An app-based language survey for Luxembourgish: regional variation and language contact

Peter Gilles
Institut fir lëtzebuergesch Sprooch- a Literaturwëssenschaft
Universitéit Lëtzebuerg
peter.gilles@uni.lu

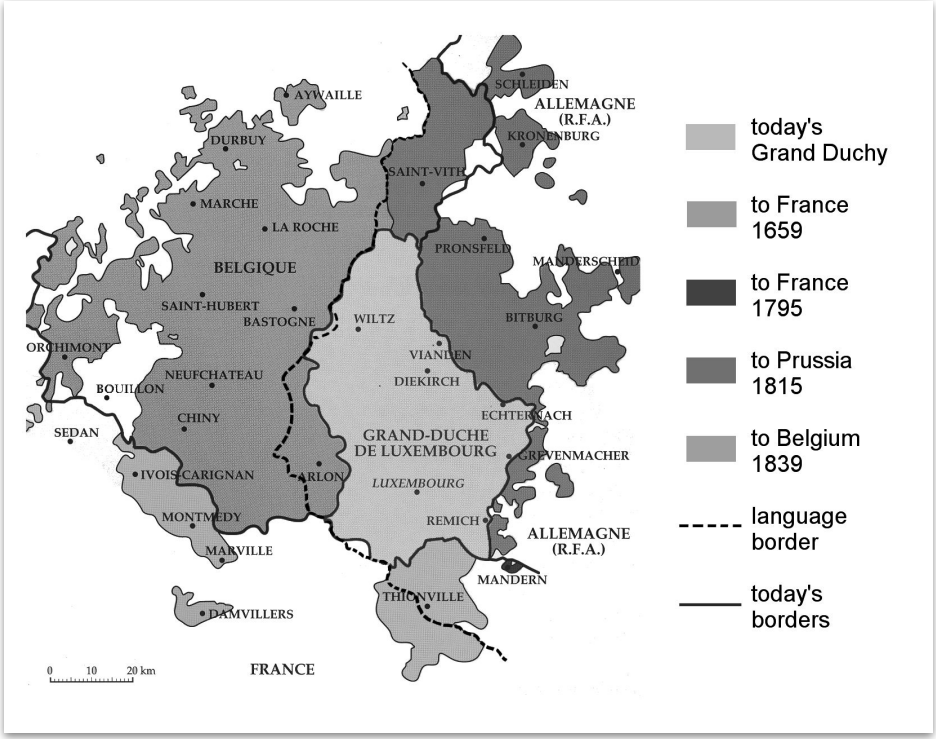
Small languages, big ideas
30.6 - 1.7.2022

Structure of the talk

1. Language situation of Luxembourg
2. Smartphone application to elicit and document language variation
3. Language variation in Luxembourgish
    a. Regional variation
    b. Language contact
    c. Internal changes/innovations
4. Conclusions
    a. Trends in the development of Luxembourgish
    b. Pros & cons of this large dataset

Luxembourg - language situation

- Location on the language border
- Development of a multilingual language situation
- Official trilingualism since 1984
  - Luxembourgish - ‘langue national’
  - Luxembourgish, French, German - administrative languages
  - French - language of the law
- Educational system
  - Alphabetisation in German
  - Subjects taught in German and French
  - Luxembourgish NOT a subject
- Immigration society
  - 634,000 inhabitants
  - 47.2 % migrants (mainly from Portugal)
  - 200,000 cross-border workers



Luxembourgish

- Originating from a Moselle-Franconian dialect
- Developed into an *Ausbau* language due to nationhood and identity symbol
- First language for ~ 55 % of the population (300,000 speakers)
- Second language for ~ 15 % of the population
- Mainly used as a spoken language, increased usage as a written language

Luxembourgish

- No actively forced standardisation
  - However, ongoing codification (orthography, lexicon)
- Highly variable language
  - Regional variation
  - Variation due to language contact with French and German
  - Internal changes/innovations
  - Increasing number of L2 speakers

Smartphone application: *Schnëssen*

- Documentation and analysis of all levels of language variation in Luxembourgish
- Audio recordings with mobile phone
  - *Schnëssen-App - Är Sprooch fir d'Fuerschung*
  - Similar Apps: *Deutschklang, Dialäkt App, English Dialects App, Stimmen, Gschmöis*
- Survey items
  - Translation from German or French: *Eine Frau hat am Freitag die schwarzen Kleider genäht.*
  - Image descriptions
  - Questions: *Wéi nennt Dir déi éischt Moolzecht vum Dag?*
  - Reading tasks: *Dass erëm d'Taxe gehéicht ginn, deet d'Jeeër haart jäizen.*
- Social information
  - Age, gender, place of birth, first language, language competencies, etc.

21:35

Schnëssen

Nach weider Opnamen ze maachen

Opnam maachen

Froebou ausfëllen

Opname lauschteren

Resultater ukucken

uni.lu UNIVERSITÉ DU LUXEMBOURG

21:35

Schnëssen

Daten zu Ärer Persoun

Liège, Belgien

Aner Lokalitéiten, déi lech sproochlech beaflosst hunn

KENG weider Aflëss

Alter

85+

Héchsten Ausbildungsofschloss

Primaire/Grondschoul

Geschlecht

Aner

Ass Lëtzebuergesch fir lech Mammesprooch?

Jo

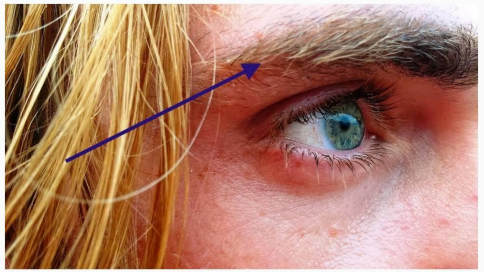
Hutt Dir eng zweet Mammesprooch?

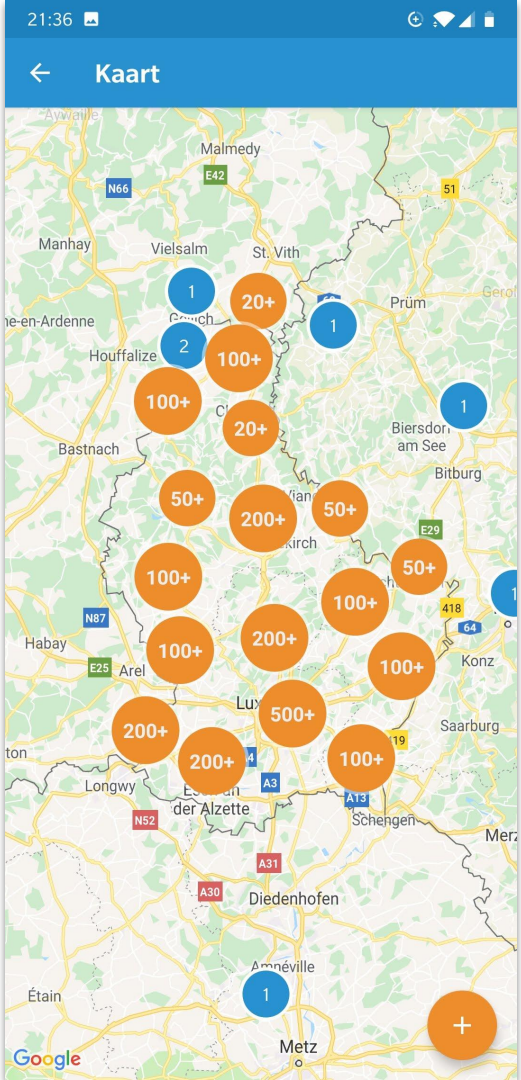
21:43

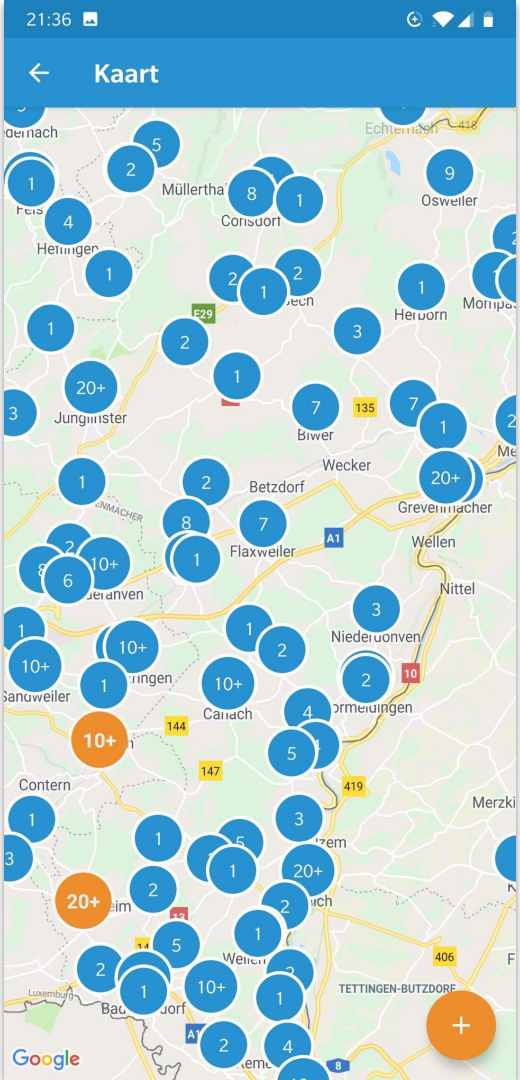
4 vu(n) 78

- Schwätzt spontan a wéi bei lech doheem
- Opnam a rouegem Raum ouni Kaméidi (Tëlee etc.)
- Haalt den Handy ca. 15cm vun lech ewech
- Maximal Ophuelzäit 10 Sekonnen

Wat ass um Bild?




21:36
Kaart
Malmedy
Manhay
Vielsalm
St. Vith
Prüm
Houffalize
Bastnach
Habay
Arel
Longwy
Diedenhofen
Étain
Metz
Google


21:36
Kaart
Müllertha
Consdorf
Osweiler
Heffingen
Junglinster
Biwer
Wecker
Betzdorf
Flaxweiler
Wellen
Nittel
Contern
Merzig
TETTINGEN-BUTZDORF
Google

21:37

Iwwersetzt dëse Saz:

Claudine Schroeders Katze will euch nicht beißen, auch wenn sie böse schaut.

45 bis 54	2018-04-25
45 bis 54	2018-04-29
25 bis 34	2018-04-30
25 bis 34	2018-04-30
45 bis 54	2018-06-03
45 bis 54	2018-06-03
15 bis 24	2018-06-07
45 bis 54	2018-06-07
15 bis 24	2018-06-09

21:38

45 bis 54	2018-04-25
45 bis 54	2018-04-29
25 bis 34	2018-04-30
25 bis 34	2018-04-30
45 bis 54	2018-06-03
15 bis 24	2018-06-07
45 bis 54	2018-06-07

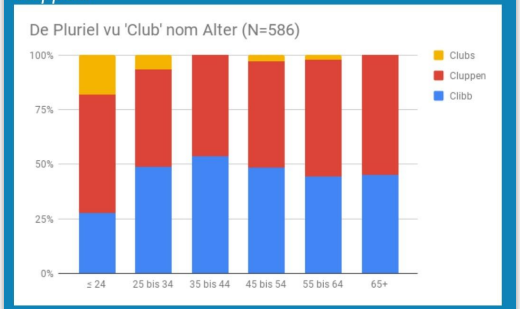
21:38

← Resultater ukucken

Clibb/Clubben a Fannen/Fans

Och haut rëm e puer fussballeresch Auswäertungen, fir den F91 och vu sproochlecher Säit z'ënnerstëtzen. An der 2. Ronn vun eise Schnëssen-Erhiewungen an der App war de Saz 'Die Fans der zwei Clubs sind aneinander geraten.' z'iwwersetzen a mir gräifen eis haut de 'Club' an d"Fannen' eraus. Am Ganzen hu mer d'Iwwersetzunge vu méi wéi 600 Leit analyséiert.

D'Haaptwuert fir d'Iwwersetzung vum 'Club' ass dann och *Club*. 50 mol ass och *Veräin* gesot ginn, nach méi seelen ass hei *Equipe*, *Team* oder *Mannschaft*. Beim *Club* ass natierlech de Pluriel interessant, well mer do jo mat de *Clibb* an de *Cluppen* (seelen och *Clubben*) zwou konkurréierend Varianten hunn. Wéi bei den aneren Analysen, déi mer hei scho virgestallt hunn, ass d'Differenzéierung nom Alter dacks charakteristesch - an sou och hei: Vum Alter 25 bis 65+ halen sech d'Varianten *Clibb* (blo) a *Cluppen* (rout) ongefeier d'Wo. Déi jong Spriecher*innen bis 24 Joer schénge *Clibb* allerdéngs ofzebauen a preferéieren éischter *Cluppen*.



En plus kann ee bei deene Jonken och den neie Pluriel op *-s* beobachten, sou dass lo och *Clubs* (giel) opdauchen. Et bleift z'ënnersichen, ob déi nei Formen direkt aus dem Engleschen kommen oder den Ëmwee iwwert d'Däitscht gemaach hunn.

Structure of the survey items

- +550 survey items
- ~ 800 variation phenomena (‘variables’) from phonetics/phonology, morphology, syntax, lexicon, language contact
- Included all phenomena of the ‘Luxemburgischer Sprachatlas’ (LSA 1963)
- Largest survey and documentation of spoken Luxembourgish

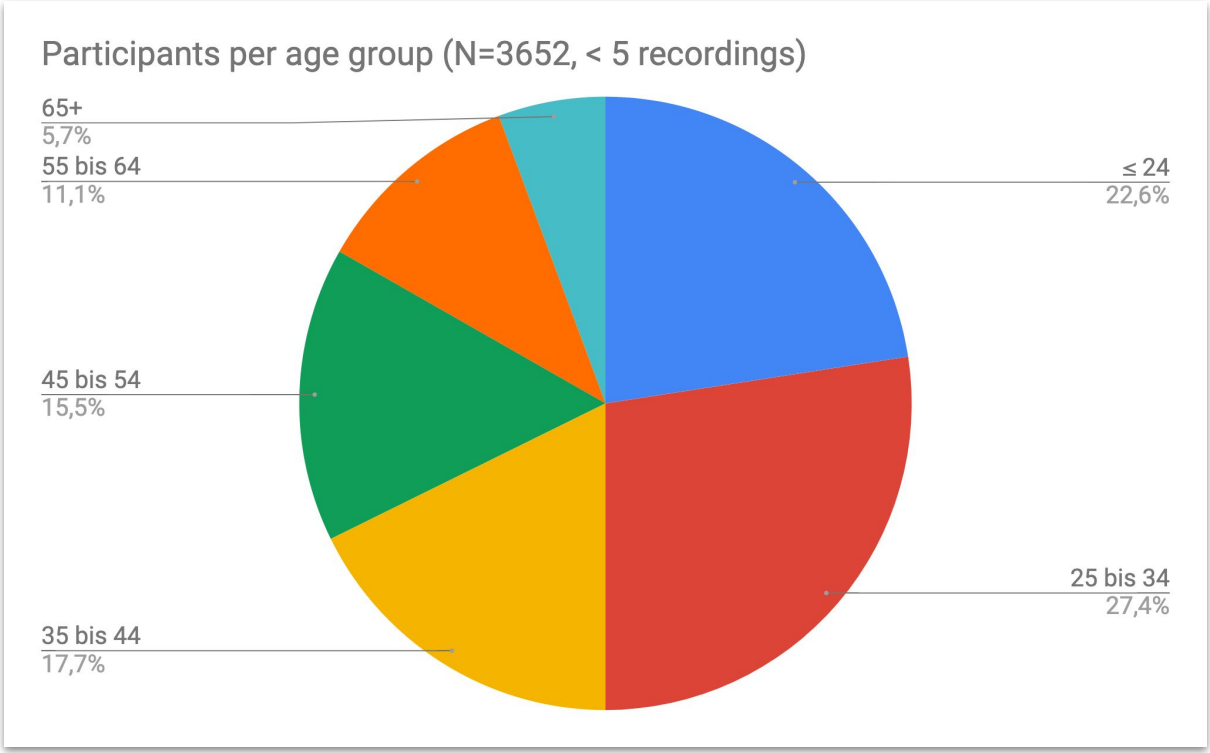
Aufbau der Items

Eine	Frau	hat	am	Freitag	die	schwarzen	Kleider	genäht.
Eng	Fra	huet	de	Freideg	déi	schw[aːʀ]z	Kleeder	gebitzt
En	Frau	hatt	e	Fregdig	d’	schw[aː]z	Kleider	gebutt
	Fr[ɔː]		um	Freddeg	di	schw[ɑʀ]z	Klegder	gebikst
			Ø	Freiden				genéit

Ta	soeur	est	géniale;	elle	raconte	souvent	les	meilleures	blagues!
Deng	Schwëster	ass	[ʒ]enial	hat	erzielt	dacks	déi	bescht	Witzer
Dein	Sëschter		[g]enial	t	verzallt	oft	di	beschte	Witzen
	Schwester			si	ziält			beste	
				se					
				et					
				‘t					

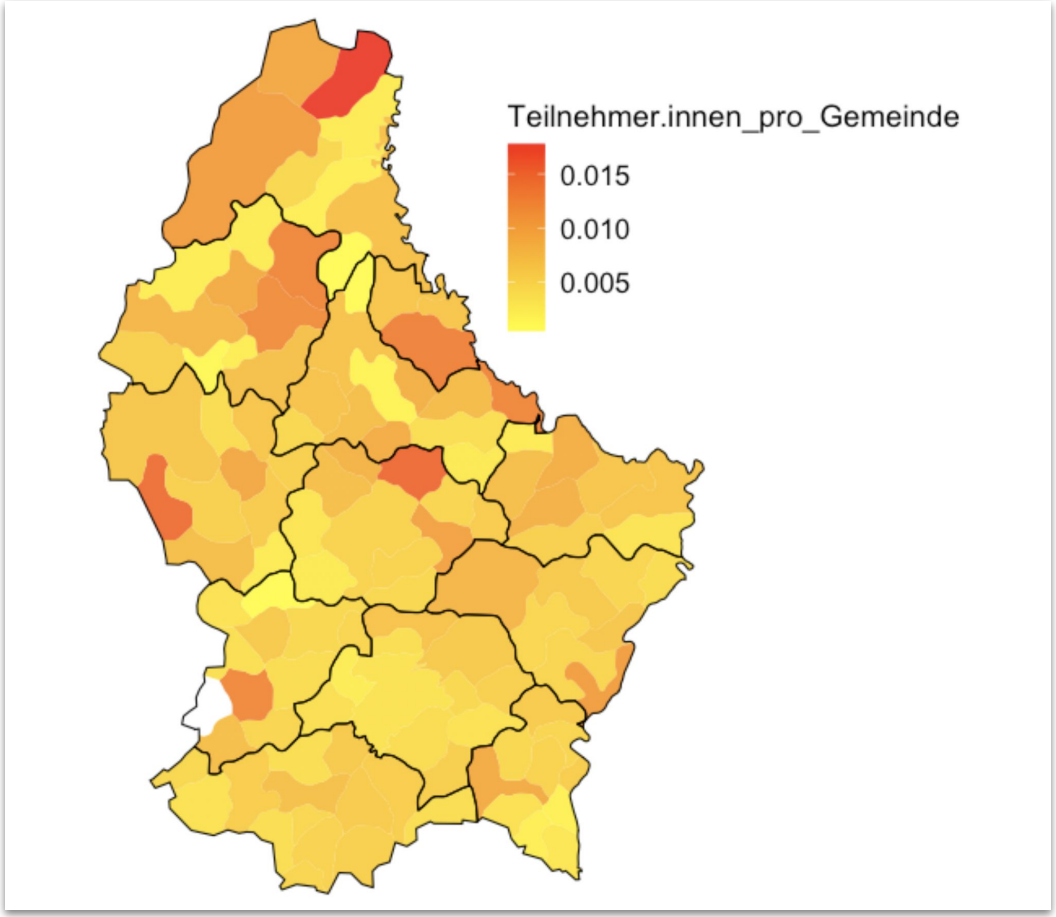
Collected Data

- Eight survey rounds since 2018
- ~ **320,000** recordings for ~ **550** items
- Up **2000** participants per survey item (mean **476**)
- Six age groups
- Decreasing rate with age
- ⅔ female, ⅓ male



Participants per Commune

- Sufficient coverage of the entire country
- 0.5 % und 1.5 % participants in relation to the number of population



Workflow of the analysis

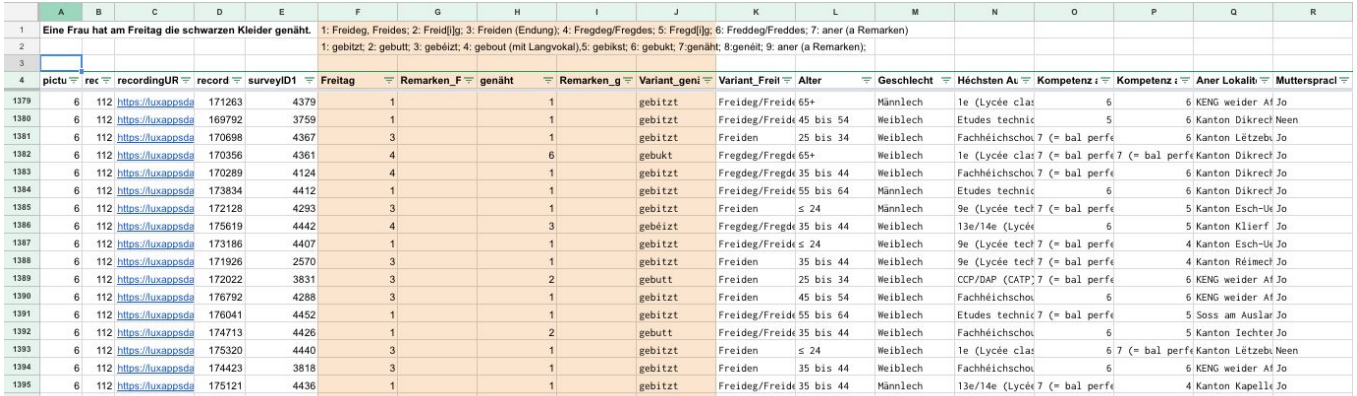
App → Preprocessing → Annotation → Analysis

- Data collection
- Geo-coding
- Archiving/Backup

- Spreadsheet per survey item
- Association with social information
- Continuous update

- Listening & annotation according to coding scheme

- Quantitative analysis
- Visualisation and map creating with R/ggplot2
- Publication in atlas (Shiny application)
- Acoustic analysis with emuR and Praat



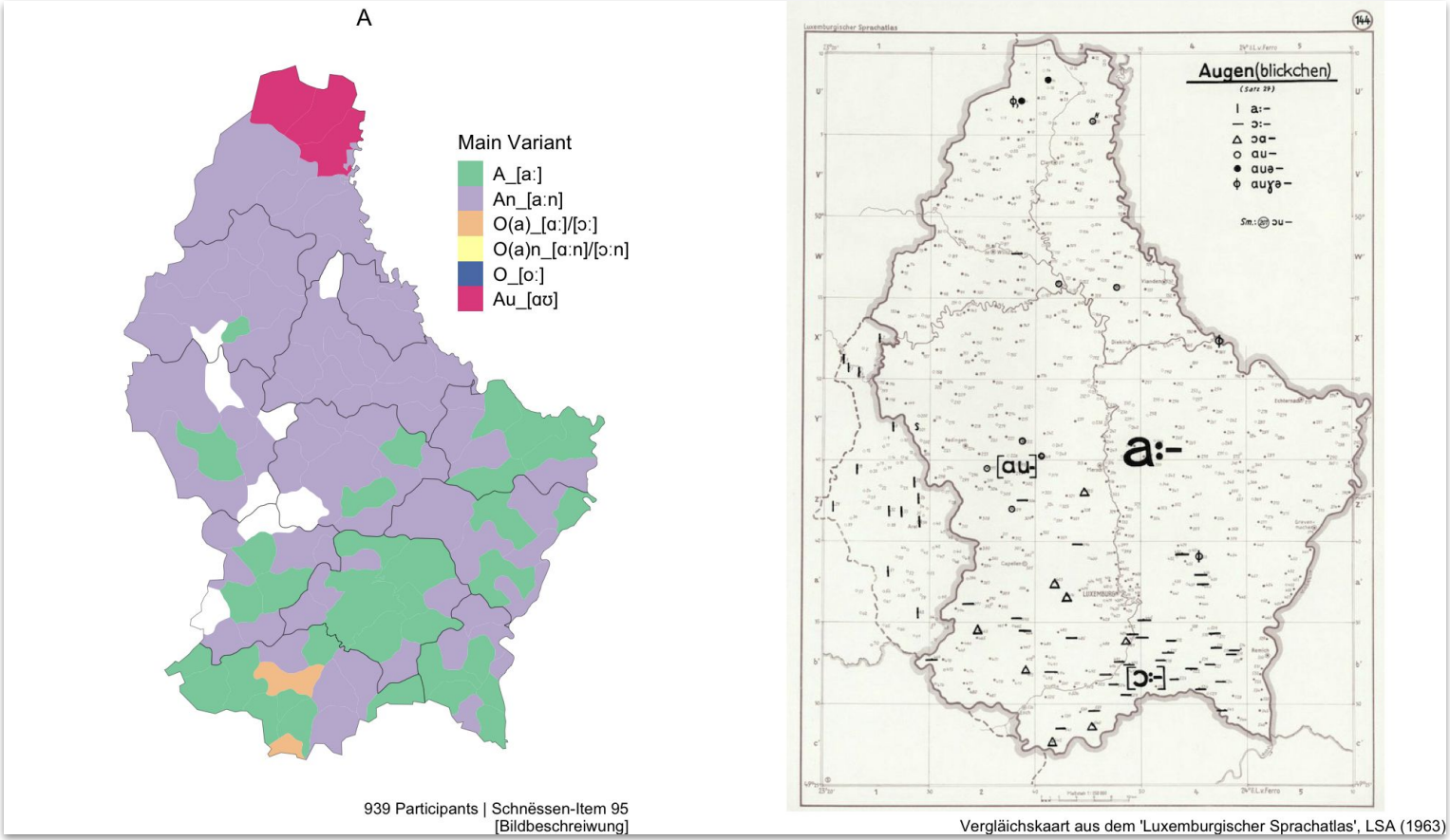
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Eine Frau hat am Freitag die schwarzen Kleider genäht.					1: Freideg, Freides; 2: Freid[i]g; 3: Freiden (Endung); 4: Fregdeg/Fregdes; 5: Fregd[i]g; 6: Freddeg/Freddes; 7: aner (a Remarken)												
2						1: gebitzt; 2: gebutt; 3: gebéizt; 4: gebout (mit Langvokal),5: gebikst; 6: gebukt; 7:genäht; 8:genéit; 9: aner (a Remarken);												
3																		
4	pictu	rec	recordingUR	record	surveyID1	Freitag	Remarken_F	genäht	Remarken_g	Variant_geni	Variant_Freit	Alter	Geschlecht	Héchsten Au	Kompetenz a	Kompetenz a	Aner Lokalit	Muttersprac
1379	6	112	https://luxappsda	171263	4379	1		1		gebitzt	Freideg/Freid	65+	Männlech	1e (Lycée clas	6	6	KENG weider Af	Jo
1380	6	112	https://luxappsda	169792	3759	1		1		gebitzt	Freideg/Freid	45 bis 54	Weiblech	Etudes technic	5	6	Kanton Dikrech	Neen
1381	6	112	https://luxappsda	170698	4367	3		1		gebitzt	Freideg/Freid	25 bis 34	Weiblech	Fachhéichschou	7 (= bal perfe	6	Kanton Lëtzebu	Jo
1382	6	112	https://luxappsda	170356	4361	4		6		gebukt	Fregdeg/Fregd	65+	Weiblech	1e (Lycée clas	7 (= bal perfe	7 (= bal perfe	Kanton Dikrech	Jo
1383	6	112	https://luxappsda	170289	4124	4		1		gebitzt	Fregdeg/Fregd	35 bis 44	Weiblech	Fachhéichschou	7 (= bal perfe	6	Kanton Dikrech	Jo
1384	6	112	https://luxappsda	173834	4412	1		1		gebitzt	Freideg/Freid	55 bis 64	Männlech	Etudes technic	6	6	Kanton Dikrech	Jo
1385	6	112	https://luxappsda	172128	4293	3		1		gebitzt	Freiden	≤ 24	Männlech	9e (Lycée tech	7 (= bal perfe	5	Kanton Esch-Ue	Jo
1386	6	112	https://luxappsda	175619	4442	4		3		gebéizt	Fregdeg/Fregd	35 bis 44	Weiblech	13e/14e (Lycé	6	5	Kanton Klierf	Jo
1387	6	112	https://luxappsda	173186	4407	1		1		gebitzt	Freideg/Freid	≤ 24	Weiblech	9e (Lycée tech	7 (= bal perfe	4	Kanton Esch-Ue	Jo
1388	6	112	https://luxappsda	171926	2570	3		1		gebitzt	Freiden	35 bis 44	Weiblech	9e (Lycée tech	7 (= bal perfe	4	Kanton Réimech	Jo
1389	6	112	https://luxappsda	172022	3831	3		2		gebutt	Freiden	25 bis 34	Weiblech	CCP/DAP (CATP	7 (= bal perfe	6	KENG weider Af	Jo
1390	6	112	https://luxappsda	176792	4288	3		1		gebitzt	Freiden	45 bis 54	Weiblech	Fachhéichschou	6	6	KENG weider Af	Jo
1391	6	112	https://luxappsda	176041	4452	1		1		gebitzt	Freideg/Freid	55 bis 64	Weiblech	Etudes technic	7 (= bal perfe	5	Soss am Auslar	Jo
1392	6	112	https://luxappsda	174713	4426	1		2		gebutt	Freideg/Freid	35 bis 44	Weiblech	Fachhéichschou	6	5	Kanton Iechter	Jo
1393	6	112	https://luxappsda	175320	4440	3		1		gebitzt	Freiden	≤ 24	Weiblech	1e (Lycée clas	6	7 (= bal perfe	Kanton Lëtzebu	Neen
1394	6	112	https://luxappsda	174423	3818	3		1		gebitzt	Freiden	35 bis 44	Weiblech	Fachhéichschou	6	6	KENG weider Af	Jo
1395	6	112	https://luxappsda	175121	4436	1		1		gebitzt	Freideg/Freid	35 bis 44	Männlech	13e/14e (Lycé	7 (= bal perfe	4	Kanton Kapelle	Jo

‘Atlas of Language Variation of Luxembourgish’

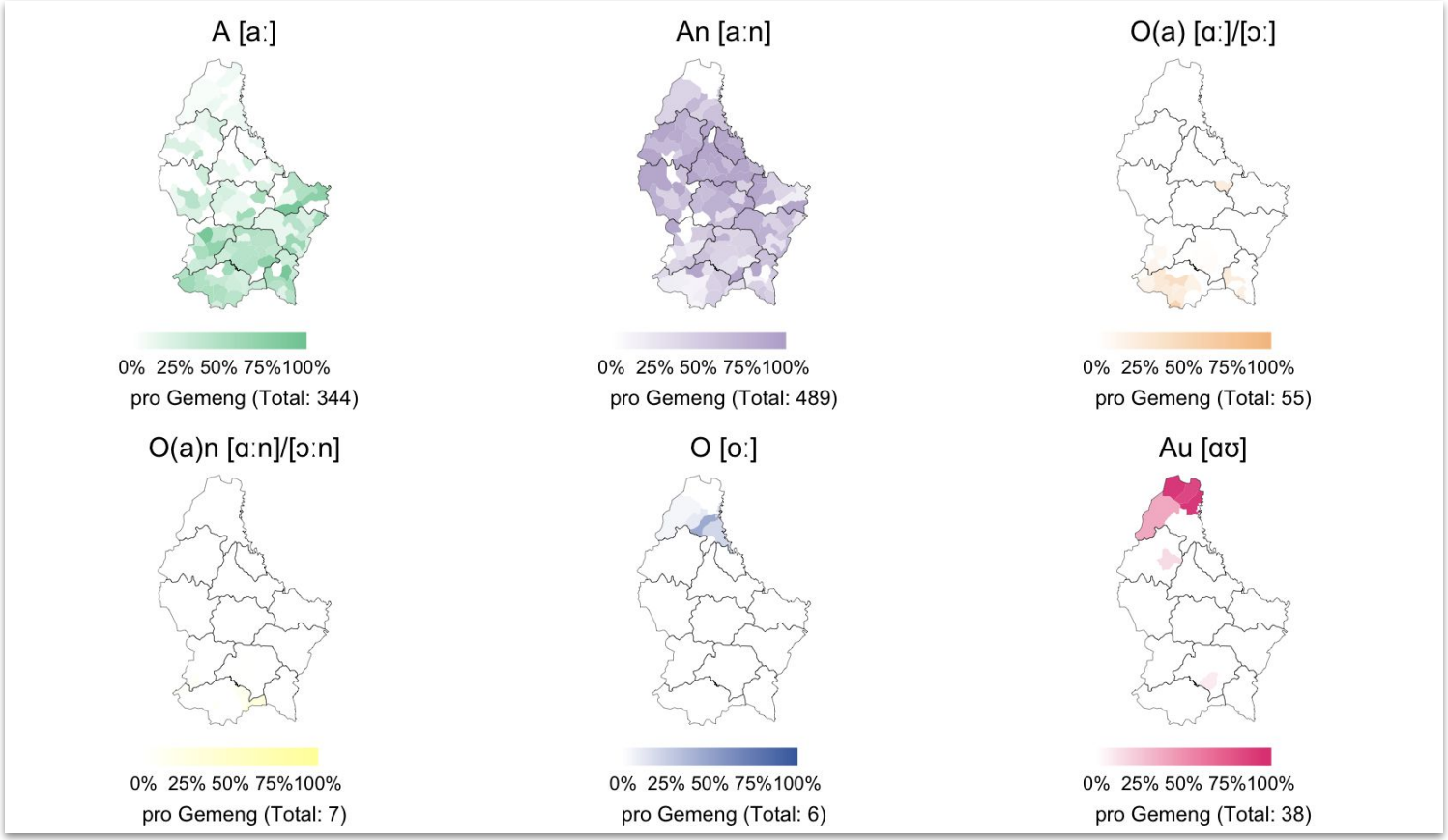
- Presentation of 600 variables
- Analysis according to
  - Regional distribution of variants
  - Age, Gender, Education, Language competencies
- Comparison with maps from the ancient ‘Luxemburgischer Sprachatlas’ (LSA 1963)
- Statistical analysis (Conditional inference trees, Random Forest)
- Audio of every recording accessible
- Available: https://infolux.uni.lu/variatiounsatlas

Regional Variation

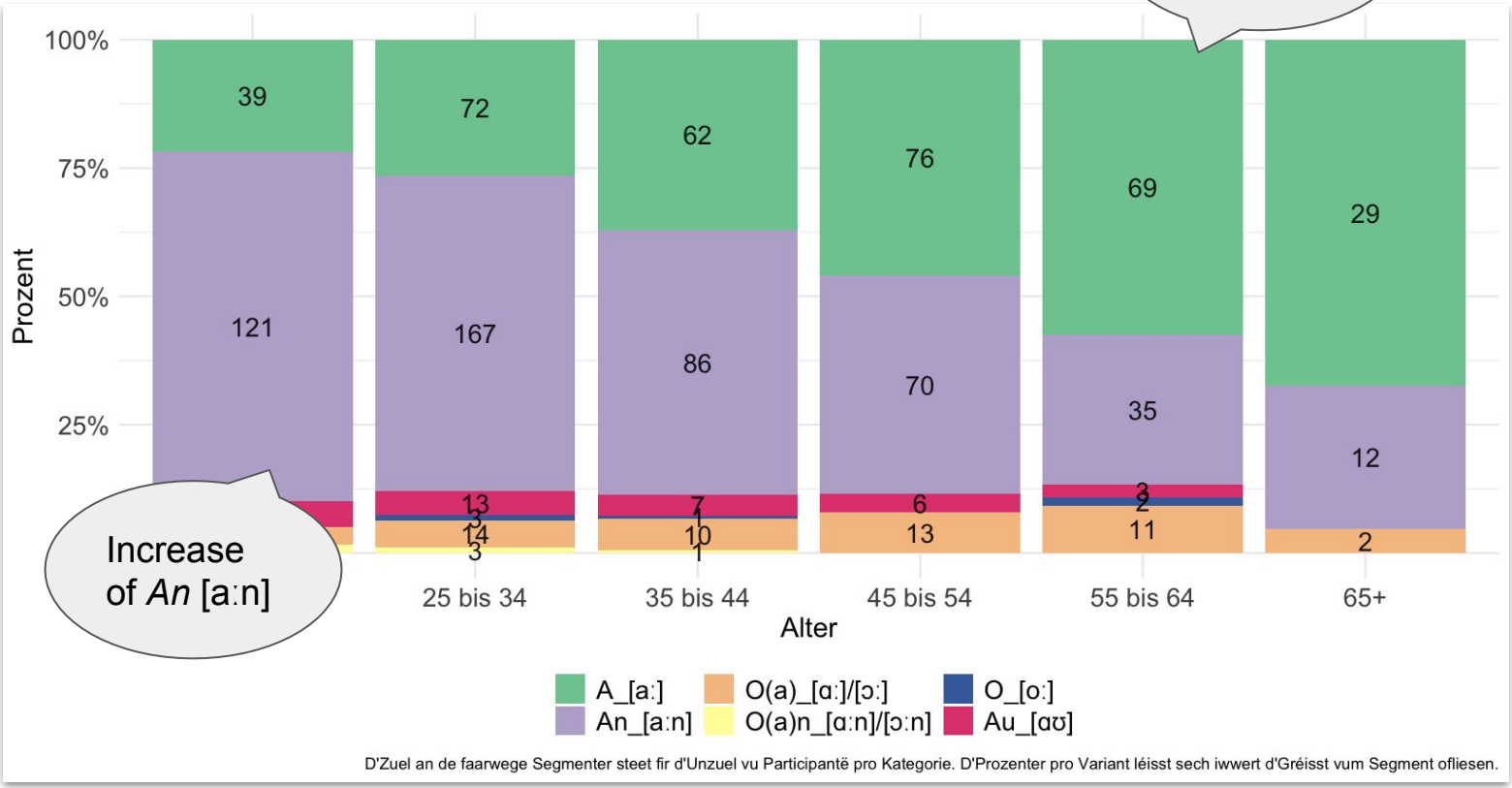
Regional Variation: *A/An* ‘eye’ (N= 938)



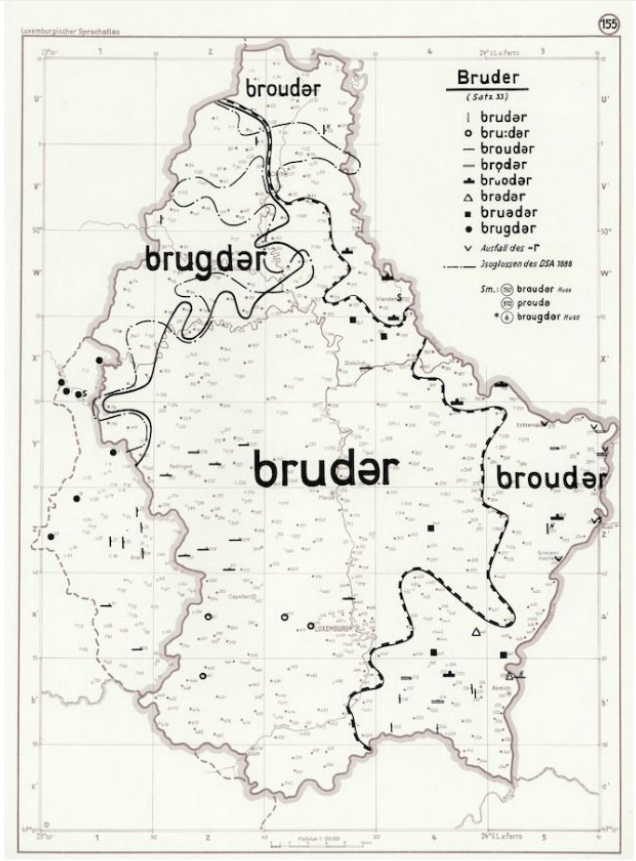
Regional Variation: *A/An* ‘eye’ (N=938)



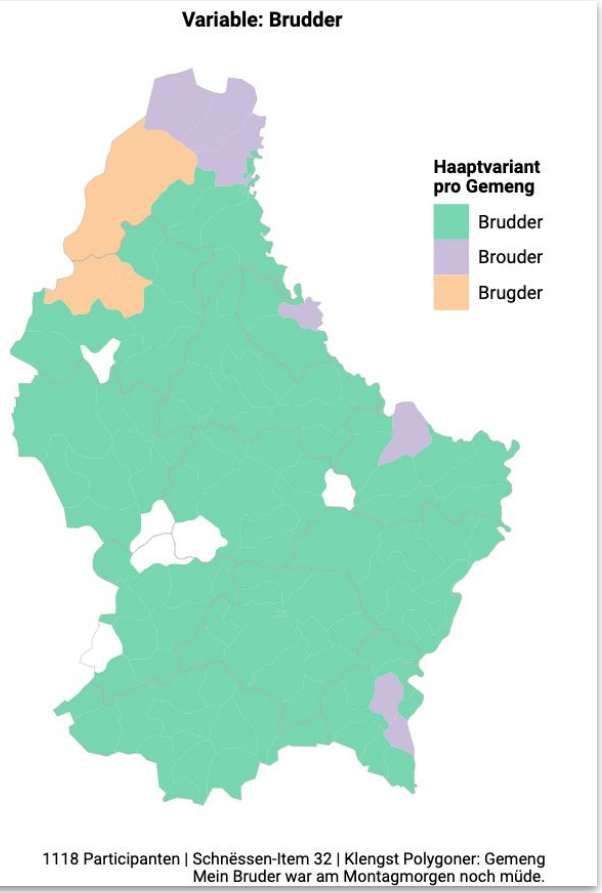
Regional Variation: *A*, *An* ‘eye’ Development in ‘apparent time’



Regional Variation: *Brudder* ‘brother’ (N=1118)



Vergläichskaart aus dem 'Luxemburgischer Sprachatlas', LSA (1963)

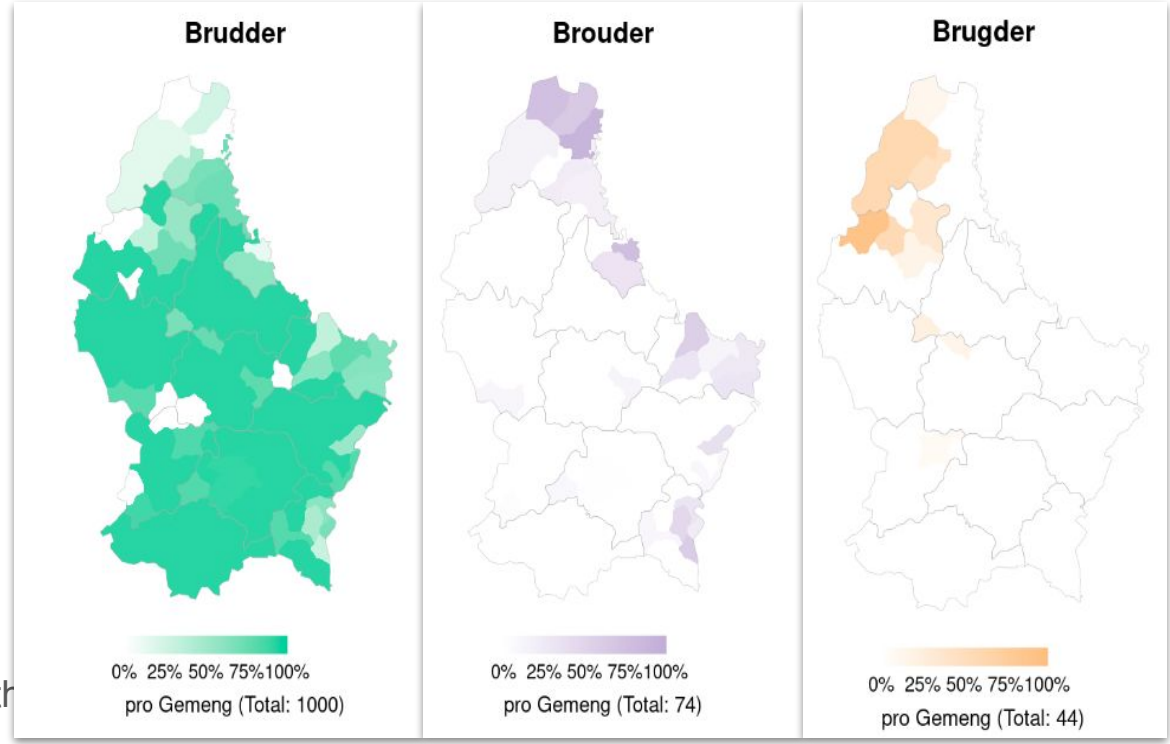


1118 Participanten | Schnëssen-Item 32 | Klengst Polygoner: Gemeng
Mein Bruder war am Montagmorgen noch müde.

Regional Variation: *Brudder* ‘brother’

- *Brudder* is spreading
- *Brouder* in the East is in decline
- *Brugder* in the North relatively stable

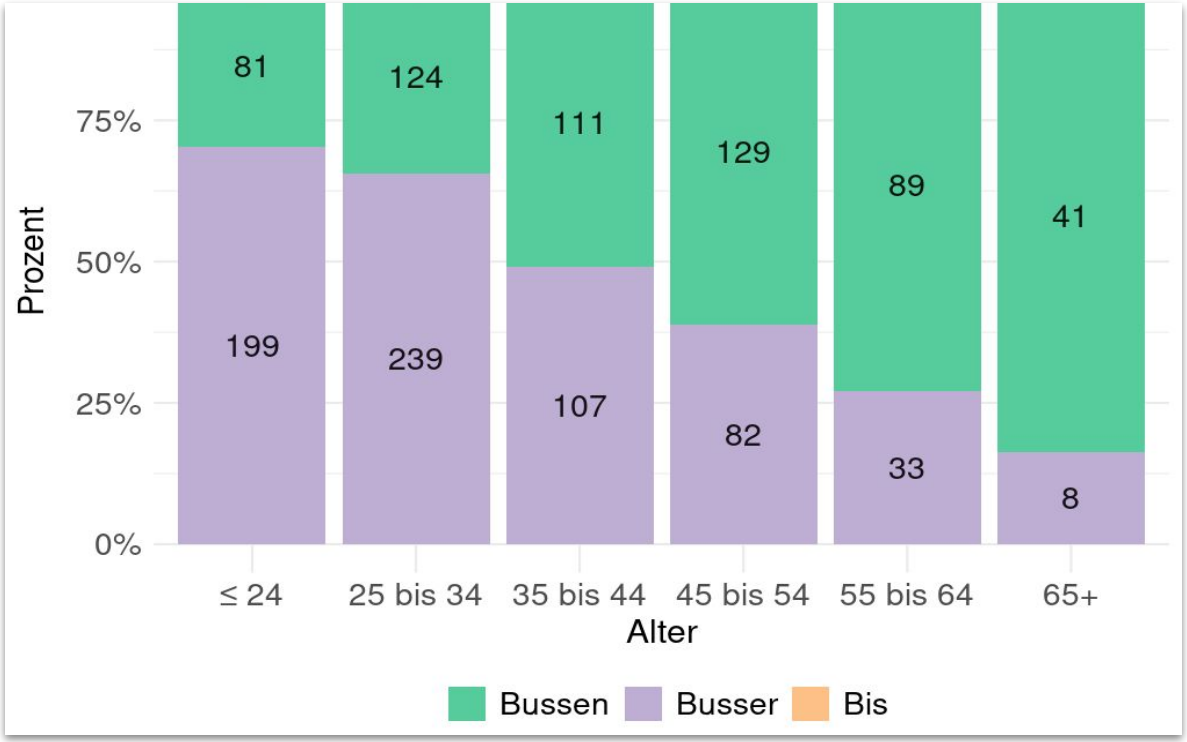
- General tendency
  - Decrease of dialect in the South and the East
  - Relative stability in the North



Morphosyntactic Variation

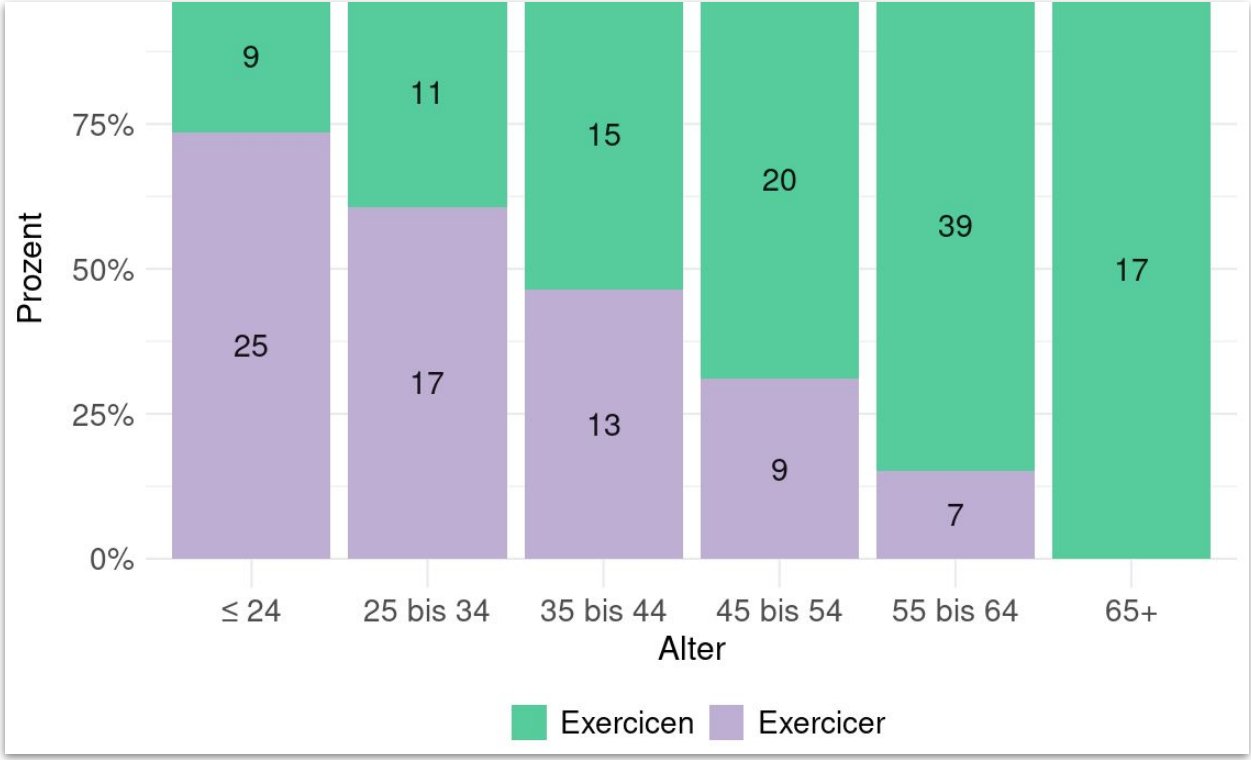
Morphosyntax: Plural *Bussen/Busser* ‘buses’ (N=1237)

- Competing suffixes *-en* and *-er* after voiceless obstruents (vgl. Entringer 2022)
- Strong increase of suffix *-er* over time
- Similar
  - *Witzen/Witzer* ‘jokes’
  - *Bicken/Bicker* ‘pens’
  - *Exercicen/Exercicer* ‘tasks’



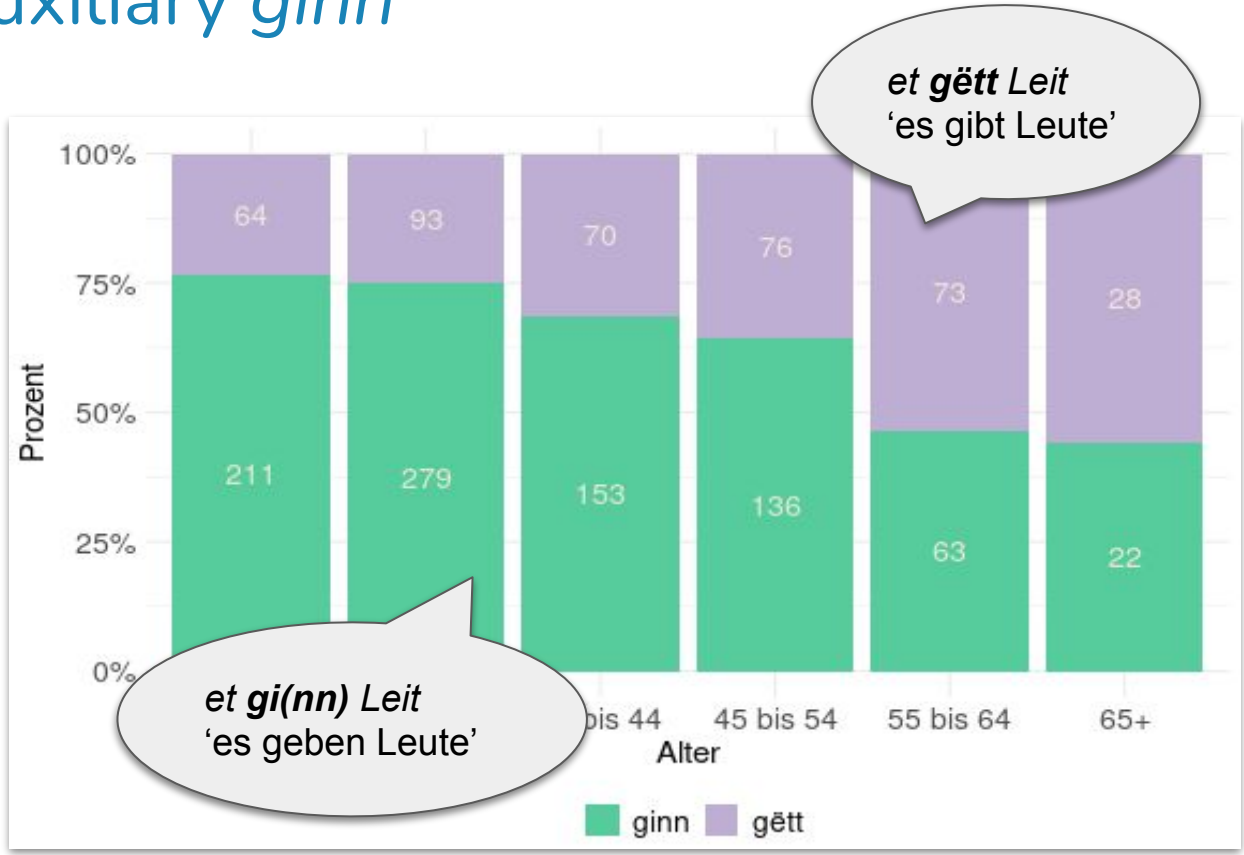
Morphosyntax: Plural *Exercicen/Exercicer* ‘tasks’

- Similar, but slightly ‘slower’ diffusion in the apparent time
- Increase of suffix *-er* is divergent to German



Morphosyntax: Auxiliary *ginn*

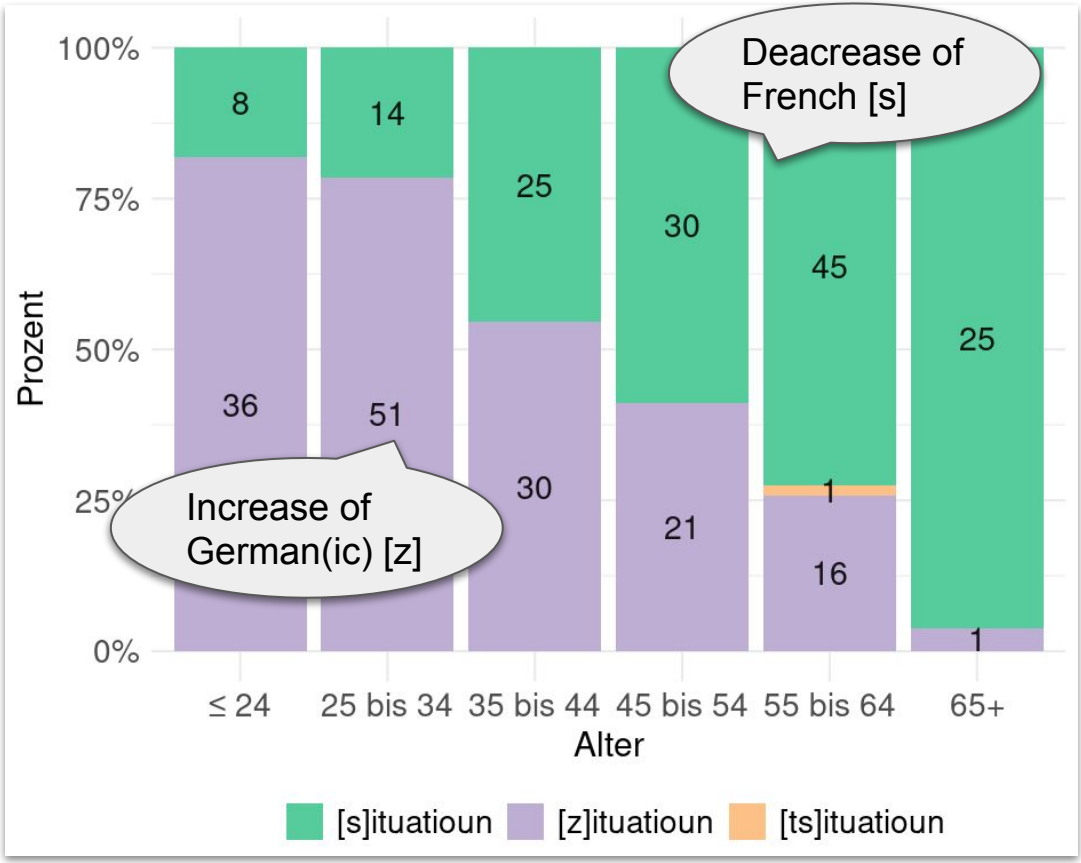
- Existential construction *et gëtt Leit* ‘there are people’/‘es gibt Leute’
- Ongoing change in number concordance
- Sg. *et **gëtt** Leit* > Pl. *et **gi(nn)** Leit*
- Divergent to German



Language Contact

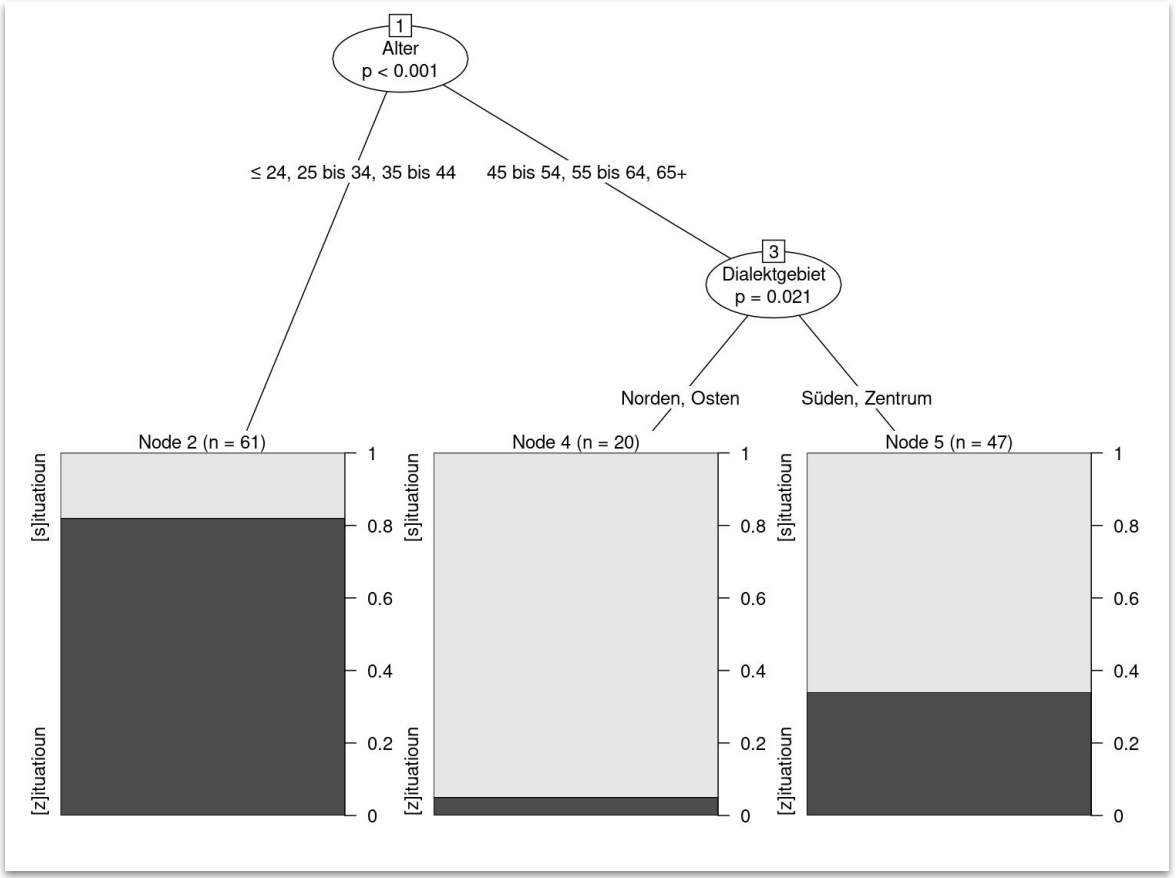
Voiceless [s] in onset: *Situatioun* ‘situation’ (N=297)

- Two options
  - French: [s]ituation
  - German: [z]ituation
- Strong increase of voiced [z] (vgl. Conrad 2017)
- Complete reversal of variant distribution
- Similar: *sensibel, sanitär, Signal, Sekretärin, Sara*



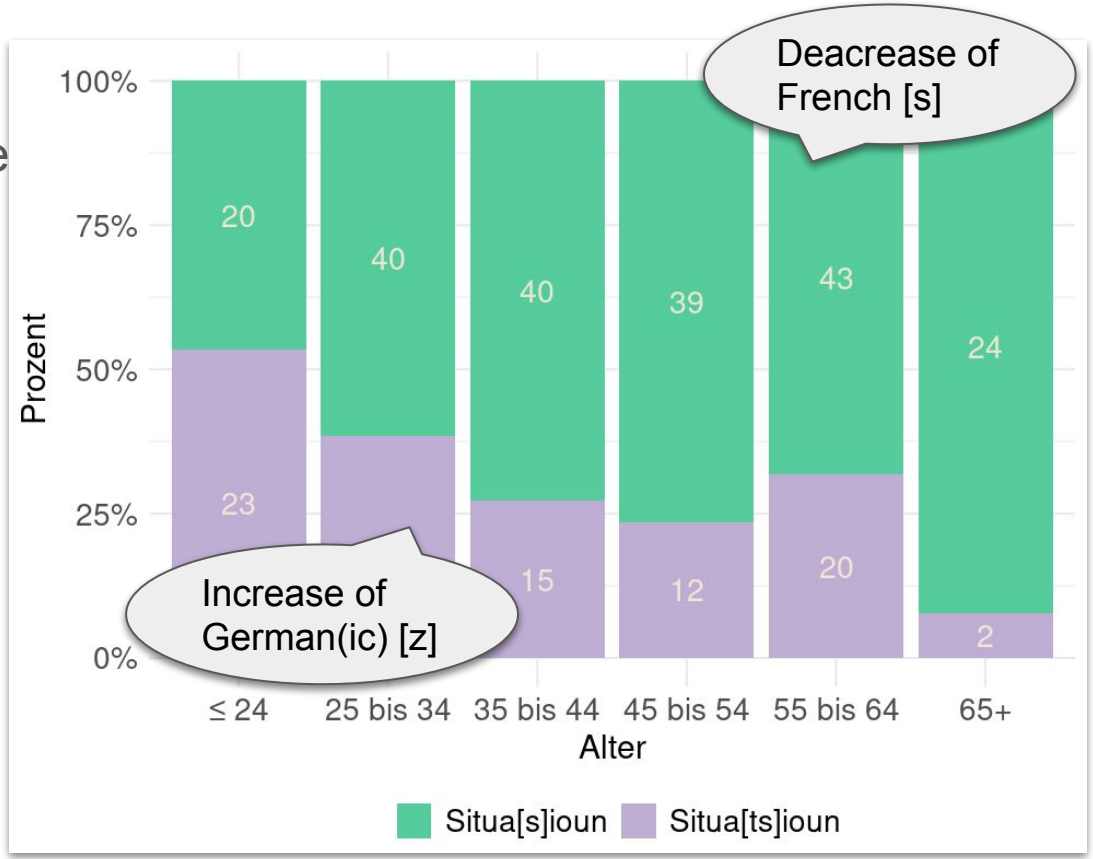
Voiceless [s] in onset: *Situatioun* ‘situation’

- ‘Conditional Inference Tree’ (vgl. Tagliamonte/Baayen 2012)
- Primary factor: age
- Secondary factor: region



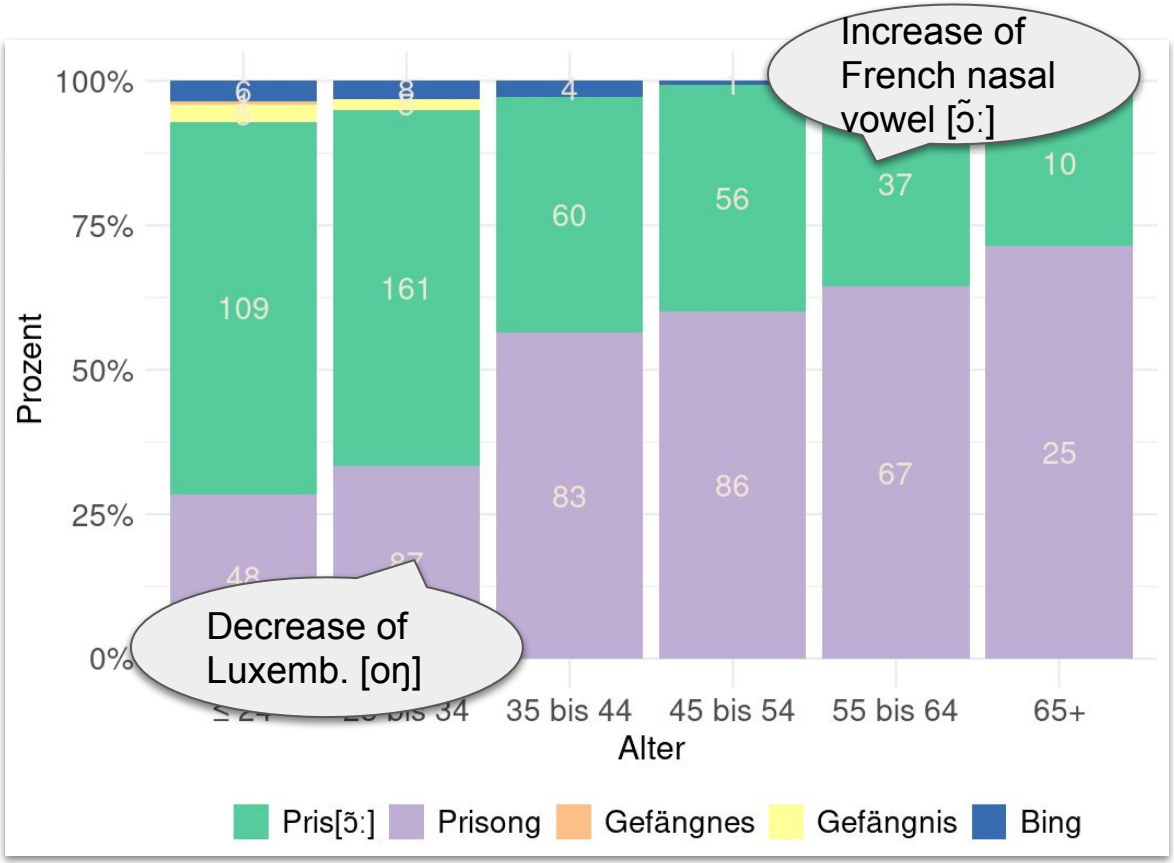
Affricate in Suffix: *Situa[s]ion* vs. *Situa[ts]ioun*

- Same trajectory as change [s] > [z], but on a lower scale
- Change is creating more variability



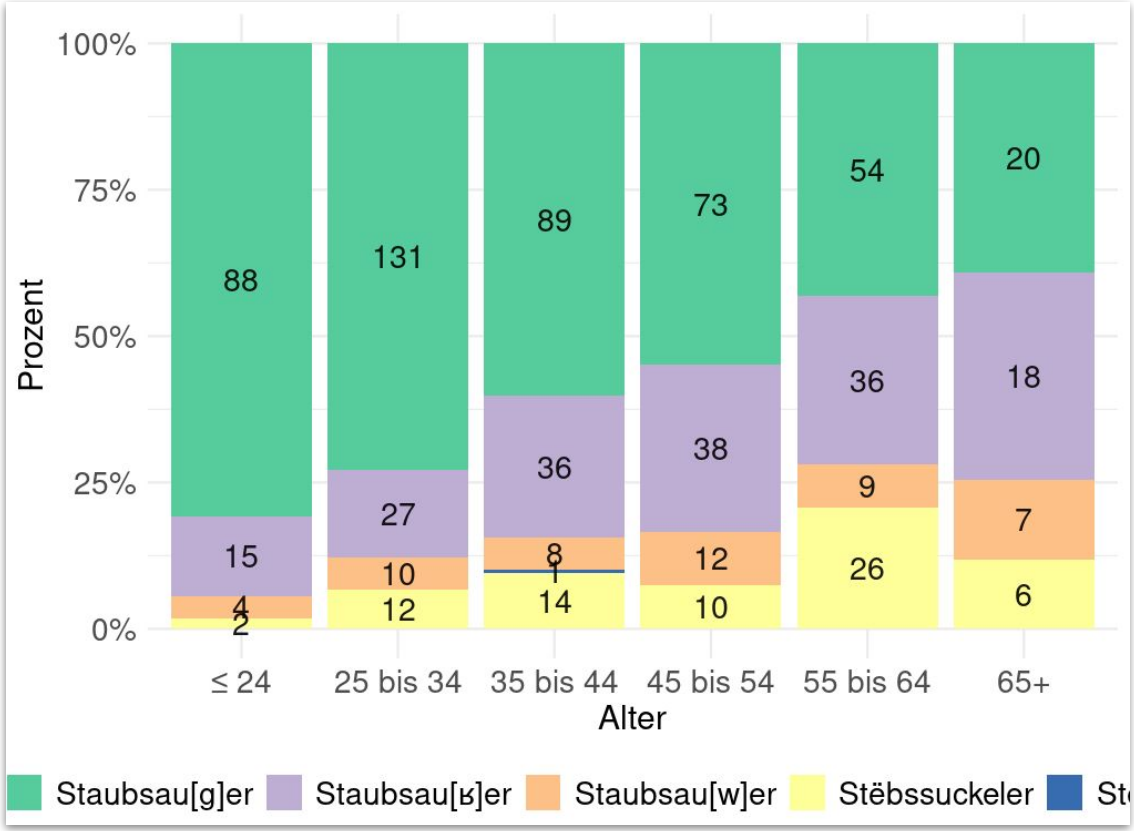
Increase of nasal vowel [ɔ̃ː]: *Prisong* ‘prison’ (N=854)

- Ongoing inference of **both** languages, German and French
- Similar: *Cartong* ‘cardboard’, *Plaffong* ‘ceiling’



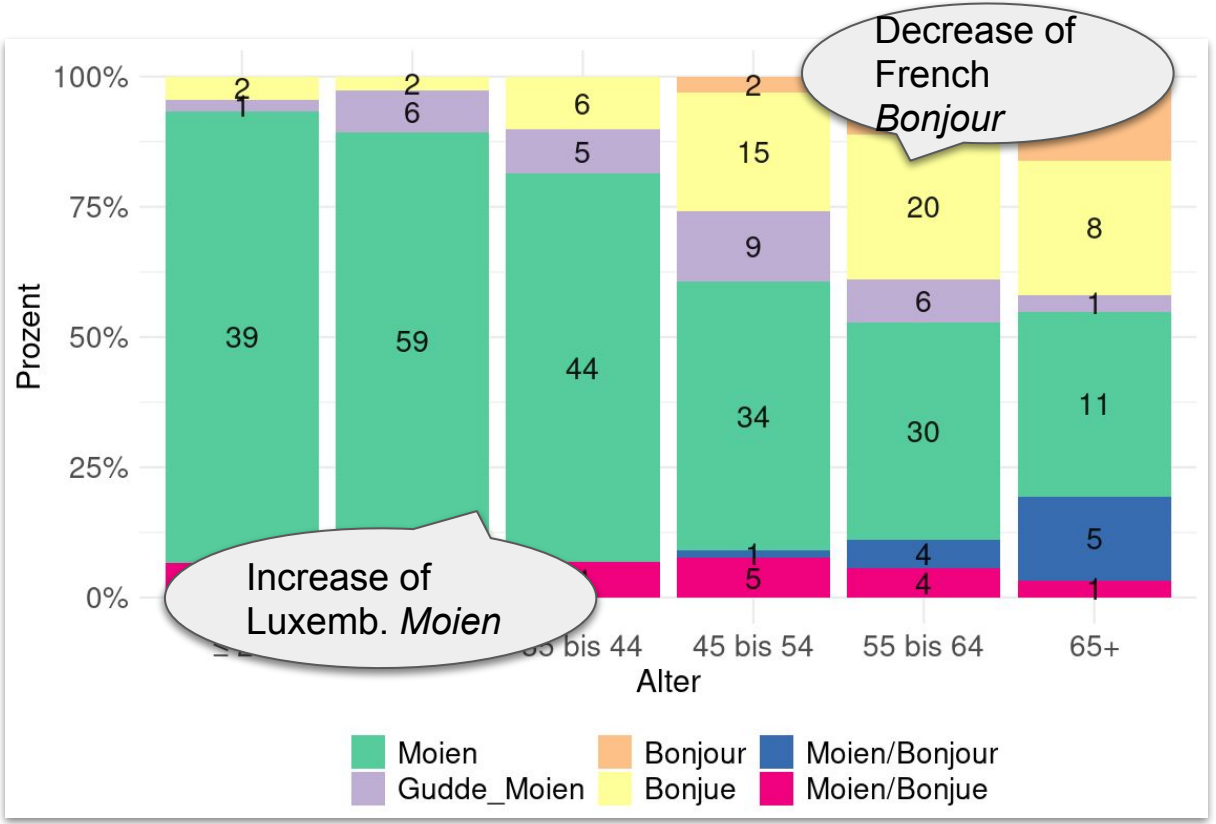
Staubsauger ‘vacuum cleaner’ (N=744)

- Decline of Luxembourgish variants *Staubsau[ʁ]er, Staubsau[w]er, Stëbssuckeler*
- Homogenisation
- Clear influence of German *Staubsauger*



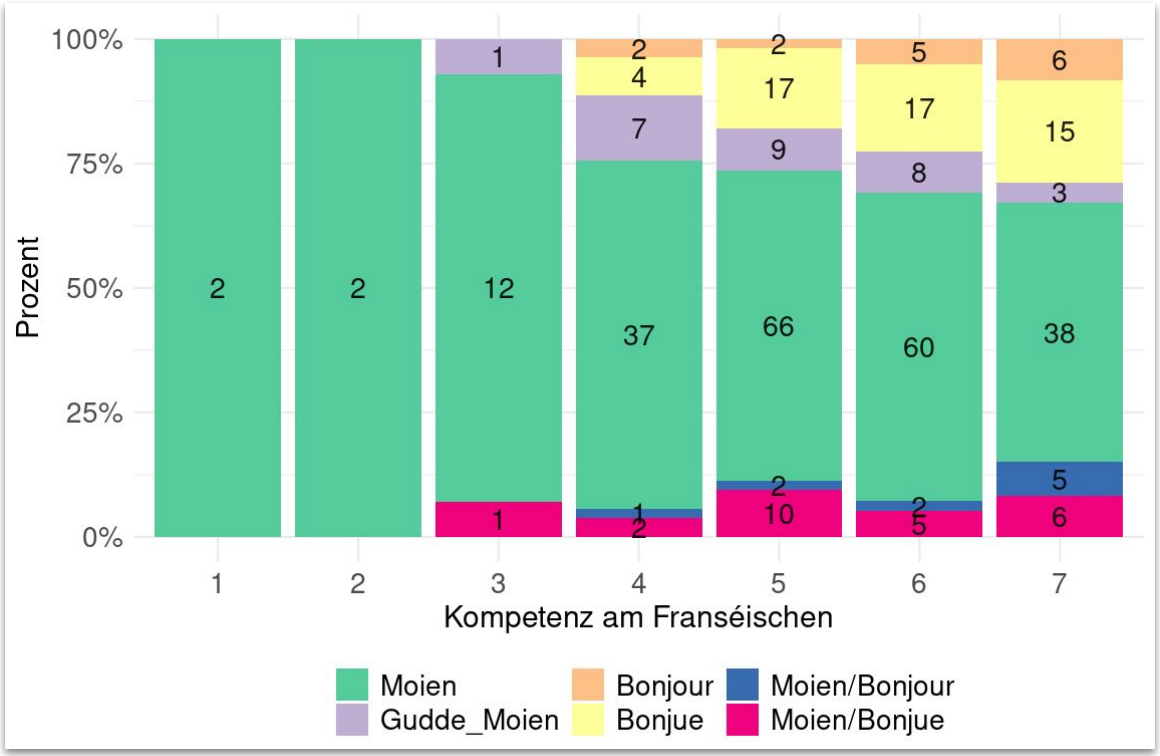
Formal greeting routine *Bonjour/Moien* (N=342)

- *Wie begrüßen Sie eine fremde Person?*
- Sign of 'Informalisation' of societal relations (?)



Formal greeting routine *Bonjour/Moien* (N=342)

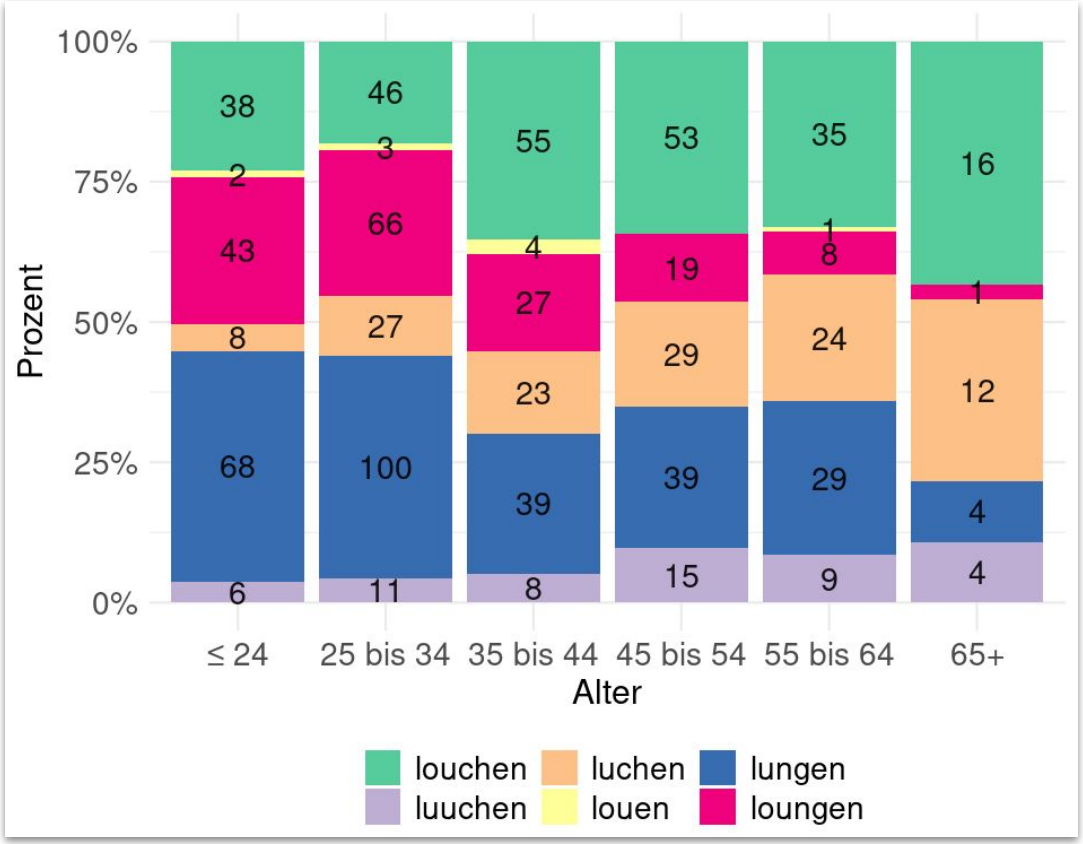
- Correlation with competence in French
- The higher the competence in French (5-7), the more *Bonjour*



Innovations

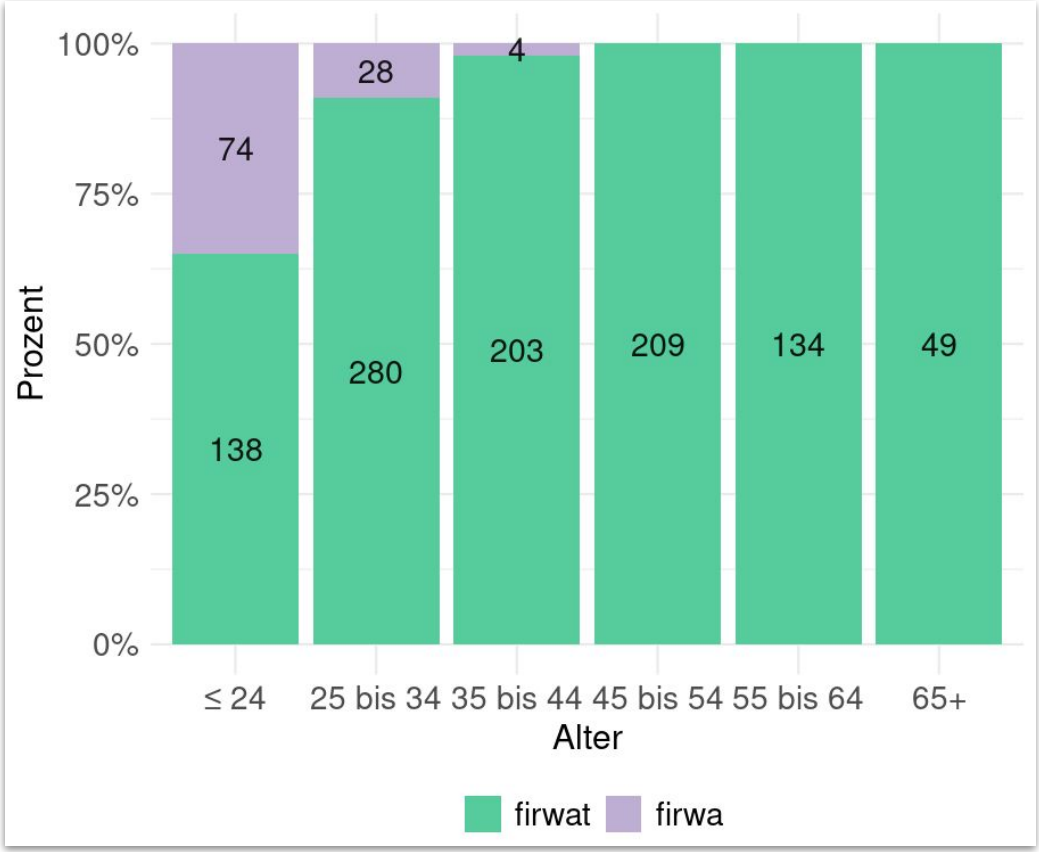
Innovation: *louchen* ‘lay’ (preterite)(N=870)

- Decrease of traditional variants *louchen, luchen, luuchen*
- Emergence of new forms *lungen, loungen*
- Analogy to similar preterites *goungen* ‘went’, *houngen* ‘hung’ (?)



Innovation: *firwat/firwa* ‘why?’ (N=1112)

- New variant *firwa* spreading fast among younger people



Discussion: Big data for a small language

- Precious corpus allowing for the first time on overview of a multitude of variation phenomena
- Dramatic changes
  - Complete reversal of variant distributions in 60 years (80/20 > 20/80)
  - Continuous decrease/increase of variants in the apparent time (no S-shape)
  - Ongoing homogenisation
- General tendencies in the analysis of apparent time
  - Further loss of regional variation
  - Increase of German interferences in phonology and lexicon
  - Decrease of interferences from French

Limitations of the corpus

- No spontaneous data
- Interferences from German or French due to the input language of the translation task
- Bird’s eye view on Luxembourgish allows a broad but quantitative substantial overview
- Missing functional perspective
    - Usage of variants for different purposes (formal/informal, accomodation to speaker etc.)
    - Passive knowledge of variants

Villmools Merci!

-

Variationsatlas: https://infolux.uni.lu/variatiounsatlas

General tendencies in the development of Luxembourgish

- Regional variation
  - Ongoing dialect levelling towards the central variety
    - Small dialects tend to decrease further (max. 25 % dialect forms in the East)
    - Variants with a wide regional spread continue to exist
  - Nördliches Dialektgebiet immer noch sehr stabil
- Sprachkontakt
  - Zunahme des Einflusses des Deutschen
  - Abnahme/Konstanz des Französischen (?)
  - specific mixture of German and French, with slight dominance of German(ic) patterns
- Innovationen / interne Wandelphänomene
  - Rapider Umbau
- Standardisierung
  - Phonetik/Phonologie: Homogenisierung
  - Morphosyntax, Lexik: Hochgradig variabel

Language contact

Hotel

Pandemie

Cent > Sent nimmt zu

Plaffong> Plaffo, Cartrong>Carton, Prisong>Prison

vulnerabel > no tendencies noticeable

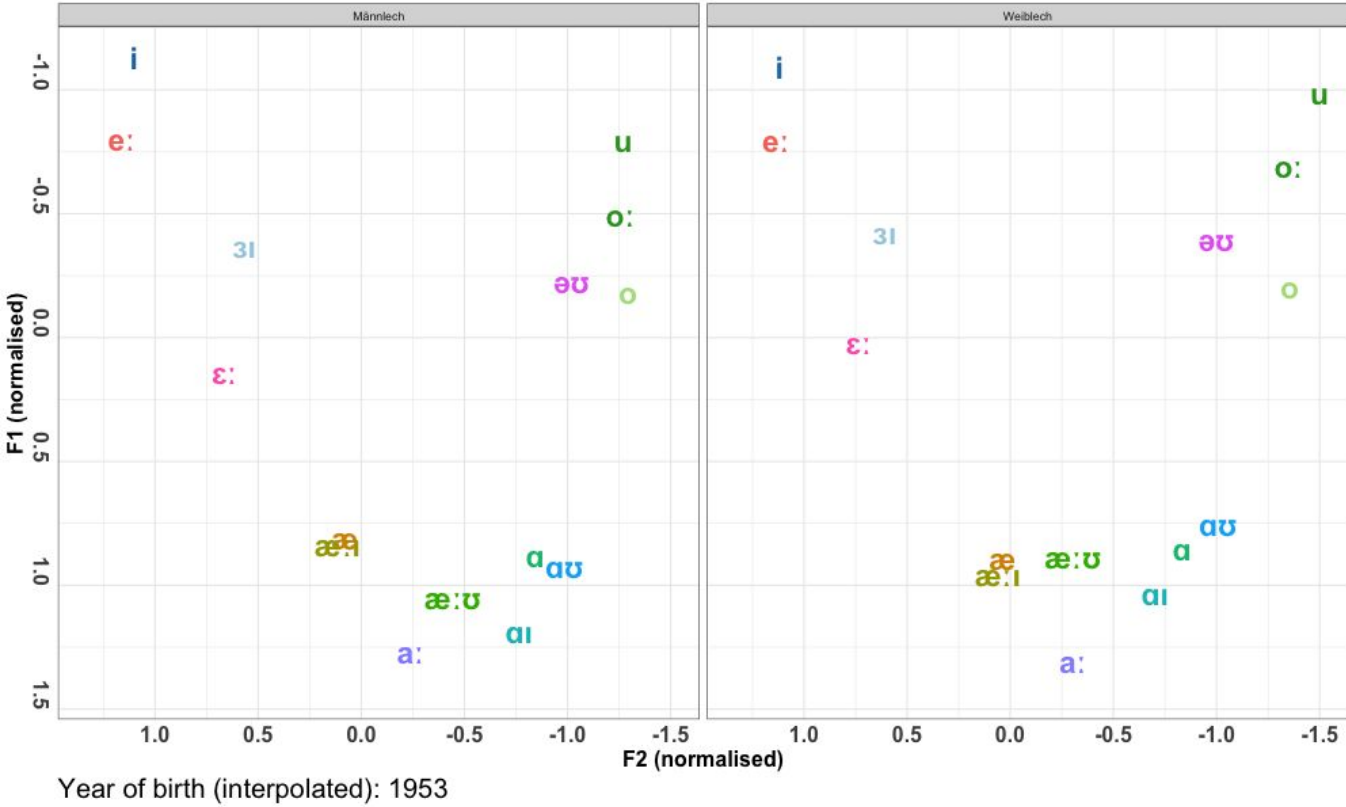
random variation on large scale; due to small networks/’families’

Phenomeener vun all linguistischem Niveau

- Phonetik-Phonologie
  - n-Reegel: *Wäi(n) beim* ‘Wein beim’, *Bee(n) gebrach* ‘Bein gebrochen’, *Ree(n) geet* ‘Regen geht’ (Gilles/Hinskens *forthcoming)*
  - Zesummefall vun ɕ an ʃ (*frech* vs. *Fräsch*) (Gilles 2019)
  - Aussprooch vu Sproochkontakt-Wierder (*sensibel*, *Situatioun*, *Persoun*, *Psychologie*) Bsp1, Bsp2, Bsp3 (vgl. Conrad 2017)
  - [j/ʑ]-Variatioun: *jäizen*, *Januar*
  - Dynamik vum Vokalsystem

Mikrofoneetik vum vokalesche Wandel (Gilles 2021)

- Rotatioune am Vokalsystem
- automatiséiert Auswäertung vun ~217.000 Vokaler
- Apparent-time Analys vu 1953 bis 2003

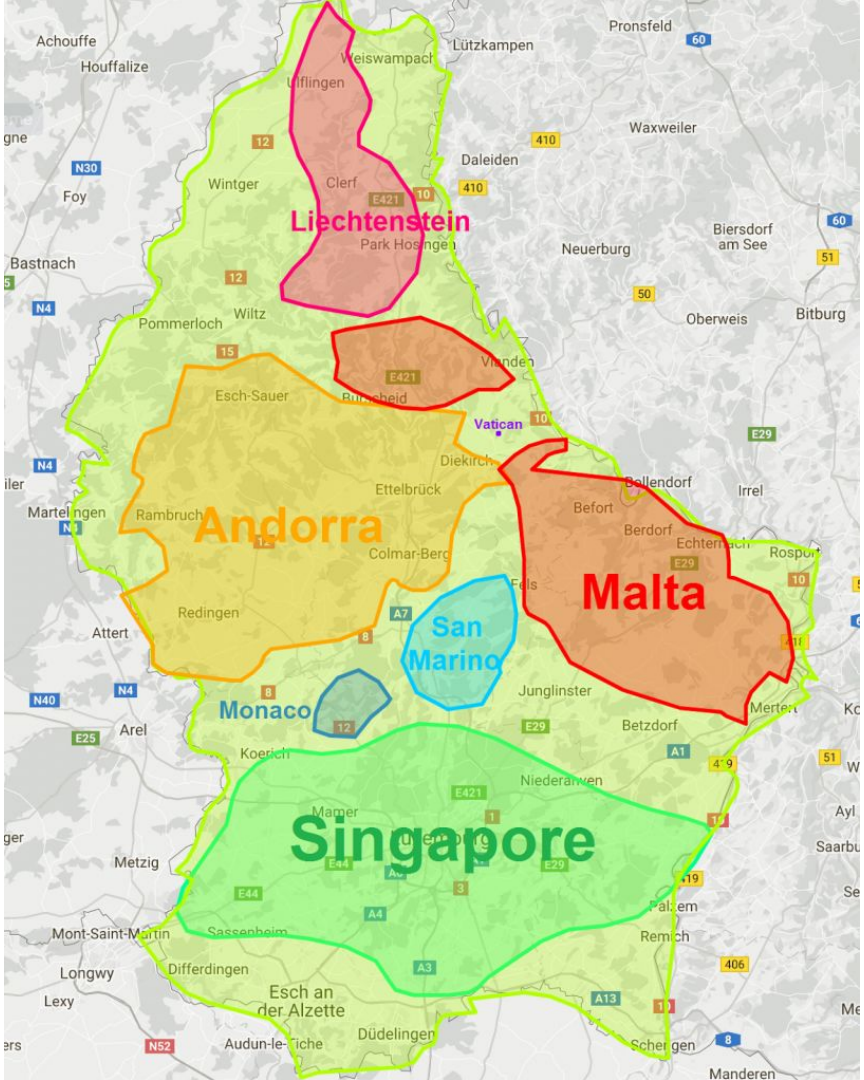
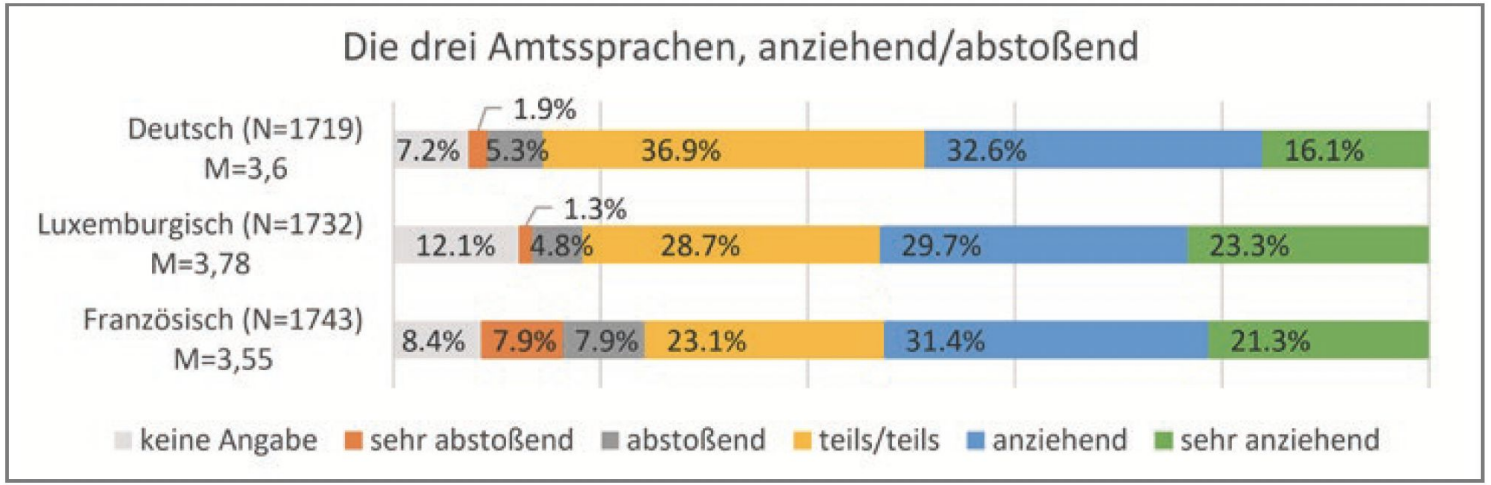


Phenomener vun all linguistischem Niveau

- Morpho-Syntax (Dissertatioun vum Nathalie Entringer)
  - Pluriel: *Club/Clibb/Clubben/Clubs*, *Walprogrammer/Walprogrammen*
  - Partizippen: *gemaach(t)*, *ka(a)f(t)*, *gefaart/gefäert*, *gewonn(en)*
  - Réckëmlautverben: *gebutt/gebitzt*, *gespiert/gespuert*, *getosch/getauscht*
  - Adjektiver
    - Superlativ (*De Claude Schmit huet dat schéinst/schéinsten/schéinstent Haus am Duerf)*
    - *dep. Adjektiver (Kanns du mir w.e.g. ee gefëllten(en) Croissant kafen.)*
    - *prenominal Adj./ Dat.Pl. (Hien huet mech mat rout/rouden Drauwen iwwerrascht.)*
  - *er*-Erweiderung bei den Adverben (*Wéi bass du elo dorop/dorobber komm)*
- Genusvariatioun bei der Referenz op weiblech Persounen *(hatt* vs. *si)* (Dissertatioun vum Sara Martin 2020)

Phenomener vun all linguistischem Niveau

- Syntax
  - staark vs. schwaach Artikelen: *dat lénkst Been/d'lénkst Been*
  - Verbcluster (*Et wier fatal, wann d'Firmaen Aarbechtsplazen ofbaue missten/missten ofbauen.*)
- Lexik
  - The usual suspects: *séier/schnell, dacks/oft, Seejomes/Ameis, Päiperlek/Schmetterling*
  - Sproochkontakt: *Poubelle/Dreckseemer, Porsche* (m./f.), *Arbitter/Schiidsriichter, Jelliskapp/Medus*
- Pragmatik
  - *Wéi begréisst Dir e gudde Frënd/eng gutt Frëndin resp. eng friem Persoun?*
  - *Wéi eng flott Wierder kennt Dir, fir aner Leit ze beleidegen?*
- Astellungsfroen
  - *Wéi nennt Dir Är Sprooch?*
  - *Schwätzt Dir e lokalen Dialekt?*


Liechtenstein
Andorra
Malta
San Marino
Monaco
Singapore
Vatican
Pronsfeld
Lützkampen
Achouffe
Houffalize
Weiswampach
Waxweiler
Daleiden
Clerf
Wintger
Foy
Bastnach
Park Hosingen
Neuerburg
Biersdorf am See
Oberweis
Bitburg
Pommerloch
Wiltz
Esch-Sauer
Ettelbrück
Diekirch
Rambruch
Martelingen
Colmar-Berg
Befort
Berdorf
Ballendorf
Irrel
Redingen
Attert
Junglinster
Betzdorf
Arel
Koerich
Niederanven
Mamer
Metzig
Mont-Saint-Martin
Sassenheim
Remich
Longwy
Lexy
Differdingen
Esch an der Alzette
Düdelingen
Audun-le-Tiche
Schengen
Manderen

Language attitudes towards French