Ignorance and Indifference: Epistemic Indefinites in Slovak

Veronika Richtarcikova

Màster: Màster Universitari en Lingüística Teòrica i Aplicada
Edició: 2012-2013
Directors: Dra. Berit Gehrke

Any de defensa: 2013

Col·lecció: Treballs i projectes de fi de màster de recerca

Programa oficial de postgrau
"Comunicació lingüística i mediació multilingüe"

Departament de Traduccion i Ciències del Llenguatge

Universitat Pompeu Fabra
Ignorance and Indifference:
Epistemic Indefinites in Slovak

Veronika Richtarcikova

Dirigit: Berit Gehrke

Barcelona, 30 d'agost de 2013
Ignorance and Indifference: Epistemic Indefinites in Slovak

Thesis presented to the department of Translation and Language Sciences in partial fulfilment of the requirements for the degree MA in Theoretical and Applied Linguistics

Author: VERONIKA RICHTARCIKOVA

Supervisor: DR. BERIT GEHRKE

External Advisor: DR. ANAMARIA FĂLĂUŞ

August 2013
Acknowledgments

I am grateful to a few amazing people that were instrumental in completion of this thesis.

Berit Gehrke, for all the help above and beyond, for teaching me to stay focused and for having confidence in me. I couldn’t have asked for a better supervisor. Anamaria Fălaş, for kindly agreeing to advise me, for generously giving her time, constructive feedback and answering my many questions.

Gennaro Chierchia, for introducing me to this area of research. In his course in Vitoria, I started listing Slovak polarity items in the margins of my densely covered notebook pages. Epistemic indefinites, this inconspicuous but hardworking part of grammar, has kept me intrigued since then. Analyzing them using the theory was a fantastic mental exercise. Dr. Chierchia said something that stuck with me and was of encouragement: "You’ll get confused. Stay calm, buckle up, and reason."

Margarita Samaniego, for taking a chance on me and my students at UNL in Loja, Ecuador, for indulging me as I was making my first baby steps into the field of linguistics. Without you, I would never have fallen in love with it.

My sister Barbi, for endlessly dissecting the Slovak data with me, for proofreading, but most of all for cheering me on. My parents, for their unconditional support.

My friends and fellow thesis writers YaFang, Yalda, Toni, Francesco, Alfonso and others, for helping me understand by patiently allowing me to explain things to them, and for the much appreciated company during the intense writing stretches. It’s been fun to share this journey with you.

Thank you.
Contents

Abbreviations and Notation ................................................................. 3
Abstract ............................................................................................ 4

1 Introduction .................................................................................. 5

2 Epistemic indefinites ................................................................. 7
   2.1 Haspelmath’s typological study and implicational map ............... 7
   2.2 Specificity ............................................................................... 9
   2.3 Knowledge of the speaker ...................................................... 10
   2.4 Free choice effect .................................................................. 11

3 A corpus study ............................................................................. 12
   3.1 Inventory of epistemic indefinites in Slovak ......................... 12
   3.2 Slovak National Corpus ....................................................... 15
   3.3 Annotation Methodology and Function Tests .................... 16
   3.4 Quantitative results and distribution of vola- and -si on Haspelmath’s map ...................................................... 25
   3.5 Qualitative observations ..................................................... 28
      3.5.1 Contexts and environments friendly to vola- and -si .... 28
      3.5.2 Ambiguous and interesting cases ................................ 32
   3.6 Conclusion of the section ..................................................... 37

4 Analysis ........................................................................................ 37
   4.1 The Alternatives-and-Exhaustification Paradigm .................... 38
   4.2 Application to Slovak ............................................................ 46
      4.2.1 Lexical entries ............................................................... 50
      4.2.2 Compatibility with modals ............................................ 51
      4.2.3 Seemingly episodic contexts ....................................... 54
      4.2.4 Imperatives ................................................................. 56

5 Discussion .................................................................................... 56

6 Conclusion .................................................................................... 61

Appendix .......................................................................................... 63
References ........................................................................................ 64
Abbreviations and notation

(i) Grammatical category labels

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>nominative case</td>
</tr>
<tr>
<td>G</td>
<td>genitive case</td>
</tr>
<tr>
<td>D</td>
<td>dative case</td>
</tr>
<tr>
<td>A</td>
<td>accusative case</td>
</tr>
<tr>
<td>L</td>
<td>locative case</td>
</tr>
<tr>
<td>M</td>
<td>masculine gender</td>
</tr>
<tr>
<td>F</td>
<td>feminine gender</td>
</tr>
<tr>
<td>NT</td>
<td>neuter gender</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>NEG</td>
<td>negator</td>
</tr>
<tr>
<td>PART</td>
<td>particle</td>
</tr>
</tbody>
</table>

(ii) Indefinite functions and features

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[SK]</td>
<td>specific known</td>
</tr>
<tr>
<td>[SU]</td>
<td>specific unknown</td>
</tr>
<tr>
<td>[S_/+]</td>
<td>(non-)specific feature</td>
</tr>
<tr>
<td>[K_/+]</td>
<td>(un)known to the</td>
</tr>
<tr>
<td>[IR]</td>
<td>irrealis</td>
</tr>
<tr>
<td>[Q_/+]</td>
<td>question</td>
</tr>
<tr>
<td>[CA_/+]</td>
<td>conditional antecedent</td>
</tr>
<tr>
<td>[IN]</td>
<td>indirect negation</td>
</tr>
<tr>
<td>[DN_/+]</td>
<td>direct negation</td>
</tr>
<tr>
<td>[CO]</td>
<td>comparative</td>
</tr>
<tr>
<td>[U-FC_/+]</td>
<td>universal free choice</td>
</tr>
<tr>
<td>[E-FC]</td>
<td>existential free choice</td>
</tr>
<tr>
<td>[∃/∀]</td>
<td>existential/universal</td>
</tr>
</tbody>
</table>

(iii) Other notation

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>downward entailing</td>
</tr>
<tr>
<td>EI</td>
<td>epistemic indefinite</td>
</tr>
<tr>
<td>FC</td>
<td>free choice</td>
</tr>
<tr>
<td>LF</td>
<td>logical form</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NPI</td>
<td>negative polarity item</td>
</tr>
<tr>
<td>SNK</td>
<td>Slovak National Corpus</td>
</tr>
</tbody>
</table>

In example sentences, indefinite pronouns are always highlighted in **boldface**. CAPITALS are used in examples to indicate sentence accent. Sometimes the behavior of two different indefinites is shown in one sentence. If both alternatives are grammatical alternatives, they are separated by a slash, e.g.

Ktosi/voťako je za dverami.

If only one of the alternatives is possible, the ungrammatical alternative follows with a star and if it is notably downgraded in acceptability, with two or one question marks, e.g.

Ktosi/*hočikto je za dverami.

Ak je niekto/??ktosi za dverami, povedz mi.
Abstract

This thesis investigates epistemic indefinites (EIs), the elements that overtly mark the 'epistemic effect', i.e. the inability to identify the discourse referent. The literature has identified various EI items across languages: German irgendein (Kratzer and Shimoyama 2002), Romanian vreun (Fălaş 2010), Italian uno qualsiasi (Chierchia 2013) and others. It is of fairly recent interest in semantics to provide a uniform solution for the properties, behavior and typology of EIs. I aim to complement this research program by providing a corpus survey and a semantic analysis of Slovak -ľova- and -si, EI items not studied until now.

While plain indefinites in Slovak (niekto) give rise to ignorance and free choice (FC) implicatures that can be easily cancelled, specialized series (ktosi or voľakoťo) are shown to carry a grammaticalized epistemic effect. I employ the Slovak National Corpus to give a description of the voľa- and -si series. By adopting a procedure from Aguilar Guevara et. al (2010), 600 randomly selected occurrences were annotated with functional labels. The following hypothesis was tested: the semantic/syntactic functions expressed by the same grammatical marker will fall into contiguous areas on an implicational map proposed by Haspelmath (1997).

The quantitative and qualitative analyses revealed the following results. (i) The functional adjacency predictions bear out with one exception of a disconnected universal-FC function. The two series are not equivalent in their distribution: voľa- was classified as 54% epistemic, 15% FC; -si was classified as 78% epistemic and 4% FC. (ii) I propose to differentiate an additional existential FC function, omitted in the original map. Following from this, (iii) the functional boundary of specific-unknown and existential-FC is identified as the most prone to dual readings. I present facts that suggest a dependency between prosodic prominence placed on the EI and the FC effect. Finally, (iv) Slovak EIs are shown to be exceptionally used in the specific-known function, which is predicted to be impossible for EIs (Aloni and Port 2010). To account for this curious behavior, I propose several pragmatic reasons evident in the corpus.

The alternatives-and-exhaustification theoretical approach (Chierchia, 2013) is adopted in the analysis of voľa- and -si. The epistemic effect is shown to be derived via a process of exhaustification of scalar and domain alternatives. Following Chierchia, I embed EIs within a wider system of polarity phenomena and contend that their behavior is determined by the types of alternatives they activate. I situate voľa- and -si with respect to the framework’s typology. A critical assessment of the theoretical perspective and how it fares in explaining the Slovak data contributes to the effort of finding a unified account of the cross-linguistic repertoire of EIs.
1 Introduction

Indefinite elements that overtly mark ignorance, i.e. the inability to identify the discourse referent, have recently gained attention in the semantic literature. Their properties and behavior are distinct from those of plain indefinites and at the same time highly complex and varied across languages that have them. In order to introduce the contrast between these specialized indefinites and plain indefinites, consider the following examples with Slovak niekto, niečo and their English equivalents somebody, something. They can give rise to ignorance and indifference implicatures:

(1) a. Nieko je za dverami.
   nie-who is behind door
   'Somebody is at the door.'
   Conventional meaning: A person is at the door.
   Ignorance implicature: The speaker does not know who it is.

(2) a. Povedz mi niečo o sebe!
   tell me nie-what about self
   'Tell me something about yourself.'
   Conventional meaning: Tell me a thing.
   Indifference implicature: The speaker does not care which specific thing.

These implicatures are of pragmatic nature, rather than part of what is asserted. As such, the following examples (in English, but the same holds for their Slovak counterparts given in (a)) show that the implicatures can be cancelled with a continuation in (b) or reinforced with a continuation in (c):

(1) b. Somebody is at the door, could you open? It is my friend.
   c. Somebody is at the door and I don’t know who it is.

(2) b. Tell me something about yourself, namely, where you work.
   c. Tell me something about yourself, it can be anything at all.

In some languages the ignorance (here epistemic effect) and indifference (free choice effect, FC henceforth) implicatures have become grammaticalized. For example, specialized epistemic indefinites (EIs) include German irgendein (Kratzer and Shimoyama 2002), Spanish algo (Alonso-Ovalle and Menéndez-Benito 2003), Romanian vreun (Fălăuș 2010) and others. Specialized FC indefinites include Spanish cualquiera (Menéndez-Benito 2010), Hungarian akárákí (Abrusán 2007), etc. An encoded conventionalized implicature is more difficult to

1 In glossing the Slovak indefinites, I use the form of marker+wh-base to indicate their morphological complexity. In the English translations, I resort to the indefinite that is grammatical and comes closest to the intended meaning.
cancel and reinforce. I will illustrate this point in the following. The Slovak 
indefinites in the sentence (3a) carry an enriched meaning roughly equivalent to 
'somebody/something and I do not know who/what'. If one tries to cancel (3b) 
or reinforce it (3c), the result is pragmatically odd.

(3)  a. **Ktosí/volakto** je za dverami.
who-si/vol-a-who is behind door
'Somebody is at the door.'
Conventional meaning: A person, and I do not know who, is at the door.

b. #Ktosí/volakto je za dverami, otvoríš? Je to Jana.
who-si/vol-a-who is behind door open is it Jane
'Somebody+epistemic effect is at the door, would you open? It’s Jane.'

c. ?Ktosí/volakto je za dverami, neviem kto je.
who-si/vol-a-who is behind door not.know who it is
'Somebody+epistemic effect is at the door, I don’t know who it is.'
(-a sense of redundancy)

The above examples suggest that *ktosi* and *volakto* could be classified as EIs. 
However, their variants also give rise to an FC effect in some contexts, as for 
example volačo in sentence (4). Just like the epistemic effect, the FC effect 
likewise appears to be grammaticalized in this item:

(4)  a. Povedz mi **volačo** o sebe!
tell me vol-a-what about self
'Tell me anything about yourself.'
Conventional meaning: Tell me a thing, I don’t care which specific thing.

b. #Povedz mi volačo o sebe! Konkrétne, kde robíš.
tell me vol-a-what about self namely where work
'Tell me anything about yourself. Namely, where you work.'

c. ?Povedz mi volačo o sebe! Môže to byť o ďom chcieš.
tell me vol-a-what about self can it be about what want
'Tell me anything about yourself. It can be whatever you want.'
(-a sense of redundancy)

When a Slovak speaker chooses this specialized variant of 'somebody/
something' over a plain indefinite, they do so in order to signal their inability 
(ignorance) or unwillingness (indifference) to identify the referent. As can be seen 
in (3) and (4) above, Slovak EIs contribute both ignorance and in some contexts, 
indifference to the semantics, as opposed to the pragmatics, of an utterance.

2 The *-si indefinites also induce the free choice effect in some contexts, but less so than
volačo. Section 3.5 gives a more detailed comparison.
This thesis is concerned with the semantics of EIs and aims to complement the research program of seeking a unified explanation of their properties. The main original contribution is introducing two Slovak EIs, lexical items not studied until now. My goals are twofold. The first goal is to give a description of their distribution, properties and behavior, using Haspelmath’s (1997) functional implicational map and adjacency predictions as a point of departure. I am especially interested in the environments in which they give an epistemic reading, a free choice reading or both. The second main goal is to demonstrate how the alternatives-and-exhaustification approach by Chierchia (2013) fares in accounting for the EIs in Slovak.

The structure of this work is as follows. First, in Section 2 I discuss Haspelmath’s seminal work and the basic concepts necessary when characterizing EIs in relation to other indefinites. In Section 3, I single out two Slovak EI series for further detailed discussion and demonstrate what is common to both of them and what is specific to each. I discuss whether the empirical evidence gained through a corpus survey supports the adjacency predictions proposed by Haspelmath. I show whether and how the patterns and the important aspects of their behavior are captured by Haspelmath’s implicational map. In Section 4, I briefly introduce the framework developed by Chierchia and then apply it to the Slovak data. I show how the approach can account for their semantic characterization and identify some of its limitations. Finally, in Section 5 I discuss remaining unresolved questions and suggest future directions of research.

2 Epistemic Indefinites

2.1 Haspelmath’s typological study and implicational map

*Indefinite Pronouns*, the seminal work on the subject by Haspelmath (1997), was the first attempt at a comprehensive overview of the properties of indefinites across languages.³ The author conducted a typological investigation of indefinites in 40 languages (Slovak not included) and showed that though the range of variation in their functional and formal properties is bewildering, it is not unconstrained. All languages in his sample were found to conform to a set of universal implicational constraints such as adjacency on a proposed map, shown in Figure 1 below. The *adjacency principle* states that an indefinite series will always express a set of functions that are contiguous on the map.

³ Haspelmath chose the cover term ‘indefinite pronoun’ for these words, even though not all are indefinite in a strict sense and not all are pronouns in a strict sense (Haspelmath, 1997:10). In order to avoid this confusion, I opted to use the term ‘indefinites’ in this thesis.
This map is an abstract categorization system for nine semantic and syntactic functions of indefinites and their arrangement graphically represents the adjacency principle. Obviously, not all distinctions are relevant and phonologically reflected in all languages. Haspelmath’s study shows that multifunctionality and overlapping of indefinites is cross-linguistically common. The majority of indefinites are used to express more than one of the functions, e.g. the English *any*-series covers all the non-specific functions. This can be explained by the fact that adjacent functions share some relevant characteristics, i.e. they have the same value on one of the binary features proposed by Haspelmath: (i) known vs. unknown, (ii) specific vs. non-specific, (iii) scalar end-point vs. no scalar end-point and (iv) in scope of negation vs. not in scope of negation. For more on the reasons behind the particular topography and distribution of features, see Haspelmath (1997:119-120). No two languages in Haspelmath’s sample had the same distribution pattern. For example, Figure 2 shows the distribution of English, Russian, Catalan and Romanian indefinites (Haspelmath 1997:65-69).
Although Haspelmath’s two-dimensional map has brought a significant improvement into our understanding of indefinites, it is not without problems. The set of only nine functions has been convincingly criticized for being too coarse grained (van der Wouden 2000), expanded to account for further subtle distinctions (Aguilar-Guevara et al. 2010) or abandoned completely (Aloni and Port 2010). However, the main generalizations by Haspelmath, namely the distinctions of indefinites and their principal semantic and syntactic functions have been generally upheld, and therefore the map serves as a good starting point for the purposes of this work.

EIs can be positioned on the map with the application of the following definition in Aloni and Port (2010): EIs are those expressions in any language that exhibit the specific unknown function [SU], and possibly others to the right, but not the specific known function [SK]. The label epistemic is certainly quite ambiguous and seems to be readily applicable to all attitudes related to knowledge, which has caused terminological confusion in the literature. However, distinctions of readings (especially between ignorance/specific and indifference/non-specific) will become essential in the course of this work. In order to elucidate how I understand EIs to be different from other specialized indefinite items, I modify the above definition slightly:

(5) EIs are those expressions in any language that convey the attitude of ignorance about the identity of the expression’s referent. They are primarily used in the [SU] function: their referent is specific but unknown. They can possibly exhibit other functions, congruent to [SU], in some less frequent contexts.

Of particular concern to my work are therefore two factors to which indefinites may be sensitive, namely specificity and knowledge. Haspelmath (1997) discusses how these factors reflect on the distribution of an indefinite across the implicational map.

2.2 Specificity

Specificity is a key concept in the semantics of reference and notoriously responsible for ambiguity in languages like English that do not overtly mark it. Consider the following example, taken from Haspelmath (1997:38). In the absence of context, both specific and non-specific readings are possible.

(6) a. He wants to buy some bicycle.
   b. (He already has one picked out in a store.)...It costs 60€. [SPECIFIC]
   c. (Any bicycle would do.)...#It costs 60€. [NON-SPECIFIC]

On the specific interpretation, a present indicative continuation with co-referenced anaphoric pronoun ‘it’ is possible (Karttunen 1976:366). The bicycle’s existence is presupposed and even if the speaker cannot identify it, it is in principle identifiable. Another way to confirm a specific reading is to paraphrase
it with an existential sentence (Heringer 1969:90): 'There is such a bicycle that he wants to buy.' This is not possible with the non-specific interpretation, in which the discourse referent’s existence is not presupposed. In (5c), the indefinite has a ‘no-matter-what’ reading and fills the [FC] function.

Haspelmath (1997) identifies contexts in which specific readings are possible and therefore specific indefinites are allowed (Table 1). As we can see, certain constructions (irrealis ‘want’, future and distributive contexts) are especially prone to the ambiguity between specific and non-specific readings.

<table>
<thead>
<tr>
<th>perfective past, ongoing present</th>
<th>'want', future, imperative, distributive</th>
<th>question, conditional</th>
<th>in the scope of negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>specific possible</td>
<td>(specific impossible)</td>
<td>non-specific possible</td>
<td></td>
</tr>
<tr>
<td>(non-specific impossible)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Contexts for (non-)specific phrases, Haspelmath (1997:45)

Another factor that determines the choice to use an EI is the speaker’s knowledge, i.e. the ability to identify the discourse referent of the indefinite.

2.3 Knowledge of the speaker

This feature applies only to specific phrases, in which the existence and therefore the identifiability of their referents are presupposed. In non-specific phrases, the referents of indefinite expressions are in principle unknown to the speaker. In the example (6) under the specific reading, the ignorance encoded in the indefinite signals that the speaker does not have more information to give:

(7) He wants to buy some bicycle. (He already has one picked out in a store.)
    #Guess what kind!

The oddity of the continuation shows that the function of the EI some here is specific and unknown [SU]. However, there are also cases in which the speaker can reasonably be expected to identify the referent and yet they choose an EI, possibly to withhold the information, to empathize with the hearer or to indicate its irrelevance in the moment. Consider the following example:

(8) Matka pokračovala: "Musím sa o volačom porozprávať s touto ťažou.
    mother continued must part about vola-what talk with this lady
    'Mother continued: I have to talk to this lady about something.'
    ...It, is important. [SPECIFIC]
    ...??Guess what! [UNKNOWN?]

4 It will be shown later that it is not necessarily the speaker’s epistemic attitude that is relevant to the interpretation. It may be another participant in the state of events. Because I am summarizing Haspelmath’s (1997) work, I will follow this simplification. Moreover, the term ‘speaker’ is to be understood as including ‘author’ for written texts.
5 Unless otherwise noted, all the Slovak data presented from this point on have been extracted from the Slovak National Corpus (SNK).
In this sentence, the mother naturally knows what she wants to talk about, however, because of her choice of the EI voľačo (‘something+epistemic effect’), it would be quite infelicitous for the children to ask about it. In cases such as these, the indefinite remains specific, but the semantic distinction between known and unknown is blurred and difficult to pinpoint. In Section 3.3.2 I will discuss what might be at play to license the use of EIs in such exceptional contexts.

Haspelmath’s (1997) summary of the possible combinations of knowledge, specificity and definiteness are presented in Table 2 below.

<table>
<thead>
<tr>
<th></th>
<th>indefinite</th>
<th>definite</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-specific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unknown to the speaker</td>
<td>known to the speaker</td>
<td>known to the speaker and the hearer</td>
</tr>
</tbody>
</table>

*Table 2: Relation between (in)definiteness, (non-)specificity and knowledge of the speaker, Haspelmath (1997:46)*

In addition to specificity and knowledge, no discussion of epistemic indefinites can overlook their capacity to convey a free choice reading in certain contexts, therefore at least a few comments on this topic are in order.

### 2.4 Free choice effect

The concept of an FC effect, originally introduced by Vendler (1967), captures the meaning that more than one member of a salient domain is an admissible option. An FC indefinite can often be substituted with a universal quantifier ‘every’, though not always. For example, in the sentence (4), replicated below, the EI voľačo triggers an FC effect with an existential quantificational force, but not a universal one. It cannot be replaced with všetko (meaning ‘everything’) without a significant change in truth conditions.

(4) Povedz mi voľačo / všetko o sebe!  
Tell me voľa-what/everything about self  
Tell me something-anything/everything about yourself.

The FC reading is easiest to come by in modal contexts, in imperatives if interpreted as a permission rather than as a command, in generic sentences in imperfective present or perfective future tense, and in hypothetical contexts (Haspelmath, 1997:49-51). In episodic contexts, EIs can fill this function if they are made prosodically prominent in a sentence, for example in (9):

(9) (Mária) Počívala a zo všetých sil sa usilovala  
Mary listened and from all strengths tried  
VOĽAČOMU porozumieť.  
voľa-what understand  
‘Mary was listening and tried her hardest to understand something, anything at all’.
Without the prosodic emphasis, the EI voleľa in this sentence gets an expected [SU] reading - the speaker's inability to identify the particular thing that Mary was trying so hard to understand. However, if stress is placed on the item, it invokes the reading of the agent’s (Mary’s) indifference as to the identity of what she was trying to understand. In other words, Mary wanted to understand at least something - no matter what. If her interlocutor said things a, b and c, it is compatible with Mary’s desire that she understand the thing a, or the thing b, or the thing c - the variation that is the signature FC phenomena. As can be seen, ignorance and indifference meanings are sometimes very difficult to disentangle and can be correlated with prosody. I will return to a more detailed analysis of how EIs interact with FC in Section 4.

3 A corpus study

We arrive at the empirical core of this thesis. In this section I will first give a brief introduction to indefinites in Slovak and the morphological markers that characterize them. There are numerous indefinite series in Slovak, but only two were selected for their specially encoded epistemic effect quality: -voleľa and -si. Next, I will outline the methodology that was adopted for the corpus annotation. The main goal of this section is to present the statistical distribution of these two series across Haspelmath’s implicational map and to determine whether the functions that these EIs serve are adjacent to each other.

3.1 Inventory of epistemic indefinites in Slovak

Slovak lacks articles but has a particularly elaborate system of indefinite reference constructions to compensate for this. Speakers of Slovak can choose among dozens of series (a term first introduced by Veyrenc 1964) of indefinites to express meanings of some-, any- and no- and other minute shades of meanings in between. As an open class, indefinites in Slovak continue to embrace new series as people use language creatively and try to bring more precision into the communication (a case in point being the emphatic ľertvie-, 'the devil knows' series). Specialized indefinite series are subject to grammaticalization, but at the same time have a relatively short lifespan and are diachronically quite unstable (Haspelmath 1997: 235, Aguilar Guevara et al. 2010).

Indefinite series consist of one member (disregarding gender and number) for each ontological category, such as person, thing, property, place, etc. Indefinites in Slovak are derived forms in which the stem, an interrogative pronoun, indicates the ontological category. This is unlike in English, where we have somebody but not somewho, but it is in line with how indefinites are derived in the majority of languages (Haspelmath 1997). The wh- stems can be attached to either a prefix (e.g. voleľ-) or a suffix (e.g. -si), an element called indefiniteness marker (a term also from Veyernc 1964). The indefiniteness marker is always an

---

6 One exception is the suffix -ever which creates indefinites whoever, whatever, etc.
‘extrafix’, i.e. the outermost affix that follows other inflectional affixes that mark case, gender or number. It is only allowed to attach once (*voľaktosi, *voľavoňakt, *ktotisi). A non-exhaustive list in Table 3 illustrates the variety of the indefinite class in Slovak.

<table>
<thead>
<tr>
<th>Indefiniteness marker</th>
<th>Ontological category</th>
<th>nie- / da(j)-</th>
<th>vlo-/-si/- / poda/-ponie-</th>
<th>ho(e(j)-</th>
<th>/ kovek/kde-/kade-</th>
<th>leda-/-soov- / malo-/-to</th>
<th>ni-</th>
<th>*\‘almost no-\‘</th>
<th>*\‘no-\‘</th>
</tr>
</thead>
<tbody>
<tr>
<td>person (-body)</td>
<td>niekto</td>
<td>vošťaktos</td>
<td>hocišťaktos</td>
<td>leškto</td>
<td>ktracko</td>
<td>malakto</td>
<td>niko</td>
<td>*\‘no-\‘</td>
<td></td>
</tr>
<tr>
<td>thing (-thing)</td>
<td>nieco</td>
<td>voščo</td>
<td>hocičo</td>
<td>lešo</td>
<td>goto</td>
<td>zasko</td>
<td>nie</td>
<td>*\‘no-\‘</td>
<td></td>
</tr>
<tr>
<td>place (-where)</td>
<td>niekde</td>
<td>vošťaktode</td>
<td>hocišťaktode</td>
<td>leškode</td>
<td>malakode</td>
<td>nikoode</td>
<td>nikde</td>
<td>*\‘no-\‘</td>
<td></td>
</tr>
<tr>
<td>time (-time)</td>
<td>niekedy</td>
<td>vošťakedy</td>
<td>hocišťakedy</td>
<td>lešakedy</td>
<td>malakedy</td>
<td>nikoedy</td>
<td>niko</td>
<td>*\‘no-\‘</td>
<td></td>
</tr>
<tr>
<td>manner (-how)</td>
<td>nejak</td>
<td>vošťajak</td>
<td>hocišťajak</td>
<td>lešjak</td>
<td>nikojak</td>
<td>*\‘no-\‘</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>property</td>
<td>nejaký</td>
<td>vošťajky</td>
<td>hocišťajky</td>
<td>lešjaký</td>
<td>nikojaký</td>
<td>*\‘no-\‘</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>selection (-which)</td>
<td>niektóreši</td>
<td>vošťaktóreši</td>
<td>hocišťaktóreši</td>
<td>lešaktóreši</td>
<td>nikoaktóreši</td>
<td>*\‘no-\‘</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>amount</td>
<td>niekoško</td>
<td>vošťakoko</td>
<td>hocišťakoko</td>
<td>lešakoko</td>
<td>nikoško</td>
<td>*\‘no-\‘</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3**: A sample of indefinite series in Slovak across ontological categories

In Table 4, note that in all of the inflected forms, the indefiniteness marker -si remains unchanged.

<table>
<thead>
<tr>
<th>Ontological category (interrogative in English)</th>
<th>Interrogative pronoun</th>
<th>M. SG</th>
<th>F. SG</th>
<th>N.SG, PL</th>
<th>M. PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>person (who)</td>
<td>kto</td>
<td>NOM</td>
<td>ktorní</td>
<td>ktorní</td>
<td>ktorní</td>
</tr>
<tr>
<td>thing (what)</td>
<td>šo</td>
<td>GEN</td>
<td>ktorní</td>
<td>ktorní</td>
<td>ktorní</td>
</tr>
<tr>
<td>place (where)</td>
<td>kde</td>
<td>DAT</td>
<td>ktorým</td>
<td>ktorým</td>
<td>ktorým</td>
</tr>
<tr>
<td>time (when)</td>
<td>kedy</td>
<td>ACC</td>
<td>ktorým</td>
<td>ktorým</td>
<td>ktorým</td>
</tr>
<tr>
<td>manner (how)</td>
<td>ako</td>
<td>LOC</td>
<td>ktorým</td>
<td>ktorým</td>
<td>ktorým</td>
</tr>
<tr>
<td>property(what kind)</td>
<td>ako</td>
<td>INSTR</td>
<td>ktorým</td>
<td>ktorým</td>
<td>ktorým</td>
</tr>
<tr>
<td>selection (which)</td>
<td>ktorý</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>amount (how much)</td>
<td>koško</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reason (why)</td>
<td>prezo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4**: Affixation in Slovak indefinites
At least six series, to my knowledge, exhibit the characteristics of a grammaticalized epistemic effect, in the narrow sense as explained in Section 2 - they signal the speaker’s ignorance and in some contexts, the agent’s indifference:

(10) **Ktosi/Volakto/Čertvieko/Bohviekto/Podakto/Poniekto** je za dverami.

who-si/vola-who/čertvie-who/bohvie-who/poda-who/pone-who is behind door

'Somebody, I don’t know who, is at the door.' #Guess who?

Out of these, two series will be studied in further detail here: -si and vola-. Although they are not the only EIs in Slovak, I have selected them for this study because of their usage frequency and therefore, availability of authentic data. Although native speakers acknowledge that the two series are (nearly) synonymous, it is important to note that -si is normally considered the ‘correct’ one and vola- is more colloquial, or archaic. The proportion of occurrences in the written text corpus reflects this fact, as will be seen in the subsequent section.

Only the indefinite forms for the ontological categories of person, thing, and property were queried in the corpus and considered in my analysis. This decision was motivated mainly by the aim to complement the existing studies of EIs in the literature, which seem to unanimously prefer data in the person-thing-property categories and to facilitate the diagnostic tests and annotation. However, it deserves mention that without empirical evidence, we cannot presume that the other categories, such as place, manner, etc. do not differ substantially in their function distribution.

To demonstrate that other ontological categories can be analyzed in much the same way, consider the following two examples.

(11) Prestala? To znamen, že volakedy si ju jedla. Prečo už quit it means that vola-when part her eat why already

not.eat pork neješ bravšovinu?

You quit? That means that you used to eat it sometime. Why don’t you eat pork anymore?’

a. ...volakedy, si ju jedla. Vtedy, si nebola vegetarianka. [SPECIFIC]

‘...you used to eat it sometime. You were not a vegetarian then.’

b. ...volakedy si ju jedla. #Hádaj kedy! [UNKNOWN]

‘...you used to eat it sometime. Guess when!’

---

7 Much of the literature on EIs also looks at determiners (e.g. some+NP, any+NP as opposed to morphologically complex forms someone, anyone that do not combine with a noun). In English, someone and some person may convey a different epistemic attitude in the same utterance. As already mentioned, Slovak does not have simple determiners per se, and to my knowledge, the expressions volakto (someone) and volákaj človek (some person) are equivalent. A survey of more native speakers’ interpretation judgments could reveal otherwise. I set this distinction aside for the purposes of this work.

'She has a brother in Germany or England. Somewhere. He showed me his photo.'

a. ...Tam voláde. Zaržba tam, dobre. [SPECIFIC]
   '...Somewhere. He earns a lot of money there.'

b. ...Tam voláde. #Hádaj kde! [UNKNOWN]
   '...Somewhere. Guess where!'

Both indefinites in the above examples can be labeled as [SU], fulfilling the specific unknown function on Haspelmath’s map. In the first test, both utterances can be continued with a pro-form which can in turn be co-referenced with the indefinite under question. Next, it is obvious both from the context and the infelicity of the continuation ‘Guess when/where!’ that the speaker cannot identify the particular time and place. The study of these expressions could potentially broaden the account of indefiniteness in an interesting way but is left to be explored in the future. From this point on, I will only focus the discussion on data from the person-thing-property ontological categories.

3.2 The Slovak National Corpus

The Slovak National Corpus (SNK) is a corpus offered to the public for research, educational, and other strictly non-commercial purposes and is provided by the Ludovít Štár Institute of Linguistics of the Slovak Academy of Sciences, Bratislava, Slovakia. The subcorpus prim-6.0 is a monolingual database of contemporary Slovak written texts and contains 1,155,742,085 tokens. The texts cover a broad range of language styles (77.8 % public media, 9.8 % literary texts, 11 % scientific and technical texts and 1.4 % unspecified).

The following queries were made in order to extract the data for the two selected indefinite series from the corpus. The search included combinations of the epistemic marker morpheme with four bases: -kto (‘who’), -čo (‘what’), -który (‘which’) and -aky (‘what property’). All four combinations were further expanded into 3 noun classes (masculine, feminine and neuter), singular and plural, and 6 grammatical cases (or fewer in case of homophony). See Appendix for a detailed list of forms queried for the two series.

The search for -si indefinites yielded 235,399 occurrences in the corpus. For -vol’a indefinites, this search yielded 7,547 occurrences. Out of these, random samples of 300 were generated for each series and annotated according to their functions.
3.3 Annotation Methodology and Function Tests

I now turn to the overview of the procedure of identifying the functional distinctions expressed by the indefinite -voľa and -si series. The starting points were the following. First, I used the original set of functional labels which combine both context (syntax) and meaning (semantics) identified by Haspelmath (1997), and I expanded it solely with the addition of an existential FC function. Secondly, I adapted a decision tree (illustrated in Figure 3), along with the relevant diagnostic tests, from Aguilar Guevara et al. (2010).

In order for an EI to qualify for a function, it must (i) be grammatical in the syntactic context that the function specifies; and (ii) have the semantic interpretation that the function specifies (Aguilar Guevara et al. 2010). The first aspect is taken for granted here because the data were retrieved from a corpus of natural language, therefore every instance must have been grammatically acceptable to some Slovak speaker. Admittedly, exceptions to this exist as even the corpus may contain ungrammatical constructions for a number of reasons. However, I chose not to discriminate my data based on this fact. With regards to (ii), the following sentence represents an instance in which the indefinite does not qualify for a function that its syntactic context would predict:

(13) Očarila ho, hoci v zrelom veku už s čošil podobnym charm him although in ripe age already with what-si similar

'Her charm him even though at that ripe old age he did not count with something like that.'

The EI čosil in this sentence would not be annotated as direct negation [DN], despite the fact that the negative marker ne on the clausal verb would tempt us to do so. Under closer inspection, the indefinite in this sentence does not exhibit the universal, but the existential meaning, as specified by the negation function.

In order to determine the semantic function of an indefinite as objectively as possible, a set of tests was applied to each EI occurrence in the corpus sample during the annotation procedure. The decision tree schematizes the order of the tests. Each branching node, starting from the top, represents a step that assigns a binary feature to the construction, e.g. [S+] for SPECIFIC or [S-] for NON-SPECIFIC. The terminal node that these steps arrive to assigns the EI’s function and its location on the adjacency map. Note that E(xistential)-FC is added to the original set of nine functions by Haspelmath. In case of an ambiguity, e.g. due to the lack of prosodic information, the item was annotated with both possible functions and each counted as 0.5 in order to keep the statistics intact. A small percentage of the data was labeled as UNCLEAR. Some of the examples that were difficult to classify are discussed in Section 3.5.
I will now explain the diagnostics in more detail and illustrate with corpus examples from both series that pass and fail each test.

The **test for specificity** asks whether a continuation with an anaphoric pronoun, co-indexed with the indefinite, is possible: ... EI ... He/She/It... Its application assigns the feature SPECIFIC [S+] or NON-SPECIFIC [S-]. For example, the utterance (14b) below is an episodic sentence, in which the speaker recounts an actual event of not understanding what was being said to her. It would be possible to continue with 'It was in French', therefore, this use of the EI passes the test and can be placed in the specific region on the map. On the contrary, in (14d) we do not have an account of the agent writing a specific thing, the object exists only in the world of his desire. The continuation 'It was a novel' is not possible and the sentence can be labeled as non-specific.

(14) a. Matka pokračovala: Musím sa o voľba porozprávať s touto párou. Mother continued: I have to talk to this lady about something. ...It is important. [S+]
b. Niekoľko mi podával ruku a vozí nad hovoril. Nerešala som mu. NE-who me give hand and VOC-what say NOT-understand PART him Someone was giving me their hand and telling me something. I didn’t understand him. ...It was in French. [S+]
c. Uvádza, že do domu podstúpia voličský dohovor, ktorý piv by potvrdzoval, že do considered that into house slip VOC-what evidence which would confirm that went to VRAȘDIN about murder. 'She considered to slip into the house some evidence, which would confirm that it was a murder.' ...#It is a weapon. [S-]
d. Od malých vekov ďalšiu mal na staranie aj voľby, napísal. from childhood much read and had urge also VOC-what write 'Since childhood he read a lot and had an urge to write something.' ...#It was a novel. [S-]
e. A naďa mi hude jablko? Mám ho azda vozíčku, darovac? and what-for me will-be apple should him maybe VOC-what give 'And for what do I need an apple? Should I give it to somebody?' ...#He is my brother. [S-]
f. Ak niekoľko tretním alebo sa na vozíčku píšam, mám iba zo zveďavosti. if NE-what blurt or PART about VOC-what ask well only of curiosity 'If I blurt out something or I ask about anything, it’s only out of curiosity.' ...#It is a math problem. [S-]
g. Dvadsať rokov sa mi nestalo, aby sa mi voľba, všimol na faru. twenty years PART me NOT happen that PART me VOC-what break-in on rectory 'It hasn’t happened to me for 20 years that anyone, would break into the rectory.' ...#He is in jail. [S-]
h. Zostavím algebru osnovu, abych som na volbo, nezabudol, make some-st outline so FART ON VOKA-what NOT FORGET
I'll make an outline so I don't forget anything.'...#It, explains the data. [S]

i. Poďvala a zo vetek sl sa usilovala volba mi, porozumie,
listened and of all strength PART try VOKA-what understand
'She was listening and tried her hardest to understand something.'...#It, surprised her. [S2]

In the same fashion, the test is applied to the -si series:

(15) a. Sadni si a potovaj ma, protože ti poviem gosi, tvrdí. Nezad to robim, ale musím. sit part and listen me because you say what-st hard NOT glad it do but must 'Sit down and listen to me because I'll tell you something hard. I don't like to do it but I must.'...#It, is important. [S+]

b. Dnes sa tu objevil algi, maj vrah, že by sa cheol už lietar, today part here sneak - some-st man said that would part want learn fly 'Today some man was sneaking around, he said that he would like to learn to fly.'...#He, seemed crazy. [S+]

c. Vietor zavíval zavjal, ako by nač chcel gosi, povedal. wind furiously howled as-if we wanted what-st say 'The wind howled furiously, as if it wanted to say something to us.'...#It, was in English. [S2]

d. Musme v tonto snare ziskan gosi, urobí, tvrdí starostka. must in this direction soon what-st do says mayor 'We must do something in this direction soon," says the mayor."...#It, is a solution to the problem. [S]

e. Budu tento Silvester emi, svojsk, osobité? will-be this New-Year's Eve your-st individual distinctive 'Will this New-Year's Eve be unique in some way?...#It, will be the fireworks. [S2]

f. Ak sa vam gosi, podobaš prihod, máli by ste myslel... if part you what-st similar happen should part part think 'If something similar happens to you, you should think..."...#It, is a car accident. [S-]

g. Nezdalo sa mu, že by bol gosi, stratil, alebo kohosi zradil. not seem part him that would part what-st lost or who-st betrayed 'He didn't think that he had lost anything, or that he had betrayed anyone...,'...#It, was his car keys. ...#It, was his friend. [S2]

h. Och, ako rád vyriec gosi, slovom. oh how difficult say what-st word 'Oh, how difficult to say something, in words!'...#It, surprised her. [S2]

Next, the test for known is applied to those sentences that passed the test for specificity. The test is quite intuitive and asks whether the speaker could possibly continue with 'Guess what/who/etc.:' and thus signal that they have more information to give about the indefinite's referent. A sentence passes this test if the speaker can reasonably be expected to identify the object in question. Consider the following examples.

(16) a. Matka pokračovala. Musím sa o volba porozprávať s touto paniou. mother continued must FART about VOKA-what talk with this lady 'Mother continued: I have to talk to this lady about something.'...(?!)Guess what! [K+]

b. Niekoľko mi podaľ ruku a volba, hovoril. Nenamnela som mu. ne-who me give hand and VOKA-what say NOT understand part him 'Someone was giving me their hand and telling me something. I didn't understand him.'...#Guess what! [K-]
a. Sadni si a potrav ma, preto e ti poviem. Nenadf to robam, ale musam. Sitt part and listen to me because you say what i hard not glad it do but must
b. Dnes sa tu obtomel someone sneaked around here, he said that would part want learn to fly. 'Today some man was sneaking around, he said that he would like to learn to fly.'

In (17a) the speaker requests that their friend have a seat because they have something to say. It is reasonable to assume that the speaker knows what they are about to say to their friend, and a continuation with 'Guess what!' would be logically acceptable. Therefore this occurrence gains the feature KNOWN, [K+]. On the other hand, the speaker in (17b) witnessed a strange, unknown man sneaking around and would not be able to identify who it was, except possibly by appearance. The continuation of 'Guess who!' would be odd and this instance is labeled as UNKNOWN, [K-].

Moving on to the non-specific region of the map, we next apply the test for universal meaning to distinguish between existential and universal interpretations of an indefinite. The test consists of paraphrasing the sentence into a universal construction 'For every/all x, ...' where x is then substituted for the indefinite expression. Schematically: ...Op (... El...) ... ⇒ ... ∀x (Op ... x...) ...

If the rephrasing is entailed within the meaning of the original utterance, the universal feature [ ∀ ] is assigned. If such a paraphrase is not entailed, the existential feature [ ∃ ] is assigned. For example, in (18c), the speaker conveys that the woman considered to slip some unspecified evidence into the house, but not more than one piece. 'For every x, she considered to slip in x into the house' is a rephrasing that is not entailed, therefore, the sentence is assigned the EXISTENTIAL [ ∃ ] feature. On the other hand, (18f) can be paraphrased in such a way. The speaker is saying that if they ask about anything, and it could be 'more than one thing, it is out of curiosity. The universal paraphrase 'for every x, when I ask x, it is out of curiosity' is entailed in the meaning of the statement, and the sentence receives the UNIVERSAL [ ∀ ] feature.

(18)c. Uva rovala, ze do domu podstr vojaka, ktori by potvrdoval, ze slo concerned that into house slip voja - some evidence which would confirm that went o vraid about murder
She considered to slip into the house some evidence, which would confirm that it was a murder.

d. Od malicka vea etal a mal nutkanie aj volvo what write from childhood much read and had urge also volva - what write Since childhood he read a lot and had an urge to write something.

8 This subnumbering is consistent with the first introduction of these sentences in (14-15).
f. Ak niečo treplia alebo sa na voľboto pýtaj, manžel sa zvedává.

if nie-what blurt or PART about VOA-what ask well only of curiosity

′If I blurt out something or I ask about anything, it’s only out of curiosity.′

⇒ for every x, when I ask x, it is out of curiosity. [v]

g. Dvadsať rokov sa mi nestalo, aby sa mi voláto vglázal na faru.

twenty years PART me NOT happen that PART me VOA-who break-in on rectory

′It hasn’t happened to me for 20 years that anyone would break into the rectory.′

⇒ for every x, it hasn’t happened for 20 years that x would break into the rectory. [v]

h. Zostavam akoši osnovu, aby som na vôlebo nezabudol.

make some-SI outline so PART on VOA-what NOT.forget

′I’ll make an outline so I don’t forget anything.′

⇒ I’ll make an outline so that for every x, I don’t forget x [v]

i. Počivala a zo všetkých slova sa usilovala voláťomu porozumieť.

listen and of all strength PART try VOA-what understand

′She was listening and tried her hardest to understand something.′

⇒ for all x, she tried to understand x [v]

Following the same procedure, the examples below illustrate the universality test as applied to the -si series:

(19) c. Vietor zrazil savújal, akoby miam chcel ľad povedať.

wind furiously howled as-if us wanted what-SI say

′The wind howled furiously, as if it wanted to say something to us.′

⇒ for all x, ...as if he wanted to say x. [3]

d. Musíme v tomto smere gošoro ľad urobiť, tvrdila starostka.

must in this direction soon what-SI do says mayor

′We must do something in this direction soon," says the mayor'.

⇒ for every x, we must do x. [3]

e. Bude tento Silvester ľudí svojský, osobity?

will-be this New-Year’s Eve what-SI individual distinctive

′Will this New Year’s Eve be unique in some way?′

⇒ for every x, will this New Year’s Eve be unique by x? [3]

f. Ak sa vám ľad podĽovbu priniesť, mali by ste mysleť...

if PART you what-SI similar happen should PART PART think

′If something similar happens to you, you should think...′

⇒ for all x, if a similar x happens, you should think [v]

g. Nezdalo sa mu, že by bol ľad, stratil, alebo kohosi kohosi zradil.

not seem PART him that would PART what-SI lost or who-SI betrayed

′He didn’t think that he had lost anything, or that he had betrayed anyone.′

⇒ for all x, he didn’t think that he had lost x [v]

h. Och, ako ľad vyrieť ľad slovom.

oh how difficult say what-SI word

′Oh, how difficult to say something in words’. 

⇒ for all x, it is difficult to say x in words [v]

At this point, those corpus occurrences that are in the existential area can be further distinguished based on the simple test for question. The [Q+] feature is assigned to an interrogative and the [Q−] to a non-interrogative construction.

(20) c. Uvažovala, že do domu podstry voňajaly dôkaz, ktorý by potvrdoval, že šlo

considered that into house slip VOA-some evidence which would confirm that went

about murder

′She considered to slip into the house some evidence, which would confirm that it was a murder.′ [Q−]

d. Od malíška veľa čital a mal mutlukanie aj voľboto napsať.

from childhood much read and had urge also VOA-what write

′Since childhood he read a lot and had an urge to write something.' [Q ]
And for what do I need an apple? Should I give it to somebody? 

The wind howled furiously, as if it wanted to say something to us.

'We must do something in this direction soon,' says the mayor.

Will this New Year's Eve be unique in some way?

The next test is for existential free choice in which we try to discern the "no matter what" reading from irrealis. As the terms themselves suggest, FC will invoke the agent's indifference meaning and irrealis will be associated with a nonfactual (subjunctive, speculative, etc.) interpretation. Admittedly, the distinction is not clear-cut. In this test, I asked whether a paraphrase using a specialized FC indefinite from the *hoci*- or *koľvek* series in place of the EI will return the same truth conditions. For example, in (22c), the woman is considering planting a piece of evidence, but not just any evidence. She is not indifferent in her choice - she would hypothetically plant only the kind of evidence that would confirm that it had been a murder. Because of the hypothetical context and the fact that substitution of voľajký with *hocijáký* sounds odd, the sentence received the irrealis [E-FC] feature. On the other hand, the EI in (22d) can be replaced with *hocičo*. The speaker describes a person who had an urge to write something or other, and did not care what. Therefore, this EI instance received [E-FC+] feature.

Turning our attention to the universal region on the map, we apply the test for negative meaning. A negative context will cause the conjunction of the two opposite statements in place of the EI to be inconsistent (...Op (a ∨ ¬a)...is inconsistent). For example, the sentence (24f) is consistent in meaning if the speaker ends up asking something or not. Therefore, it receives NON-
NEGATIVE [Neg+] feature. On the other hand, sentence (24g) would not be consistent if in the past 20 years someone would break into the rectory. Therefore, the context is recognized as NEGATIVE and receives a [Neg+] feature. For such sentences, Haspelmath’s (1997) map also makes a distinction between direct and indirect negation. The negative marker in (24g) is in the matrix clause and the EI occurs in the subordinate clause and therefore the INDIRECT [D-] feature is added to the annotation. An example of DIRECT [D+] negation is the sentence (24h).

Following the same procedure, the examples below illustrate the negation test as applied to the -si series:

From the sentences that turn out non-negative, we move down the decision tree and apply a test that identifies the universal free choice function. According to Aguilar Guevara et al. (2010), the test for FC is identical to the above test for negative meaning, but instead of consistency, we look for informativeness. If ...Op(a ∨ ¬a)... is informative, the sentence receives a FREE CHOICE, here called ‘universal’ [U-FC+] feature, and if ...Op(a ∨ ¬a)...is not informative, it will receive a [U-FC-] feature. As a more intuitive explanation, this test asks whether the agent is indifferent as to the identity of the referent, i.e. whether we get the ‘no matter what’ reading. Let us consider the following examples.
In (27f), an antecedent to a condition that provides two contradictory options is not very informative to the hearer who would not know in which situation they should follow the speaker’s advice expressed in the consequent of the statement. However, in the example (27h), the reformulation is in principle informative and the interpretation we get is likewise that of the speaker’s difficulty to say anything, no matter what - the signature FC effect. And finally, the self-explanatory test for conditional context applies a [CA+] feature to an indefinite that occurs in a conditional antecedent:

This concludes the detailed explanation of the annotation procedure. The above selection of examples, chosen from the corpus and subjected to the battery of tests for the purpose of illustration, resulted in the following function labels:

(30) voi’a-series

a. Matka pokračovala: ‚Musíš sa o nácenné porozprávať s touto párou.‘
   'Mother continued: ‘I have to talk to this lady about something.’'

   'Someone was giving me their hand and telling me something. I didn’t understand him.'

c. Uvažovala, že do domu podstrel vojáka, ktorý by potvrdzuval, že slovo
   'She considered to slip into the house some evidence, which would confirm that it was a
   murder.'

d. Od malého veľmi rýchlo anéž napsal.
   'Since childhood he read a lot and had urge also to write something.'
To summarize, the following table lists the tests that correspond to each node on the decision tree in Figure 3 and the feature labels they assign:
Table 5: Diagnostics

3.4 Quantitative results and distribution of *voľ* and *-si* on Haspelmath’s map

The graphs in Figure 4 summarize the quantitative results of the corpus analysis. They show that while the intuitions that both specialized series encode an epistemic effect and mainly serve the [SU] function were borne out, their distribution is not equivalent. The *-si* EI is more narrowly dedicated to the reading of ignorance, with about four out of five occurrences annotated as [SU]. The two other functions with the highest percentages were [IR] of 5% and [E-FC] of 4%. The [DN] and [CO] functions were not represented by this series at all. The *voľ* EI is a slightly more multitasking series, serving the [SU] function about half the time while being quite comfortable in the neighboring regions as well. The second highest percentage after [SU] was the existential [E-FC] function with 15%, followed by [IR], [Q] and surprisingly, [SK] at around 8%.
In order to briefly comment on the presence of vořa- and -si in the functions other than [SU], I will for a moment return to the typological study by Haspelmath (1997). His map in principle offers 95 geometrically possible combinations of functions that any one indefinite series could express. His survey, however, attested only 37 different combinations. A stipulation that 58 unattested combinations are not merely accidental gaps led Haspelmath to propose further constraints on the distribution of indefinites. One of these principles is that it is impossible that the function [(U-)FC] and [CO] be combined with [SK]. As predicted, the [CO] function was not represented in my samples. However, both series were labeled as [U-FC] in a small percentage of cases. This finding of a gap poses a problem to the congruency hypothesis or to the arrangement of the map in general. Alternatively, the insignificant numbers could simply be attributed to incorrect assignment due to only one annotator, or a fluke on the part of the data source.

9 The justification for the constraint is highly informal and speculative: "The principle can probably be explained by the need to distinguish between meanings that are too different - the emphatic functions 'comparative' and 'free choice' are just too different semantically from 'specific-known' to be expressible in the same way" (Haspelmath, 1997:77).

10 The annotation task, even if employing strict diagnostics, relies on interpretation judgments that are not always black and white. In a very similar annotation study of an English corpus of indefinite uses by Aguilar-Guevara et al. (2010), the inter-annotator agreement scores ranged from 52% to 62%.
It is also worth pointing out that corpora as a source of data provide only positive evidence of the most common constructions. We would not be justified in saying that what is not present is ungrammatical. In natural speech, people might use *voľa* in contexts of comparative and *-si* in comparative and direct negation, the functions not attested in the sample. However, these may not be common enough to make it to the sample of 300 occurrences surveyed here. For example, the following artificial examples (not from the SNK) are not fully deviant in the right context:

(32) a. Košice sú krajšie ako akšsi/voľjakaš inš mestš. [CO]
Košice are more-beautiful than some-si/voľa-some other cities
*Košice is more beautiful than some other cities.*

b. Čosi sa jej nepáši. [DN]
what-si PART her NOT.please
*She doesn’t like something.*

Considering the above evidence, the prevalence of ambiguities, the sometimes impossible task of teasing interpretations and functions apart led me to reconsider Haspelmath’s original map. It seems appropriate to discard the idea of crisp boundary lines between the functions and rather represent them as concentrations of different intensities spanning across non-delineated, overlapping as opposed to contrasting, functional regions (see Figure 5 below). The concept of a gradating intensity, as opposed to a monolithic outlined region, further captures the intuition behind adding the word *primarily* to the definition of EIs in (5). The corpus analysis results showed that the indefinites in fact have a primary function in which they are used more frequently than in others. Whether the concentrations tend to bleed out into neighboring functions or are in the middle of converging on just one function is a question of the direction that the grammaticalization of specialized indefinites takes in a language. It will remain an interesting task to be resolved through a diachronic corpus study.

*Figure 5: Concentrations of *voľa* and *-si* on the Haspelmath’s map*
3.5 Qualitative observations

3.5.1 Contexts and environments friendly to voľa- and -si

In this descriptive section I will take a closer look at the data gained from the SNK and lay the groundwork for the later semantic analysis. I will focus on the environments where the ignorance and free choice readings were brought about. I will describe some of the characteristics of the two surveyed Slovak indefinites along the five most significant parameters of variation found in the literature on EIs: (i) degree of ignorance, (ii) type of evidence, (iii) interaction with modals, (iv) interaction with plurality and (v) interpretation in imperative constructions.

The concept of the degree of ignorance will be discussed in more detail in the Analysis Section 4, however, a descriptively relevant difference between voľa- and -si deserves a brief mention. I will introduce the examples and preliminary observations here, without offering explanations. Degree of ignorance is used to describe whether all or only some domain members are possible referents of the indefinite. If it is consistent with what the speaker knows that any and all possibilities in the context could satisfy the existential claim, with no exceptions, such an indefinite will convey total ignorance. In a context where the speaker knows that some domain members are excluded but at least two possible values remain, a partial ignorance indefinite can be used. Where do Slovak EIs fall with respect to this feature? Let us take a look at the following examples:

(33) Niekto mi podával ruku a voľa ho vovoril. Nerozumela som mu.  
Someone was giving me their hand and telling me something. I didn’t understand him.

(34) Dnes sa tu obmietať ako si, vravel, že by sa chcel učiť lietať.  
Today some man was sneaking around, he said that he would like to learn to fly.

In (33), the EI voľa could refer to anything, no restriction, and remain consistent with the speaker’s knowledge. At the same time, however, the context is quite compatible with the speaker’s considering only a subdomain of topics, for example excluding the possibility that the comment was about the weather or about food. Such reading is quite accessible in the corpus sample of the voľa-series and I conclude that it is a preferred-partial ignorance EI (but compatible with total ignorance). In (34), the domain must be completely unrestricted. The use of akoší would be very odd if all men were not possible referents. For example, if the speaker already had a subdomain of only male coworkers in mind:

(34)b. Dnes sa tu obmietať volajaky/akoší muž, jeden z tvojich kolegov a vravel, že by sa chcel učiť lietať.  
Today some-some/some-si man was sneaking around, one of your coworkers, and he said that he would like to learn to fly.'
The additional fact that inserting čosi in (33) results in a preferred total ignorance reading, making the clarifying continuation odd, further supports the intuition that -si is a total ignorance EI:

(33)b. Nieko to podāval ruku a vočľavo/čosi hovoril ...ale nebo to o jedle.  
'Someone was giving me their hand and telling me vočľavo-what/si ...but it wasn’t about food.'

These facts seem to indicate that contexts of total ignorance are friendly to -si and contexts of partial ignorance are friendly to vočľava, at least tendentially.

Another criterion of description of an EI-friendly context is type of evidence. First mentioned by Alonso-Ovalle and Menéndez-Benito (A&M henceforth, 2003) in describing algún and further explored by Aloni and Port (2010), type of evidence captures the idea that an indefinite is sensitive to different methods of identification, for example a direct perceptual access to the individual despite not being able to identify them by name. With the help of the following examples, one adapted from A&M and the other from the SNK sample, I offer an observation that neither vočľava nor -si are odd when the speaker can see and point to the object. This is unlike Spanish algún but the same as English some.

(35) a. Look! Some professor is dancing on the table! (A&M, 2003:4)
   b. #Mira! Algo profesor está bailando encima de la mesa!
   c. Pozri! Akši/Volajaky profesor tancuje na stole!

(36)a. Kočša povedal: 'Volajaky pán ide k nám.' VISUAL ACCESS POSSIBLE
   coachman said vola-some man goes to us
   'Coachman said: "Some man is walking towards us."'
   b. (Pozri!) Ktosi odhna odo dver mačku Micku. VISUAL ACCESS POSSIBLE
   look who-si shoes from door cat Micka(name)
   '(Look!) 'Someone is shooing the kitty cat away from the door.'

However, the following examples suggest that although using the EIs in situations with visual access to the referent is possible, it is not expected:

(37) a. Ktosi na zakerne kopol do boku. NO VISUAL ACCESS ASSUMED
   who-si me viciously kicked into side
   'Someone viciously kicked me in the side.'

(38) a. „Je neskor! Zvolal ktorši z veliteľov. NO VISUAL ACCESS ASSUMED
   is late cried which-SI of commanders
   "It is too late!" cried one of the commanders.'
   ?Bol vysoký.
   was tall
   'He was tall.'
In the above situations, the EI leads the hearer to assume that the speaker did not see the referent; in (37) the speaker only felt the kick and in (38) he/she only heard the cry. It would be odd to describe the individual by their physical characteristics, for example their height. Otherwise, another indefinite variant (equivalent to 'one') would have been chosen by the speaker:

(37) c. Jeden ma zakerne kopol do boku.  
    'Someone viciously kicked me in the side.'  
(38) c. ‘Je neskoro! ’ zvolal jeden z veliteľov.  
    'It is too late!' cried one of commanders.

Let us now look at voľa- and -si’s interaction with modality. Fălăuş (2009) shows that Romanian vreun is not grammatical in positive non-modal contexts nor in the direct scope of deontic modals (those concerned with permissions, obligations, etc.). Aloni and Port (2010) note that Italian un qualche is grammatical both in positive non-modal contexts and under deontic modals. German irgendein is also used under deontic modals (A&M 2010). How about Slovak EIs? In positive non-modal contexts, such as in (33) and (34), they are grammatical and convey epistemic effect. With respect to deontic modals, I offer the following examples to illustrate where voľa- and -si stand:

(39) Aby mohol zmysluplne a tvorivo vypoveda o sebe a o svete, musí v životy vohľa naštudovať, ale aj si nazír.  
    ‘In order to meaningfully and creatively state something about himself and the world, he must study something, but also live it.’

(40) Musíme v tomto smere oškoro čosi urobí,” tvrdí starostka.  
    ‘We must do something in this direction soon,” says the mayor.’

In (39), the speaker says that a person must study something in order to speak meaningfully and creatively. That something may be specific and unknown to the speaker, however, the default reading is something, no matter what - non-specific. In (40), the mayor expresses a need to do something about a pressing problem. She could be referring to a specific thing, however, the reading of a non-specific free choice is again the easier interpretation. To sum up, under the scope of a deontic modal of obligation, both EIs are grammatical; they allow for the epistemic reading but an FC reading is preferred.

Turning our attention to the interaction with plurality, the literature makes a distinction between EIs that trigger an epistemic effect only in the singular form, like Spanish algún, and others that serve this function in both singular and plural, like German irgendein (A&M 2010). We can see in (41) and (42) below
that the epistemic characterization holds for the plural forms of voľa- and -si. For both series, ‘Guess who?’ and ‘Namely’ continuations are infelicitous in plural, signaling ignorance.

(41) Gazda, gazdiná a ešte akši ľudia stáli v pivore a farmer, farmer and also some-si people stood in front-room and hládeli otvorenými dverami dovnútra. #Boli to Marta a Jožo. looked open door inside were it Marta and Joey

‘The farmers and also some people stood in the front room and looked in through the open door. It was Marta and Joey.’

(42) (Sestra v kaplnke:) ‘Ze vraj aj zrak sa tu voľaktorym vrátil.’ sister in chapel supposedly also sight partic here voľa-whom returned

#Konkrátne Marte a Jožovi. concretely Martha and Joey

(A sister in a chapel:) ‘They say that some even regained their sight here. Namely, Martha and Joey.’

Imperative constructions were found in the corpus sample of the voľa-series only. In each of the following examples, a strong FC effect comes about (recall from Table 1 that imperative construction is not conducive to a specific reading):

(43) Poď, dušička, zobni si voľzo, nech neodpadne. come soul peck partic voľa-what so not faint

‘Come, darling, eat something, anything, so that you don’t faint.’

(44) Objednaj voľzo na pitie. order voľa-what for drinking

‘Order something, anything, to drink.’

(45) Pomoc! Pomoc! Držte ho voľaktol! Držte ho!

help help hold him voľa-who hold him

‘Help! Help! Someone, anyone, hold him! Hold him!’

Why is the voľa-series used in imperatives while -si is not? And more generally, why does voľa- induce the indifference no-matter-what reading more often than the -si series, as the statistical results revealed? Two possible factors could be responsible for this distinction, one related to etymology and one to lexical competition and interpretive economy. First, the indefiniteness marker for the voľa-series is morphologically related to such words as voľa (‘will’), voľnosť (‘freedom’), voľba (‘choice’). A documented etymology of this word is not available but it seems apparent that there is a connection to its being more comfortable in the [FC] region on the map. Another plausible factor in this distinction is the possibility that -si, being a total variation EI, competes with specialized morphologically marked FC indefinites. If the hearer has a completely free choice of satisfying the command with any of the domain members, an indefinite from the -kôvek or hoci-series (see Table 3) would express this function directly instead of borrowing from a dedicated EI series like -si. According to Aloni (2010), emphatic FC uses come with a high cost for the
interpreter, who must take an extra step to integrate the encoded ignorance and arrive at the implicature of indifference. Therefore, for the purpose of economy, languages develop specialized morphology to express FC meaning (such as Slovak -koľek or hoci-). Their availability in the relevant contexts blocks the use of EIs, for which the marginal FC function is more difficult to process.

Finally, I would like to offer one more observation. The indefinite series overlap on the map and all share the main purpose of referring with only subtle differences in meaning that are evidently and notoriously difficult to tease apart. Therefore, it is not surprising that **stylistic choices** would play a role in a speaker’s use of a specialized indefinite variant. In the following example (46), the speaker uses three different indefinites in succession, and likely does not intends to convey three different epistemic states about the same state of events.

(46) McMurphy na mňa zmlurkne, pridusene mi **čosi** povie, **volažo** mi McMurphy at me wink stiffly me what-si says VOLA-what me rozpíra, **nieťo** mi hovorí povedla tej gumenej hadice. talks NIE-what me speaks next-to that rubber hose 'McMurphy winks at me, in a stifled voice he says something to me, he talks to me about something, he tells me something next to that hose.'

Obviously, the speaker is stylistically repeating the same information for emphasis and effect. He/She achieves it through using nearly synonymous words, for the indefinites as well as the verbs.

### 3.5.2 Ambiguous and interesting cases

While the procedure of applying the diagnostics was intended to provide an objective way to discern the features of an indefinite in context and eventually its functional label, in many cases it proved to be rather inconclusive on the one hand, or lead to some unexpected results on the other. In this section, I will present some examples that were especially difficult to classify and whose classification was surprising for not fitting the predicted epistemic mold.

First, in both series, the [SU]/[E-FC] functional boundary was the most prone to inducing two possible interpretations. The reason is that a stress placed on the indefinite item can enrich the meaning with a 'no matter what' implicature. Unfortunately, prosodic information is not, and could not be, provided in the SNK (a corpus of written texts), however, this phenomenon still informs us about the capacity of EIs to induce FC effect under some conditions. The following are two examples illustrating this functionally ambiguous behavior. I offer a plausible scenario below that would bring out each reading:

(47) Dohodli sa, že (a) **volaktory**/ (b) **VOLAKTORY** priateľ im bude agreed PART that VOLA-which/VOLA-WHICH friend them will-be sprostredkovať im. intermediary 'They agreed that (a) some+epistemic effect / (b) some+FC effect friend will be their intermediary.'
a. unstressed: the speaker’s ignorance/specific unknown reading -
Your two friends, Silvia and Tom, decided to start a business. They want to export Slovak products to Spain, because there they have a friend, Carlos, who will help them. You don’t know this person and you are hearing about him for the first time. Later you talk about it with your family, who ask you: But how do Silvia and Tom want to sell products in Spain if they live in Slovakia? You say...

b. stressed: the agents’ indifference/non-specific FC reading -
Silvia and Tom cannot agree on whom to hire as their intermediary. They agree that it will be one of their friends but Silvia insists he must have an MBA and Tom would take whoever is willing to take a low salary. After a long debate they finally agree to take on any friend at all, as soon as possible. Later you talk about it to your family, who ask you: And which friend did Silvia and Tom decide to hire, have they agreed on someone already? You say: No....

Likewise, in the -si series the stress seems to play the same role. The statement in (48) is about a woman who must know something particular but unknown about men in (48a) or, under stress, she must know at least something (the bare minimum of things to know), no matter what, in (48b):

(48) Muži majú zvláštné nadanie veriť, že dva a dva môže byť aj päť, aj tri. Nuž ona o tom musí (a) čosi / (b) ČOSI vedieť — vedieť also three well she about that must what-si WHAT-si know after-all veľa videla a veľa pochula. much saw and much heard

'Men have a special talent to believe that two and two may be even five, or three. Well she must know (a) something + epistemic effect / (b) something + FC effect about it because she has seen and heard a lot.'

At this point, I offer the above as mere intuitions that would require an empirical study to confirm whether there really is an intonational focus + FC effect relationship. The interaction of prosody and an indefinite’s function is developed in Fălaş (2013) and cannot be pursued in more detail in this thesis.

Another point of ambiguity that arose during the corpus sample annotation was brought about by what could be called a flavor of epistemicity. It is the interference of a weak epistemic effect that the vola- and -si items contribute even if they occur in non-[SU] functions. This can be quite easily seen in interrogative constructions that received the label [Q]. In the following example (49), (a) is taken from the corpus sample and (b-d) would be viable alternatives available to the speaker who wants to express the core meaning of ‘something’. Their approximate interpretations are provided and as can be seen, the EI adds its distinct ignorance flavor to the mix:
(49) a. Ani neuvážujete o ĉemi podobnom?
   even not. consider about what-ĉi similar
   'You don’t even consider something similar?'
   - something the speaker cannot identify exactly and possibly even
     assumes that the hearer might not be able to do so either.

b. Ani neuvážujete o nieţom podobnom?
   - something specific and known to the hearer

c. Ani neuvážujete o niţom podobnom?
   - anything at all, negative expectation

d. Ani neuvážujete o ćomkolvek podobnom?
   - anything, non-specific, no matter what

The following are two more examples where the indefinite serves a non-[SU]
function, in particular [Q], but a similar epistemic interference can be observed:

(50) Aký zver, aký netvor mohol ĉosi taká napáchať?
   what animal what monster could what-ĉi like-that inflict
   'What monster could do something(+speaker’s ignorance) like that?'

(51) Ale nezdá sa vam voľajky smutný?
   but not. seem part you voľa-some sad
   'But doesn’t he seem to you somewhat(+speaker’s ignorance) sad?'

Let us now turn to the curious cases in which Slovak EIs apparently undergo
an identity crisis. The first definition of EIs introduced in Section 2.2 predicts
that they would not be possible in the function [SK] (Aloni and Port 2010).
However, exactly that happens in 8% of the data for the voľa-series and in 3% for the -si series. Below I present typical examples of this phenomenon and offer
three stipulations for why EIs are appropriate for such an unusual function. In
(52), the indefinite refers to something that the speaker would reasonably be able
to identify - the very item that they want to ask the cleaning staff:

(52) Chem prenajal voľajka/akoľ auto.
(i) 'I want to rent VOĻA-some/some-ĉI car.' [FC] preferred
   In this sentence (not from SNK), the speaker who is also the agent would be satisfied with
   any available car. Non-specific reading.
(ii) Jana chcela prenajal voľajka/akoľ auto.
   'Jane wanted to rent VOĻA-some/some-ĉI car.' [SU] preferred unless stressed
   In this sentence, the preferred reading is that the speaker is ignorant about the specific car
   Jane wants to rent.

Examples such as these show that the relationship between the functional reading of the EI
and its sensitivity to the attitude verb is not very clear at this point.

---

11 Some EIs in other languages, e.g. Romanian, are not licensed in the scope of the verb want.
This is because want (as opposed to other attitude verbs like hope) lacks an epistemic
component and is compatible with prior knowledge of the embedded proposition (Scheffler
2008, Fălăuş 2010). As seen in (i-ii) below, the Slovak EIs are grammatical in combination
with the verb want and can exhibit [SK] function as above, but possibly also [FC] and [SU]
functions.

(i) Chem prenajal voľajka/akoľ auto.
   'I want to rent VOĻA-some/some-ĉI car.' [FC] preferred

(ii) Jana chcela prenajal voľajka/akoľ auto.
   'Jane wanted to rent VOĻA-some/some-ĉI car.' [SU] preferred unless stressed

Examples such as these show that the relationship between the functional reading of the EI
and its sensitivity to the attitude verb is not very clear at this point.
‘Vrátili sa už ľudia, ktorí tu upratovali dnes ráno? Chcem sa ich náhle opýtať.’

‘Have the people who had cleaned here this morning returned? I want to quickly ask them something.’

A question ‘What?’ from the hearer would be pragmatically odd in this dialogue even though the speaker has the information. By the choice of an EI instead of a plain indefinite as well as adding the adverb quickly, they signal that they do not wish to elaborate or that more information is irrelevant/of no concern to the hearer at the moment.

In the next two examples, something other than irrelevance seems to be at play. Again, the speaker can reasonably be expected to identify the referent of the indefinite. In (53) it is the observation immediately following that people in the sign of Libra behave a certain way, and in (54) it is the sanctions on money laundering cases mentioned right before.

(53) Všímom si čo tebe ušlo. Ľudia v znamení Váhy nikdy nepostrehnú odľahčení na seba.

‘I noticed something that you missed. People in the sign of Libra always fail to notice references to them.’

(54) ... na ich (pripady prania spinávých peňazí) postih nemohli byť pripravené ani štátna správa, ani jej pracovníci. Už počujem množstvo argumentov, prečo voláťo takéto nie je možné.

‘...nor the government nor its staff could be properly prepared for their (cases of money laundering) sanctions. I can already hear much arguments as to why something like this is not possible.’

Why does the speaker choose to use the specialized indefinites in these two situations? Although not ignorant themselves, it is possible that they express empathy with the ignorance of someone else, e.g. the hearer (see point (i) in the Discussion Section 5), and so act out of a conversational cooperation. In the example (53) the hearer had failed to notice something and the speaker seems to be sensitive to this in his choice of an indefinite with an epistemic flavor, as if putting themselves in the hearer’s shoes. In (54), the speaker is knowledgeable about the topic on which he is speaking, namely the government’s preparedness for sanctions on money laundering cases. However, he is speaking about a third, ignorant, party that is raising objections and chooses to use an indefinite volaťo for what seem to be empathetic reasons.
In the final selection of examples, I present probably the most typical reason for when an EI is used despite the knowledge of the speaker:

(55) „Aha. Ten muže sa bude uchádzať u vášho strýka o miesto. Aha that man part will apply at your uncle about post
Volajte vám poradím, sľúdna, ale nesmiete ma vyzraditi. Povedzte vola-what you advise miss but not can me blurt-out tell strýkovi, aby nenajmal Texa." uncle so-that not hire Tex
'Look. That guy will ask your uncle for a job. I will advise you something, miss, but you must not tell anyone. Tell your uncle not to hire Tex.'

(56) Stalo sa však čosi pre ma neuveriteľne: nový občiansky happened part however what-si for me unbelievable new citizen preukáz vystavili na príslušnom oddelení polície na počkani! card issued on appropriate department police on waiting Pri všetkých tých nervákov čosi potešujúce. next-to all those nerves what-SI pleasing
'Something incredible happened to me: they issued a new ID card at the police department while I waited! For all those nerves, something that made me glad.'

In the above, the speakers immediately proceed to give more information on the referent. In (55) the speaker first gives a hint on what will be said next - an advice about something. He/she then dispenses the advice, namely, not to hire Tex. In (56), the speaker gives a hint about the something that happened to him - it was incredible. He then goes on to describe it further. In both sentences, it seems that the speaker is trying to create a certain sense of suspense or alert the hearer to pay attention to what they are about to say next.

To sum up, there are at least three possible pragmatic explanations for the marginal [SK] function of EIs: irrelevance in the conversation, empathy with the hearer, and the intention to obscure the knowledge at the particular moment of utterance. In a conversation, speakers make choices between various indefinites with the aim to capture an intended fine shade of meaning of 'some thing I can identify but do not want you to ask me to do so at this moment’. A Slovak system of indefinites has the complexity to allow this kind of precision, or its very close approximation. Although this exceptional behavior is curious, the proportion of [SK] uses was very low in my corpus sample and each case still arguably retained the flavor of epistemicity in some more or less obvious fashion. Therefore, I assume that it does not place the identification of vola and -si as EI items in jeopardy.
3.6 Conclusion of the section

In this section, I focused on describing the distribution, interpretation and pragmatic and syntactic environments friendly to the Slovak EIs, drawing from authentic examples found in the Slovak National Corpus. Several findings that have emerged from this survey can be highlighted.

The two series were shown to have some distinct characteristics with regards to interpretation. While the voľa- series is consistent with both partial and total variation EI (33), -si is a total variation EI (34). They are similar, however, in that they both suggest no visual access to the referent, but are compatible with it (35-38). Under the scope of deontic modals, the FC interpretation becomes preferred (40). They both retain the [SU] reading in positive non-modal contexts and in plural form (41-42). The imperative construction was found only for the voľa- series and shown to carry a strong FC interpretation (43-45).

Ambiguity, the cause of which was attributed to a prosodic emphasis, was found mostly in the region between [SU] and [E-FC] function, see examples (47-48). Both series seem to contribute a special layer of epistemic meaning to interrogative constructions, as contrasted to other alternative indefinite variants used in the same utterance (49-51). Counter to intuition, they were both found meddling in the [SK] function, see examples (52-56).

It is important to emphasize that the observations offered in this section were based on data gained only from a written-text corpus. Haspelmath’s (1997) hypothesis of contingency on a functional map, which provided the foundation for my study, was based on data from grammar reference books. These sources are prescriptive in nature and arguably limited in usefulness for describing how native speakers use any given language. Furthermore, the present study only represents written and not spoken language. It is possible that epistemic indefinites are used differently in the spoken register and that the distribution across functions would look differently if an analysis of a spoken corpus were conducted. Had I chosen a different source of data, the ambiguities and the difficulties of identifying the correct interpretations might not have arisen. Furthermore, supplementing my intuitions with native speaker acceptability judgments, carefully probed in contexts, would likely illuminate the instances where the functional classification was nearly impossible. Future research in this vein will hopefully enrich the picture given in this work.

4 Analysis

Having scratched the surface of the task of describing the behavior of Slovak EIs in natural speech and locating them on Haspelmath’s functional map, I now turn to a formal analysis of their semantics. In this section, I will present a theoretical perspective on EIs proposed by Chierchia (2013) and show how the two newly introduced Slovak series fit within it. I have selected this paradigm in part because it is a fairly recent one but becoming quite prominent in the field.
It builds on many insights from such work as Kadmon and Landman (1993), Krifka (1995), Kratzer & Shimoyama (2002), Chierchia, Fox & Spector (2012) among others. Its attraction lies in that it is quite simple in its basic concept but addresses a wide range of facts across languages attested in the literature, including the dimensions of variation discussed in the previous section. Its main "philosophical" thesis is that the underlying logical structure of language is tightly interfaced with our natural and spontaneous ability to run computations and draw inferences. The main "linguistic" thesis is that there is a unified explanation for the complete repertoire of the polarity phenomena, to which the items of interest in this thesis, EIs, are assumed to belong.

The idea of a radical unity of the polarity system is appealing for its promise of simplicity and elegance, however, this theory also comprises a rather sophisticated system of constraints and binary parameters. In the following paragraphs, I provide a synopsis of some of the complexity. I limit the discussion to those claims of the theory that have a direct role in understanding EIs and specifically the characteristics introduced in the previous section. I begin by showing the motivation for the claim of interconnectedness of polarity items and explaining the basic concepts of activated alternatives and their exhaustification. The introductory section 4.1 serves as a foundation for the subsequent application of the analysis to the Slovak data in 4.2. Finally, the last section 4.3 is devoted to some open questions and unresolved issues.

4.1 The Alternatives-and-Exhaustification Paradigm

This approach embeds EIs within a wider framework of polarity and FC phenomena. Chierchia argues that all of the polarity sensitive items (such as universal FC items: English any, Italian qualunque, existential FC items: German irgendein, weak NPIs: English ever; strong NPIs: English in weeks; emphatic NPIs and minimizers: lift a finger; and N-words: Italian nessuno) are in fact based on one indefinite form with an identical logical import. As its name implies, the cornerstone characteristics of this theory have to do with the notions of alternatives and of exhaustification. In this section, I will discuss them in turn.

Alternatives are taken into account during the interpretation of a sentence that contains an indefinite item. To illustrate this intuition, consider the example (57). Through a disjunctive phrasing, the hearer is offered a choice of two desserts. Logically, upon hearing this offer, he/she is entertaining the following alternatives: he/she could have one of the desserts, he/she could have the other one (these comprise the domain alternatives), or he/she could have them both (this is the scalar alternative). Assuming cooperation on the part of the speaker, Gricean reasoning12 would lead the hearer to arrive at the interpretation that

12 The FC interpretation does not result from standard Gricean reasoning. The domain alternatives must be 'pre-exhaustified' with the addition of 'only' to each one, e.g. {only ice cream, only cake}, as explained later in the present section. See Fox (2007, Ch.4) and Chierchia (2013, Ch.2) for a more detailed discussion of (neo-)Gricean accounts of how the FC and scalar implicatures come about and why grammar, not relevance, defines the domain alternatives.
he/she could have ice cream, he also could have cake, but he may not have both. The FC inference turned the original disjunction into a conjunction-like reading:

\[(57)\]
\[a. \text{You may have ice cream or cake.} \]
\[b. \text{assertion: } \diamond ( \text{have ice cream } \lor \text{have cake}) \]
\[c. \text{domain alternatives: } \{ \diamond \text{have ice cream}, \diamond \text{have cake}\} \]
\[d. \text{scalar alternatives: } \{ \diamond \text{(have ice cream } \land \diamond \text{have cake })\} \]
\[e. \text{FC interpretation: } \diamond \text{have ice cream } \land \diamond \text{have cake } \land \neg \diamond \text{have both} \]

\text{Key: the symbol } \diamond \text{ is used to express a modal of possibility.} 

The following example (58) differs from the above in that it contains an indefinite in place of the disjunction in (57). However, we arrive at an identical FC interpretation. Both (57) and (58) assert a disjunction (under the scope of a modal), but by drawing an inference, they turn out to convey a conjunction (with scope over the modal). Let us assume that the sentence is uttered in a restaurant. The hearer, a customer, is reading the dessert menu and is probably fortunate to have more available options to choose from: ice cream, cake, muffin, pudding, etc. Having any one of them would comply with the permission:

\[(58)\]
\[a. \text{You may have a/any dessert.} \]
\[b. \text{assertion: } \diamond ( \text{have ice cream } \lor \text{have cake } \lor \text{have muffin } \lor \text{etc.) } \]
\[c. \text{interpretation:} \]
\[\text{free choice implicature: } \diamond \text{have ice cream } \land \diamond \text{have cake } \land \]
\[\diamond \text{have muffin, etc.} \]
\[\text{scalar implicature: } \land \neg \diamond \text{any two or more at the same time} \]

The English \textit{a} and \textit{any} yield an FC reading just like disjunction. Moreover, without further context, the sentence in (58) is clearly identical in meaning whether \textit{a} or \textit{any} are used. However, Chierchia claims that these two indefinites have at one point divorced by \textit{any} taking an adjectival suffix. Along with the morphological change of \textit{any} came the specialized meaning and distinct truth conditions i.e. a grammaticalized inference, leading to its more constrained distribution. The crucial distinction between them is that the plain indefinite \textit{a/an} is able to activate the subdomain alternatives when motivated by context or other pragmatic reasons. On the other hand, a specialized and morphologically marked indefinite \textit{any} is doing it obligatorily, due to its lexical semantic content. The item \textit{any} requires that the domain alternatives be considered by the hearer and factored into the meaning. In support of the fact that the ordinary scalar indefinites \textit{a/an} bring on the alternatives and the FC implicature optionally is that they can be easily called off with a \textit{namely} continuation, as in (59a). The same continuation after the item \textit{any} in (59b) is odd:

\[(59)\]
\[a. \text{You may have a dessert, namely this cheesecake.} \]
\[b. \#\text{You may have any dessert, namely this cheesecake.} \]
Other indefinites that have a grammaticalized inference, such as for example EIs, also obligatorily activate alternatives. And once the alternatives are activated, they must be resolved, i.e. factored into the interpretation. This is done through the process of **exhaustification**. This term describes an inference calculation, much like Gricean reasoning, through which the hearer arrives at the enriched, informationally strengthened interpretation that only the stated assertion is *exhaustively* true. All the non-entailed alternatives must as a consequence be false. The exhaustification is triggered and brought about by a covert syntactic device, an **exhaustification operator** $O$.\(^{13}\) It is akin to overt *only* and in complementary distribution to it. $O$ is an operator that combines with a set $C$ of propositional alternatives and a proposition $p$. Its semantics, formally shown in (60) below, says that $p$ and all its entailments from the set $C$ are the only true members of the set $C$.

\[(60)\]  
\[O_C(p) = p \land \forall q \in C [q \rightarrow p \subseteq q],\]
where $p \subseteq q$ means $p$ entails $q$.

Importantly, the FC implicature comes from the assumption, due to Fox (2007), that the domain alternatives are considered individually. They are pre-exhaustified obtained by applying the operator $O$ recursively: first to each alternative, and then to the whole assertion again. This means that for example the domain alternative 'have ice cream' in (57c) turns into the proposition 'have *only* ice cream', or in other words, 'have ice cream and not have cake'.

Let me illustrate how such a computation works on an example. In our offer of dessert in a restaurant (58), the interpretation we would like to derive is that all the available desserts, i.e. all the domain members, are equally valid choices for the hearer. If we assume a situation in which the dessert menu includes only two items for the sake of simplicity, the calculation would go as follows:

\[(61)\]  
\[a. \text{ You may have any dessert.} \]
\[b. \text{ LF: } O \circ \exists x \in D \ [\text{one } +SC,+D]^{14} (x) \land \text{dessert } (x) \land \text{have } (you, x)]\]
\[\text{where } D = \{\text{ice cream, cake}\}\]
\[c. \text{ ASSERTION: } \Diamond (p \lor q)\]
\[\text{ KEY: } \Diamond = \text{you may, } p = \text{have ice cream, } q = \text{have cake}\]
\[d. \text{ PRE-EXHAUSTIFIED DOMAIN ALTERNATIVES: } \{O \circ p = p \land \lnot q,\]
\[O \circ q = q \land \lnot p\}\]
\[\text{(in prose, 'you may only have an ice cream and nothing else, you may only have a cake and nothing else')}\]

\(^{13}\) The origins of the idea of exhaustification can be traced to Groenendijk and Stokhof’s (1984) seminal work on the semantics of questions and answers, and their notion of strong exhaustiveness. They in turn expanded on the alternatives semantics work on interrogatives by Hamblin (1973) and Karttunen (1976).

\(^{14}\) Chierchia (2013) posits a feature checking approach to the (in-)activeness of scalar (SC) and domain (D) alternatives. The feature set has a default '-' setting. The operator $O$ assigns a value '+' to either or both features, depending on its restrictions. The EIs under discussion in this work have always-active alternatives and for the sake of clarity, I will omit this element in the LF representations. The reader can assume it is always there.
This is the signature property of the FC phenomena, which plays an important role in the epistemic effect as well (argued to be a speaker-oriented free choiceness). Through a process of recursive exhaustification, the logical structure of a modal disjunction results in a distributed modal conjunction:

\[(62) \diamond (a \lor b \lor c \lor \ldots) \rightarrow \diamond a \land \diamond b \land \diamond c \land \ldots\]

One of the most important consequences of this analysis is that a logical contradiction comes about by exhaustifying in contexts where the activated alternatives are entailed by the assertion. Chierchia and others who subscribe to the thesis of a tight interrelationship between logic and grammar by Gajewski (2002), take this to point to the reasons why such contexts are ungrammatical for a particular polarity or FC item.\(^\text{15}\) To illustrate, *any* in episodic contexts is unconditionally false in an affirmative statement (63a) and rescued when under negation (64a):

\[(63)\]
a. I had *any* dessert.

b. **Assertion**: have ice cream \lor have cake

\[O \exists x \in D \{\text{dessert}(x) \land \text{had}(I, x)\}, \text{ where } D = \{\text{ice cream, cake}\}\]

c. **Domain alternatives**: I had ice cream, I had cake

d. **Scalar alternatives**: I had ice cream \land I had cake

e. **Exhaustification**:

\[\exists x \in \{\text{ice cream, cake}\}[\text{dessert}(x) \land \text{had}(I, x)] \land
\neg \exists x \in \{\text{ice cream}\}[\text{dessert}(x) \land \text{had}(I, x)] \land
\neg \exists x \in \{\text{cake}\}[\text{dessert}(x) \land \text{had}(I, x)] = \bot\]

\[(64)\]
a. I didn’t have *any* dessert.

b. **Assertion**: \(\neg\)have ice cream \lor have cake

\[O \neg \exists x \in D \{\text{dessert}(x) \land \text{had}(I, x)\}, \text{ where } D = \{\text{ice cream, cake}\}\]

c. **Domain alternatives**:

{I didn’t have ice cream, I didn’t have cake}

d. **Scalar alternatives**: I didn’t have ice cream \land I didn’t have cake

e. **Exhaustification**: vacuous because all the alternatives are entailed

\[\neg \exists x \in \{\text{ice cream, cake}\}[\text{dessert}(x) \land \text{had}(I, x)] = \textbf{the original assertion}\]

---

\(^{15}\) See Chierchia (2013, Ch.1, Section 3) for a summary of the argumentation and differences between various sources of ungrammaticality.
The alternatives in (63c,d) and (64c,d) introduced by any must be factored into the meanings of the assertions. This is done through the presence of the exhaustification operator O in the structure (63b and 64b). As the definition (60) states, the application of the operator O makes all the non-entailed alternatives false, i.e. in the logical calculation, O negates these alternatives (shown in 63e and 64e). Therefore, once we apply O to the assertion in (63a), the result is contradictory: 'I had a dessert from the set \{ice cream, cake\}, but I did not have ice cream and I did not have cake'. This sentence is contradictory and ungrammatical. When applying the operator O to (64a), the alternatives are found to be entailed and as such cannot be eliminated. The calculation does not have any effect and returns the original assertion. The ungrammaticality was avoided thanks to negation, a downward entailing operator.

Examples such as these make a strong argument that logic and analyticity play a central role in grammar and even determine whether a sentence is grammatical or not. Relevant to this thesis, logical inference calculation likewise reveals whether an EI can or cannot be felicitously used in a given context. Following closely Chapter 5 of Chierchia (2013), I will now turn to the discussion of existential free choice items, to which EIs are said to belong.

Under Chierchia’s theory, EIs are a particular manifestation of free choiceness. Namely, they are an instance of morphologically marked existential-FC\textsuperscript{16} items, which include constructions such as Italian un \textit{N} qualsiasi (where \textit{N} stands for 'noun') or German \textit{irgendein N}. There is a parallel between their requirement of active domain and scalar alternatives, a requirement of a variation among possibly satisfying values (neither can be followed by 'Namely') and a parallel between how the FC and the EI computations arrive at a modal conjunction implicature. The EIs are, in other words, interpreted just like FCIs, except that they convey that as far as the speaker’s knowledge is concerned, all (or at least two, depending on further restrictions) domain members are possible options for the referent’s identity. To reiterate, the basic assumption is that EIs are existentials that have identical properties to FCIs and to ordinary indefinites embedded in FC environments (e.g. imperatives: \textit{pick a card}) and so can be integrated under the same analysis. Let us consider an example with an English EI item, singular some. The reading of (65) that we are interested in is a specific unknown reading, in which the speaker knows that Susan is marrying a particular doctor, but he/she does not know which one:

\begin{align*}
(65) \text{a. Susan must marry some doctor.} \\
\text{b. LF: } & O \quad \exists x \in D \left[ \text{one (x) } \land \text{ doctor (x) } \land \text{ marry (Susan, x)} \right] \\
& \text{where } D = \{ \text{Dr. House, Dr. Jekyll} \}
\end{align*}

\textsuperscript{16} In Chierchia, EIs fall under the umbrella of E-FC. This differs from the Haspelmath map used in the previous section, in which EIs and E-FCs were two distinct labels, the former for a specific unknown ignorance reading and the latter for a non-specific indifference reading. From this point on, I adopt Chierchia’s assumption that their logical import and treatment is identical (some problems with this assumption are discussed in Section 5). Whenever the specific/non-specific reading distinction becomes significant, I will use the term 'indifference' in lieu of FC to avoid confusion.
(for the sake of simplicity assuming a small world with only two doctors)

c. assertion: □ (p ∨ q)

key: □ = must, p = marry Dr. House, q = marry Dr. Jekyll

d. pre-exhaustified domain alternatives:

{O□p = □p ∧ ◊q, 
O□q = □q ∧ ◊p}

e. scalar alternatives: \{□(p ∧ q)\}

f. interpretation:

\[ □(p ∨ q) ∧ ◊(□p ∧◊q) ∧ ◊(□q ∧◊p) ∧ ◊□(p ∧ q) = □(p ∨ q) ∧ ◊p ↔ ◊q ∧ ◊□(p ∧ q) = ◊p ∧ ◊q ∧ ◊□(p ∧ q) \]

In sum, when we form and parse sentences, we factor in contextual or lexically encoded alternatives; optionally so with plain indefinite terms but obligatorily with EIs and other specialized items. However, the picture is not that simple, otherwise there would be only one type of an EI and that is clearly not the case. Chierchia argues that the variation among them (and among indefinites in general) is a result of each item’s grammaticalized sensitivity to a small set of parameters. This set of parameters explains the existence and behavior of different types of EIs and can be summarized along two dimensions: (i) the indefinite’s interaction with different modes of exhaustification, and (ii) the types of alternatives that get factored in. In the following paragraphs, I will briefly comment on the kind of cross-linguistic variation exhibited by EIs, and how the above parameters capture their basic properties.

The dimension of mode or strength of exhaustification concerns the item’s parallel acceptability as a negative polarity item (NPI). For instance, German *irgendein* loses its FC reading in negative contexts and acts as an NPI (66), but Italian *un N qualsiasi* does not (67).

(66) Niemand konnte *irgendein* Buch lesen.  
noone could *IRGEND* book read  (German)  
'Noone could read any book.'

(67) * Nessuno ha potuto leggere *un* libro *qualsiasi*. (NPI reading banned)  
Nobody could-
PERF read a book *QUALSIAI* (Italian)  
(Chierchia, 2013, Ch.5)

Chierchia attributes this behavior to a distinction between modes of exhaustification. *Irgendein* is sensitive to a weak exhaustification, whereas *un N qualsiasi*, a ‘pure’ FCI, is sensitive to a strong exhaustification with ‘proper strengthening (PS)’ (the operator O_PS requires that the result of the calculation properly entails the original unexhaustified assertion). In downward entailing contexts the assertion entails its alternatives, therefore, those EIs that are sensitive to the strong exhaustifying operator O_PS, are prevented from serving as NPIs (See Chierchia, 2013, Ch.5 for more detailed argumentation).

17 Such sensitivity can be reflected in the item’s morphology, e.g. through a series-creating affixation (see Table 3) that may encode specialized epistemic or FC semantic content.
Another relevant distinction between EIs has to do with the range of variation/uncertainty across the domain of quantification, tied to the size of an item’s exhaustifiable alternatives that get factored in. Chierchia (2013, Ch.5) illustrates this with two Italian EIs, *uno qualsiasi* and *un qualche*. Both convey uncertainty on the part of the speaker that is normally associated with the epistemic effect, and thus the infelicity of the ‘namely’ continuation in (68).

\[(68) \text{?? Gianni ha portato un libro qualsiasi / un qualche libro vale a dire Meaning and Grammar.} \]
\[(68) \text{'Gianni has brought some book, namely Meaning and Grammar.'} \]

However, these two items do not convey the same type of ignorance. According to Chierchia, the variation requirements associated with the *uno qualsiasi* series are stronger than those associated with *un qualche*. This can be shown in a hide-and-seek scenario that has been proposed as a good way to distinguish the degrees of variation (A&M 2010). In a simple set up, we suspect that John is hidden in a room of a house, but we know that he cannot be in the master bedroom (because it is always locked). *Un qualche* can be used in this situation, but *un qualsiasi* cannot:

\[(69) \text{Gianni puo essere in una stanza qualsiasi / una qualche stanza di quella casa.} \]
\[(69) \text{'Gianni may be in any room in that house.'} \]

EIs like *un qualsiasi* are infelicitous in a hide-and-seek scenario with a house with one room locked because they require that all alternatives be acceptable referent candidates, without exceptions. As total ignorance EIs, their lexical constraint is to activate and exhaustify alternatives that include all the subsets of the domain. On the other hand, partial ignorance EIs like *un qualche* are felicitous in the hide-and-seek scenario because they allow the speaker to consider some, but not necessarily all the alternatives (rooms in the house) as possible candidates (hiding spots). This reading derives from their lexical choice to activate only ‘small’ domain alternatives, otherwise called ‘singleton alternatives’, a conclusion reached by A&M (2010) and taken up in more detail in the next section. Finally, there is also a distinction between partial variation EIs that are compatible with total variation context, like German *irgendein* (Aloni and Port 2010) and Spanish *algún* (A&M 2010), and EIs that are anti-total variation, represented by Romanian *vreun* (Fălaş 2009). In all cases, the type of lexically activated alternatives is assumed to be responsible for the variation behavior.

Table 6 below summarizes the main dimensions of variability discussed above and shows where some EIs in selected languages fall with regards to the various
parameters proposed by the theory. The intention of the table is not to set up a taxonomy but rather to illustrate that a set of switches (lexical choices between different modes of exhaustification and types of alternatives) determine the item’s behavior. Although to explore the nuances of this proposed generative matrix would take us too far afield and the reader is encouraged to turn to Chierchia (2013) for a more detailed explanation, the overview of the matrix I have given is relevant to this brief introduction to the theory. First, it illustrates how the proposed analysis accommodates the cross-linguistic uniformity and diversity of the polarity system via the available combinations of parameters. Second, it offers a concrete goal in my present undertaking. Assuming this uniform approach, in the following section I will categorize the two Slovak EI series with respect to the table, describe the underlying parameters that determine their behavior and demonstrate the inference calculation.

<table>
<thead>
<tr>
<th>Types of alternatives</th>
<th>Modes of exhaustification</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Only truth conditions count for exhaustification</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>i. Truth conditions + implicatures + presuppositions count for exhaustification (O*)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. Proper strengthening (OPS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree-alternatives</td>
<td>E, koli bhii, ek bhii, give a hoot</td>
<td></td>
</tr>
<tr>
<td>(Emphatic NPIs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>E*, sleep a wink</td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>O, Ever</td>
<td></td>
</tr>
<tr>
<td>D-alternatives and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ-alternatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Pure NPIs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O, Mai</td>
<td></td>
</tr>
<tr>
<td>Exhausted</td>
<td>O, irgendein, un quasiassi NP</td>
<td></td>
</tr>
<tr>
<td>D-alternatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Total variation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>³-FCI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhausted Singleton</td>
<td>O, irgendein, un quasiassi NP</td>
<td></td>
</tr>
<tr>
<td>D-alternatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Partial variation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>³-FCI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhausted</td>
<td>O, any</td>
<td></td>
</tr>
<tr>
<td>D-alternatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ Reduced scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(γ-FCI)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Some comments on notation: Question marks denote that such items have not been attested yet. The capital E is an ‘even’-like (as opposed to ‘only’-like) operator, not relevant to EIs. The various super-/subscripts are simply conventions - the basic exhaustification mechanism of exclusion of non-entailed alternatives is preserved.

Table 6: ‘Generative matrix’ - Parameters and their effect on the types of polarity items. Chierchia (2013), Chapter 5, highlighting mine.
4.2 Application to Slovak

It is now time to delve into the semantics of voľa- and -si in light of the alternatives-and-exhaustification theory. I will focus on the two dimensions of variation described above in order to locate the EIs in Chierchia’s polarity matrix: constraints on the types of alternatives they activate (and the resulting range of freedom), and on the mode of exhaustification (compatibility with the NPI function). I will then demonstrate how the framework captures their behavior as described in Section 3.5.

First, let us begin with a very difficult to discern range of variation (total vs. partial). The hide-and-seek scenario unfortunately proves to be an inconclusive test for the Slovak EIs. Recall that the scenario involves John hiding in a house where one of the rooms is known to be locked and therefore is out of the alternative pool. When describing this state of affairs, in Slovak we need to use the indefinite built on the 'which' stem (70b), instead of 'what kind' (70a) that would normally translate into *some*. However, an intuition that I share with a few consulted native speakers dictates that a plain indefinite ([NeIE]+which) sounds the most natural in this situation. If the domain of \{x: x is a room in this house\} is explicitly stated, the voľa-series becomes slightly more felicitous (70c).

(70) a. Ján je v(o) akejsi volajakej nejakej izbe.
    John is in some-VO/A who say some room
    'John is in some room.'

b. Ján je v(o) ktorejsi volaktorej niektorej izbe.
    John is in which-VO/A which some room

c. Ján je v(o) ktorejsi volaktorej izbe v tomto dome.
    John is in which-VO/A which room in this house

Without conducting a larger native speaker interpretation survey, we cannot conclude that the hide-and-seek scenario test straightforwardly reveals any obvious distinction between the series. However, in the descriptive section 3.5, I offered two corpus examples that led me to the tentative observation that voľa- is an EI with a preferred partial ignorance interpretation but also compatible with total ignorance. In contrast, -si prefers a total ignorance interpretation. Furthermore, -si requires that its domain of alternatives be as wide as possible and unrestricted by context. The illustrative data are repeated here:

(71) Nieko mi podával ruku a volalo hovoril. Nerozumela som mu.
    Nie-who me give hand and vola-say him not understand
    Someone was giving me their hand and telling me something. I didn’t
    understand him.’

-where the reference of volalo could be satisfied even if some of the alternatives in D={x: x is a thing to say} are for some reason excluded
Dnes sa tu obáivali a kysí muž, vzravel, že by sa chcel ucíť lieta.

Learn fly

*Today some man was sneaking around, he said that he would like to learn to fly.*

-where preferred reading is that the reference of a kysí man could be satisfied with any of the alternative candidates in D= \{x: x is a man\}

...?? It was (not) one of your coworkers.

The task ahead is to derive this distinction from their semantic content, i.e.

derive (preferred) total and (preferred) partial variation from the differences in the sets of alternatives the EI items activate. I will take the second sentence, as the indefinite + NP makes the distinction a little clearer, and closely follow the insights of A&L (2010) as well as Falauş (2009), which state that the partial variation effect is derived from an alternative set restricted to singletons:

(73) a. SIMPLIFIED LF: sa tu obáivali (i) a kysí / (ii) volajaky

\[\exists x \in D \exists x (x) \land \exists x (x) \land \text{sneak around} (x)\]

where D=\{x: x is a man\}

for the sake of simplicity, we assume that there are only three men in the domain, \{John, Peter, Luke\}

b. INTERPRETATION:

Some man was sneaking around, and as far as the speaker knows it could have been (i) any man at all (complete ignorance), or (ii) any man from a limited set of possibilities, e.g. it couldn’t have been your male coworkers because you know they were at a conference (partial ignorance\(^{19}\)).

C. PRE-EXHAUSTIFIED DOMAIN ALTERNATIVES

(i). NON-MINIMAL: \{O\_s (p ∨ q), O\_s (p ∨ r), O\_s (r ∨ p)\}

\[\{\_s (p ∨ q) \land \_s (p ∨ r) \land \_s (r ∨ p)\}\]

(ii). MINIMAL (SINGLETONS): \{O\_s p , O\_s q, O\_s r\}

\[\{\_s p \land \_s q \land \_s r\}\]

key: p=John was sneaking around, q = Peter was sneaking around,

r = Luke was sneaking around

\[\_s = \text{as far as the speaker knows, it is necessary that...}\]

---

18 I place an epistemic modal in the LF, although it does not overtly appear in the sentence. For the motivation behind this and more details, see section 4.2.2.

19 While it seems here that there is a strict distinction, the two series are contrasted like this only to demonstrate how the derivations of total and partial readings differ. Remember that the voľa- is not excluded from complete ignorance use.
Along the dimension of types of alternatives (the rows on the Table 6), this property places the -si series alongside the indefinites irgendein, un qualsiasi NP, un NP qualsiasi, un oarecare, un quelconque, and the voľa- series alongside irgendein, álga n, vreum and un qualche. However, it bears restating that the total vs. partial variation distinction between the two series is not as apparent as it would need to be for a confident placement. Chierchia (2013, Ch.5) alludes to the possibility that the total and partial variations come in further degrees and this could well be the case. The blurriness of the readings they give could also be due to, as mentioned earlier, the fact that we are looking at two items in the flux of a grammaticalization process. They might become stabilized on one or the other type of variation over time. Perhaps the horizontal dimension - the mode of exhaustification - will be more illuminating. Let us now turn our attention to it.

According to Chierchia, some E-FC items lose the ignorance and indifference effect despite the dedicated morphology, when they are under negation and generally in downward entailing (DE) contexts (such as in scope of doubt, in if-antecedents, etc.). This loss may lead to reverting back to a plain NPI-like interpretation (e.g. German irgendein), attributed to a weak mode of exhaustification, or ungrammaticality/unacceptability of the NPI reading (e.g. Italian un N qualsiasi, cf. example (67)), attributed to a strong mode of exhaustification in the presence of an operator O_Ps. A third possibility is if prosodic focus intervenes and a marginal 'not just any x but some special x' interpretation becomes available, retaining a trace of the original flavor of the indefinite. What happens to Slovak EIs under negation? The following examples were found in the corpus sample:

(74) a.Ozarila ho, hocí v zrelom veku už s voľa- podobným charm him although in ripe age already with what-si similar neraťal.

    NOT.CONT 'She charmed him even though at that ripe old age he did not count with something like that.'

    b. SIMPLIFIED LF: 'He didn’t count with something-si like that.'

    O a[exists for D [something like that (x) ∧ ¬ count (he,x)]]

    where D= {x: x is an event similar to her charming him}

    assuming that the domain only contains three items

In this example and those that follow, I simplify the calculations for the sake of readability. I leave out the scalar alternatives and the derived scalar implicature in the interpretation.; the reader should assume them to be always present.
c. INTERPRETATION: $\Box s \ (p \lor q \lor r) \land \Diamond s \land \Box q \land \Box r$

**KEY:**
- $p =$ he didn’t count with $x_1$
- $q =$ he didn’t count with $x_2$
- $r =$ he didn’t count with $x_3$
- $\Box s =$ as far as the speaker knows, it is necessary that...
- $\Diamond s =$ as far as the speaker knows, it is possible that...

The reading we get is that there was something specific with which the man did not count, and as far as the speaker and the agent know/consider relevant at the moment of the utterance, it could have been anything out of a domain consisting of things *like that* - things similar to a previously mentioned situation (her charming him). Due to the insertion of a covert epistemic modal, we get the correct interpretation of ignorance, as opposed to an NPI reading, even though the verb is under negation. The same is true for the voľa- series, for which I offer an artificial sentence, for the lack of a suitable corpus example:

(75) a. Nikto nečíta voľajakú knihu.

*nobody not.reads voľa-some book.*

'Nobody reads some unknown book.'

The above sentence conveys the reading that there is a specific book that the speaker cannot identify, it is consistent with the speaker’s knowledge that it could be any book in the domain, and nobody reads that one book. With the help of focus, the sentence can also acquire an emphatic ‘just’ FC interpretation: ‘Nobody reads just ANY old book, everyone reads special books’.

b. Nikto nečíta VOĽAJAKÚ knihu.

The EI voľa- does not take an NPI role under negation either. As a strict agreement language, an indefinite *nijakú* would be used for this purpose:

c. Nikto nečíta nijakú knihu.

*ní-who not.reads ní-some book.*

INTERPRETATION: Nobody reads any book(s).

Taking this property of not behaving like an NPI into consideration (the dimension of strength of exhaustification - the columns on the generative matrix, see Table 6), these properties place the *-si series alongside the indefinites *un NP qualsiasi, un oarecare, un quelconque*, and the voľa- series alongside *algun* and *un qualche*. The following table visually summarizes the above tentative conclusions:
Before moving on to further detailed analysis, it is useful to summarize the discussion up to now and provide the semantics of the two Slovak EIs. In light of Chierchia’s approach adopted here, both voľa- and -si are presented as existential elements that obligatorily activate scalar and domain alternatives. The lexical entries are given in (76a) and (77a) and the sets of alternatives in (76b-c) and (77b-c). Like all indefinites, Slovak EIs obligatorily activate scalar alternatives, those that are responsible for the implicature of ‘no more than one’. The scalar alternatives (SC-ALT) include subsets of D that comprise more than one member. Like all polarity items, they also obligatorily activate domain alternatives (D-ALT). The D-ALTs of the series voľa-, as opposed to the series -si, have to be minimal. In other words, they are restricted to singleton set u, a proper subdomain of D. The semantics of voľa- and -si are therefore identical except for the type of D-ALTs they activate.

(76) a. || voľa || = λPλQ [ ∃x∈D [ one(x) ∧ P(x) ∧ Q(x) ] ]
   b. || voľa || ∈ SC-ALT = {λPλQ [ ∃x∈{u} [ one(x) ∧ P(x) ∧ Q(x) ]; u∈D} 
   c. || voľa || ∈ SC-ALT = {λPλQ [ ∃x∈D [ n (x) ∧ P(x) ∧ Q(x) ]; one < n ]} 

d. VERIFYING MODELS:

<table>
<thead>
<tr>
<th>Modes of exhaustification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of alternatives</td>
</tr>
<tr>
<td>Exhaustified D-alternatives</td>
</tr>
<tr>
<td>(Total variation 3-FC)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Exhaustified Singleton D-alternatives</td>
</tr>
<tr>
<td>(Partial variation 3-FC)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
| Table 7: ‘Generative matrix’, only E-FCs, placement of the Slovak EIs

4.2.1 Lexical entries

...
For sentences with a partial variation vola- indefinite, the model M5 is excluded as a verifying model, because it has no variation at all. All other options are admissible, including the total variation M4.

\[
\text{(77) a.} \quad \text{\textit{\textit{\textit{\textit{-si} =}}} \lambda P \lambda Q \left[ \exists x \in D \left[ \text{one}(x) \land P(x) \land Q(x) \right] \right]}
\]
\[
\text{b.} \quad \text{\textit{\textit{\textit{\textit{-si} \text{ ALT} =}}} \{ \lambda P \lambda Q \left[ \exists x \in D' \left[ \text{one}(x) \land P(x) \land Q(x) \right] \text{;} D' \subseteq D \} \}
\]
\[
\text{c.} \quad \text{\textit{\textit{\textit{\textit{-si} \text{ SC-ALT} =}}} \{ \lambda P \lambda Q \left[ \exists x \in D \left[ \text{n}(x) \land P(x) \land Q(x) \right] \right] \text{;} \text{one} < n \}
\]

\text{d. VERIFYING MODELS:}

\[
\begin{array}{cccccc}
\text{*M1} & \text{*M2} & \text{*M3} & \text{M4} & \text{*M5} \\
\text{partial variation} & \text{partial variation} & \text{partial variation} & \text{total variation} & \text{no variation} \\
\text{falsifying model} & \text{falsifying model} & \text{falsifying model} & \text{falsifying model} & \text{falsifying model} \\
\text{w1 a} & \text{w1 a} & \text{w1 a} & \text{w1 a} & \text{w1 a} \\
\text{w2 a} & \text{w2 b} & \text{w2 c} & \text{w2 b} & \text{w2 a} \\
\text{w3 b} & \text{w3 a} & \text{w3 c} & \text{w3 c} & \text{w3 a} \\
\end{array}
\]

The sentences that contain the indefinite -si are typically only compatible with total variation models like M4 in (77d), in which each alternative is the referent in one of the possible worlds.

I now set to explore how the framework and the proposed semantics of the two Slovak EIs explain their behavior and interpretations in various contexts.

\textbf{4.2.2 Compatibility with modals}

One of the key attributes of E-FCs is that they are grammatical in the scope of overt modal auxiliaries (Chierchia, 2013, Ch.5). Given the sensitivity characterization of the Slovak EIs with respect to the parameters and typology proposed by Chierchia, their compatibility with modals is not surprising. The presence of modality, in fact, plays an important role in avoiding a logical crash between the scalar and the FC implicatures that would otherwise be a necessary result of the exhaustification process.

Modal constructions were attested in the corpus samples. Below are two examples with an epistemic modal of possibility scoping over -si (78) and vola-(79), and of a deontic modal of obligation/necessity over -si (80) and vola- (81).

\text{(78) a.} Teraz sa súce prpad zdá jednoduch, no pod povrhom sa mäže skryvať aj čo-šťie hlubšie.  

\text{can hide also what-si deeper}  

‘Although now the case seems simple, something deeper may also be hiding beneath the surface.’

b. \text{Simplified LF: ‘There may be something-si hiding beneath the surface.’}  

\text{O} \text{ } \exists x \in D \left[ \text{one}(x) \land \text{thing}(x) \land \text{hiding beneath the surface}(x) \right]  

assuming a domain that contains only three items:

\text{c. ASSERTION: } \diamond ( p \lor q \lor r )  

d. \text{INTERPRETATION: } \diamond ( p \lor q \lor r ) \land \diamond p \land \diamond q \land \diamond r  \quad - \text{speaker’s ignorance}
In the following sentence, we again get an [SU] ([E-FC] if stressed) reading:

(79) a. Alebo naopak, reč možno byť o voľajakom končikovi, záhube, or conversely talk can be about voľa-some hobby interest ktoréj sa venujú. which PART dedicate
'Or, conversely, the talk may be about some hobby, interest, to which they are dedicated.'

b. simplified LF:
\[ O \exists \times_2 D [\text{one}(x) \land \text{a hobby they are devoted to}(x) \land \text{topic of the talk}(x)] \]
where D = \{x: x is a hobby\}

c. assertion: \[ p \lor q \lor r \]
d. interpretation: \[ p \lor q \lor r \land (\neg p \land \neg q) \lor (\neg p \land \neg r) \lor (q \land r) \]

When under the scope of a possibility modal, both EIIs result in an ignorance reading due to the logical calculation returning an FC-like distributed modal conjunction.

As far as deontic (obligation, permission and similar) modals are concerned, I repeat the examples in (39-40). These sentences receive dual readings. The [SU] interpretation is possible (especially for the -si series which seems more ignorance inducing, as the corpus survey showed) but [E-FC] is preferred for both:

(80)a. Musime v tomto smere čoskoro čosi urobit,” tvrdí starostka.
must in this direction soon what-si do says mayor
"We must do something in this direction soon,” says the mayor.’

b. simplified LF: 'We must do something-sl.'
\[ O \Box \exists_2 D [\text{one}(x) \land \text{thing}(x) \land \text{do}(we,x)] \]
where D = \{a, b, c\}

c. assertion: \[ p \lor q \lor r \]
d. interpretation: \[ p \lor q \lor r \land (\neg p \land \neg q) \lor (\neg p \land \neg r) \lor (q \land r) \]

(i.) indifference - preferred: We must do something, any alternative thing to do is an admissible option as far as the mayor is concerned.
(ii) ignorance: We must do something specific, but at the moment, the mayor is unaware of what it is.

(81)a. Aby mohol zmysluplne a tvorivo vypovedať sebe a
about world must in life VOĽA-what study but also PART live
In order to meaningfully and creatively state something about himself and the world, he must study something, but also live it.’

b. simplified LF: 'He must study something-VOĽA.'
\[ O\exists \exists_2 D [\text{one}(x) \land \text{thing}(x) \land \text{study}(he,x)] \]
where D = \{a, b, c\}

c. assertion: \[ p \lor q \lor r \]
d. interpretation: \[ p \lor q \lor r \land (\neg p \land \neg q) \lor (\neg p \land \neg r) \lor (q \land r) \]

(i.) indifference - strongly preferred: He must study something or other, and any alternative (possibly with exceptions) would do.
(ii) ignorance - odd in the context, but possible: He must study something specific, but the speaker is unaware of what it is.
There is a difference between the felicity of -si and voľa- under the deontic modals. I will now use artificial examples to contrast them with the Italian data provided in Chierchia (2013, Ch.5, original gloss). Recall that I have placed -si next to un NP qualsiasi and voľa- next to un qualche in the matrix. We should expect them to be parallel to these items as far as their preference for an indifference versus an ignorance interpretation.

(82) a. Devi leggere un libro qualsiasi dalla lista di letture.             (Italian)
   'You must read a book whatever from the list of readings'

Chierchia explains: 'The sentence states that there is some obligation to read a book, which can be satisfied by reading any book in the relevant domain. No exception. The felicity of this example requires total variation (i.e. total freedom of choice) across the (indexically supplied) domain of quantification' (2013, Ch.5, Section 1). Although the Slovak -si series has been classified as conveying total variation (or some degree thereof), the deontic modal in the translated sentence (82) does not necessarily cancel the specific epistemic effect nor does it invoke the non-specific 'whatever' FC effect:

b. Musíš potešiť niektorú knihu z tohto zoznámu.
   must read some-si/which-si book from this list

c. [SU] INTERPRETATION: 'You must read a specific book from this list, you are not free to choose but at the moment I do not have more information on which book it is.'

In determining whether un qualche is parallel to voľa-, I adapt the following example taken again from Chierchia (2013, Ch.5, original gloss). It shows that the FC reading is indeed easier with voľa-. However, when it is substituted with a -si EI, the ignorance specific-unknown reading returns, indicating that the two series behave differently under deontic modals.

(83) a. Devo comprare una qualche assicurazione sulla vita.             (Italian)
   'I must get some life insurance'

b. Musím kúpiť niektorú záujmovú poistenie.
   must buy some-si/which-si life insurance

c. [E-FC] INTERPRETATION (PREFERRED FOR VOĽA-): 'I must buy some life insurance and (almost) any possible life insurance will satisfy this obligation.'

d. [SU] INTERPRETATION (PREFERRED FOR -SI): 'I must buy some specific life insurance but I don’t know which exactly.'

As we can see, deontic modals allow for dual readings. First, as we would expect from EIs, the specific/ignorance reading says that all the alternatives are consistent with the speaker’s current knowledge. Second, the presence of a modal also makes available the non-specific/indifference reading that says that
all life insurance policies are consistent with the speaker’s desires. How does ignorance become indifference? One explanation, discussed in passing in Ch. 5 of Chierchia (2013) is the possibility of two different scoping construals of the modal in the logical form of the sentence, as I show in the following:

(84) I must buy some life insurance.
   a. O □S [some life insurance, □ [I buy tₖ]]                  IGNORANCE
   b. O □ [I buy some life insurance]                          INDIFFERENCE

The two readings are correctly predicted with these two logical forms. The narrow scope construal in (84a) has the EI outside of the scope of the deontic modal. We must place a second null, speaker-oriented epistemic modal □S (which I will discuss in the next subsection) in the construction in order to avoid a logical crash of the indefinite (similarly to (63)). This logical form says: There is one life insurance (exhaustified scalar alternatives), I cannot identify it (exhaustified domain alternatives) and I must buy it. It is compatible with my beliefs (covert epistemic modal) that it be life insurance a, or b, etc. The construal of a wide scope over the indefinite in (84b) gives us an FC reading, not a speaker-oriented ignorance. The outcome of the exhaustification here is a reading in which there is a life insurance, I must buy it, and it is compatible with my obligation that I get the insurance a, insurance b, etc. In effect, it allows me to choose the life insurance. There may be pragmatic and discourse related reasons for why an EI would prefer one logical construal over another and a speaker-oriented modal base over an agent-oriented one. However, this is an undeveloped point in Chierchia.

4.2.3 Seemingly episodic contexts

So far, I have given many examples with an overt modal present in the construction, either may or must. However, EIs are attested as grammatical even if no modal is around. For example, there is no modal in the following statement (85a) which would be calculated as contradictory. The strategy to overcome this problem and rescue this statement that should not be contradictory, is to stipulate the existence and presence of a covert speaker-oriented modal □S (85c):

(85) a. Some man was sneaking around.
   b. # O [Some man was sneaking around] = ⊥
   c. ✓ O □S [Some man was sneaking around]

21 Consider there are only two men in the domain \{A, B\} and there are two alternative propositions, a and b. Exhaustifying the domain alternatives will contradict the exhaustification of the scalar alternatives: (a>b) ∧ ¬(a∧b) ∧ ¬(a∧b) ∧ ¬(b∧a) ∧ (a∧b) = (a>b) ∧ ¬((a∧b) ∧ ¬(a>b) ∧ ¬(b∧a) ∧ (a∧b) = 1

In prose, ‘if the man A was sneaking around then also the man B was sneaking around and not both were sneaking around’ is a contradictory statement.
An insertion of a covert modal is supported by the intuitive modal rephrasing
of the sentence into 'For all the speaker knows (or cares), or in all the worlds
compatible with the speaker’s belief (or desire), it might have been any man'.
This solution is further motivated by the consideration that languages have a
rich system of modalities that are not overtly expressed with a dedicated
morphology. Some examples are genericity, infinitives used with a goal oriented
bouletic interpretation (e.g. 'That ice cream is to die for') and, most pertinently
here, assertions that convey information about the speaker’s epistemic state. In
other words, a 'null/covert/phonologically unrealized assertoric modal' (along the
line of Kratzer and Shimoyama 2002) must be present in the logical form of an
episodic sentence containing an EI. In the course of the exhaustification process,
the covert modal, just like its overt relative, gets distributed over the conjunction
of alternative propositions and we are left with exactly the existential free choice
effect that we interpret (detailed out in (73)).

Both ignorance and indifference readings arise due to the covert modal’s
presence in an LF, depending on whether we consider the speaker’s worlds or the
agent’s worlds, in addition to an extra element, unclear at the moment, which
decides between an ignorance, therefore specific, and an indifference, therefore
non-specific reading. For example, the following sentence was labeled as
ambiguous between [SU] and [E-FC] in the annotated corpus sample, but the
covert modal is not sufficient to explain the distinction in readings:

(86) a. V aplnej tme sa pokazali volašo uvideli.
in complete darkness part tried vola-what see
‘they tried’ gets a ‘they said: we must’ interpretation and an agent
oriented necessity modal is inserted into the logical form

b. interpretation:  O □ ∃x.D [one(x) ∧ thing(x) ∧ try to see(they,x)]

(i) SPECIFIC UNKNOWN/IGNORANCE: ‘There was something specific they
tried to see, but as far as the speaker knows, it could have been
anything.’

(ii) NON-SPECIFIC FREE CHOICE: ‘They tried to see something, but as far as
the agents cared, it could have been anything.’

The theory therefore makes use of a covert modal operator, typically an
epistemic, speaker-oriented one, but a non-epistemic/bouletic, agent-oriented one
is also possible. This modal is necessary to rescue statements that would
otherwise be logically contradictory. Such implicit modalization can result in an
ignorance reading, or in some contexts also an indifference reading. The details of
this stipulation leave many open issues to explore and clarify and I will mention
a few in Section 5.
4.2.4 Imperatives

The imperative construction is attested in the corpus only for the voľa-series. Even though the series -si is not completely ungrammatical in imperatives, the two are not equivalent in acceptability when used in the same sentence. This could stem from the etymology of voľa’s morphology or from interpretative economy and competition with the specially marked indifference series (e.g. -kôvek or hoci-), as discussed in section 3.5.1. The following is an example from the corpus with the corresponding calculation:

(87) a. Poď, dušička, zobni si voľa/o voľa/o voľa/o voľa/o/osi osi osi osi osi nech neodpadne.

'Come, darling, eat something so that you don’t faint.'

The item ósi, being a total-ignorance EI, requires that all things are possible candidates, which is pragmatically odd. Voľa/o allows for this wide domain to be limited and excludes all the non-edible things and all the edible things outside of the salient context of a kitchen, for example. Following Chierchia who assumes that imperatives are very close to modals and as such work in rather similar ways, the LF and the interpretation can be shown as follows:

b. O □! ∃x ∈ D [one(x) ∧ thing(x) ∧ eat(you,x)]
where D = {x: x is a thing}

c. INTERPRETATION: □! (p ∨ q ∨ r) ∧ (¬p ∧ q) ∨ (¬p ∧ r) ∨ (¬q ∧ r)
You must eat something and at least two of the contextual and appropriate domain members are possible options that will satisfy the command.

This concludes the analysis of the EIs’ semantics in different contexts using the alternatives-and-exhaustification framework and brings me to propose their lexical entries. In doing so in the following subsection, I continue to use Chierchia’s notation.

5 Discussion

Chierchia’s parametrized alternatives-and-exhaustification analysis approach provided a way to situate the Slovak EIs in a broader cross-linguistic spectrum of polarity items. Working within this framework and probing voľa-š and -si’s behavior with respect to a limited set of parameters, I was able to propose their most important lexical properties and as a consequence, account for their interpretations of ignorance or indifference and their distribution. Naturally, this introductory work inevitably leaves us with some loose ends, unanswered questions and also interesting new directions for further investigation. I mention a few of them in this section.
(i) Distinction between specific and non-specific readings

If we adopt Chierchia’s system, we treat EIs as E-FCIs. They have obligatorily active scalar and pre-exhaustified domain alternatives. Through sentential exhaustification in DE and modal contexts, they bring out an implicature which conveys free choice: (almost) every domain alternative is an admissible referent candidate as far as the speaker’s knowledge or the agent’s desires are concerned. However, the system does not satisfactorily predict the distinction between specific ignorance vs. non-specific indifference readings. Furthermore, neither does it capture well whether that ignorance or indifference is of the speaker or of another participant in the state of affairs. These are separate dimensions of variation that do not necessarily correlate. To clarify, sentences usually may have a clearly preferred reading of a speaker’s ignorance or of an agent’s indifference. But in principle, there is also the speaker’s indifference, the agent’s ignorance, and if the speaker is not the same as the agent, what would prevent the indefinite to pick up on both of their attitudes and create other combinations (see Table 8)? Consider an example from the SNK that I annotated as non-specific agent’s indifference [E-FC] function, because it is the preferred reading:

\[(88)(=14d) \text{'Since childhood he read a lot and had urge to write something'}\]

<table>
<thead>
<tr>
<th>speaker</th>
<th>agent</th>
<th>paraphrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>ignorance</td>
<td>ignorance</td>
<td>Something specific that neither the speaker nor the boy can identify precisely.</td>
</tr>
<tr>
<td>indifference</td>
<td>indifference</td>
<td>He wanted to write anything, neither the speaker nor the boy care what.</td>
</tr>
<tr>
<td>ignorance</td>
<td>indifference</td>
<td>?? The boy would have been happy to write anything, but the speaker (perhaps with more retrospective knowledge) knew he wanted to write something specific although he could not identify it further.</td>
</tr>
<tr>
<td>indifference</td>
<td>ignorance</td>
<td>?? The boy wanted to write something specific but unknown to him, while the speaker does not care what.</td>
</tr>
</tbody>
</table>

Table 8: Possible, even if implausible, reading combinations

Would the picture become even more complicated if we added another participant? For example, the above sentence could be changed to: Since childhood he read a lot and his parents told him to write something. Now we have the speaker, the boy who reads a lot, and his parents; and each of these individuals can have an epistemic attitude of either ignorance or indifference. It seems that if an indefinite can pick up the epistemic attitude of a speaker and of the agent, perhaps it should have the capacity to convey any of the other participants’ attitudes. I am not aware of literature that lays out the constraints in play that would dictate the final preferred reading.
In Chierchia’s formalization, the orientation of the modal base is represented in the logical form with a subscript, S for speaker and A for agent. Chierchia (2013, Ch.5, Section 6) notes that $\Box_A$ appears to be bouletic in nature (a modal that expresses what is possible or necessary given someone’s desires) and as such it is compatible with total variation indefinites that express FC. On the other hand, the modal $\Box_S$ is compatible with partial variation only and the ignorance reading (e.g. un NP quahingque requires total variation and clashes with $\Box_S$).

While these are of descriptive value, there are no principled predictions of when the modal would receive which subscript. If the assignment of the subscript is dependent on the indefinite’s interpretation, we run into a circularity issue: an indefinite receives $\Box_S$ in the LF because it has a speaker-oriented reading and it has a speaker-oriented reading because the modal in its LF is $\Box_S$. Secondly, it doesn’t explain why a total variation- si series gets the [SU] reading with no overt modal present. As noted in Fâlauş (2010), recent literature on E-FCIs and modals (cf Port 2010) casts doubts on the total vs. partial variation stemming from an inherent lexical property that would carry into all environments. The provided evidence includes convincing examples containing German irgendein which seems to take on one or the other characterization from the interaction with the context of occurrence. Moreover, Chierchia (2013, Ch.5, Section 3.3) notes that the degree of variation might not even be a binary parameter (one with two values: total and partial) but rather a scale of different strengths (e.g. strong total, weak total, etc.). From what we know now and from seeing that the popular hide-and-seek scenario by A&M (2010) is not an ideal test at least for the Slovak data (see Section 4.2 for how the strength of variation is extremely difficult to evaluate), this parameter is not a reliable basis for predicting the speaker vs. agent orientation. The modal orientation of the reading and the nature of its relationship with the indefinite is certainly a point that deserves further study. A “stipulation free account of the interactions between the compatibility of modals with various kinds of EIs and their respective lexical properties” (2013, Ch.5, Section 6) is an open issue for Slovak as much as for any other language studied to date.

With regards to the problem of (non-)specificity, a well accepted idea by Quine (1956) is that specific and non-specific readings derive from different scope relations to the existential quantifier that binds the variable corresponding to the ambiguous NP. An example of scope implemented in the formalization used in this thesis is shown in (84). A covert assertoric modal has been proposed that intervenes, rescues an otherwise contradictory LF, and brings out the speaker’s epistemic state. However, little is known about what restraints this insertion and prevents overgenerating. Nor can we predict when this covert modal appears, other than a rather circular reasoning of ‘when the interpretation calls for it’. The specific and the non-specific readings are intuitively distinguishable, however, not explained fully through the exhaustification derivation.
(ii) Universal readings

EIs are considered to be existential quantifiers under Chierchia. However, in this work I presented some facts where Slovak EIs received a universal reading. Recall that [U-FC] occurrences, however uncommon, posed a problem for the adjacency principle by Haspelmath (1997) as well. Because the comparative function was not represented at all, the conditional and universal free choice functions were no longer connected which resulted in an unwelcome gap on the implicational map (see Figure 5). In addition to the examples presented in the Corpus Analysis section, e.g. (30i) and (31h), I offer the following as an illustration.

(89) U nás nie je zvykom, ako u vás v meste, volakomu vykař.

‘It is not customary here, as it is in the city, to address someone with the formal form of ’you’.’

In this sentence, the speaker conveys that for every person outside of the city, it is not customary to address that person with a formal form of ’you’. It could be very happily substituted with kazdanu (meaning ’everyone’) and retain the same truth conditions. The relevant diagnostics would steer us away from labeling it as [SU] or [E-FC] - it cannot be co-indexed with a personal pronoun and it does not lend itself to a ’there is’ paraphrase. Therefore, we must conclude it is universal, but is it free choice?

(iii) Atypical uses of the Slovak EIs

In the corpus sample, I came across some instances of curious contexts that suggest that there is more to the semantics of EIs than a free choice effect. I will abstain from an in-depth discussion here due to the fact that these uses deviate from the main focus of the thesis. As already mentioned, I have chosen to look at data in which EIs act as determiners. In other uses, some shown below, they would not be captured by the alternatives-and-exhaustification framework as intuitively without some modifications. First, in (90-92), the EI -si stands alone and activates some sort of a scale or a contextually salient amount value.

(90) Je čosi po 22.00 hodine.

‘It’s a little over 10pm.’

(91) Beloch, mušského pohlavia, vek ľosi nad tridsať, vyska meter

‘Of white race, male sex, age a little over thirty, height 1.80 m.’
(92) Zrejme bude o čoťi menej aj cestujúcich z Londýna.  
"It seems there will be somewhat fewer travelers from London."

In (93-95), the EI appears to act as a modifier, not necessarily of an NP. It enriches what could easily be said without it, modifying the adjective with a specific unknown degree adverb, in the vicinity of 'a little' in (93). In (94-95), however, this does not hold, rather it conveys something akin to notability.

(93) Sú akysí urahani, zdá sa, že to potrebujú."
   Are some-sí weary, seems PART that it need
   'They are sort of weary/weary in a notable way. It seems that they need it.'

(94) Vyzařoval tam akysí pokoj, miernosť, atmosféra oddychu.
   emanated there some-sí peace, gentleness, atmosphere rest
   'Some sort of/notable peace, gentleness and an atmosphere of rest emanated there.'

(95) Nuž, voľajaky je vám ďudny.
   well VOĽA-some is you weird
   'Well, he is weird in some (notable) way.'

(iv) Identification method

The frequent example in this thesis, first given as (15b), tells of some strange man that was sneaking around. What is interesting about this sentence is that the speaker’s ignorance is of a very particular kind. He is probably knowledgeable enough about the strange man to identify him easily for example in a line-up. However, the indefinite picks up on the one characteristic of the strange man of which the speaker does not have knowledge - his name (or even more likely, his personal acquaintance with the speaker). There are various types of knowledge (as was discussed briefly in Section 3.5.1 in connection to visual access and type of evidence). The Chierchia system adopted here does not incorporate types of knowledge, which at the same time averts adding too much complexity to the theory and fails to account for a real phenomenon to which language seems to be sensitive. An alternative proposal for an analysis of indefinites, Aloni (2001), builds on precisely this observation that a speaker’s knowledge is not a binary feature but can be characterized depending on the identification method chosen by context.

Aloni puts forward the following scenario to clearly illustrate the intuition:

(96) a. In front of you lie two cards face-down. One is an ace of spades, the other is an ace of hearts. You know that the winning card is the ace of hearts, but you don’t know whether the ace of hearts is the card on the left or the card on the right.

b. You know which card is the winning card.    (Aloni, 2001:16)
Under the premise of (96a), the statement (96b) is true depending on whether the identification method is by kind (TRUE) or by position (FALSE). It is easy to see that in real communication, we evaluate and interpret what indefinites refer to in a similar fashion. The following example is one of many facts in the corpus sample that makes this obvious:

(97) Dostal odmenu za to, že vykopal v zamrznutej zemi \textit{volajkamu} received reward for it that dug-out in frozen earth \textit{voľA}-some staraniu tuľákovi Jozefovi hrob.

old vagrant Joseph grave

‘He received an award for digging out a grave in the frozen earth for some old vagrant named Joseph.’

In this situation, the speaker has a lot of identifying characteristics available for the indefinite’s particular referent – he was old, he was a vagrant, and we know his name (analogous to knowing which is the winning cards, considering only the characteristic [hearts, spades]). Despite having plenty of identifying information, the EI’s presence in the sentence communicates a certain level of uncertainty as to his identity (acquaintance being roughly analogous to the characteristic [left, right]). It is hardly imaginable that the domain of possible candidates for this referent would be very wide. Facts like these perhaps lend themselves more intuitively to a conceptual covers analysis than an alternatives-and-exhaustification analysis. A more systematic comparison of strengths, weaknesses, and mutual compatibility of different approaches to Slovak EIs promises to be part of a fruitful and interesting forthcoming project.

6 Conclusion

This thesis has contributed to the study of the semantics of Slovak, a language that has been underrepresented in the formal linguistic area of research. The main focus of my work were two indefinite series with a grammaticalized epistemic effect, \textit{voľa}- and \textit{-si}.

I accomplished the two initial goals. First, I approached the topic by giving an introductory descriptive survey of epistemic indefinites in Slovak, using the Slovak National Corpus. I showed that by large, their behavior conforms to the nine semantico-syntactic functions proposed in the work on indefinites by Haspelmath (1997). With an exception of a disconnected universal free choice function, their functional distribution falls on congruent areas of his implicational map. There a few differences between their distribution, for example, the \textit{voľa}- series is more multi-tasking and more disposed to serving the free choice function. The \textit{-si} series, on the other hand, is more dedicated to the specific unknown - ignorance function.

22 This concept is termed \textit{conceptual cover} (CC) and defined as a set of functions from possible worlds to individuals. See Aloni (2001) for details and formalization.
Through a qualitative description of the readings and typical constructions of the two EIs, I was able to uncover their characteristics along the main parameters mentioned in the literature. This description was used as a springboard for accomplishing my second goal, their semantic analysis. I implemented an alternatives-and-exhaustification framework by Chierchia (2013) that works on the premises that grammar is in its essence a logical inferential apparatus and that all scalar and polarity phenomena are radically unified in that they all activate alternative domain and scalar propositions. I adopted his assumption that ignorance is an instantiation of a speaker oriented existential free choice, a point in which Chierchia differs from Haspelmath. By probing in what ways Slovak EIs are sensitive to the parameters in this framework, I drew conclusions about where they stand in comparison to EIs in other languages. At the same time, I endeavored to provide a critical analysis of this new theory. I identified several questions with regard to the extent to which it captures all of the empirical data, most notably the lack of explanation for the preference of ignorance vs. indifference interpretations. I suggested throughout that various distinctions between the two series are mere conjecture at this time and can be resolved through complementing the corpus survey with native speaker judgment questionnaires.

Looking at epistemic indefinites in various languages, as I did in Slovak, will help us understand what these items have in common, how they vary, and, most importantly, what the explanatory principles behind their behaviors are. Answers to these research questions will uncover important pieces to the puzzle of how languages convey ignorance and indifference. I hope my small contribution was a step in this direction.
Appendix

Corpus Queries

The following tables list the items searched in the Slovak National Corpus in order to extract the data for analysis.

(i) -voľa indefinites

Table 7: SNK query for the voľa- series indefinites

The result yielded 7547 occurrences.
Out of this concordance, a random sample of 300 was generated and annotated.

(ii) -si indefinites

Table 8: SNK query for the -si series indefinites

The result yielded 235,399 occurrences.
Out of this concordance, a random sample of 300 was generated and annotated.
Bibliography


