Andrzej Łachwa
SEMIOTICS OF ADDRESS PHRASES

The theory of the automated information search in natural language texts often poses problems, which requires thorough research on this language, conducted partly from such points of view that are not applied in traditional linguistics. An IT specialist conducting this research would mainly notice those characteristics of natural language texts that can be used when building the devices he/she is interested in, in particular search systems which are members of the class of the so called Natural Language Understanding (NLU) systems (cf. Studnicki 1985, Łachwa 1986).

The object of this research, of which a fragment will be presented below, is the language phenomenon of some characteristic referential expressions occurring in certain types of written utterances, namely expressions referring to language utterances by specifying the place of these utterances in a certain system (e.g. a text). A great number of such referential expressions can be observed, for instance, in the corpus of legal texts. Some of them are parts of referential phrases (cf. Studnicki *et al.* 1983, manuscript). We have to stress, however, that only some referential phrases contain the said expressions and that these expressions are sometimes parts of other language expressions.[1] The aim of this study is to describe the said phenomenon in such a way as to provide a sufficient theoretical basis for developing the methods of automated interpretation of referential expressions, i.e. for formulating algorithms indicating those places in a corpus which the interpreted expression refers to. This operation leads to finding the utterance which the given language expression (containing the interpreted referential expression) refers

---

[1] As regards legal texts, this can be validation formulas or referential expressions.

to, i.e. utterances which are in strong semantic relation with the utterance containing this phrase (cf. Łachwa 1986; Studnicki *et al.* manuscript).

A study of the available materials has shown that the referential expressions discussed here are structurally and functionally similar to the expressions called, in the common meaning of this word, addresses, i.e. to the expressions used to define our place of residence, place of work, etc. We believe that the problems related to the structure and use of the latter are much clearer than those related to the expressions that are the proper subject of this article and that addresses are but a particular case of a more general phenomenon, we aim to develop a relevant theory based on the analysis of the meaning of the word *address* and the expressions of which addresses, in their 'common' meaning, are composed. Therefore, this part of the research will be the focus of this article.

The informal structures included in this theory (the theory of the phenomenon of indicating a place in certain physical spaces by using referential expressions referring to the labels of these places and the structures of these spaces) and the terminological apparatus introduced here will be used to analyse certain referential expressions occurring in the corpus of legal texts.

<div align="center">*</div>

<div align="center">*   *</div>

1. In common use, in contemporary Polish [as well as in contemporary English – transl. note], *adres* (an address) is usually either an EXPRESSION or an INFORMATION. For example, in the sentence *Can you read this address?* address is an expression, while in the sentence *This address is no longer valid* the meaning of address is a certain information. These are not the only meanings or ways of using the word *address*,[2] but they are the only ones relevant to us.

---

[2]It is claimed (Szymczak 1982; Doroszewski 1958: 29) that the word *adres* in Polish usually means a 'place of residence' or a 'place of work'. This would be the supposed meaning of this word (further referred to as the object meaning) in the sentence *Daj mi znać, gdyby twój adres uległ zmianie* [*Let me know, if your address changes*]. Consequently, if it really meant the place of residence (or work), the content of this sentence should be similar to *Daj mi znać, gdybyś zmienił miejsce zamieszkania* [*Let me know, if your place of residence changes*]. But sometimes someone's address may change while the place of residence does not. Thus, the word *adres* does not have an object meaning here, and neither does it in expressions such as *adres miejsca pracy* [*work address*], *adres zamieszkania* [*address of residence*], *adres budynku* [*address of a building*], etc. (otherwise, these expressions would not make sense). Let us now consider sentences such as: *Pod tym adresem nikt nie mieszka* [*Nobody lives at this address*], *Wysłałem*

2. Every expression that can be called an address expresses an information, also referred to as an address. For example, in the sentence *Read this address and remember it*, we are firstly referring to a certain expression (address-expression) and then to the information contained in it (address-information). It is worth noting that there are also such expressions for which it would be counterintuitive to call them addresses, yet which contain an address-information, for example in utterances such as this one: *Our offices are located* IN MIODOWA NUMBER 7 ON THE GROUND FLOOR AND AT NUMBER 144 IN THE OUTBUILDING. *Please tell these addresses to your colleagues.* If we follow our language intuition, we will not call the capitalised expression an address. We will rather say that it expresses or contains addresses of those offices, i.e. that these are addresses-information. This is what the second sentence of this utterance refers to.

3. The subject of our deliberations will be both those expressions that are called addresses in contemporary Polish [and English] and those which are not called that, even though they in fact express an information-address. Thus, it will be convenient to adopt two terminological conventions: a regulating one and a constructing/structural one. We can agree to use the word *address* in this article solely in the second meaning of the two described in point 1, i.e. for naming an INFORMATION which is called an address in the common Polish language. Furthermore, we shall agree that every REFERENTIAL EXPRESSION OF THE Polish language containing one or more address in the meaning adopted under the previous agreement will be called an ADDRESS PHRASE. An example of an address phrase is the expression capitalised in the utterance analysed in point 2.

4. Address phrases refer to a certain kind of PLACES in physical SPACE.[3] By

---

list na twój adres [*I sent a letter to your address*]. Indeed, these sentences speak of a place of residence, but the word *adres* was used with pronouns *pod* and *na* and we are not able to build sentences, in which the word a*dres* would be used without a pronoun, while referring to a place. Thus, the above sentences are phraseologisms, and they are the ones that have the object meaning.

The word *adres* is also used in phrasemes such as *powiedzieć coś pod czyimś adresem* [say something about someone], *zły adres* [wrong address], *pomyłka w adresie* [an error in address], etc. The word is also sometimes used to refer to an official letter containing wishes or congratulations addressed to an eminent person or a high official, usually given to this person in a ceremonial way. Moreover, this word occurs in terms composed of multiple words in official language and in the languages of various scientific disciplines, where its meaning is defined by specific terminological agreements and from where it sometimes spreads to everyday language.

[3]Some claim that an address is composed of two types of information: one concerning a person or institution, and the second one the location of the former. However, we

physical space we mean such places as a town, district, country and places in the sense of limited parts (fragments, areas) of these spaces. As we explain further in this article (cf. point 7), only some of these places in relevant physical spaces can be the denotata of address phrases, and these are the only ones that are of interest to us. Consequently, the spaces in question will from now on be treated as consisting of a FINITE number of places.

5. An address expressed by an address phrase always specifies the place in several 'steps', by gradually 'reducing the space'. For example, the address can specify, in this order: a country, a town in this country, a street in this town, a house in this street, and an apartment in this house. Thus, first of all the address specifies a place among OTHER PLACES in the same space, and all these places (a finite number of them) form a characteristic structure in this space. A place contains other places, of which each one contains yet other places, etc. In other words, we are talking about a set of places in the collective (mereological) sense, i.e. a set in which the relation of 'being an element' is transitive. This relation partly orders the places in the space, and thus gives the space a structure, in the language of graph theory called a TREE. The top of this tree represents the entire space, the remaining nodes represent the places in the space, and the links correspond to the relation of 'immediate constituency'.[4] Secondly, an address is composed of some portions of information, which can be ordered from the most general to the most detailed ones. This structure of information represents a certain structure of objects (places) which the information refers to. It concerns in particular a STRING of places of which every subsequent one is contained in the previous one. The string leads to the place which is the final element and at the same time the target indicated by the address, i.e. the denotatum of the address.

6. The places that we are talking about form certain TYPOLOGICAL GROUPS[5] in physical space. For example, there are groups of towns, groups of streets, groups of districts, etc. The principle of this division is that each two places, one being an element of the other, belong to two different typological groups. On the other hand, the tree structure of space divides all places represented in it into groups which are immediate constituents of the place located 'higher' in the structure. To distinguish these groups from the previous

---

do not consider this interpretation correct.

[4]When talking about elements B, C of a collective set A, we say that B is an immediate constituent of C, if $\forall D \in A \ (B \in D \in C \rightarrow D = B \lor D = C)$.

[5]A group in common meaning, i.e. a certain number of units forming a single separate whole.

ones, we will call them STRUCTURAL GROUPS. In a given physical space, a typological group can include elements of different structural groups (for example, a typological group of buildings in a town is composed of buildings which are elements of each street and square of the town, i.e. buildings belonging to different structural groups of this space). The opposite is also possible – a structural group may include elements belonging to various typological groups (streets and squares of a district in a town may form a structural group in this town, including elements of two different typological groups – streets and squares).

7. The denotata of address phrases are places such as towns, houses, streets, apartments, villages, etc. These are places meant for people: their places of residence, work, leisure or other activities. However, they are not of the same type as seats in a theatre, bus or train [translator's note: in Polish both types are referred to with the same word – *miejsce*] or a place in a queue, i.e. they are not intended for one person to lie, sit or stand; such places will not be indicated by addresses (we do not call the information written on a ticket to the theatre or train an 'address').[6] Moreover, the places referred to by addresses are LABELLED in the given space in a special way – just as streets, squares, houses, apartments, housing estates, etc., are labelled.

8. Places are labelled if there is a language sign (simple or complex) attributed to each of them (arbitrarily or based on a custom). The sign will be called a LABEL. Each of the labels carries one or several portions of information, which either specify the membership of the place in a relevant typological group, or distinguish the place in a given structural group. We will call this information, respectively: TYPOLOGICAL and SPECIFIC information. Each label carries no more than one typological information and at least one (usually only one) specific information. For example, the label *Kraków* expresses only specific information, while the label *ulica Dietla* [Dietel Street] expresses both typological and specific information. The first of the aforementioned conditions does not preclude the possibility that the place belongs to several typological groups, as a place can have more than one label, and these labels can carry different typological information.

9. The portions of information discussed above are represented in labels DISJUNCTIVELY, which means that if a label expresses more than one portion of information, it can always be divided into separate parts, of which each carries one portion of information. We will call these parts TYPOLOGICAL COMPONENTS and PROPER COMPONENTS respectively. A typological

---

[6]It seems, however, that expressions of this kind are one of the examples of the phenomenon for which we are building a theory here.

component is a generic name or an abbreviation of the name. The proper component can be a number, a letter, a combination of a letter and a number, a proper name or – with some generalisation – a part of a proper name, without the generic name contained in it (as the generic name, in this case, is the typological component of the label). For example, if we say that the typological component in the label *Pałac Staszica* is the word *pałac* [palace], and the proper component is the word *Staszica*, then it will be a simplification, as a proper name is not composed of separate words understood as separately denoting signs. More specifically speaking, a label can be attributed[7] a generic name identical in shape (or sound) to one of the inscriptions (or sounds) being parts of it, if only the place to which the label is attributed is the designatum of this generic name.

10. Place labels in a given space do not have to be unambiguous in this space. They are also usually not unambiguous within their typological group. As it turns out, however, they must be unambiguous within each structural group. For instance, house numbers are usually unambiguous only on a single street, and street names are often unambiguous only in a given town. The reason for this unambiguity is that an address is always perceived as UNAMBIGUOUS INFORMATION in a given space. For example, when someone says that they live in *ulica Krakowska 17*, and we refer the address phrase to the physical space of Kraków City, we expect that it is unambiguous in this space. It seems irrational to assume that there are two streets called *Krakowska* in the same city or two buildings numbered 17 on the same street. There may be some doubt, however, as to what space is concerned in a given case. This should always be clear from the context, circumstances or our knowledge. A phrase such as *ulica Pawia, 2nd floor, apartment no. 7*, in turn, would rather not be called an address, or at least would be deemed incomplete, as it does not identify any specific building in Pawia street. Further in this article, we will focus only on complete and unambiguous addresses in corresponding spaces.

---

[7]It is not, however, a simple task. For example, the labels *Osiedle Piastów, Osiedle Złotego Wieku, Plac Matejki, Plac Św. Ducha, Plac Wiosny Ludów* can be attributed generic names *osiedle* [housing estate, district] and *plac* [square], while the labels *Zielona Góra, Srebrna Góra* and *Babia Góra, Nowy Targ* and *Długi Targ, Nowa Kamienica* and *Szara Kamienica* cannot be attributed the names *góra, targ, kamienica* [mountain, marketplace, tenement] respectively, as some of the places with these labels are not designata of these names (some are proper names of towns or streets). As it turns out, determining whether a label is complex (carrying more than one elementary portion of information) or not and distinguishing the components of a complex label may be a difficult task.

11. As we have mentioned, an address specifies a place by defining a string of places leading to the place (cf. point 5). In other words, an address specifies a place by indicating several places forming a string, and each of these indications is a portion of information. An address phrase expresses the information by referring to labels. For example, the phrase *w najwyższym budynku stolicy na ostatnim piętrze, na które dochodzą windy samoobsługowe* [in the highest building in the city, on the last floor that can be reached by self-service elevators] is not an address phrase, as opposed to the phrase *w PKiN w Warszawie, na XII piętrze* [in the Palace of Culture and Science in Warsaw, on the 12th floor], although both these phrases indicate the same three-element string of places in the physical space of Poland, each expressing three portions of information. Naturally, we are aware of the fact that it is possible to find or build phrases for which it is not clear whether they are address phrases or not. In this article, however, we focus only on TYPICAL addresses, i.e. phrases that can be considered addresses without any doubt (however, see point 25).

12. The denotata of an address phrase are naturally the places the addresses of which it expresses. But each phrase refers also, at least to some extent, to each of the places forming relevant strings in the related space. For convenience, we will call these places PRE-REFERENTS of an address phrase, while the denotata of this phrase will be called FINAL REFERENTS. According to this terminological agreement, the last pre-referent in a string is the final referent.

13. From the pragmatic point of view, an address or an address phrase refers to two spaces: the one with which the creator associates it and the one with which the recipient associates it. Naturally, these two spaces are usually not identical, but they are always similar enough in the fragments to which the phrase refers that we can deem them undistinguishable. A space related to an address phrase will be further called an ADDRESS SPACE.

14. The structure of an address phrase, in particular the structure of the fragment we are interested in, always results from the shape and content of the address phrase to which it is related. For example, the phrase *Kraków, Pałac Pod Baranami, pokój 7* is related to the space, in which a city, building and numbered apartments have been distinguished, and the phrase *Kraków, Rynek Główny 28, room 7*, having the same final referent, is related to a space, in which a city, a square in this city, numbered buildings in the square and numbered apartments in the building have been distinguished. Sometimes, however (not only to indicate a relevant space, but also to retrace its structure), apart from the address phrase itself, we have to take into

account the circumstances in which it was used: the time, the place, the sender and the recipient (the association of the term *address* with the terms *addressee of an utterance* and *place of utterance*, which suggests itself here, is absolutely incidental).

15. Further in this article, it will still be easier to treat certain parts of an address phrase as elementary information and call them MORPHEMES.[8] They will include, first of all, conjunctions, punctuation marks and expressions acting as conjunctions in a given situation. We will simply call all elements of this group CONJUNCTIONS. Another group of morphemes are the smallest parts of address phrases corresponding to the components of labels of the places that are pre-referents of these phrases, thus at the same time corresponding to the specific information carried by these labels. The morphemes belonging to this group will be called PROPER morphemes. The third group of morphemes will include the generic names of the aforementioned places. We will call them GENERIC morphemes. We should add that a generic morpheme does need to have a counterpart among the generic components of place labels in a given space. It is enough that such a name acts as a generic name of a relevant place. The correspondence between proper and generic morphemes of labels of places that are pre-referents of these phrases is a PARADIGMATIC RELATIONSHIP in a broad meaning. This relationship includes the relations between various inflectional and conjugational forms of words and various forms of other signs which make up expressions, as well as the relations between an abbreviation and the full form, a number and the corresponding numeral, and sometimes even synonymy or paraphrasing.[9]

16. Apart from morphemes, address phrases may contain other language signs. However, from our perspective, these signs have nothing to do with the interpretation of these phrases, therefore they will not be discussed here.

17. For convenience, we will treat some morphemes or morpheme sequences as ELEMENTARY REFERENTIAL parts of the address phrase. This regards proper morphemes or sequences of morphemes containing proper morphemes and referring to single pre-referents of an address phrase. We will call these parts of address phrases ATOMS. Each atom contains not more than one generic morpheme. We can also assume that one atom can contain no more than two proper morphemes (for it is hard to imagine a label including more than two proper components) . The information expressed by an atom

---

[8]The term 'morpheme' will be used in the technical meaning defined above.

[9]Due to the broad understanding of the paradigmatic relationship, finding a label corresponding to the analysed morpheme or a sequence of morphemes in each case may sometimes prove to be a complicated task (for an automaton).

will be called ATOMIC. Taking into account the aforementioned restrictions, an atomic information will be composed of one, two or three portions of information.

18. An address phrase is a verbalised information about the paths to be taken by the address space tree, related to this phrase in order to reach a given place in this space. If we look at address phrases this way, every phrase of this kind can be considered a description of a step-by-step procedure of moving around in a given space. Every atom of an address phrase can be treated as an instruction defining a single step. Each step can be made only once the previous ones have been completed. As we can see, atomic information is not autonomous – pieces of information are bound to each other in various ways. In particular atomic information can be DIRECTLY DEPENDENT, one from the other, and MUTUALLY INDEPENDENT. Let a_1, a_2, ... a_k represent atomic information contained in an address phrase and let A_1, A_2, ... A_k represent the corresponding pre-referents of this phrase in the address space related to this phrase.[10] We will say that information a_i is directly dependent on information a_j, if place A_i is spatially contained in place A_j and no A_p among the remaining pre-referents of this phrase fulfils the condition $A_i \not\subseteq A_p \not\subseteq A_j$. We will say that the information a_{i_1}, a_{i_2}, ..., a_{i_p} selected from a_1, a_2, ... a_k is mutually independent, if none of the places A_{i_1}, A_{i_2}, ..., A_{i_p} is contained in any of these places. For example, in the address phrase *w Krakowie na Kazimierzu przy ul. Szerokiej 7 oraz przy ul. Miodowej 5 i 14* [*in Krakow, in Kazimierz district, 7 Szeroka Street and 5 and 14 Miodowa Street*], the information expressed by the atom *na Kazimierzu* is directly dependent on the information expressed by the atom *w Krakowie*, and the information expressed by the atom *14* is directly dependent on the information *przy ul. Miodowej*; the information expressed by the atoms *przy ul. Miodowej* and *przy ul. Szerokiej* is mutually independent, etc. The concepts of direct dependency and mutual independence can be extended by including complex information, i.e. portions of information composed of more than one piece of atomic information. For example, in the analysed phrase, the information expressed by the biatomic phrase *przy ul. Szerokiej 7* is directly dependent on the information expressed by the atom *na Kazimierzu* and, at the same time, on the information expressed by the phrase *w Krakowie na Kazimierzu*. The information expressed by the phrases *przy ul. Szerokiej 7* and *przy ul. Miodowej 5 i 14* is, in turn, mutually independent.

19. As we mentioned at the beginning (see points 2 and 3), an address phrase

---

[10]Sometimes two pieces of atomic information a_n, a_i or a greater number of pieces of information correspond to the same pre-referent – see point 23.

can express several addresses, without being composed of address phrases corresponding to these addresses, whereas an address phrase expressing a single address does not necessarily have to be a sequence of atoms corresponding to the sequence of its pre-referents, forming a string in the related address space (like, for example, in the following phrase: *w budynku nr 72 w Krakowie przy ul. Dietla* [*in the building no. 72 in Krakow on the Dietla Street*]). The order of atoms and the use and selection of conjunctions depend, to a large extent, on stylistic considerations decided by the author of an address phrase. It is governed, however, by at least two rules. According to the first rule, sequences of atoms and conjunctions placed next to each other (further called syntagmas), expressing mutually independent information, must be separated by conjunctions. However, conjunctions can also appear between syntagmas expressing directly dependent information. The second rule (more general) is that the structure of an address phrase should ensure unambiguous interpretation of a relevant address space. We can assume that this rule is fulfilled, at least in typical address phrases.

20. An address phrase is an expression composed of atoms and conjunctions, and each atom is a morpheme, a pair or a trio of morphemes, of which at least one is a proper morpheme, and no more than one is generic. The FORMAL STRUCTURE of this phrase can be illustrated by placing conjunctions in angle brackets, proper and generic morphemes in parentheses and atoms in square brackets, leaving other signs in the phrase without special distinction (e.g. pronouns). The marker of the formal structure thus constructed will be called an f-MARKER. For example, the formal structure of the following address phrase:

(1) *w Krakowie przy ul Brackiej 2, Brackiej 23 oraz przy placu Matejki 15* [*in Krakow on 2 and 23 Bracka Street and on 15 Matejko Square*] is represented by the following f-marker:

(2) *w* [(*Krakowie*)] *przy* [(*ul.*) (*Brackiej*)] [(*2*)] <,> [(*Brackiej*)] [(*23*)] <*oraz*> *przy* [(*placu*) (*Matejki*)] [(*15*)]

21. The structure of atomic information expressed by an address phrase, called in short a STRUCTURE OF INFORMATION, differs from the formal structure of the phrase. A simple method to show the structure of information is to attribute a marker of this structure, further called the f-marker, to the address phrase. We build the marker using a_i, $+$ and $\rightarrow$. Each a_i represents an atomic information carried by the i-th atom of the given address phrase (i.e. the atom which is in the i-th position in this phrase), each '+' represents a conjunction, and the arrows connect some of the a_i symbols. In particular, if an arrow connects a_i with a_j, then the information represented by the

symbol a_j (toward which the arrow points) is directly dependent on the information represented by a_i; if, however, arrows point from a_i to a_{j_1}, $a_{j_2}, \ldots a_{j_p}$, then the information represented by symbols $a_{j_1}, a_{j_2}, \ldots a_{j_p}$ is mutually independent. For example, the structure of information of the address phrase (1) is represented by the following *i*-marker:

(3)

$$a_1 \quad a_2 \quad a_3 \ + \ a_4 \quad a_5 \ + \ a_6 \quad a_7$$
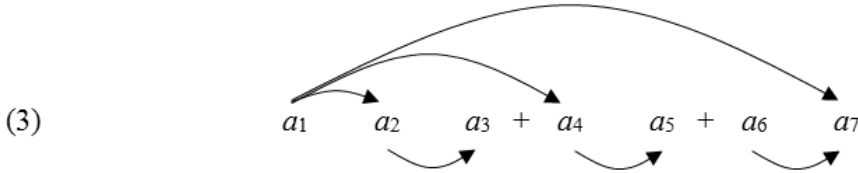
22. Let us now consider an address phrase specifying the same strings of places as (1), but in a slightly different way:

(4) *w Krakowie pod numerami 2 i 23 przy ul. Brackiej, a także przy placu Matejki 15*
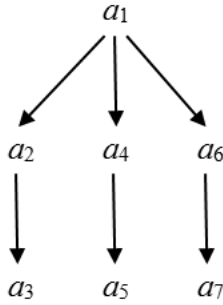
The formal structure and the structure of information of this phrase is shown by the following markers:

(5) *w* [(*Krakowie*)] *pod* [(*numerami*) (*2*) <*i*> [(*23*)] *przy* [(*ul.*) (*Brackiej*)] <*, a także*> *przy* [(*placu*) (*Matejki*)] [(*15*)]

(6)

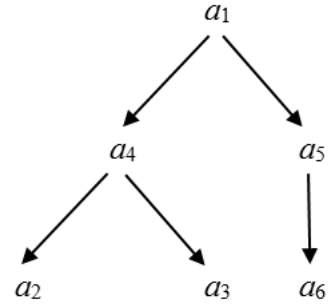$$a_1 \quad a_2 \ + \ a_3 \quad a_4 \ + \ a_5 \quad a_6$$
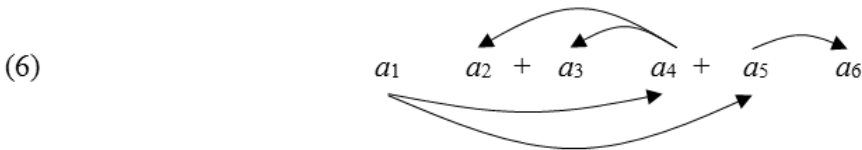
We can see the similarity of the structures of information in the address phrases (1) and (4) by presenting the relevant *i*-markers, i.e. markers (3) and (6), as trees (the symbols of conjunctions are omitted):
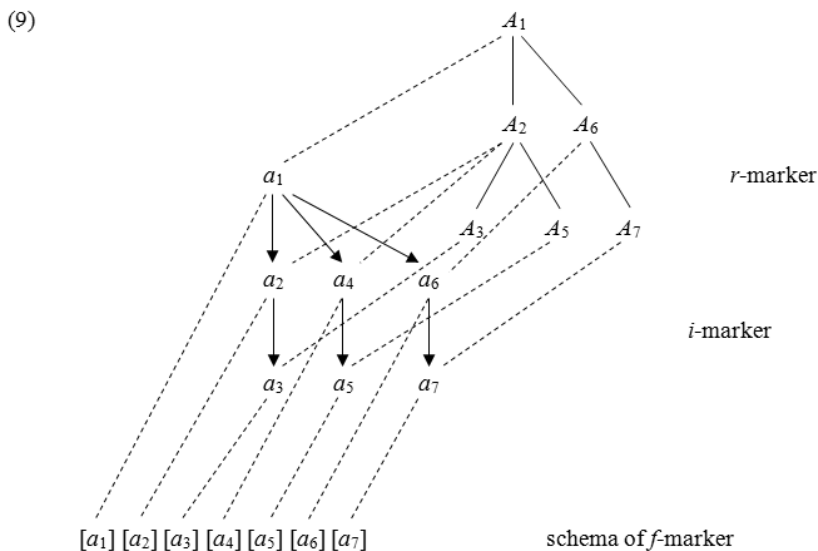
(7)

(8)

23. An address phrase specifies a certain structure of places (pre-referents) in the related space. The structure, here called the STRUCTURE OF REFERENTS, can be isomorphic to the structure of information – in fact, it always is, if each pre-referent has exactly one atom corresponding to it in the phrase. On the other hand, these structures may also be merely similar, not isomorphic – it is so, if two or more atoms correspond to some of the pre-referents. The structures of referents of address phrases (1) and (4) are identical; as regards phrase (4), its structure of referents and structure of information are isomorphic, while in phrase (1) the structure of referents differs from the structure of information (for example, it has a different number of elements). Diagram (9) shows the relations between the formal structure, the structure of information and the structure of referents of the address phrase (1). The formal structure of this phrase is represented in the diagram by the image of its f-marker, which is a sequence of symbols $[a_i]$ replacing the sequence of atoms (conjunctions and other signs of the phrase will not be shown in the diagram). The structure of referents of phrase (1) is represented by an r-MARKER, composed of symbols A_i representing pre-referents corresponding to atomic information a_i (see point 18). If a referent corresponds to several pieces of information a_{i_1}, a_{i_2},... a_{i_p}, then it is marked in the r-marker with the lowest of the indexes $i_1\ i_2, \ldots, i_p$. The edges of the r-marker represent the segments of the relation of 'being an element' in the given address space (see point 5).

(9)

A_1

A_2 A_6

a_1 *r*-marker

A_3 A_5 A_7

a_2 a_4 a_6

a_3 a_5 a_7 *i*-marker

$[a_1]$ $[a_2]$ $[a_3]$ $[a_4]$ $[a_5]$ $[a_6]$ $[a_7]$ schema of *f*-marker

24. So far, we have been neglecting all grammatical differences. But the difference between the singular and plural forms of generic names which are found in address phrases is important to us. An atom in which a generic morpheme is in the singular form[11] carries some additional information – namely that the phrase contains atoms which are in mutual independence relation with the former and correspond to the pre-referents of the same type as the former and not containing generic morphemes. Apart from that, we have neglected the differences between conjunctions, although conjunctions may vary in the 'strength' of connecting, and these differences may be of importance to us. Let us consider two address phrases:

(10) *w Krakowie przy ul Długiej 2 i przy ul. Miodowej 5 oraz przy ul. Kruczej 7 w Warszawie* [*in Krakow on 2 Długa Street and on 5 Miodowa Street as well as on 7 Krucza Street in Warsaw*]

(11) *w Krakowie przy ul Długiej 2, a także przy ul. Miodowej 5 i przy ul. Kruczej 7 w Warszawie* [*in Krakow on 2 Długa Street as well as on 5 Miodowa*

---

[11]When we are dealing with an abbreviation of a generic name, it is often impossible to recognise the plural form.

Street and 7 Krucza Street in Warsaw]

If we treated the conjunctions in these phrases as indistinguishable, we would not be sure whether they indicate the *Miodowa* street in Krakow or the one in Warsaw (referring to the structure of the given address space would not help, as there is a street with that label in both cities). The fact that we do not have these doubts proves that the differences between conjunctions must be taken into account in our discussion. However, there is not enough place here to present the details of relevant modifications.

25. Finally, a few words about unusual address phrases. We often encounter expressions which, as our language intuition suggests, contain addresses, although their structure differs from what we have said so far about the structure of address phrases.[12] These are in particular those expressions in which, instead of several atoms with the construction described above, corresponding to the same pre-referent, there is a pronoun in place of a proper morpheme in the second atom or one of the further atoms (e.g. *przy ulicy Pawiej 8, a także przy tejże ulicy pod numerem 15* [*on 8 Pawia Street and on the same street no. 15*]). This case is called pronominal substitution. Proformal substitution also seems possible here.[13] In our opinion, none of these two types of substitution violate the theoretical constructions proposed above. These substitutes of proper morphemes will be called NON-PROPER MORPHEMES.

*

*   *

Let us now pass over the generalisation of our deliberations (the generalisation would involve isolating the introduced terms and constructions from the model on which they are based) and check whether the theoretical approach presented above can be used to interpret expressions which are the main subject of this study, and if yes, then how they can be used. As we have said at the beginning of this article, we are interested in certain referential expressions used in legal texts. Without going into any further detail (which

---

[12]We can also assume that natural language gives us the possibility to create expressions departing ever further from the proposed type, but which can still be identified as containing addresses, up to those as to which we would have doubts whether they contain addressed or not.

[13]The theory of substitution was introduced and developed by Ronald Harweg (1978).

would be necessary for adjusting the theory to a new model), we propose an analysis of one example.

In the first article of a certain law we find a phrase referring to utterances from the corpus of legal texts. The addresses of these utterances are marked by the expression (*), which is a part of this phrase.

(*) *w §§1, 2 i 5 artykułu drugiego niniejszego rozdziału, w art. 7, 9 i art. 15 §7 rozdziału następnego, a także w rozdziale XII, art. 123–125* [*in paragraph 1, 2 and 5 of the second article of this chapter, in article 7, 9 and article 15 paragraph 7 of the next chapter, as well as in chapter XII, articles 123–125*]

The formal structure of this address phrase (we now apply the terminology of our theory) will be shown by the following f-marker:

(**) w [(§§)(1)] <,> [(2)] <i> [(5)] [(artykułu) (drugiego)] [(niniejszego) (rozdziału)] <,> w [(art.) (7)] <,> [(9)] <i> [(art.) (15)] [(§)(7)] [(rozdziału) (następnego)] <, a także> w [(rozdziale) (XII)] <,> [(art.) (123–125)]

In this marker, there are two non-proper morphemes (*niniejszego, nastepnego*). Instead of information about the proper components of the labels of chapters, these morphemes carry information about the positions of these chapters in relation to the position of the utterance containing the address phrase. Moreover, the last morpheme of the phrase carries information about two proper components of labels of two articles. These articles are the limits of the language unit which is a sequence of articles. The whole unit is, in this case, the pre-referent of the address phrase and the atom containing this 'unusual' morpheme defines this unit. The structure of information of the address phrase (*) is represented by the following i-marker:

(***)
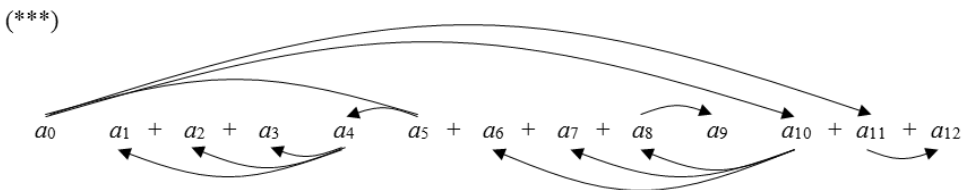
The symbol a_0 means a non-verbalised information about the address space related to phrase (*), i.e. the space in which we will find the referents of this phrase. In this case, the space is the text of the law to which the analysed phrase belongs.

Bibliography

1. Doroszewski, Witold, ed. (1958) *Słownik języka polskiego. Vol. I.*

Warszawa: PWN.

2. Harweg, Roland (1977) "Substitutional Text Linguistics." In *Current Trends in Text Linguistics*, ed. Wolfgang U. Dressler, 247-259. Berlin: W. de Gruiter.

3. Łachwa, Andrzej (1986) "Postulat naturalności przedmiotowej." Materiały na I Konwersatorium „Sztuczna inteligencja", cz. I, 91-99. Warszawa: Polskie Towarzystwo Cybernetyczne.

4. Studnicki, Franciszek (1985) "Wykładnia prawa w jej aspekcie komputacyjnym." *Studia Filozoficzne* 2-3: 93-104.

5. Studnicki, Franciszek, Polanowska, Beata, Stabrawa, Ewa, Fall, Jarosław M. and Andrzej Łachwa (1983) "Zautomatyzowane rozwiązanie odesłań występujących w tekstach prawnych." *Studia semiotyczne* XIII: 65-90.

6. Studnicki, Franciszek, Stabrawa, Ewa, Fall, Jarosław M. and Andrzej Łachwa (manuscript) *Sprawozdanie z wyniku projektu badawczego "Anaphora".* Kraków: Katedra Informatyki Uniwersytetu Jagiellońskiego.

7. Szymczak, Mieczysław, ed. (1982) *Słownik języka polskiego. Vol. I.* Warszawa: PWN.