

Leon Koj
THINKING AND INNER SPEECH

Originally published as "Myślenie i mowa wewnętrzna," *Studia Semiotyczne* 7 (1977), 227–248. Translated by Kinga Bulsiewicz.

INTRODUCTION

a. The most important problem of semiotic research boils down to the attitude of speech and language towards thought. Deciphering this connection shall allow us to understand the impact of speech upon all human actions. Unfortunately the influence of speech and language on thought is a very complex, multilevel and difficult issue. From many possible layers of this problem I would like to analyze one. I wish to cast some light on the hypothesis that one type of human thought is verbal. At the same time I do not wish go into consideration whether there actually exist any other types of thought. I shall also not consider the mechanisms governing verbal thought or what are its functions among all processes of mental activities. My task shall be limited: I wish to more specifically formulate the hypothesis about the verbal character of a type of human thinking, subsequently to draw out certain simple consequences to this hypothesis. These consequences may help to prove the hypothesis.

b. I have the following reason to limit my analysis to making the above described hypothesis more precise. The disputes regarding the thought perceived as internalized speech are very abstruse. The reason behind that is: the said disputes originate from the multitude of information that may influence the outcome of the hypothesis. Some researchers use introspective data to explain the verbal character of some types of thought. Others prefer the data collected with a miograph. Yet another group of researches analyzes the data gathered through observing people with impaired speech and hearing. There are also some who find research on brain damage and the resulting speech and mental function defects to be most valid. Data gathered by

developmental psychology are taken under consideration. The hypothesis is also backed by general preconceptions about the purposefulness of functioning of living organisms: it is more economical to try out the designed action using internalized tools than to act it out thoroughly, wasting both time and energy. Unfortunately, combined data from all these fields is inconclusive. Since each of these fields provide data both in favor and against the hypothesis.

Another thing causing the complexity are misunderstandings. There is no agreement as to how internalization of speech works. Are we considering Watson's micromovements of the organs, or the changes in electric potential of the speech organs? Are only changes in electric potential and chemistry of the brain considered, or maybe undefined physical manifestations of words?

Similar or even greater terminological differences surround the key term "thinking" which is known for its numerous meanings.

The inconclusiveness of certain facts may be caused by misunderstandings. The said facts fit in with one way of understanding the hypothesis, while another way of understanding it may cause the same facts to become irrelevant and prove nothing.

The strong dependency of the connection between thinking, internalized speech and epistemological concepts is another circumstance making it harder to come to less contradictory conclusions. Some of the proposed solutions reference nominalism, others conceptualism. Sometimes epistemological arguments introduce elements that transcend the empirical and have the form of a postulate.

Due to the above described complex theoretical situation it seems appropriate to divide the problem into smaller, simpler and mutually exclusive parts. It also seems that the right thing to do is to more precisely formulate the hypothesis. The possibility to prove the hypothesis shall remain secondary.

c. The notion to specify common statements is often considered dubious. Such specified statements are often accused of transforming the original intent of the utterance so that the final result bears a completely different content than the original. Such issues may also occur herein. The basic method of analyzing statements is through determining the language in which the statements will be again formulated. This language usually has only a limited number of initial concepts which must suffice to express a terminologically richer common utterance.

There is no other way, statements formulated in an isolated, limited language are one-sided and their content is simplified. Despite these limitations, or maybe because of them, specification is so useful. Since there are only

few terms (concepts) present in the isolated language they may assume a symbolic form. This allows for a particular statement to be written down precisely and without any misunderstanding. Any extra comments, interpretations or additions present in the common language lose their influence on the process of identifying the relationship between isolated statements. This last moment seems most crucial: the distillation of language and symbolic recording of statements allow us to disregard the verbal, constantly fluctuating expressions and apparent additions which in fact change the recorded expressions to such an extent that the consequences of statements change drastically. Therefore distilling a part of a language and remaining strictly within its boundaries shall be my primary method of analysis herein.

In the available literature on the subject the hypothesis is justified as follows: numerous different facts are mentioned for and against the connection between thinking and internalized speech. These facts are presented in such a way to prove the said connection or to prove some other positive thesis. Such explanations will always be flawed. It is impossible for a general and unrelated thesis to be based on detailed and particular data. Therefore I shall still analyze the correctness of the hypothesis, however I apply a different approach. I shall begin with making an assumption about a hypothesis (or more precisely a group of hypotheses) and I shall make it my starting point without providing explanations. From there I shall venture to analyze the consequences of these hypotheses. If these consequences prove to be true, then the hypothesis will be confirmed. If the consequences prove to be false than at least a part of the hypothesis will be disproved. The whole argumentation for or against the hypothesis will be limited to determining its logical outcomes. This course of action significantly limits the scope of research, because presupposed and simple hypotheses may not be used to prove everything that could have a connection to the analyzed hypothesis about verbal nature of thoughts. Most of the known data cannot be expressed in the language which I shall use to formulate the hypothesis and its consequences.

FORMULATION OF THE HYPOTHESIS

Contrary to what intuition may suggest the hypothesis about the connection between thought and internalized speech does not consist of one sentence. Several questions need to be answered as a part of this hypothesis. Well, thinking is supposed to, in one way or another, be connected to inner speech. This is one of the elements of the hypothesis. Additionally I need to determine how does internalized speech work — the basic issue here is how similar it is to communicative speech. Only this similarity may give

grounds for calling certain inner processes 'speech'. This, however, requires a proper characteristic of communicative speech. The hypothesis about the connection between thinking and internalized speech consists of three sets of sentences which I shall present in the following order. The first set characterizes communicative speech. One can easily imagine how complicated a full characteristic of communicative speech might be. It is completely impossible to present it here in full. I shall provide a part of it comprising of three simple theorems. The lack of syntactic description of both the surface and deep level of the communicative speech is an obvious limitation that will also influence the characteristic of inner speech to which the second set of sentences will be dedicated. This second set states the similarity between inner and communicative speech. The third set of sentences touches the issue most important for this text: it establishes the connection between inner speech and thinking.

a. Firstly I need to determine the concepts which I shall use hereafter. Choosing these concepts is at the same time a way to more precisely define the scope of analysis. As I previously declared I shall try to limit the subject matter, therefore I shall consider only the necessary concepts without which it would be impossible to formulate the hypothesis. I want to establish the connections between thinking and inner speech, so I need these two concepts symbolized respectively by T and I . Since inner speech ought to resemble communicative speech, the last one therefore requires its own group of concepts. I shall utilize the concept of using an expression (a piece of writing, a sound, a movement or concaveness/convexedness etc.) as a sentence sign or in some cases as a naming sign. Their respective symbols are S_s and S_n .

The basic difference between the signs of communicative speech and inner speech is the fact that we can inter-subjectively, articulately and distinctly create and receive; whereas one cannot say the same about inner speech yet. To underline this difference I shall introduce the concept of noticing N . With inner speech we do not use signs understood in the normal way (that is the inter-subjectively perceived signs), but their respective equivalents. When I will be describing inner speech I shall use the concept of sign equivalents. Therefore I shall introduce yet another concept — sign equivalent — which will be symbolized by a vertical line | on both sides of a symbol.

While formulating a hypothesis about the relationship between thinking and inner speech I need to also consider the time and the person. It is possible that sometimes things happen in a certain way and sometimes in another. Some things always coincide, others occasionally. For example, when we use inner speech we sometimes also notice the signs we create inside.

Therefore our concepts (T, I, S_s, S_n, N) must correspond to time and person. They must assume the predicate form like $T(t, v...)$, $I(t, v...)$, $S_s(t, v...)$, $S_n(t, v...)$, and $N(t, v...)$ — where t is a time variable and v is a person variable. The concept of an equivalent $|$ does not correspond to anything, it is a function. This concept does not have to correspond, because it will occur in the context of the already corresponding concept $I(t, v...)$.

These instruments need to be supplemented with nonspecific tools from the field of logic. It would be convenient to assume that I can use a part of the set theory (the set membership symbol) in addition to the widely understood functional calculus. I shall also borrow two concepts of a part. One concept of a part will be C and it is purely syntactic in character, another, more general is P which I shall use when I will want to underline that one time segment is a part of another time segment. I am not going to characterize these borrowed concepts. Another addition to conceptual instruments are the names of expressions or more precisely type variables for expression sets. Variables $\emptyset, \emptyset', \emptyset''$ etc are type variables for a set of sentence names; a, a', a'' etc are type variables for a set of naming expressions' names; finally η, η', η'' etc are type variables for a set of arbitrary expressions' names. Names for types of expressions I shall formulate in a well known way, by listing the variables and using corner quotes.

All theorems included in the analysis of the hypothesis will be formulated only with the use of the above listed concepts. It is clear that limiting the number of instruments results in narrowing the possibility of expressing interesting and important theorems. However, this limited possibility to describe a wide assortment of facts and phenomena may also mean that with use of such a small (compared to the size of the problem) set of instruments it will be possible to precisely say things which are usually formulated in a very sloppy and trivial manner.

b. The goal is to consider the connections between inner statements (whatever they really are) and thinking. If such statements exist they are utilized to think about objects and phenomena. Therefore they are utilized much in the same way as signs in communicative speech. While speaking about signs in the context of thinking I shall refer to their semantic not syntactic aspect. The concept of sentence sign $S_s(t, v...)$ will be complemented by sentence sign's name and an occurrence which draws attention. This means that a full notation of a sentence sign looks as follows: $S_s(t, v, \emptyset, p)$. This symbolic notation will be read like this: in time t for a person v sentence \emptyset is the sign of the occurrence p . For similar reasons the notation of the naming sign needs to be expanded to the following form: $S_n(t, v, a,$

y). This symbolic notation means as much as: in time *t* for a person *v* name *a* is the sign of the object *y*. In the case of inner statements the notation $I(t, v, \dots)$ requires that the name of the innerly expressed creation will be added. Therefore the final form of the concept is $I(t, v, |\eta|)$ which will be read as: in time *t* a person *v* innerly expresses the equivalent of η i.e. $|\eta|$. The concept of noticing is obvious. We notice signs (we do not discuss other things, because they are of no interest to us). Therefore I propose this notation $N(t, v, \eta)$ which reads: in time *t* a person *v* notices an expression η . Finally to the concept of thinking. I make the assumption that we are thinking not about objects but occurrences. About objects we simply think, we assign certain notions to them. Such thoughts are judgments — sentence equivalents. As a result of that I shall complete $T(t, v, \dots)$ with sentence signs: $T(t, v, p)$ or $T(t, v, z(x))$ where $z(x)$ is function with the variable *x* and the predicate *z*. $T(t, v, p)$ should be read as follows: in time *t* a person *v* thinks that *p*.

c. Now I shall characterize the concept of a sentence sign and provide similar explanations for inner speech. I shall assume that the concept of sentence signs is extensional. Moreover I shall assume that if the signs are identical and utilized by the same person at the same time then such signs are equal.

$$A1. \emptyset = \emptyset' \rightarrow [S_s(t, v, \emptyset, p) \equiv S_s(t, v, \emptyset', p)]$$

I shall also assume that if within the whole time *t* one expresses \emptyset and at the same time expresses \emptyset' , then \emptyset and \emptyset' are identical. To paraphrase: within one time only one sentence can be expressed.

$$A2. S_s(t, v, \emptyset, p) \wedge S_s(t, v, \emptyset', p) \rightarrow \emptyset = \emptyset'$$

The above assumption may be considered true only if the time *t*, when \emptyset and \emptyset' were used, will be properly comprehended. If this time be too long then it would be obviously possible to use two different signs. Therefore the sentence sign utilization time will be defined as a very short time, so short in fact that the whole time *t* is the time required to utilize the sign \emptyset . It is the time of the actual utilization of the sign without any pauses or moments of consideration, which are not the utilization (by the speaker or recipient) itself. Such a specification of time is given by

$$A3. S_s(t, v, \emptyset, p) \rightarrow \sim \bigvee_{t_1} [S_s(t, v, \emptyset, p) \wedge t_1 \neq t \wedge t_1 P t]$$

I shall not describe sentence signs any further. Other theorems might have described the process of communication between a few people. Inner speech, however, does not have a communicative function, and the goal of the three initial theorems was to provide a formula for internalized speaking, so any additional assumptions about sentence signs are not necessary.

Having established certain formulas in A1-A3 I can now describe inner

speech. If it is in fact a type of speech it has to be similar to communicative speech as much as possible, also in the context presented in A1-A3. To ascertain this similarity, at least to some small degree, I shall assume that inner speech will fulfill similar theorems as those described in A1-A3.

Most importantly I shall assume that internalized speech is as extensional as communicative speech.

$$\text{B1. } \emptyset = \emptyset' \rightarrow [I(t, v, |\emptyset|) \equiv I(t, v, |\emptyset'|)]$$

Seemingly B1 is different than A1 — it states that signs are the same, therefore their spoken equivalents are equal. I can make B1 more similar to A1 by making one additional assumption:

$$\text{B1'. } \eta = \eta' \equiv |\eta| = |\eta'|$$

B1' states that a signs' equivalents are equal when signs are equal. It is also true for a reverse statement. Since two communicative signs cannot be spoken at the same time I shall assume the same about internalized speaking:

$$\text{B2. } I(t, v, |\emptyset|) \wedge I(t, v, |\emptyset'|) \rightarrow \emptyset = \emptyset'$$

Making inner speech and communicative speech alike may have profound consequences since it may determine the linear character of thinking connected with internalized speaking.

Finally, I shall analogically define time which is necessary to internally express sign equivalents:

$$\text{B3. } I(t, v, |\emptyset|) \rightarrow \sim \bigvee_{t_1} [I(t, v, |\emptyset|) \wedge t_1 Pt \wedge t_1 \neq t]$$

This does not mean, however, that the time of internalized expression should be identified with the time needed to say a sign. At this point in both cases the time in question is the moment necessary to express signs or their equivalents. In the light of the statements made thus far these times may or may not be identical — this is yet to be established.

d. Now I shall analyze the core of the hypothesis in question which links at least one type of thinking to inner speech. The simplest way to formulate this concept is: thinking about something implies innerly saying some sentence:

$$\text{(a) } T(t, v, p) \equiv I(t, v, |\emptyset|)$$

The above formula is unfortunately unacceptable. It may be, however, divided into two parts which I shall analyze separately:

$$\text{(b) } T(t, v, p) \rightarrow I(t, v, |\emptyset|)$$

$$\text{(c) } I(t, v, |\emptyset|) \rightarrow T(t, v, p)$$

(b) states that the consequence of thinking p is speaking the sentence equivalent \emptyset . In this case the letter " p " is a variable, so according to the simple rules governing variables it can be substituted with any sentence and

nothing besides p will change. This would mean that the consequence of thinking would always be expressing the same sentence equivalent \emptyset . If \emptyset type sentences were a part of the common language then each and every person thinking about anything would have to speak \emptyset type sentences of the Polish language. For instance each Hindu person would have to speak Polish sentences if \emptyset would be Polish sentences. Such consequences of (b) are unacceptable. The difficulty lies in the lack of any connection between the innerly spoken sentence and the occurrence that is thought about at that moment. A similar thing will happen in (c) when \emptyset is substituted with a sentence. This leads to the following consequence: inner expression of the equivalent \emptyset allows for thinking about anything. Moreover (c) prompts further subjects for consideration. According to (c) inner speech always leads to thinking. Therefore there is a postulate hidden in (c) that inner speaking (which is the issue herein) is not just some random internalized speech that may happen while learning some meaningless syllables, as Ebbinghaus did. Formula (c) demands that inner speech results in thinking. As a result (c) implicitly requires a division of inner speech into a different type where only one of them is relevant to the subject of this article. This part of (c) seems to be correct and it should reappear in the analysis below.

Regardless of the above mentioned problems to establish the connection between internalized speaking and thinking I need to establish the connection between the occurrence that is thought about and the sentence referencing it. One can assume that when \emptyset is a sign of p then a sufficient equivalence occurs between thinking and inner expressing of the equivalent \emptyset , thus:

$$C(1) S_s(t, v, \emptyset, p) \rightarrow [I(t, v, |\emptyset|) \equiv T(t, v, p)]$$

It should be noted that the equivalence in (1) occurs under the condition that in a given moment t person v (as the speaker or the recipient) uses the sentence \emptyset as a sign of p . This condition is relatively rarely fulfilled. We often think without speaking at the same time, nor listening to or reading anything. Therefore C1 does not give much information about internalized speech and thinking. We have no idea what are relations between speech and thinking in cases when no sign is being used. One may consider the following additions of rather limited importance:

$$C1' \bigvee_p [T(t, v, p)] \equiv V\emptyset [I(t, v, |\emptyset|)]^1$$

$$C1'' I(t, v, |\emptyset|) \rightarrow \bigvee_{pt_1} [T(t, v, p) \wedge S_s(t_1, v, \emptyset, p)]$$

At the same time the following version is unacceptable:

¹Hypotheses which numbers are followed by an apostrophe are partial hypotheses which I shall not use in the course of further analysis.

$$(d) \ V_t S_s(t, v, \emptyset, p) \rightarrow \bigwedge_{t_1} [I(t_1, v, |\emptyset|) \equiv T(t_1, v, p)]^2$$

The above states that if \emptyset was used as a sign of p , then innerly uttering an equivalent of \emptyset always leads to thinking about p . This would introduce the postulate of synonymy of all sentences. Such a postulate would be very far removed from reality.

C1 may be a source of numerous doubts. These arise mostly when it comes to the time necessary to think and to utter a sign. While writing about B3 I showed that times necessary to think (innerly verbalize) and to utilize a sign do not have to be equal. This matter was settled in C1. Statements made by numerous physiologists allow me to assume that thinking and inner verbalization takes less time than a communicative expression of a given thought. It may be assumed that often the thought precedes the utterance. This, however, does not completely exclude situations when we think as slowly as we speak. Taking the above into account it can be concluded that C1 informs us about this equivalence of thinking and inner speech which occurs when appropriate signs are utilized. The condition of equivalence is not fulfilled as often as I have suspected. For the time being it is impossible to acknowledge more frequently occurring conditions of the equivalence introduced by C1.

Assumption C1 states that under some conditions thinking is accompanied by inner speech. C1 does not, however, straightforwardly connect inner utterances with thinking. C1 allows for a situation when one inner utterance is accompanied by numerous thoughts. In such a case the relationship between internalized speech and thoughts would be rather accidental and loose. If some type of thought is to have a verbalized character then a straightforward attribution of thought to the appropriate innerly spoken statement must be assumed. C2 is such an assumption (definition of $\overline{t_1'v}$ will be provided below):

$$\begin{aligned} \text{C2. } & I(t, v, |\emptyset|) \rightarrow (T(t, v, p) \wedge T(t, v, q) \rightarrow \\ & \rightarrow [T(t, v, p) \stackrel{t_1'v}{=} T(t, v, q)]) \wedge \\ & \wedge \{ [T(t, v, p) \stackrel{t_1'v}{=} T(t, v, q)] \rightarrow [T(t, v, p) \equiv T(t, v, q)] \} \wedge \\ & \wedge \{ [T(t, v, p) \equiv T(t, v, q)] \rightarrow (p \equiv q) \} \end{aligned}$$

C2 has the following meaning: when v innerly utters \emptyset and at the same time thinks about p and about q then those thoughts about p and about q are identical. The fact that these thoughts are identical results in their equivalence, equivalence of thoughts means that the thought about occur-

²Sentences enumerated with the use of letters only (ex. (a)) have been rejected.

rences coincided. The two last statements describe the identical character more precisely, that is there are no differences between them, there is also no difference in their subject.

In C2 there is the symbol " $\underset{t_1 v}{=}$ " meaning that thoughts are exactly identical. I define this symbol as follows:

$$D1. p \underset{t_1 v}{=} q \equiv [T(t_1, v, p) \equiv T(t_1, v, q)]$$

D1 comes from certain intuitive conclusions: one perceives two occurrences as identical when one thinks about them at the same time. This definition is based on cognitive realism.

It is impossible to tightly connect thinking to internalized speaking. Many expressions have synonyms. I shall assume that synonyms — being equivalents — are connected to the same thoughts. In case of a full identification of thinking with internalized uttering of equivalents of synonymous sentences thoughts on one hand should be identical (because of synonyms), and different on the other (because they would be identical with inner utterances of different form). Such full identification combined with the theory of the existence of synonyms would lead to contradiction. It is hard to reject the existence of synonyms. There are also no grounds to reject the conclusion that synonyms evoke the same thoughts. The only thing left is to reject the theorem that thinking and inner speech are identical. This solution has another advantage, it allows for other inner speech internalized activities within the boundaries of thinking. It is widely known that one aspect of Piaget's psychological theory is based on the conviction that internalized motions are a part of thinking. Therefore acknowledging the existence of synonyms I must assume that thinking and inner speech are equivalent but not identical. This assumption must be true for both the whole internally uttered sentence and for its parts and elements. The context of innerly spoken sentences works in a similar way. The reason behind this generalization is simple: not only might whole sentences have synonyms, but also parts of a sentence, its context and generally every utterance. Therefore the following is correct:

$$C3. S_s(t, v, \emptyset, p) \rightarrow \{(\eta C \emptyset \vee \emptyset C \eta \vee \emptyset \in \eta \vee \eta \in \emptyset) \rightarrow \\ \rightarrow \sim [T(t, v, p) \underset{t_1 v}{=} I(t, v, |\eta|)]\}$$

So an inner utterance is not identical to the thought about a respective occurrence. These loose relations between thinking and inner speech have also another aspect. One thought may be expressed through different utterances. Moreover, interchanging between synonymous utterances should not in any way lead to a change of the original thought. This last conviction cannot be

expressed in our language thus I can only accept its consequences formulated in C5. Unfortunately C5, which will be analyzed hereafter, is far less obvious than the conviction written above.

The assumption C3 together with C1 lead to interesting problems which, however, will not be analyzed herein. But they do need to be at least mentioned. Since inner speech (due to the existence of synonyms) cannot be identified with thinking, it therefore only accompanies discursive (verbal) thinking, a question arises: what is discursive thinking and why is it always accompanied by inner speech? What part does this speech play in thinking? If I identified thinking with internalized nonverbal reactions to events (for instance the tensing of relevant muscles) then what is the part of inner speech in thus understood thinking? I shall not answer these questions herein.

Internalized speaking should be an equivalent of communicative speaking. I have given a few theorems which emphasize their similarity. But I have not expanded upon the subject. Now I shall complement the hypotheses B1 to B3 by attributing inner speech to communicative speech. I shall assume that while one talks out loud, listens or reads certain inner processes take place which I shall call inner speech. In the case of people who have the knowledge of the given language, inner speech governs communicative speech. The opposite happens when one receives a message. Additionally, when one uses signs (as a speaker or recipient) then one notices those signs. This is obvious in the case of reception. Less so in the case of speaking, but true nevertheless. We control the statements we utter and upon making a mistake we correct ourselves. All above observations allow for the following sentence:

$$\text{C4. } S_s(t, v, \emptyset, p) \rightarrow I(t, v, |\emptyset|) \wedge N(t, v, \emptyset)$$

Theorem C4 conclusively settles that the time necessary to utilize a sign is the same as the time necessary to innerly utter an equivalent of a relevant sentence. This equivalence of times may cause doubt to arise. I want to also add in reference to what I have already remarked upon in the context of C1, that due to A3 and B3 the time necessary to utilize a communicative sign and the time required for an inner utterance of that sign's equivalent is the shortest period possible which allows these two things to happen. For example: if a person uses 30 seconds to utter a statement, but from these 30 seconds 20 seconds are used for a pause (the person is silent) then I shall conclude that this person actually uses 10 seconds to utter the sentence. In a such perceived short time in which a sign is used, according to C4, I always refer to the denotement of a sign. This does not exclude the possibility that within the described pauses one once again speaks about the denotements

or thinks about something completely different.³

Seemingly C4 is an assumption that does not provide much new information about the connection between communicative speech and inner utterances (and consequently thinking). However, the truth is that many elements of C4 should be considered. One perceives a sentence on the level of the surface structure. It is not certain, however, that the same structure is innerly verbalized and not for instance the deep structure. To avoid discussing this difficult subject I shall assume that that sentences \emptyset are governed by the rules of logical syntax. I shall also assume, as Lakoff did, that deep structures are logical by nature. Therefore in the analysis herein I shall discuss a simplified variant, where deep structure is identical with the surface structure.

Another concern brought about by C4 is of the philosophical kind. C4 is not philosophically neutral, it shows preference towards the negation of nominalism. One of the variants of nominalism says that the only realities are things and reactions to these things. Mental reactions are limited to verbal reactions. C4_i and C₁ postulate that beside the fragments of reality, signs and verbal reaction there is also thinking about a given fragment of reality. I made the assumption that thinking and verbal reactions are not one and the same thing, therefore C4 does not work in favor of nominalism (in one of its variants).

One other noteworthy aspect is the fact that C4 requires a proper definition of the concept of the sign. According to C4 thinking p is a pre-requirement for \emptyset to become the sign of the occurrence p . In light of C4 as a child, who reacts to p by uttering \emptyset and who has yet to internalize speech, it does not utilize signs in the way understood herein. Because using signs in the way that is specified in this article requires verbal thinking which is equal to internalized speech. The latter is unavailable to a child. A child cannot utilize a sign in the sense described in C4. According to C4, internalized speech is required to fully master signs.

Coming back to the issue of time necessary to innerly utter sentence equivalents I must make one more assumption. The time necessary for communicative and inner speech has been assumed as equal, which is the biggest simplification introduced herein. This simplification may raise objections. Without much doubt, however, I can assume the following: provided that discursive thinking lasts shorter than uttering a communicative sentence

³More precisely this does not refer to redundant sentences. The analysis in this article concentrates on utterances made in a language with a low redundancy ratio (see information about deep and surface structures).

then similarly less time would also pass for an inner utterance of the same sentence:

$$C5. S_s(t, v, \emptyset, p) \wedge t_1 Pt \wedge T(t_1, v, p) \rightarrow I(t_1, v, |\emptyset|)$$

It is generally known that we cannot perform a certain action and at the same time refrain from performing it. It is also impossible for people to be excited about something and at the same time completely emotionally restrained in respect to one and the same thing. When the processes of excitement and restraint in relation to a performed activity are beginning to reach a similar level, then a person starts to experience that as a discomfort. People call it a problematic situation, a dissonance or a lack of balance. And they feel like Buridan's ass.

This antagonism between excitement and restraint is of crucial importance to this analysis. Signs evoke various reactions. Some of them are reactions to the signs themselves. Reactions of people who do not know the given language are limited. People who know the given language have those kinds of reactions and additional reactions to the phenomenon or thing to which the sign refers. Among these reactions to the denotation there is also the one of interest to me that is a discursive thought reaction which is closely connected with inner speech. The moment one reacts to a denotation with inner speech, according to A3 and B3, one is incapable of any other verbal reaction at the same time, i.e. is incapable of reacting to the sign itself. Therefore if one reacts both to the denotation and the sign, the reactions to the sign should be nonverbal which can be often observed. Moreover, if one reacts with discursive thinking to the denotation of the sign then verbal reactions to the sign itself are blocked. This blocking mechanism is more general than the previously described inability to produce two different inner utterances at the same time. The previous mechanisms allowed a situation where the sign in use was the denotation and that this sign referred itself.

In this borderline situation there would be no separate reaction for the denotation and for the sign. The denotation and the sign would merge together. But the most usual situation is when the denotation and the sign differ. In these usual situations a pattern emerges — the verbal reaction to the sign is being blocked. This common reflex in relation to signs and the antagonism between blocking and stimulation leads to the lack of verbal reaction to a sign (the reaction is to the denotation) while it is utilized, even if the sign itself is the denotation.

The conclusion is that one cannot at the same time react with inner speech (or thinking) to the sign or its denotation. Other reactions also lead to this conclusion. Every sign has its equivalent — ex. proper circumduction,

proper definiens — which has a different structure than the sign itself. Now I shall describe a particular situation: a person thinks simultaneously about the sign and the denotation of this sign — that means the sign and the denotation are identical, or the denotation is a part of the sign or its context. However, if a sign has a non-identical synonym (and it has already been assumed here that it must have one) than there should be a possibility to insert this synonym in place of the original sign. The thought connected with the inner utterance of a statement changed through replacement should be identical to the one connected with the original statement. According to the provisionally made assumption the original statement was the denotation and something was thought about it (ex. about its shape), so after the statement gets replaced the thought should stay the same (replacing of synonyms). This, however, is impossible because the change of the statement causes the change of the subject of thought. The change in the subject of thought changes the thought itself (see C2). So now I have reached a contradiction. Therefore the assumption is wrong: I cannot assume that one ever reacts verbally to a sign when this sign is being utilized.

In the herein used restricted language I cannot formulate the above described notions about the antagonism between blocking and stimulation, and about the inability to replace synonyms. So the results of these deliberations cannot serve me as assumptions for further analysis. In this case I shall use the conclusions from the previous paragraph as my assumption:

$$\text{C6. } T(t, v, p) \wedge I(t, v, |\emptyset|) \wedge (\eta \in \emptyset \vee \emptyset \in \eta \vee \eta C \emptyset \vee \emptyset C \eta) \rightarrow \\ \rightarrow \sim \{I(t, v, |\emptyset') \wedge [I(t, v, |\emptyset') \equiv T(t, v, z(\eta))]\}$$

According to C6 when one thinks about p and utter $|\emptyset|$ one does not determine anything about \emptyset , neither its parts nor its elements nor its context. One does not verbally react to \emptyset , or any of its parts, when one utilizes \emptyset . In short C6 is the conclusion of the deliberations written down above.

In the assumption C3 I have introduced the concepts of identical occurrences which were defined in D1. That definition states that two thoughts are identical when and only when thinking one of those thoughts must lead to the thinking of the other and vice versa. In D1 one I referenced a possibility of thinking about some other thought. I shall secure this possibility with a separate assumption which will confirm that thinking about thinking is accompanied by inner speech. This will be written down as follows:

$$\text{C7. } S_s(t, v, \emptyset, p) \rightarrow \bigvee_{t_1} [T(t_1, v, T(t, v, p) \wedge \bigvee_{\emptyset} I(t, v, |\emptyset'))]$$

Lastly I will point out a certain obvious thing. Noticing a sign does not always lead to inner speech. One may thoughtlessly look at a piece of writing or listen to someone speaking and not have any inner utterances or thoughts

on that matter:

$$C7'. \sim \bigwedge_t [N(t, v, \eta) \rightarrow I(t, v, |\eta|)]$$

I shall not use C7' in my further deliberations. Its sole purpose is to illustrate that in such a restricted language one is able to formulate further theorems unrelated to the already accepted theorems. They shall not be expanded upon because my analysis has a finely declared purpose.

The above defined assumptions provide the general characteristics of the use of signs (A1-A3), they show that inner speech is similar to the use of signs (B1-B3), and finally in C1 to C7' they assign inner speech to thinking and also establish the relationships between these actions and the utilization of the sentence sign and the noticing of a sign. The provided characteristics — in the light of the richness and vastness of the subject matter — are very limited and present a simplified version of only these aspects which are absolutely crucial in describing the relations between inner speech and thinking.

THE CONSEQUENCES OF THE HYPOTHESIS

At this point it is difficult to say how probable are the assumed theorems. I did try to choose them in such a way that they seem least suspicious. Such care does not guarantee that they are actually true, however, efforts to explain them usually do not go beyond such care. To establish their worth I need to derive at least a few of their consequences. Maybe some of them will allow me to connect the subject of this article with other issues, maybe the conclusions reached will be viable for an empirical verification.

C1 and C4 lead to:

$$(1) S_s(t, v, \emptyset, p) \rightarrow T(t, v, p)$$

If I use C1 and B3 I can derive:

$$(2) S_s(t, v, \emptyset, p) \rightarrow \{T(t, v, p) \rightarrow \sim[t_1 \neq t \wedge t_1 Pt \wedge I(t_1, v, |\emptyset|)]\}$$

The result of C5 and (2):

$$(3) S_s(t, v, \emptyset, p) \rightarrow \{T(t, v, p) \rightarrow \sim[t_1 \neq t \wedge t_1 Pt \wedge T(t_1, v, p)]\}$$

The theorems C4 and C2 lead to:

$$(4) S_s(t, v, \emptyset, p) \rightarrow \{T(t, v, p) \wedge T(t, v, q) \rightarrow \\ \rightarrow [T(t, v, p) \underset{t_1 v}{=} T(t, v, q)]\}$$

The consequence of C2 is:

$$(5) I_1(t, v, |\emptyset|) \rightarrow \{[T(t_1, v, p) \equiv T(t_1, v, q)] \rightarrow (p \equiv q)\}$$

C4 and (5) result in:

$$(6) S_s(t, v, \emptyset, p) \rightarrow \{[T(t_1, v, T(t, v, p)) \equiv T(t_1, v, T(t, v, q))] \rightarrow \\ \rightarrow [T(t, v, p) \equiv T(t, v, q)]\}$$

Based on (6), (1), C4, C1, C6 and D1 I can assume that:

$$(7) S_s(t, v, \emptyset, p) \wedge (\eta C \emptyset \vee \emptyset C \eta \vee \eta \in \emptyset \vee \emptyset \in \eta) \rightarrow \\ \rightarrow \sim [T(t, v, p) \underset{t_1 v}{=} T(t, v, z(\eta))]$$

Conclusions (1), (3), (4) and (7) are identical with the assumptions I made in another essay *Semantics and pragmatics* (see pages 36-37, theorems 7.1, 7.2, 7.3 and 7.5). Since these assumptions lead to the transparency of signs and the theorem of the psychological inability to reach semantic antonymies, the hypothesis presented herein results in various theorems which have not previously been analyzed together. This is a rather interesting result and it may lead to a joint analysis of other subjects which have only been considered separately so far. The above deliberations also help to better understand the rule of transparency.

COMMENTS ON THE GENERALIZATION OF ACHIEVED RESULTS

Throughout the article I have analyzed the internalized uttering of statements which were expressed through the variable \emptyset present in all theorems I have presented herein so far. It also happens that a person utters names or other lexical classes. It seems that the above results may be easily adjusted to describe the uttering of names. The reasoning behind this line of thought is simple: it is agreed that name — like sentences — form semantic relationships with fragments of reality. In case of other utterances, for instance logical connectives, the situation is not that obvious. For utterances other than statements and names it would be difficult to formulate an equivalent of the assumption C1, that is the equivalent of the condition which establishes the semantic connection between the utterance and object serving as the subject of thinking.

While establishing the connection between the inner uttering of a name and thinking I will not consider the specific characteristics which differentiate a name from a sentence. So the A1-A3, B1-B3, C1-C7 may be repeated without any significant changes, because they do not provide any characteristic of sentences. Unfortunately some changes must be made, I shall present them in example A1:

$$A1. \emptyset = \emptyset' \rightarrow [S_s(t, v, \emptyset, p) \equiv S_s(t, v, \emptyset', p)]$$

A1 is easy to grasp: the same sentence in a given time relates to the same occurrence.

Now let's try to do the same with a name:

$$(a) a = a' \rightarrow [S_n(t, v, a, y) \equiv S_n(t, v, a', y)]$$

In case of (a) a difficulty arises which I shall illustrate with the following paronomasia:

(b) A diamond is not always a diamond.

In (b) the word "diamond" is used in two different meanings, the first one as a gem, the other as a figure having four sides of equal length forming two acute angles and two obtuse angles. If a makes the assumption that the time of utilizing a name is the time of utilizing a sentence then a at the same time means y and y' and $y \neq y'$, where y is a gem and y' is a figure. In light of the assumption I have made about the time of utilization of names an attempt to use A1 for names turns out to be false. But is this assumption necessary? There is a lot in favor of this assumption. A name does not perform a fully independent function (apart from in rather unusual situations when it constitutes a sentence fragment). This lack of independence can be seen in the fact that when utilizing the name the speaker directs the recipients attention towards an object, the recipient is then forced to think something about the object. The recipient must at least think that this object exists, is placed somewhere or the speaker wants to draw attention to it. However, when one thinks about something then this thought is in a form of a sentence which might have only been uttered innerly. Therefore one should and probably must treat a name as a part of a sentence. This leads to the conclusion that the name means something as long as the sentence it belongs to is being utilized. So my assumption seems to be correct and it does not — so it seems — cause any major problems. Names described in $A1^n$ — the equivalent to A1 — must be additionally relativized by their place in the sentence. In such a case the equivalent of A1 does not raise doubt.

$$A1^n. a_{\emptyset i} = a'_{\emptyset i} \rightarrow [S_n(t, v, a_{\emptyset i}, y) \equiv S_n(t, v, a'_{\emptyset i}y)]$$

In $A1^n_n$ the lower index next to a and a' means that these both expressions are positioned on the i -th place in the sentence. According to $A1^n$ expressions positioned in the same place, in the same sentence and designate the same object.

A real problem arises when I try to formulate an equivalent for C1. If I were to leave the relativization introduced by $A1^n_n$ the equivalent of C1 would look like this:

$$C1^n_n. S_n(t, v, a_{\emptyset i}, y) \rightarrow [I(t_1 v, |a_{\emptyset i}|) \equiv \underset{s}{V}(t, v, s(y))]$$

According to $C1_n$ the time in which a means y is the same with the time of uttering $a_{\emptyset i}$ and the time of thinking about y .

In $C1_n$ the time necessary to think y which is the same as the one necessary to internally utter \emptyset would be identical with the time necessary to

innerly say $a_{\emptyset i}$. That would lead to a situation when the time necessary to innerly utter the whole sentence \emptyset would be the same as the time necessary to utter a part of \emptyset i.e. the name $a_{\emptyset i}$. This short reasoning shows that the time relations in C1 are far more complex than in C1ⁿ.

The difficulty highlighted above repeats in several other places therefore it is impossible to simply transfer onto names the rules I have set for sentences. This problem is too complicated to be dealt with herein and I shall leave it for some other article.

PROSODY AND TRANSPARENCY OF SIGNS

a. The rule of transparency is very often mentioned while deliberating signs. However, it is formulated in various ways. The core idea of the rule remains the same: when one uses a sign one thinks about its denotation; however when one thinks about the sign one does not consider the sign's denotation and the sign loses its denotative character. Herein I use the following way of formulating this rule:

$$(8) S_s(t, v, \emptyset, p) \rightarrow T(t, v, p)$$

$$(9) S_s(t, v, \emptyset, p) \wedge (\eta C \emptyset \vee \emptyset C \eta \vee \eta \in \emptyset \vee \emptyset \in \eta) \rightarrow \sim \sqrt_s [T(t, v, s(\eta))]$$

In light of the above presented analyses it seems appropriate to consider the rule of transparency and certain objections that are being held against it. Giving the rule more thought seems important, because transparency is not some marginal trait of signs. For instance, Ajdukiewicz uses this rule to differentiate the associative theory of sign from the intentional theory of sign. Morris uses it to define what a sign is. Husserl also values this rule.

b. A group of facts which seem to contradict the rule may be the situations described in stylistics, meaning the influence of prosodic elements on the meaning of an expression. It is generally accepted that raising one's voice may change the meaning of an expression. The way one interprets and perceives a poem is strongly influenced by homophony which is created through rhythm, momentum and onomatopoeias. These common truths may be used against the transparency of signs. Primarily I may discard all definitions of the rule which state that while using a sign one does not notice it at all, as if one did not see it (Schaff). Still there is the possible contradiction between the less radical versions of the rule and the above described prosodic phenomena. I shall present now the doubts raised against the rule of transparency which arise from the effects of the external form of a sign on the user of the sign. While utilizing a sign (as the speaker or

recipient) a person reacts to the denotation of the sign, but also to the length of the sign, its homophonic quality, the tone of voice and so on. Some of these reactions are internalized like microtensing of muscle etc. The said reactions may influence human behavior, for example how a person will solve a problem which has been signalized by the sign. Therefore, the way a statement is formulated, its prosodic qualities among others, may impact on the attitude and the inner and outer actions of the recipient. One does not react merely to the denotation but also to the sign itself, one type of the reaction to the sign itself are the changes in the trail of thought. In this situation is it not so that reactions to the sign itself (with the exception of the reactions to the sign which replace the reactions to the denotation of the sign) should be described as thinking about the sign? If this was the case then the transparency rule would be false: one would be thinking about the sign and its denotation at the same time. This may be the cause of why Jakobson lists the metalanguage function among the functions of a language. In light of the analysis herein the correctness of the transparency issue boils down to the question: can thoughts about the sign itself, which are not equivalents of thoughts about the denotation of a sign, be seen as thoughts according to the line of reasoning presented in this article? That is thoughts accompanied by internalized utterances about the sign. Only in such a case one can reject the rule of transparency. In short: the rule of transparency excludes the possibility of verbal thought about a sign when the said sign is being utilized.

c. The premise of the rule of transparency — subsequent elements of the hypothesis of the connection between inner speech and thinking — were chosen in such a way that they were not to be dismissed at the very start. This care for initial probability should not be the only reason behind accepting a hypothesis. If that was the case proving a hypothesis would be pure speculation. Therefore it would be useful if it was possible, while analyzing the hypotheses, to reach such conclusions which could be later proved or disproved through experiments.

It seems that just with a few additional assumptions the rule of transparency can be proved with an experiment. The rule states that while using a sign one thinks about its denotation. These thoughts are accompanied by inner speech. At the same time while utilizing the sign one does not think about the sign itself and does not innerly say anything about it. The way to prove this would be to create a situation where a person needs to simultaneously think about the sign and its denotation. However, it is hard to determine what people are actually thinking. I would have to assume that

it is possible to determine what people are thinking by assessing their visible reactions. Uttering of unknown sentences may serve as a way of measuring thoughts. Similarly performing actions according to verbal orders — a person must understand the order to perform the task so that a person must have thought about the content of the order. To disprove the rule of transparency in an experiment the test subject must be provoked into thinking two things at the same time. The problem is that the person in question is not capable of simultaneously saying two sentences which would be reactions to the two thoughts. That would be physically impossible (see A3). This person may only be asked to produce one sentence which would refer to one thought. But the test subject may have nonverbal reactions to a spoken stimulus. Appropriate reactions would be to prove that this person understood the order (verbal stimulus) i.e. prove a second thought. If the person would be able to simultaneously speak about the order and carry it out, it would disprove the rule of transparency. This is how I imagine an experiment could look like.

The test subject receives several orders to perform some nonverbal actions like pressing buttons. When the person performs these tasks t (the time of performing them correctly) is being recorded. Later the person receives other instructions: the person must describe in a formalized way the characteristics of pieces of text that he sees. The content of these texts are previously received instructions (about pushing the buttons). The person must perform the second set of orders as fast as possible — again the time t_1 would be measured. The third part of the experiment would be: the person has to verbally describe and at the same time perform the orders. The time t_2 of performing both tasks at once would also be measured.

There are two possible outcomes: 1) time t_2 would be the same length as times t and t_1 ; or 2) time t_2 would be longer than one or both t and t_1 . What would the first option mean? I think that it would only mean that in the shortest possible time of understanding an instruction (and thinking about what this instruction informs about) people can also think about the content of the instruction, of course, if they can describe it properly. In short if t_2 and t or t_1 are identical then it is possible to think about the sign and its denotation at the same time. Thus the rule of transparency would be disproved and with it the premise for the hypothesis about the connection between thinking and inner speech.

The other possible outcome (if t_2 is longer than t or t_1) is less obvious. On one hand this result would be in line with the following consequence of the rule of transparency:

$$S_s(t, v, \emptyset, p) \wedge T(t, v, p) \wedge T(t_1, v, z(\emptyset)) \rightarrow t \neq t_1,$$

because if t is not the same as t_1 and does not even partially coincide then t together with t_1 creates a period of time longer than t_2 . However this result would not be very valuable as proof for the hypothesis, because there is no way to determine (on the basis of the rule of transparency) how much longer t_2 should be from t or t_1 .

There are various ways of looking at the lengths of times t_2 . The most obvious way is to treat t and t_1 separately and consider t_2 to their sum. This is a very natural line of reasoning and it also provides a hint as to the length of the periods t_2 . According to A3 and B3 t and t_1 are very short periods of time, so short in fact that they are fully "occupied" by the action of formulating thoughts. No other mental verbal action may be fitted into periods of time devoted to formulating thoughts.

The other way to consider this issue is to assume that periods t_2 are shorter than the sum of t and t_1 , but are longer than t or t_1 themselves. I have already established that the time of uttering a sentence and the time of formulating an appropriate thought are identical when:

- a) the surface structure of a sentence and the deep structure are identical,
- b) the sentence is not redundant (at least not very redundant) (see the comments on C4).

Orders given to the test subject would be formulated in a natural language. The redundancy level in such languages is about 70%. The above presented rule of transparency does not apply to natural languages. Therefore I must make more assumptions, which will allow me to prove it with the help of natural languages. I shall assume that thinking is not redundant and it takes seven times less time to think something than to utter it. This would mean that $t_2 = t + (t_1 - 0.7t_1)$. I shall also point out that t_2 also encompasses the time of latency. Similar latency periods are hidden within t and t_1 . By adding t to $(t_1 - 0.7t_1)$ I also add two periods of latency (expressing the exposition to instructions) which is unnecessary. I need to subtract one of those periods. If t_{lat} is the time of latency then the whole time of performing both actions should not be shorter than:

$$t_2 = t + [(t_1 - t_{lat}) - 0.7(t_1 - t_{lat})]$$

The above presented assumption is rather imprecise. It allows me, however, to determine that the true time t_2 should fulfil the following condition:

$$\{t + [(t_1 - t_{lat}) - 0.7(t_1 - t_{lat})]\} \leq t_2 \leq (t + t_1)$$

The remarks I have made outline a path which could lead to proving the rule of transparency and its premises which form the hypothesis about the connection between a certain type of thinking and inner speech.

I believe that my approach to proving the rule of transparency goes beyond previously conducted research (see Maruszewski and Nowkowska, *Próba eksperymentalnego badania "przezroczystości słowa dla znaczenia"*, *Studia Psycholog.* 1970, 10). On the other hand, it is not free from omissions and simplifications. These are connected to the redundancy of language and its supposed consequence — the shorter period of verbal thinking. This requires further research.