1. INTRODUCTION

In Romanian there is more than one way to express the middle voice construction and this can be considered a very significant language internal variation that needs to be studied. The claim that the two middle voice constructions cover the same semantic spectrum, as they were taken to be by previous linguists and traditional grammarians, is not acceptable (Avram Mioara 1997, Daniliuc Laura & Radu 2000, Calude 2003). This paper offers an analysis of the two middle constructions *se* and *isi* using data from the narratives and newspaper genre. Although the existence of two Romanian middle markers has been attested (Kemmer 1993) little has been written about it.

This paper is an attempt to analyze the middle voice constructions *se* and *isi* in Romanian by investigating their usage and semantic characteristics using corpus data. All the papers written till now have one main flaw; the papers do not use any statistically representative data base. The present paper takes an empirical data based approach to language and so to the middle voice phenomena finding main semantic and cognitive differences between the two middle constructions. Secondly, it discusses major usages that were overlooked by the previous non-empirical and based on introspection approaches to the notion of middle voice in Romanian. In order to check the validity of intuition based analysis such as the ones done by previous grammarians this paper will also include a section of intuition based tests given to native speakers of Romanian. The purpose is to check whether the speakers’ intuitions are similar to the results found in the corpus.

1.1. Aim and Material

The following are the objectives of this paper:

1. Define the lexical associations (collocations): what are the tendencies for the target grammatical construction *se* and *isi* to co-occur with particular words? In other words, what are the semantic domains of the verbs with which these two constructions occur? I will look at two different genres, fairy tales and newspaper genre the politics section. At first a micro-analysis will be provided. By micro-analysis I mean that genre internal examination of the two middle
constructions is going to be done. Secondly, a macro-analysis meaning differences between genre are going to be attested for. When necessary reference to word order or animacy of the subject etc will be introduced.

2. Grammatical associations: identify contextual factors associated with the structural variants. I.e. what are the main semantic and syntactic differences between the middle constructions se and isi?

3. Intuition data: what does it show us? Is it more similar to the findings in fairy tales or in newspapers? Are the intuition judgments a better example of the prototypical usages and the corpus an elaboration?

**The language material:**

**Self-built written corpora:**
- a. Newspaper genre (Romania Libera) politics: 138,745 words, 18,925 types
- b. Fairy tales: 96,393 words, 12,361 types

**On the spot Intuitive judgment test** of speakers: 24 native Romanian speakers were asked to form sentences that immediately jump into their minds using the se and isi constructions. (2 sentences per construction).

The computer program used for this paper is MonoConc (Michael Barlow 2002).

1.2. The adopted framework

**Middle voice:**
Lyons' definition of middle voice “the middle voice indicates that the action or state affects the subject of the verb or his interest” sounds intuitively correct. I am going to adopt Kemmer's definition of the middle voice (1993). Middle voice refers to events in which there is a low level of distinguishability among participants. This lack of differentiation leads to lesser degree of elaboration of the event. Impersonal, spontaneous, passive-like constructions are encompassed in the middle voice field.

All these constructions share the property of not having defined distinguishable participants and of designating fully elaborated events.¹

The following sentences exemplify the se and isi constructions in Romanian.

(1) A doua zi [[se]] tot gandea Aleodor unde sa se ascunda
   the second day MM still thought Aleodor where to hide
   The next day Aleodor continuously thought about where to hide

(2) El [[isi]] aminteste de iubirea lui.
   He MM remembers of the love his.
   He remembers his love (he recollects)

2. THE FREQUENCY DISTRIBUTION OF SE AND ISI ACROSS GENRES

Before discussing the middle constructions of Romanian and the semantic domains they cover, it is important to examine the distribution of se and isi across genres as represented by the two corpora.
Two general tendencies are evident in the data. *Se*-construction and *isi*-construction are more frequently used in the fairy tales genre than in the newspaper genre. Secondly, the *se*-construction is more frequently used than the *isi*-construction.

Table one illustrates the raw frequency and the occurrence of *se* and *isi* once every how many words. In the newspaper genre the *middle* construction *se* occurs once every 126 words; whereas the *isi* middle construction occurs only once every 1101 words. In the fairy tale corpus the middle construction *se* appears once every 60 words; whereas the *isi*-construction occurs once every 446 words.

<table>
<thead>
<tr>
<th>Corpora</th>
<th>Number of words</th>
<th>Raw</th>
<th>Frequency</th>
<th>Frequency per word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Romania Libera Corpus</td>
<td>138745</td>
<td>1094</td>
<td>126</td>
<td>126.8236</td>
</tr>
<tr>
<td>Fairy Tales</td>
<td>96,393</td>
<td>1588</td>
<td>216</td>
<td>60.70088</td>
</tr>
</tbody>
</table>

How can one explain these differences in the distribution of *se* and *isi* across genres?

This results point to one major tendency and this is that the frequency of *se* and *isi* is correlated to the functions covered by these two constructions. *Se*-middle and *isi*-middle although both part of the middle voice semantic field have very different functions. *Se*-middle functions as the reference point with respect to which the event is calculated: *se* profiles the event. Moreover, it focuses the attention on the change of state the subject undergoes. Whereas, *isi*-middle marker signals that the object (landmark) is understood as being located in the dominion of the subject (trajector). I.e. the subject becomes the reference point. The relation between the landmark and the trajector is one of abstract possession. Hence, *isi* delimits the dominion where the object is related to the subject’s dominion, his sphere of control. Another way to express possession in Romanian is to use the internal possessive construction marking the possessum with genitive case. Hence, in some cases the speaker can choose between these two constructions leaving more room for variation. This is why *isi*-construction appears less frequently than *se*-construction.

3. PREVIOUS ANALYSIS OF THE MIDDLE VOICE IN ROMANIAN AND ITS LIMITATION

**Mioara Avram (1997)**

The following are the functions of *se* and *isi* as interpreted by Avram (1997)

**Objective-reflexive**: (meaning that the subject is the doer of the action and also the object on which the action is done on)
- El *se* (ACC) imbraca/He dresses
- Eu *imi* (DAT) amintesc/ I remember.

**Reciprocal-reflexive**: 
- Noi *ne* (ACC) certam des/ We fight often
- Ei *isi* (DAT) dau palme/They slap each other.

**Dynamic reflexive**: (the pronoun marks the intense participation of the subject)
El se (ACC) ruga de iertare/ He was begging to be forgiven.
El isi (DAT) inchipuie ca ne-a convins/ He imagined himself that he convinced us.

**Passive – Reflexive:** (the subject which is identical with the reflexive pronoun undergoes the action which is done by someone else)
Biletele se (ACC) vand la casa/The tickets are sold at the ticket bar.
Eu m-am(DAT) nascut iarna/ I was born in winter.

**Impersonal – reflexive** (without any passive sense)
Se doarme bine in pat? / Does one sleep well in the bed/
La ei se mananca mult/ At their place one eats well.

**Eventive reflexive:** (expresses the change into another state)
El s-a imbolnavit/He got sick.

Previous analysis did not take into consideration real data. Moreover, they have looked at isi and se in one package. In other words, as one can see they have divided the functions of se and isi and showed examples of both isi and se in with the same semantic interpretation. Hence, they did not consider the two constructions to be different except of their case marking, one is Accusative and one is Dative. But no, semantic difference between them was mentioned. Hence, isi and se were considered to have a similar semantic distribution. This paper will show that this is not at all the case. Not only does the text distribution of se and isi differ but the main difference is in the semantics of the collocations with which they appear. Moreover, as one can see se and isi were always considered to be reflexive in meaning. However, I will show that there are main differences between se and isi middles. The semantic domain that they cover although overlapping in some cases, is different. In other words, the distributions of the most frequent verbs they occur with are different. Secondly, isi is not yet so grammaticalized in its nature because it cannot yet appear with spontaneous, impersonal and passive like constructions. Moreover, the most important difference between se and isi is their semantic function in the sentence. Se directs access to the event; whereas isi marker pulls the object into the subject’s dominion, its sphere of control. These type of constructions where also named by some linguists external possession. Another extension of isi-middle construction which was not discussed in any descriptive grammar is to a causative construction. All these differences were ignored by previous analyses.

4. THE MIDDLE VOICE AS REPRESENTED IN FAIRY TALES

4.1. Se-middle construction
As shown in the following graph the se-construction is frequently used with spontaneous event verbs and movement verbs including translational motion verbs, non-translational motion verbs and change in body posture. What is interesting to see is that although in all the grammars the se-construction is the one which exemplifies the true reflexive with grooming verbs such as ‘wash oneself’, ‘dress oneself’ in a fairy tale corpus where one would suspect this kind of grooming verbs and body action verbs to be very persistent se almost never occurs with them. In the very infrequent cases where
se did occur with a grooming/body care verb regional language was used. By regional I mean dialects of the country side such as ‘se gâteste’ – cook oneself which means make oneself beautiful by dressing nicely. The fact that se-construction do not appear as frequently with grooming and body action verbs in this corpus is not due to the low frequency of these kind of verbs in the corpus. A rough check has been done to verify that this will not skew the data. I will exemplify each and one of the semantic verbal groupings.

Diagram1: Se fairy tales distribution of verbs

**Spontaneous Events:**
(3) a. Dar cand ninsoarea [[se]] inteti si se transforma in viscol
  but when snow MM got stronger and MM transform in storm
b. Cand [[se]] trezi fata isi lua ocheanul fermecat
  When MM wake up girl....
  When the girl woke up

**Speech action:**
(4) Apoi i [[se]] adresa lui Aladin astfel: -Eu sunt
  Afterwards 3SG MM address to Aladin in this way: - I am
  Afterwards he addressed Aladin in this way:-I am
(5) Maiestatea voastra...N-am stiu…”, [[se]] scuza piticul.
  Majesty your no-have know, MM excuse the dwarf
  Your Majesty, I did not know the dwarf excused himself

**Cognitive middles:**
(6) A doua zi [[se]] tot gandea Aleodor unde sa se ascunda
  the second day MM still thought Aleodor where to hide
  The next day Aleodor continuously thought about where to hide
**Emotional middle:**
(7) si [[se]] temea ca ii va ajunge capul in par.
   And MM scared that ............
   He was scared that............

**Perception:**
(8) Fata imparatului [[se]] tot uita cu drag la el si l-a sarutat.
   Daughter the emperor MM still looked with care/love at him and kissed him
   The daughter’s emperor was looking at him with love and kissed him

**Existential:**
(9) iar cand se trezi, vazu ca [[se]] afla pe o pajiste foarte frumoasa,
   and when MM woke up, saw that MM exists on a meadow very beautiful
   and when she woke up she saw that she is on a very beautiful meadow

**Translational Motion:**
(10) Bunicuta Holle ii porunci sa [[se]] duca sub streasina casei
   Grandmother Holle ordered her to MM go under roof house-GEN
   Grandmother Holley ordered her to go under the roof of the house

**Change in body posture:**
(11) In vremea asta, Gringoire [[se]] ridica anevoie de jos.
   in time this, Gringore MM stood-up difficult from down
   In this time, Gringore hardly raised from down

**Grooming verbs:**
(12) degetel isi destepta fratii, le spuse sa [[se]] imbrace....
   Little finger woke up his brothers and told them to MM dress.
   Little finger woke up his brothers and told them to dress

(13) mai apoi sa [[se]] gateasca si ea, sa fie cat mai frumoasa ...
   and afterwards MM cook and her, to be as beautiful as possible
   She dressed up in order to be as beautiful as possible

**Activity verbs:**
(14) Nimeni nu [[se]] juca cu el.
   No one not MM played with him.
   None played with him

If there is a second participant in the argument structure of a sentence with se usually it is not obligatory or it forms a frequent collocate with the verb as shown in sentences (15) and (16) below.

**Collocates:**
(15) Ei [[se]] sorbeau din ochii
    They MM sipped from eyes
    They looked at each other intensively (meaning they liked each other a lot)
The looked at each other strongly. (Usually implying that they like each other a great deal)

(16) Ei [[se]] indemnau la treaba
They MM stimulate to work
They stimulated each other to work

Usually, the argument structure of *se* is a one participant one. The Subject of the sentence can be animate or inanimate. In the case of an inanimate subject the most common use is to express spontaneous events. An animate subject such as human most frequently appears in translational motion events, emotional, cognitive middles. However, in a fairy tale genre animate and abstract notions can also appear as subjects of emotional, translational motions and even change in body posture. In most of the cases the word order is not the canonical SVO but VOS.

In conclusion, all the above mentioned sentences show that *se*-middle functions the reference point to the event. It focuses on the change in state the subject undergoes.

4.2. *Isi* – construction
The *isi*-middle construction occurs most frequently with speech verbs, cognitive verbs, and intentional perception verbs. Although in grammars examples of grooming verbs are always given using the *se*-middles, when looking into a corpus, body action and grooming verbs are much more frequent in the case of the *isi* middle than in the expected *se*-middle. Verbs related to movement of the body, be it motion with reference to a path or non-translational verbs are not as common with *isi*-middle as they are with *se*-middle.

Diagram 2: *Isi* fairy tales distribution of verbs
As one can see what the following examples have in common is that the isi functions as a dominion indicator. It pulls the object into the subject’s dominion, into its sphere of control. Some examples of the Data and relevant comments:

**Cognitive middles:**

(17) El [[isi]] aduse aminte de stiuca, pe care o chema ...
    He MM brought mind of the pike…..
    He remembered the pike

**Speech verbs:**

(18) Frollo [[isi]] zise ca avea sa-i iasa din minti de iu ...
    Frollo MM says that had to go out of his mind
    Frollo said to himself that he has to get out of his mind

In the case of isi although ‘ducе’ - go is a common verb it never has a translational meaning – the meaning that it has is much more metaphoric – it usually means to carry – to carry his own destiny or to carry his life in the middle of these people – in other words to live among them.

(19) Erau saraci lipiti pamantului si abia [[isi]] mai duceau traful
    Were poor glued ground and hardly MM go the living
    They were really poor and hardly carrying their living

When we have non-translational verbs then it is a very figurative usage such as:

(20) [[isi]] lasa in jos crengile pline de fructe
    MM let in down the brenches full with fruits
    the tree let his branches down to give the people its fruits (the tree bended)

A lot of verbs with isi have as their helping verb to give – they gave themselves a meeting meaning they – to make an appointment with someone

(21) Apoi [[isi]] dadu sufletul in bratele sotului sau. (spontaneous event)
    And then MM gave soul in the arms the husband her.
    She gave her soul in the hands of her husband – meaning to succumb

(22) dar apoi [[isi]] dadu seama ca totul era in zadar (cognitive middle)
    but then MM give justification that everything is in vain
    but then he understood/recognized that everything was in vain

4.2.1. Properties of isi-construction

The object (possessum) can be body parts, inalienable things such as kinship relations, alienable things such a house, a flute and also abstract nouns. Hence, the relation is one of abstract possession. By using isi-marker objects that are not inherently part of the subject's dominion are brought into its dominion.
Inherent possessed object:
(23) El [[isi]] freca *(mainile) cu disperare.
He MM rub *(hands) with desperation
He rubbed his hands with desperation.

(24) [[isi]] scarpina capul Bolek
MM scratched the head Bolek
Bolek scratched his head

Kinship expressing object:
(25) Frumoasa, [[isi]] ajuta familia dupa puterile ei.
The beauty MM helps the family after power her
The beautiful girl helps her family as much as she can

Non-inherently object pulled into the dominion of the subject by isi-marker:
(26) El [[isi]] pierdu toata ziua.
He MM lost the whole day
He lost his whole day

4.2.2. The nature of the possessor
Grammatically it always is the subject of the sentence. Semantically, it usually has to be an animate argument. However, in figurative language an inanimate possessor is acceptable.

(27) a. Batranetea [[isi]] spune cuvantul
Oldness MM says the word
Oldness says its word. (used when someone feels that he got old)
b. Viata nu [[isi]] are rostul
Life no MM has reason.
Life has no reason.

4.2.3. Extension to the causative domain
Causative construction using the middle morphology isi:

(28) El [[isi]] supara nevasta.
He MM makes angry wife.
He makes his wife angry.

(29) El [[isi]] promoveaza familia (pe diverse functii)
He MM promotes family (in different functions)
He makes his family promote (to different functions)

Hence, the lexical associations of these two middle constructions vary a great deal. The verbal classes with which these constructions appear which were always dealt with in the same bundle can highlight the semantic differences between se-middle and isi-middle. In isi-middle the subject of the sentence has more control over the event
because in most of the cases except figurative expressions the subject is animate or human and this is why activity verbs and speech verbs are more frequent but not at all spontaneous events; whereas in the case of *se* the subject of the sentence can be both animate and inanimate so a lot of spontaneous events both with animate and inanimate occur.

### 4.3. COLLOCATION RESULTS OF SE AND ISI CONSTRUCTIONS

One other interesting thing to check is what are the most frequent collocates with these two constructions. This would be interesting to compare with the results obtained in the intuitive test. This kind of comparison might help us to gain a better understanding of what happens in the minds of the speaker- cognitive entity that is responsible for the data base i.e. the mind that produces and interprets the text. I will come back to this point trying to provide some hypothesis.

The most frequently used collocation with *se*-construction is with the translational motion verb ‘duce – go’ in its different conjugation (duca, duce,duc) etc.

**Collocation to the right of the word *se*:**

<table>
<thead>
<tr>
<th>Freq</th>
<th>T-Score</th>
<th>Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>3.029123</td>
<td><em>se</em> duse go</td>
</tr>
<tr>
<td>29</td>
<td>3.004913</td>
<td><em>se</em> uita look</td>
</tr>
<tr>
<td>15</td>
<td>2.666022</td>
<td><em>se</em> facu do</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td><em>se</em> duca will go</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td><em>se</em> afla exists</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td><em>se</em> intoarse come back</td>
</tr>
</tbody>
</table>

**Collocations to the right of the word *isi*:**

<table>
<thead>
<tr>
<th>Freq</th>
<th>T-Score</th>
<th>Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>3.029123</td>
<td><em>isi</em> lua take</td>
</tr>
<tr>
<td>10</td>
<td>3.004913</td>
<td><em>isi</em> facu make</td>
</tr>
<tr>
<td>10</td>
<td>2.666022</td>
<td><em>isi</em> dadu give</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td><em>isi</em> vazu see</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td><em>isi</em> spuse say</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td><em>isi</em> zise say</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td><em>isi</em> intreba ask</td>
</tr>
</tbody>
</table>

Using the collocation program the following are the results:

**Score method: T-Score**

<table>
<thead>
<tr>
<th>Freq</th>
<th>T-Score</th>
<th>Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3.029123</td>
<td><em>se</em> duse go</td>
</tr>
<tr>
<td>10</td>
<td>3.004913</td>
<td><em>se</em> uita look</td>
</tr>
<tr>
<td>8</td>
<td>2.666022</td>
<td><em>se</em> facu do</td>
</tr>
</tbody>
</table>
ANNE-MARIE HARTENSTEIN

Score method: T-Score
Freq T-Score Collocation
________________________
3 1.676591 isi facu do
3 1.676591 isi pierdu loose
3 1.676591 isi dadu give
3 1.658104 isi lua take
3 1.639618 isi vazu see
2 1.368931 isi veni come
2 1.368931 isi ducea go

Let’s keep these results in mind when analyzing the results of the intuition tests.

5. THE MIDDLE VOICE AS REPRESENTED IN THE NEWSPAPER GENRE
   (POLITICS SECTION)

5.1. Se-construction
Diagram 3 illustrates that the most common usage of the *se*-middle construction in newspapers is in impersonal constructions. The next semantic domain is the emotional verbs but they are so much less frequent than the impersonal constructions. Hence, in the case of the newspaper genre there is less variety in the usage of *se*-construction.

(30) Asa [[se]] si explica de ce, fie si intr-un regim ...
   This MM and explanation why…..
   This is how it can be explained why…..

(31) La Sala "Horia Lovinescu" [[se]] vor juca: "Miresele capitanului sau
   in hall "Horia Lovinescu" MM will play "Miresele capitanului
   In the hall "Horia Lovinescu" the play named "Miresele capitanului will be played.

Emotional middle (desiderative):
(32) Multi pensionari [[se]] plang ca se trezesc cu sume imputate pe ...
   Lots of pensioners MM cry that MM wake up with sums
   Lots of pensioners complain that they find themselves with ridiculous sums.
As one can see, *se* became so grammaticalized in its nature that it can be also used in impersonal constructions, *se*-passive constructions, spontaneous events etc. Here, again one can see that the function of *se*-marker is to focus the attention of the conceptualizer towards the event and the change of state the subject undergoes, a very different function than *isi*-marker.

5.1.2. Isi construction

Diagram 4 shows that there is a main difference between *se* and *isi* also in the newspaper genre. Here again *isi* appears with mental and cognitive verbs and the majority with activity verbs. Moreover, the main difference between *se* and *isi* middles
in this genre is the type of the subject. In the case of *isi*-middle the subject is animate or human; whereas in the case of *se*-middle construction abstract subjects or inanimate subjects are the most frequent ones. Hence, even if the kind of newspaper genre asks for more impersonal construction one can still see that it is *se* that is chosen to express them and not *isi*.

**Collocations with *se***:
It can be seen from the list of collocations that impersonal constructions are the most frequent ones. ‘li se’ – to them se- in this case one could continue a sentence such as To them it is told that ……..

<table>
<thead>
<tr>
<th>Freq</th>
<th>T-Score</th>
<th>Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>8.666824</td>
<td><em>se va</em> it will</td>
</tr>
<tr>
<td>75</td>
<td>8.333703</td>
<td><em>se afla</em> it exists</td>
</tr>
<tr>
<td>36</td>
<td>5.671161</td>
<td><em>se vor</em> they will</td>
</tr>
<tr>
<td>33</td>
<td>5.509864</td>
<td><em>se poate</em> it can be</td>
</tr>
<tr>
<td>27</td>
<td>4.860742</td>
<td><em>se face</em> is made</td>
</tr>
<tr>
<td>16</td>
<td>3.860244</td>
<td><em>se pare</em> it seems</td>
</tr>
<tr>
<td>9</td>
<td>2.901348</td>
<td><em>se numara</em> it is counted</td>
</tr>
<tr>
<td>9</td>
<td>2.879426</td>
<td><em>se fac</em> made</td>
</tr>
<tr>
<td>9</td>
<td>2.846542</td>
<td><em>se stie</em> it is known</td>
</tr>
<tr>
<td>9</td>
<td>2.769813</td>
<td><em>se putea</em> it is possible</td>
</tr>
</tbody>
</table>

**Collocations with *isi***:

<table>
<thead>
<tr>
<th>Freq</th>
<th>T-Score</th>
<th>Collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>3.953372</td>
<td><em>isi va</em> he will</td>
</tr>
<tr>
<td>12</td>
<td>3.337250</td>
<td><em>isi vor</em> they will</td>
</tr>
<tr>
<td>4</td>
<td>1.951175</td>
<td><em>isi propune</em> he proposes</td>
</tr>
<tr>
<td>4</td>
<td>1.918624</td>
<td><em>isi face</em> to make oneself</td>
</tr>
</tbody>
</table>

6. A MACRO-ANALYSIS OF *SE* AND *ISI* (INTRA-GENRE)

The main differences in the usages of the middle construction in the two genres of fairy tales and Newspaper (politic section):

1. In the newspaper genre impersonal middles with the *se*-construction are extremely frequent. This is not the case in fairy tales.

2. In fairy tales *se*-construction were used in spontaneous events, motion verbs, and cognitive verbs etc.
3. In the newspaper genre the *isi*-construction was mostly used with activity verbs mental and cognitive verbs; whereas in fairy tales the variety of the semantic mappings is much larger. There are some similarities in both genres no emotional middles and spontaneous middles are expressed with *isi*. However, in the case of the newspaper genre body actions, grooming verbs non-translational verbs are not available; whereas in fairy tales genre this is not the case.

4. Although there are main differences in the semantics of the verbs in the two genres, one tendency remained unchanged and this is that *se*-middle constructions occur more often than *isi*-construction in both genres. This was discussed in section 2.0.

In conclusion, although there are major differences between the two genres, one can learn a lot about the function of *se* and *isi* markers by approaching the phenomenon of middle voice from a usage-based model to language. Namely, although both constructions were considered to have the same functions called reflexive pronouns by previous grammarians, we found that there is a difference in their function in the sentence. *Se*-marker directs the access to the event; whereas *isi*-marker pulls the object into the dominion of the subject.

7. THE INTUITION JUDGMENT TEST AND ITS COMPARISON TO THE CORPUS RESULTS

Twenty-four native speakers of Romanian were asked to build two sentences each using *se* and *isi*. The subjects were of different genders and ages (age 30 and above). This aspect was not taken into account because we do not believe that this is an important factor for our present study. The following sections illustrate the results of this intuitive test.

7.1. *Se* – middle construction

<table>
<thead>
<tr>
<th>Verb</th>
<th>translation</th>
<th>matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duce</td>
<td>Go</td>
<td>13</td>
</tr>
<tr>
<td>Plimba</td>
<td>Walk</td>
<td>2</td>
</tr>
<tr>
<td>Culce</td>
<td>go to sleep</td>
<td>3</td>
</tr>
<tr>
<td>Mearga</td>
<td>walk, go</td>
<td>1</td>
</tr>
<tr>
<td>odihneste</td>
<td>Rest</td>
<td>1</td>
</tr>
<tr>
<td>Scoala</td>
<td>wake up, rise</td>
<td>1</td>
</tr>
<tr>
<td>Invarte</td>
<td>turn around</td>
<td>1</td>
</tr>
<tr>
<td>Intoarce</td>
<td>come back</td>
<td>1</td>
</tr>
<tr>
<td>straduieste</td>
<td>trying hard</td>
<td>2</td>
</tr>
<tr>
<td>pregateste</td>
<td>Preparing</td>
<td>2</td>
</tr>
<tr>
<td>poate vorbi</td>
<td>can talk</td>
<td>2</td>
</tr>
<tr>
<td>Face</td>
<td>it gets (make)</td>
<td>1</td>
</tr>
<tr>
<td>razgandesc</td>
<td>change the mind</td>
<td>1</td>
</tr>
<tr>
<td>Joaca</td>
<td>play</td>
<td>1</td>
</tr>
</tbody>
</table>
Speakers were asked to intuitively construct sentences on the spot with *se*-middle construction and *isi*-middle construction, most of them immediately came up with a collocation of ‘*se+ motion verb*’. Most of the times it was a translational motion verb but non-translational verb were also present. The semantic choice of the verbs is much more similar to the once found in the fairy tale genre than the one in the politics. However, interesting enough impersonal middles with *se*, as found in the newspaper genre, were also frequently present in the speakers’ intuition. The most striking finding was the fact that 13 people out of 24 have constructed sentences with *se* using exactly the same collocation: se duce – se+go. The speakers were not tested at the same time or in the same room. They were also not from the same family. Hence, these findings cannot be a coincidence. They must show that speakers have this expression entrenched in their minds. Hence, basically the subjects prefer more frequent verb frames with *se* as opposed to more rare ones in the newspaper genre. I.e. the fairy tales verb frames are

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meanings</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crede</td>
<td>think himself</td>
<td>1</td>
</tr>
<tr>
<td>imprietesc</td>
<td>become friends</td>
<td>1</td>
</tr>
<tr>
<td>Uita</td>
<td>look at</td>
<td>1</td>
</tr>
<tr>
<td>intampla</td>
<td>Happen</td>
<td>1</td>
</tr>
<tr>
<td>Lupta</td>
<td>fight (a war)</td>
<td>1</td>
</tr>
<tr>
<td>mandresc</td>
<td>Proud</td>
<td>1</td>
</tr>
<tr>
<td>Ascund</td>
<td>Hide</td>
<td>1</td>
</tr>
<tr>
<td>foloseste</td>
<td>it is used</td>
<td>1</td>
</tr>
<tr>
<td>Pare</td>
<td>it seems</td>
<td>1</td>
</tr>
<tr>
<td>Cere</td>
<td>it is needed</td>
<td>1</td>
</tr>
<tr>
<td>Curete</td>
<td>clean up</td>
<td>1</td>
</tr>
<tr>
<td>cantareste</td>
<td>to weight</td>
<td>1</td>
</tr>
<tr>
<td>Numara</td>
<td>Count</td>
<td>1</td>
</tr>
<tr>
<td>Spala</td>
<td>wash oneself</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>
more frequent. That's what the subjects prefer as well. So, we have the expected frequency effect. Thus, people choose frequent structures which they hear more and then incorporate them into their permanent memory.

(33) [[Se]] foloseste aspiratorul pentru curatenie.
   MM using vacuum cleaner for cleaning
   The vacuum cleaner is used for cleaning

Thus, the intuition of the speakers fits much more the data based results than the direct reflexive examples, that where always presented as the most prototypical ones, in traditional grammars.

7.2. Isi middle construction
The following table illustrates the results of the experiment for the isi-construction.

<table>
<thead>
<tr>
<th>Verbs</th>
<th>Translation</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plimba</td>
<td>Walk</td>
<td>3</td>
</tr>
<tr>
<td>pregateste</td>
<td>Prepare</td>
<td>3</td>
</tr>
<tr>
<td>Coase</td>
<td>Sew</td>
<td>2</td>
</tr>
<tr>
<td>Doreste</td>
<td>Desire</td>
<td>1</td>
</tr>
<tr>
<td>ia ramas</td>
<td>say good bye</td>
<td>1</td>
</tr>
<tr>
<td>Ara</td>
<td>Plough</td>
<td>1</td>
</tr>
<tr>
<td>mangaie</td>
<td>Pat</td>
<td>1</td>
</tr>
<tr>
<td>Linge</td>
<td>Lick</td>
<td>1</td>
</tr>
<tr>
<td>Spala</td>
<td>Wash</td>
<td>2</td>
</tr>
<tr>
<td>inchipuie</td>
<td>imagine</td>
<td>1</td>
</tr>
<tr>
<td>imagineaza</td>
<td>imagine</td>
<td>1</td>
</tr>
<tr>
<td>provoaca</td>
<td>Provoke</td>
<td>1</td>
</tr>
<tr>
<td>cumpara</td>
<td>Buy</td>
<td>3</td>
</tr>
<tr>
<td>Pune</td>
<td>Put</td>
<td>1</td>
</tr>
<tr>
<td>Cauta</td>
<td>look for</td>
<td>1</td>
</tr>
<tr>
<td>face/reface</td>
<td>make/remake</td>
<td>7</td>
</tr>
<tr>
<td>Pazesc</td>
<td>Protect</td>
<td>1</td>
</tr>
<tr>
<td>asuma raspundere</td>
<td>take responsibility</td>
<td>1</td>
</tr>
<tr>
<td>schimba</td>
<td>Change</td>
<td>2</td>
</tr>
<tr>
<td>Imparte</td>
<td>share/divide</td>
<td>1</td>
</tr>
<tr>
<td>Duce</td>
<td>Go</td>
<td>2</td>
</tr>
<tr>
<td>Spune</td>
<td>Say</td>
<td>2</td>
</tr>
<tr>
<td>Propun</td>
<td>Propose</td>
<td>1</td>
</tr>
<tr>
<td>Termina</td>
<td>Finish</td>
<td>1</td>
</tr>
<tr>
<td>Lauda</td>
<td>Boast</td>
<td>1</td>
</tr>
<tr>
<td>scarpina</td>
<td>Scratch</td>
<td>1</td>
</tr>
<tr>
<td>Poarta</td>
<td>Wear</td>
<td>1</td>
</tr>
<tr>
<td>Creaza</td>
<td>Create</td>
<td>1</td>
</tr>
<tr>
<td>asteapta</td>
<td>Wait</td>
<td>1</td>
</tr>
</tbody>
</table>
Ia  |  take breakfast  |  1  
---|------------------|-----
**Total**  |  |  **47**

Diagram 6: *isi* middle distribution of verbs (intuition results)

Not as in the case of *se*-construction, the sentences with *isi*-middle construction provided by the speakers were part of a variety of semantic fields. There was no strong collocation like in the case of *se*. The strongest collocation was with the verb to do ‘*isi* + face’ (face – do/make) which is a very vague verb. By vague I mean that it can participate in different semantic environments as shown in the following examples:

(34) Fiecare om *isi* face socoteli.
    Every person makes calculations.
    Every person makes his own calculations

(35) Ea *isi* face lectile.
    She makes homework(PL)
    She makes her homework

(36) *Isi* face griji degeaba.
    MM makes problem in vain
    He worries in vain.

Here again as in the case of the fairy tales if translational motion verbs are used the example with ‘duce – go’ is metaphorical.

(37) Fiecare om *isi* duce povoara lui.
    Every man carries heaviness his.
    Every person carries his own problems and life.
In conclusion, after looking at the abovementioned results of the intuition test and comparing them to the two types of corpora, one cannot remain indifferent to the cognitive entity that is responsible for the data base i.e. the mind that produces and interprets the text. Hence, one needs to provide more elaborate explanation as to why our intuition is closer to fairy tales although, we are most of our lives more exposed to newspaper style of genre and not fairy tales. Moreover, it would be interesting to take a better look at the individual differences of the talkers. I could notice that the sentences each individual constructed were part of his own experience. Hence, different activations win even in the case of intuitive introspective judgments of the speakers.

8. CONCLUSION

In conclusion, this paper has shown that a corpus-based research is crucial in order to gain a better understanding of the voice system in Romanian. Se and isi middles which were considered as one and the same construction namely, Romanian reflexives, have been shown to be very different in the semantic domain that they cover. Moreover, an extension of the isi-middle construction to external possessions and causative constructions has been observed. It has been shown that the functions of these two middle markers differ. Namely, se directs the access towards the event; whereas, isi-marker functions as a dominion indicator, making the subject of the sentence be the reference point with respect to which the object of the sentence in understood. Although there are main differences in the semantics of the verbs in the two genres, one tendency remained unchanged and this is that se-middle constructions occur more often than isi-construction in both genres. It was shown that these results point to one major tendency and this is that the frequency of se and isi is correlated to the functions covered by these two constructions. Finally, I argued that subjects prefer more frequent verb frames as opposed to rare. I.e. the fairy tales verb frames are more frequent. That's what the subjects prefer as well. So, the expected frequency effect is observed.

9. DISCUSSION AND POINTS FOR FURTHER RESEARCH

This paper contributes to the study of the Middle Voice in Romanian because it is an attempt to work with real data. There are two main views as to how scientific work should be done. One view is taken by Chomsky and other generative grammar linguists who believe that a deductive strategy which first makes up a minimalist theory and then sees if it can account for what happens in real language is the way to do real scientific work. However, as Derwing (1973) nicely puts it: “Chomsky’s philosophy of linguistics is imbued with the idea that theories are tested by inquiring whether the data at hand are or are not compatible with a transformational-generative description; rarely does one find a transformationist exposing the basic tenets of TGG to falsification in every possible way by exposing them…….”(1973:237). All the system is build on the linguist’s intuition which as Chomsky says is what the naïve speaker really has in his brain but is not consciously aware of it.

I subscribe to the second view of doing linguistics and this is that theories are developed from their own observational basis. Each hypothetical concept arises out of a
broad base data collection. Hence, “substantial importance is given to the actual use of
the linguistic system and a speaker’s knowledge of this use” (Langacker 2000:1, 2)
I subscribe to the strategy in cognitive grammar that “rules” can only arise as
schematizations of overtly occurring expressions. However, far this abstraction may
precede the schemas that emerge spring the soil of actual usage. Furthermore, Talmy’s
view that Linguistic forms can direct the distribution of one’s attention in a certain type
of pattern is also one I am adopting. Hence, as a first step to get to a real insight to how
the Middle Voice works in Romanian, I collected my data from a corpus. Results of the
most frequent lexical associations with the two middle constructions in Romanian were
obtained. I have showed how these results differ from the ones that Traditional
grammarians have used because they just trusted their own intuitions without paying
attention to language in use of the ‘naïve speaker’. Hence, in order to get answers as to
how people conceptualize one has to first look at real data in usage. This paper was the
first step of data collection and analysis. However, the second step that needs to be
taken is to provide answers as to why these differences in the distribution of se and isi
middle construction occur. In other words, what are the motivations for the tendencies
illustrated by this paper? In order to do so one has to precede the research in the
following way. We will look at verbs which can only appear with one or the other
middle voice construction and see what the semantic differences between the verbs per
se and the events that they illustrate are? (deponents). In this case we will also focus on
verbs which while semantically they seem to be part of the same semantic grouping (for
example cognitive verbs), do not take the same middle voice construction.

(38) a. El [[isi]] aminteste…..
    He MM remembers…
    He remembers (something)
b. El [[se]] gandeste…
    He MM thinks….  
    He thinks (about something)

However, this is not enough, in order to really hypothesis what conceptual
structures the middle constructions in Romanian have, one has to find out what the
conceptual lines people follow when they use the middle voice constructions are? In
order to do so an experimental study has to be done. My next project is to conduct an
experimental study which will be able to investigate what kind of schemas the two
different middle constructions invoke in the people’s mind. One such way, is to show
native speakers of Romanian non-verbal video clips and let them explain in their own
words what they have seen. The video-clips will make use of action verbs that can be
used both with se –middle construction and with isi-middle construction. While
explaining, native speakers will use one or the other construction. We will try to analyze
the main differences in the conceptualization of the speakers that brought them to
choose one expression over another and of course the grayer area where no clear cut
usage was found. If necessary, after the experiment interviews are going to be
conducted. In these interviews the speakers will be asked why they have decided to
choose one or the other construction. And if they hesitated, then what was their process
of thinking that made them hesitate but still finally decide on one of the constructions
over the other.
This experiment might also be able to provide more elaborate explanation as to why our intuition is closer to fairy tales although, most of our lives we are exposed to newspaper style genre and not fairy tales.

Finally, it would also be interesting to take a better look at the individual differences of the talkers.

In conclusion, the meanings of se and isi markers are also defined by the paradigmatic oppositions they enter in. Hence, analyze cases in which one construction can replace another. What makes the speaker unconsciously choose one over the other?

Finally, I believe that in order to provide a complete and fine grained analysis of the middle voice in Romanian one should also take into account the different case markings of se and isi markers. isi is marked dative; whereas the se is marked accusative. Langacker (1991, 1993), Wierzbicka (1988), Cienki (1993) all discuss how the meaning of case can be captured systematically. According to Langacker a case comprises a network of related senses pertaining primarily to semantic role, and takes some role archetype as its prototypical value. Cienki also argues that cases highlight relations between entities. One of the entities in the relationship is identified as more prominent by being marked with a case form. In my opinion, as well, case markers are meaningful elements that are combined with nominals to specify the nature of their involvement in a clausal process. The dative case is a grammatical expression of the role of the affected person not so the accusative case. Hence, this parameter should also be taken into account in order to add another piece to the PUZZLE called the MIDDLE VOICE.

NOTES:

1 This paper does not address the question if there is any need to really put a clear cut distinction between the reflexive and the middle constructions. Or if one should see the reflexive construction as being an instance of periphrastic middles (Shibatani 2003)

2 The verbal classes are also adopted from Kemmer 1993.

REFERENCES:

Barber, E.J.W. 1975. Voice, beyond the Passive. BLS 1, 16.24