: <http://www.pandorabots.com/pandora/talk?botid=f5d922d97e345aa1> (retrieved 25.05.2015)
- Barsky 1997 – *Barsky R.F.* Noam Chomsky: A Life of Dissent. Cambridge: MIT Press, 1997.
- Chinese Text Project. The Analects [www – Chinese Text Project. The Analects](http://ctext.org/analects). Available at: <http://ctext.org/analects> (retrieved 30.06.2014)
- Rubin 1915 – *Rubin E.* Synsoplevede figurer. Copenhagen: Gyldendal, 1915.
- Talmy 1972 – *Talmy L.* Semantic structures in English and Atsugewi: Ph.D dissertation. Department of Linguistics, University of California, Berkeley, USA, 1972.
- Talmy 1983 – *Talmy L.* How language structures space // H.L.Pick, L.P.Acredolo (eds.). Spatial Orientation: Theory, Research, and Application. New York: Plenum Press, 1983.
- Talmy 2000 – *Talmy L.* Toward a Cognitive Semantics, vol. I. Cambridge: MIT Press, 2000.
- Whorf 1956 – *Whorf B.L.* Language, Thought, and Reality // J.B.Carroll (ed.). Cambridge: MIT Press, 1956.