

# The Lexicological and Lexicographical Problem of Lexical Variation

Working Group 4  
Vienna Meeting Presentation  
11 February 2015

## **Group Members:**

Chris Mulhall  
Anna Havinga  
Karlien Franco  
Martina Ivanová

# Overview of Italian Phrase Variation Strand

Research Parameters	Corpus and Linguistic Sample	Specifics
<b>Units of Analysis</b>	Noun Phrase (NP) Idioms Verb Phrase (VP) Idioms	Compare variability trends within VP and NP and across both categories
<b>Type of Variation</b>	Word Class Analysis	Content Word Variation [Verbs/Nouns/Adjectives] Determining inter and intra-group variation patterns
<b>Empirical Linguistic Context</b>	Paper and Electronic Dictionaries  <a href="mailto:cmulhall@wit.ie">cmulhall@wit.ie</a>	Bilingual Italian-English Dictionaries [Renaissance Period – to present]  Monolingual Italian Dictionaries [corpus age]

# Task Objectives

<b>Theoretically- Grounded</b> [LEXICOLOGICAL]	<b>Objective 1</b>	To theorise (and explain) about the presence of variable units in some phrases and not others.
	<b>Objective 2</b>	To classify the frequency of content word variation forms
<b>Empirically- Contextualise</b> [LEXICOLOGICAL- LEXICOGRAPHICAL]	<b>Objective 3</b>	To explore the role of corpora in prioritising or stigmatising certain variant forms
	<b>Objective 4</b>	To assess the diversity (or similarity) of the approach to dealing with lexical variation in paper and electronic dictionaries in an Italian (-English) context.

# Areas for Consideration: An Italian Context

- Examine differences in the coverage of variation between (a) monolingual and bilingual Italian dictionaries and (b) contrast approaches in paper and electronic dictionaries
- Chart the evolving trends in recording variable phrases, in particular idioms, in Italian dictionaries since the Renaissance
- Explore the possible marginalisation of Southern Italian phrasal variants in predominantly Northern Italian published dictionaries

# Timeline of Proposed Research Tasks

Time Line	Task Area	Task activity
<b>March 2015 - May 2015</b>	Lexicological: Documentation and Analysis of Content Word Variation in Phrases	March '15 - Verb variation April '15 - Noun Variation May '15 - Adjective Variation
<b>June- July 2015</b>	Theoretical (Linguistics)	Classification of Content Word Variation types for the purposes of developing a typological classification
<b>August 2015</b>	Lexicographical: Assessing Lexicographical approaches to	<b>Development of a chronologically comparative matrix of Italian lexicographical</b>

*How much **variation** can a  
standard language take?  
How much **standard** can a  
versatile language take? (cf. Löffler  
2004: 25)*

# Initial ideas for projects

- Comparative analysis of dictionaries (e.g. *Duden* [German orthographic dictionary] vs. *Variantenwörterbuch* [dictionary on variants of German], *OED* vs. *Duden*, *Österreichisches Wörterbuch* [Austrian dictionary] vs. *Duden* etc.) - to establish how dictionaries deal with variation at the moment
- Compilation/Use of versatile corpora as the basis for dictionaries - to find variants and classify them
- Digitize dictionaries - to present a lot of information clearly

# Example

## ***Duden* (2006) - orthographic dictionary**

- *the* standard dictionary for German – for everyday life, “easy to use for everybody” (Duden 2006: Vorwort)
- presents variants used in Germany as the norm, while marking others as regional (e.g. Austrian, Swiss etc.)

## ***Variantenwörterbuch* (2004) - dictionary on variants of German**

- used by people interested in variation – it won't be found on every bookshelf
- lists all standard German words, which have more than one variant in the German speaking area (Austria, Germany, Switzerland, Liechtenstein, Luxembourg, East Belgium, and South Tyrol)



# Example

## **Duden (2006) - orthographic dictionary**

- abbekommen [no marking, even though it is very uncommon in Austria - see *Variantenwörterbuch*]
- abbeuteln (*bayr. österr. für abschütteln*) [not listed in the *Variantenwörterbuch*]

## **Variantenwörterbuch (2004) - dictionary on variants of German**

- abbekommen st. V./hat: **1. D-nord/mittel;** ↗ ABKRIEGEN CH D (ohne südost), ↗ MITKRIEGEN D, ↗ ABHABEN D-nord/mittel >einen Teil von etw. bekommen< [example sentence] [...] **2. CH D-nord/mittel;** ↗ ABKRIEGEN CH D (ohne südost) >(etw. Nachteiliges) erdulden müssen< [example sentences] - **In A selten** [my emphasis]

To examine the inclusion of **frequency** information of lexical **variants** in onomasiological dictionaries and its **methods of presentation** to the user

# Frequency of lexical variants

how is frequency presented in onomasiological dictionaries?

- general vocabulary in Dutch dialect dictionaries of Limburgian and Brabantian dialects:
  - frequency can be inferred from database
- Dictionary of Bavarian dialects:
  - frequency inferred from database

⇒ not presented to user directly

- Opoeteren, Opglabbeek, Lottum, Grubbenvorst en Blerick.
- bestemoer:** Well.
- oma:** freq. Zuid.Oostlb., verspr. Oost.Zuidlb., Noord.Oostlb. en Centr.Maaslds.; ook in Kerkrade Wb. 2 en in Weert, Tungalroy, Hunsel, Stevensweert, Maastricht, Neerpelt, Venlo, Well en Wanssum.
- omaaetje** (*omaake*): Middelaar en Well.
- bonne** (fr.): Tongeren, Hasselt, Gutshoven, Heers, Kortessem, Gorsen en Sint-Truiden.
- bonne-bonne** (fr.) (*bobonne*): Kuringen.
- bonne-tje** (*bommeke, bonneke*): Maastricht, Hoepertingen, Voort en Wellen.
- bonne-ma** (ook *bomma*): verspr. Maaskemp. en Demerkemp.; ook in Sittard, Gronsveld, Maaseik, Mechelen-aan-de-Maas, Lanaken, Maastricht, Bilzen, Tongeren, Achel, Kaulille, Zepperen, Kortessem, Herk-de-Stad, Nielbij-St.-Truiden en Jeuk.
- bonne-maatje** (*bonne-maake*): Ulbeek.
- bonne-meetje** (*bonne-meeke*): Ulbeek.
- bonne-maman** (fr.) (ook *bomma-ma*): Reppel, Maastricht, Tongeren, Mal, Neerpelt, Kaulille, Helchteren, Kuringen, Hasselt, Paal, Ulbeek, Sint-Truiden, Montenaken en Borlo.
- dikke ma:** Borlo.
- dikke mam:** Neu-Moresnet, Kuringen, Gutshoven en Heers.
- grand** (fr.) (*gram*): Hasselt.
- grand-ma** (ook *gramma*): Valkenburg, Margraten, Echt/Gebroek, Maastricht, Ulbeek en Hoepertingen.
- grand-maman** (fr.): Neu-Moresnet, Limbricht, Valkenburg, Maastricht en Hamont.
- grand-mère** (fr.) (*grammeer*): Neu-Moresnet, Sint-Pieters-Voeren, Ulestraten, Valkenburg, 's-Gravenvoeren, Maastricht, Sint-Pieter, Diepenbeek, Riksinggen, Tongeren, Hasselt en Wellen.
- groot, de** -: verspr. Kleverlds.; ook in Zutendaal en Geulle.
- grootje:** verspr. Noord.Oostlb., Horns en Maaskemp.; ook in Pey, Weert, Hushoven, Neeritter, Meldert, Venlo, Middelaar en Gennep.
- grootma:** Bocholt, Lanaken, Vlijtingen, Overpelt, Alken, Gorsen en Sint-Truiden.
- grootmaatje** (*grootmaake*): Sint-Lambrechts-Herk.
- grootmam:** verspr. Lonerlds.; ook in Heerlen, Wijlre, Rosmeer, Vlijtingen, Rijkhoven, Nieuwerkerken, Sint-Truiden en Aalst-bij-St.-Truiden.
- grootmama** (ook *grootmamme*): Neeroeteren, Borgharen en Sint-Truiden.
- grootmem:** Lanklaar, Lanaken en Valkenburg.
- groot-mère** (*grootmeer*): Lanaken.
- gross** (du.) (*groos*): Kerkrade, Ubachsberg, Heerlen, Schaesberg en Jabeek.
- grossmama** (du.) (*groosmama*): Heerlen en Welkenraedt.
- gross-mam** (*groosmam*): freq. Oost.Zuidlb.; ook in Kerkrade Wb. 2 en in Nuth/Aalbeek.
- gross-moeder** (*groosmodder, groosmoeder*): freq. Ripuar. en Oost.Zuidlb.
- grote ma:** Bocholt, Bree, Veldwezelt, Riksinggen, Kuringen, Paal en Borlo.
- grote mem:** Kuringen en Veldwezelt.
- grote moe:** Kuringen, Paal, Tessen-derlo en Linkhout.
- grote moeder:** Lanaken.
- grote moei:** Meldert.
- grote moene:** Grote-Brogel, Ellikom, Peer en Hechtel.
- ma:** Stevoort.
- mam:** Neu-Moresnet, Bilzen, Mal, Scholen, Zonhoven, Sint-

lexical variants  
concept 'grandmother'

locations where  
variant occurs

sometimes some  
frequency  
information

# Work plan

- how is frequency presented in other dictionaries?
  - create overview of onomasiological dictionaries in Europe
- what is the best way?
- contribute to research on lexical variation

# How does this contribute to research on lexical variation?

- frequency as a proxy of entrenchment (Langacker 1987, Schmid 2007):
  - frequency of the variant
    - entrenchment of the concept
- frequency as an explanandum of lexical variation

# Overview of Slovak strand

Research Parameters	Corpus and Linguistic Sample	Specifics
<p><b>Units of Analysis</b></p>	<ul style="list-style-type: none"> <li>- lexemes (idioms, multiword expressions)</li> <li>- morphemes</li> </ul>	<ol style="list-style-type: none"> <li>1. Lexemes varying with respect to their phonological, orthographic, morphological and word-formation characteristics</li> <li>2. Multiword units/idioms showing capacity to vary in one or more parts</li> <li>3. Morphemes varying with respect to their formal characteristics</li> </ol>
<p><b>Empirical Linguistic Context</b></p>	<p>Paper and Electronic Dictionaries</p>	<p>Monolingual Dictionaries</p> <ul style="list-style-type: none"> <li>- Dictionary of Slovak Language Root Morphemes (2012)</li> <li>- Dictionary of Contemporary Slovak Language (2006, 2011)</li> <li>- Valency Dictionary of Slovak Verbs on Corpus Basis (2014)</li> <li>- Dictionary of Slovak Multiword Expressions (in preparation)</li> <li>- Dictionary Idioms in Slovak Language (in preparation)</li> </ul>

# Areas for Consideration: Slovak context

- Compile a consistent and noncontradictory theoretical basis for postulating the variation
- Examine theoretically the differences between variation and other phenomena (e.g. synonymy, transformations)
- Idioms: distinguish transformations (being independent idioms usually) from variants (of a single idiom)



# Areas for Consideration: Slovak context

- Comparative analysis of dictionaries – dynamic aspects in processing the variation  
(e.g. *Dictionary of Contemporary Slovak Language* (2006, 2011) vs. *Short Dictionary of Slovak Language* (2003))
- Establish how different dictionaries deal with variation in the span of time
- Real usage offers an ever-changing and constant display of new variants, some of which may become stable in due time

# Areas for Consideration: Slovak context

- Major desiderata for a formal approach to the study of variation:
  - use of corpora as the basis for dictionaries – to find variants and classify them
  - frequency as important factor for processing the variants (Sketch Engine Difference)

# Example

Sketch Engine Search  in Help

user: martinai corpus: Omnia (SNK 6.0-juls-all + Dictum + Iura + Bruna + Wanda 1.0.20) 2,34 G (#654) Search  in Omnia (SNK 6.0-juls-all + Dic

Concordance  
Word List  
Word Sketch  
Thesaurus  
Find X

## telefonický/telefónny

Omnia (SNK 6.0-juls-all + Dictum + Iura + Bruna + Wanda 1.0.20) 2,34 G (#654) freqs = 23,077 | 54,957

**telefonický** 6.0 4.0 2.0 0 -2.0 -4.0 -6.0 **telefónny**

X Sb	24,386	57,871	0.2	0.2	Z X	6,193	16,829	0.0	0.1	X Z	3,317	14,484	0.0	0.1
urgencia	<a href="#">62</a>	0	5.7	--	odpočutý	<a href="#">15</a>	0	5.6	--	0800	<a href="#">23</a>	<a href="#">121</a>	6.1	7.5
rozhovor	<a href="#">4,498</a>	<a href="#">170</a>	6.1	1.3	nahráný	<a href="#">9</a>	0	3.1	--	905	<a href="#">8</a>	<a href="#">113</a>	2.7	6.1
kontakt	<a href="#">2,816</a>	<a href="#">591</a>	5.8	3.5	odpočúvaný	<a href="#">47</a>	<a href="#">18</a>	6.6	4.5	02	<a href="#">8</a>	<a href="#">212</a>	1.8	6.3
infolinka	<a href="#">79</a>	<a href="#">55</a>	5.5	4.3	prerušený	<a href="#">33</a>	<a href="#">29</a>	3.0	2.7	158	<a href="#">16</a>	<a href="#">758</a>	3.0	8.4
hovor	<a href="#">1,227</a>	<a href="#">698</a>	6.7	5.8	skrátенý	0	<a href="#">52</a>	--	3.0	911	0	<a href="#">66</a>	--	5.4
linka	<a href="#">1,200</a>	<a href="#">4,963</a>	5.1	7.1	prednastavený	0	<a href="#">11</a>	--	3.4	0903	0	<a href="#">109</a>	--	5.6
záznamník	<a href="#">12</a>	<a href="#">184</a>	3.1	6.2	prevádzkujúci	0	<a href="#">13</a>	--	3.5	09610	0	<a href="#">23</a>	--	5.7
prístroj	<a href="#">28</a>	<a href="#">1,027</a>	0.1	5.3	udaný	0	<a href="#">13</a>	--	3.6	918	0	<a href="#">49</a>	--	5.7
ústredňa	<a href="#">25</a>	<a href="#">1,507</a>	3.3	8.7	predvolený	0	<a href="#">15</a>	--	3.7	907	0	<a href="#">63</a>	--	5.8
prípojka	<a href="#">8</a>	<a href="#">674</a>	0.8	6.9	ip	0	<a href="#">65</a>	--	3.8	021	0	<a href="#">34</a>	--	5.9
číslo	<a href="#">101</a>	<a href="#">22,122</a>	-0.1	7.7	predplatený	0	<a href="#">33</a>	--	3.8	159	0	<a href="#">115</a>	--	6.1
zoznam	<a href="#">8</a>	<a href="#">3,689</a>	-2.7	6.1	volaný	0	<a href="#">26</a>	--	4.1	0961	0	<a href="#">11</a>	--	6.1
hovoriča	0	<a href="#">80</a>	--	5.4	komutovaný	0	<a href="#">10</a>	--	4.2	0850	0	<a href="#">46</a>	--	6.2
slúchadlo	0	<a href="#">410</a>	--	6.0	poskytujúci	0	<a href="#">108</a>	--	4.5	903	0	<a href="#">111</a>	--	6.4
automat	0	<a href="#">1,789</a>	--	7.5	pobočkový	0	<a href="#">17</a>	--	4.6	041	0	<a href="#">116</a>	--	7.0
búdka	0	<a href="#">1,953</a>	--	8.9	isdn	0	<a href="#">26</a>	--	4.6	+421	0	<a href="#">165</a>	--	7.0

frequency for the second lemma

# Variants in collocation type of dictionary

- adjectives with the same root morpheme and different suffixes – as parts of adjective + noun collocations
- frequency as the basis for processing the variants:
- *telefonická urgencia*
- *telefonický/telefónny rozhovor, telefonický/telefónny kontakt*
- *telefónna/telefonická linka, telefónny/telefonický záznamník, telefónny/telefonický prístroj*
- *telefónna búdka, telefónny automat, telefónne slúchadlo*

# Research Timeline: Slovakian Context

- **March - June 2015**

- theoretical issues of linguistic variation, linguistic variation and language-particular versus language-universal properties, semasiological and onomasiological perspective in variation, microvariation and macrovariation, synchronic and diachronic variation
- presentation of paper *Variation and variants in Dictionary of Slovak Language Root Morphemes* - WORD-FORMATION THEORIES II, TYPOLOGY AND UNIVERSALS IN WORD-FORMATION, Košice (Slovakia) 26-28 June 2015

- **June - July 2015**

- usage-based approach towards linguistic variation, role of corpora, lexical variation and lexicography, descriptive and explanatory adequacy in investigation of variation

- **July - December 2015**

- linguistic variation in multiword expressions, variation and synonymy, typology of variants
- presentation of paper *Typology of Variants in the Dictionary of Multiword Expressions of Slovak* - The tenth annual meeting of the Slavic Linguistics Society (SLS), Heidelberg University (Germany), 4-6 September 2015
- preparation of chapter *Variation and multiword expressions* for the publication *Multiword Expressions in Slovak (in comparison with other languages)* (Ološtiak et al 2015)