Zeitschrift: Annalas da la Societad Retorumantscha

Herausgeber: Societad Retorumantscha

Band: 136 (2023)

Artikel: Ju sa begia sche quai è romontsch - rätoromanische Spracherwerbskorpora und das Problem der Sprachzuweisung

Autor: Cathomas, Claudia

DOI: https://doi.org/10.5169/seals-1061893

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 10.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Ju sa begia sche quai è romontsch – rätoromanische Spracherwerbskorpora und das Problem der Sprachzuweisung

Claudia Cathomas *(Institut für Kulturforschung Graubünden)*

Abstract
Rimnadas da datas linguisticas davart la lingua dal mintgadi èn in novum per la retoromanistica. En ils davos onns èn vegnids lantschads dus projects che s'occupan cun l'acquisiziun da la lingua rumantscha a l'intern ed ordaifer il territori rumantsch. Quellas datas pon esser interessantas per differentas direcziuns da perscrutaziun linguistica, sco per exempel l'analisa dal contact linguistic e dal process da midadas linguisticas e da la substituziun da la lingua. L'annotaziun da linguas en in corpus n'è dentant betg adina evidenta. Questa contribuziun duai dar ina curta survista davart ils dus projects e las difficultads metodologicas en l'annotaziun e l'analisa da fenomens da contact en corpus linguistics rumantschs.

lingua rumantscha / Romansh language – acquisiziun da la lingua / language acquisition – linguistica da corpus / corpus linguistics – code-switching – linguas minoritaras / minority languages – language tagging – emprests linguistics / borrowing – Tuatschin

1. Einführung

Die Verwendung und Weitergabe von Minderheitensprachen im Kreis der Familie ist für deren Bestehen von grosser Wichtigkeit, vor allem ausserhalb des Sprachgebiets. Spätestens seit der Forschung Fishmans (1991, 2012) wird davon ausgegangen, dass der transgenerationale Sprachgebrauch in der Familie der wichtigste Faktor für die Vitalität und den Erhalt einer Minderheitensprache ist. Dabei scheint die Wahl der Strategien in der Verwendung der Sprachen in der Familie den Erfolg in Bezug auf den Erhalt einer Minderheitensprache zu beeinflussen (cf. dazu De Houwer 2017, Schwartz 2020). Den Fragen, ob und auf welche Weise das Rätoromanische innerhalb und ausserhalb des Sprachgebiets an die nächste Generation weitergeben wird, konnte aufgrund fehlender Daten lange Zeit nicht nachgegangen werden. Erst in den letzten Jahren wurden im Rahmen von einzelnen Projekten Erhebungen sehr umfangreicher Sprachdaten durchgeführt, die die rätoromanische Alltagssprache auf möglichst naturalistische Art und Weise dokumentieren. Diese Art von Sprachkorpus stellt ein Novum für die Rätoromanistik dar, und die dadurch entstehende Datengrundlage bietet viel Potenzial für diverse linguistische Studien, z. B. zu Sprachkontakt- und Sprachwandelphänomenen oder zum Prozess des Spracherwerbs von Minderheitensprachen.

In rätoromanischen Familien ist die Alltagskommunikation oft von Einflüssen verschiedener Sprachen und Varietäten geprägt. Die Spuren des lange andauernden intensiven Kontaktes mit dem Deutschen lassen sich in der gesprochenen rätoromanischen Sprache deutlich erkennen.

Dem Sprachkontakt in der Rätoromania haben sich bereits einige sprachwissenschaftliche Studien gewidmet. Weinreich hat das Rätoromanische bereits in den 1950er-Jahren in einer Pionierstudie zum Sprachkontakt untersucht (Weinreich 1951 [2011], 1953), es folgten von den 1970er- bis zu den 1990er-Jahren weitere soziolinguistische Studien, wie die von Bernard Cathomas (1977), der die Zweisprachigkeit von Rätoromanen in Chur beschrieben hat, und Solèr (1983), der sich mit dem Sprachwandel in der surselvischen Gemeinde Lumbrein und später zusammen mit Ebneter und Toth (1989) mit dem Sprachwechsel im Domleschg und am Heinzenberg auseinandergesetzt hat (cf. auch Solèr 1999). Auch neuere Studien nahmen sich dem Thema an. Marietta Cathomas zum Beispiel analysierte in einer Fernsehproduktion die Sprache rätoromanischer Jugendlicher (M. Cathomas 2011) und im Rahmen des Nationalfondprojektes *sms4science* wurden rätoromanische SMS auf Sprachkontaktphänomene hin untersucht (C. Cathomas 2015, C. Cathomas et al. 2015). Die erwähnten Studien beziehen sich teilweise auch auf einzelne Ton- und Videoaufnahmen, doch grosse Sprachkorpora der gesprochenen und spontanen rätoromanischen Alltagssprache fehlten bis anhin ganz.

Die jüngsten Bemühungen der Erhebung grosser Sprachkorpora können deshalb als lang ersehnter Ausgangspunkt für die Dokumentation und Analyse verschiedenster Aspekte der Alltagssprache rätoromanischer Varietäten angesehen werden. Gleichzeitig ergeben sich für Forschende in der Bearbeitung eines solchen Datenvolumens etliche methodische und konzeptuelle Herausforderungen. Die folgenden Ausführungen geben einen kurzen Einblick in zwei Sprachdokumentationsprojekte zum rätoromanischen Spracherwerb. Im Anschluss soll auf die Problematik der Sprachzuweisung eingegangen werden, welche die Bearbeitung beider Sprachkorpora betrifft.

2. Sprachkorpora zum rätoromanischen Spracherwerb

2.1 Das Tuatschin-Projekt

Im Rahmen des Nationalfonds-Projekts «The morphosyntax of agreement in Tuatschin: acquisition and contact» (SNF Projekt Nr. 159544, Leitung durch Sabine Stoll und Michele Loporcaro, Universität Zürich), das sich mit Fragen der grammatischen Kongruenz im Prozess des Spracherwerbs auseinandersetzt, wurden in einem Zeitraum von drei Jahren (2017 bis 2019) zwei Arten von Sprachdaten zur surselvischen Varietät Tuatschin erhoben. Im Zentrum standen monatliche Aufnahmen von sechs tuatschinsprachigen Kindern im Alter von zwei bis vier Jahren über eine Zeitspanne von einem bis zwei Jahren hinweg. Die audiovisuellen

Aufnahmen wurden von den jeweiligen Müttern gemacht und sollen den kindlichen Spracherwerb in möglichst natürlichen Situationen festhalten (ca. 250 Stunden Aufnahmen). Zusätzlich wurden Gespräche unter Erwachsenen in Tuatschin aufgenommen (ca. 50 Stunden Aufnahmen), welche idealerweise gewöhnliche Alltagssituationen (z. B. Gespräche bei Tisch, während Autofahrten oder während Tarockpartien) möglichst naturalistisch widerspiegeln sollen.

Dieses grosse Volumen an Daten wurde auf verschiedene Arten aufbereitet. Die Gespräche wurden transkribiert, wofür vorgängig eine einheitliche Rechtschreibung des Tuatschin festgelegt wurden musste. Daraufhin wurden sie ins Deutsche übersetzt. Einen wichtigen Schritt in der Aufbereitung der Sprachdaten stellt die grammatische Annotation dar. Hierbei wurden jedem Token im Korpus unterschiedliche Eigenschaften zugewiesen. Die Annotation beinhaltete die Ebenen Lemma (die Grundform eines Wortes auf Tuatschin und Englisch), Wortklasse (*Part of Speech,* POS), Morphosyntax und Sprache. Auf der Basis von manuellen Annotationen wurde ein automatisches POS-Tagging spezifisch für das Tuatschin entwickelt (cf. Walther & Sagot 2017). Zu den Teilnehmenden wurden des Weiteren Angaben zu Alter, Wohnort und Hauptsprachen sowie Einschätzungen des eigenen Sprachgebrauchs und Spracheinstellungen erhoben.

Die Tuatschin-Korpora stellen die erste ausführliche Dokumentation der Alltagssprache dieser rätoromanischen Varietät dar. Auf der Grundlage dieser Daten und weiterer Erhebungen wurden auch grammatikalische Aspekte untersucht (Maurer-Cecchini 2017) und eine Grammatik des Tuatschin erarbeitet (Maurer-Cecchini 2021). Aufgrund der besonderen regionalen Sprachsituation stellt das Tuatschin auch für die Untersuchung von Sprachkontaktphänomenen ein attraktives Forschungsobjekt dar. Die Einflüsse von anderen Sprachen auf die rätoromanische Alltagssprache sowie das Zusammenspiel der einzelnen Varietäten Tuatschin und Sursilvan wurden im Rahmen des Projekts «Code-Switching im Tuatschin», das durch das Institut für Kulturforschung und den Universitären Forschungsschwerpunkt Sprache und Raum der Universität Zürich gemeinsam finanziert wurde, untersucht (cf. dazu C. Cathomas 2021a).

2.2 Pilotstudie Diaspora

Im Jahr 2021 wurde in Vorbereitung auf einen Antrag für ein Sinergia-Projekt beim Schweizerischen Nationalfond eine Pilotstudie zum Spracherwerb und Spracherhalt des Rätoromanischen in der Diaspora durchgeführt (C. Cathomas 2021b). Ziel der Studie war unter anderem die Erprobung des Erhebungsinstruments LENA (cf. Wang et al. 2017 und

www.lena.org). Es handelt sich hierbei um ein Aufnahmegerät, das speziell für Sprachaufnahmen bei Kindern im Spracherwerbsalter entwickelt wurde. Im Rahmen der Pilotstudie trugen sechs Kinder aus rätoromanischen Familien über eine Woche lang dieses Aufnahmegerät auf sich. Dabei wurde im Voraus jeweils eine Woche gewählt, die als möglichst typisch charakterisiert wird und somit den Alltag abbilden kann. Bei LENA handelt es sich um ein kleines und leichtes Gerät, das unauffällig an den Kleidern des Kindes befestigt werden kann und so über den ganzen Tag Audioaufnahmen generiert. Hierdurch entstehen authentische Aufnahmen des Familienalltags, wobei nicht nur die Sprache des Kindes, sondern auch die Sprachumgebung des Kindes dokumentiert wird.

Die Sprachkonstellationen in den Familien, die an der Studie teilgenommen haben, waren hierbei unterschiedlich (nur ein Elternteil rätoromanischsprachig, beide Elternteile rätoromanischsprachig, unterschiedliche Varietäten). Die Wohnorte der Familien sind auf drei Punkte der Deutschschweiz aufgeteilt: Zürich, Bern und Romanshorn.

Im Zuge der Aufnahmen wurden anhand dieser Methode rund 450 Stunden Sprachdaten erhoben. Neben den durch die LENA-Geräte aufgenommenen Daten wurde mit den Eltern in den sechs Familien eine soziolinguistische Befragung durchgeführt. Der Fragebogen sollte verschiedene Aspekte zu Sprachverwendung, Spracheinstellungen und Charakteristika des mehrsprachigen Alltags erheben und abklären, welcher Bedarf nach Hilfsmitteln zur Vermittlung des Rätoromanischen und nach Möglichkeiten zur Sprachverwendung besteht.

Die Daten der Pilotstudie wurden noch nicht transkribiert und deshalb nicht systematisch analysiert. Erste Einblicke zeigen jedoch verschiedene interessante Mechanismen und Situationen im sprachlichen Alltag der untersuchten Familien. Ein spannender Aspekt ist hierbei beispielsweise die Reaktion der Rätoromanisch sprechenden Eltern auf das auf Deutsch antwortende Kind. Bei den Aufnahmen fällt immer wieder auf, dass in solchen Situationen die vom Kind auf Deutsch gemachte Aussage im Rätoromanischen wiederholt wird, bevor das Gespräch fortgeführt wird, wie folgender Ausschnitt (1) illustriert:

(1) k1: *do duets chli bisse* ('hier juckt es ein wenig')
v1: *nua?* ('wo?')
k1: *do* ('hier')
v1: *cheu?* ('hier?')
k1: *do am Fuess und do am Fuess* ('hier am Fuss und hier am Fuss')
v1: *aha*
k1: *ehm, d'Lina isch döt. drum bi de Lina hets chli bisst*
('ehm. Lina ist da, bei Lina hat es ein wenig gejuckt')

v1: *mhm*
k1: *de het si müesse, ehm, ehm, döt uf dem Stueu si*
('dann musste sie, ehm, ehm, dort auf dem Stuhl sein')
v: 1 *mhm*
[Husten, Stimmen in der Ferne]
v1: *lu lein nus dus ir a letg plaunsiu?* ('wollen wir dann langsam ins Bett?')
k1: *nei bim Stueu duets chli bisse, do* ('nein beim Stuhl juckt es, hier')
v1: *gie, leu mordi. Lu savessen nus dus lu ir a letg, buc? Hm, tgei manegias? Ir si a trer en il pigiama, e lu meins a cucher* ('ja, dort juckt es. Wir zwei könnten dann ins Bett gehen, nicht? Hm, was meinst du? Raufgehen und den Pyjama anziehen, und dann gehen wir schlafen')
k1: *nei ich wot mit dir no spile* ('nein, ich möchte noch mit dir spielen')
v1: ***was?***

Bei manchen Eltern scheint ein solches Verhalten intuitiv stärker ausgeprägt als bei anderen. Weitere Beispiele solcher teils unbewusster Strategien zur Förderung des Rätoromanischen sind Diskussionen um rätoromanische Ausdrücke oder das Fragen des Kindes nach einer rätoromanischen Version des Gesagten. Diesbezüglich ergibt sich die interessante Forschungsfrage, wie sich solche Verhaltensweisen auf die Sprachwahl und das Sprachverständnis des Kindes auswirken können (z. B. analog zur Forschung von Schwartz 2020).

Beispiel (1) zeigt am Schluss auch den Fall eines Wechsels ins Deutsche beim rätoromanischsprechenden Vater. Neben diesen kompletten Wechseln ins Deutsche kommen bei allen Familien auch viele Fälle von inserierten deutschen Ausdrücken im Rätoromanischen vor, wie sie für das Alltagsromanische typisch sind. Eine Analyse solcher Muster könnte sich möglicherweise in Hinblick auf die Identifikation der Bedingungen für einen möglichen Sprachwechsel innerhalb der Familie als sehr lohnenswert erweisen. Der Sprachwechsel resultiert offenbar oft auch aus spezifischen lexikalischen Hürden, wie das fehlende spezifische Vokabular im Rätoromanischen (cf. dazu auch vergleichbare Ergebnisse zu den Sprechenden des Tuatschin, C. Cathomas 2021a). Die befragten Eltern gaben dazu im Fragebogen an, solchen lexikalischen Schwierigkeiten in der Vermittlung des Rätoromanischen an ihr Kind täglich zu begegnen, einerseits weil das Rätoromanische oft an den familiären Rahmen gebunden ist und deshalb spezifisches Vokabular in anderen Bereichen fehlt, andererseits weil die Dominanz des Deutschen oft die deutschen Begriffe präsenter erscheinen lässt, weshalb sie nach ihrer eigenen Einschätzung schneller zugänglich sind. Der folgende Ausschnitt zeigt einen solchen typischen Fall (2), der auch in der Familie auftritt, in der fast ausschliesslich Rätoromanisch gesprochen wird.

(2) k5: *co ha quel num?* ('wie heisst dieser hier?')
v5: *manegias quel cheu? L'antilopa? ah quel cheu? Co dian ins Stinktier... Spetga cheu sto il papi mirar suenter... ah, quei ei in scunc. Quel lai ir las detgas toffas.* ('meinst du dieses Tier hier? Die Antilope? Ah dieses hier? Wie sagt man Stinktier... Warte hier muss Papa nachschlagen... ah, das ist ein *scunc*. Der furzt ziemlich')
k5: *scunc. Quel lai ir toffas.* ('*scunc*. Der furzt.')
v5: *gie, in scunc.* ('Ja, ein *scunc*')
k5: *papi, jeu less mirar* ***kippers*** ('Papi ich möchte mir Kipper anschauen')
v5: *also gie, lu lein mirar* ***kippers*** ('also ja, dann lass uns Kipper anschauen')
k5: *gie. kippers e* ***betonmischers.*** ('ja, Kipper und Betonmischer')
v5: *gie, sas aunc quei* ***betonmischer*** *dad ier* ('ja, kannst du dich noch an den Betonmischer von gestern erinnern')

Beispiel (2) illustriert die Situation, in der ein Elternteil sich mit dem Kind ein Buch anschaut und lexikalische Hürden antrifft (das Suchen nach dem rätoromanischen Äquivalent für 'Stinktier'). Der Vater löst das Problem mit dem Nachschlagen im (wahrscheinlich digitalen) Wörterbuch. Das Verwenden der darauffolgenden deutschen Wörter *Kipper* ('Kipplaster') und *Betonmischer* zeigen Fälle, bei denen das deutsche Wort in der Alltagssprache bereits integrierter zu sein scheint und keine rätoromanische Entsprechung mehr gesucht wird.

3. Die Herausforderung der Sprachzuweisung

Die erwähnten Sprachkorpora können als wichtige Grundlage für die Erforschung von Sprachkontaktphänomenen und Sprachwandelprozessen bei Minderheitensprachen angesehen werden. In der Bearbeitung der Daten, die als Grundlage für die Analyse des Zusammenspiels der verschiedenen Sprachen und Sprachvarietäten dienen sollten, ergaben sich jedoch verschiedene methodische Schwierigkeiten.

Wenn in Korpora Sprachkontaktphänomene untersucht werden sollen, wird häufig auf ein Sprachtagging zurückgegriffen, d. h. für jedes Element im Korpus (meist auf der Wortebene) wird die jeweilige Sprache annotiert. Dieser Ansatz geht davon aus, dass Sprachen und Varietäten grundsätzlich voneinander getrennt werden können. Elemente aus anderen Sprachen werden hierbei entweder als integriert, d. h. dem Sprachsystem zugehörig, angesehen oder als *Code-Switching*, d. h. als Element, das einem anderen Sprachsystem angehört.

Das Spannungsfeld zwischen der spontanen Entlehnung (Code-Switching) und der integrierten Entlehnung wurde in der Forschung breit diskutiert (u. a. von Matras 2009, Gardner-Chloros 2009, Adamou 2016, Poplack 2018). Die Sprachkontaktforschung geht davon aus, dass die In-

tegration eines Wortes in eine andere Sprache immer einen graduellen Prozess darstellt (cf. dazu z. B. Matras 2009: 110). *Nonce-Borrowings* können nur spontan verwendet werden und durch ihre Einmaligkeit keinen Einfluss auf die Entlehnungsprozesse in der jeweiligen Sprache haben; sie können sich, bei vermehrtem Gebrauch durch mehrere Sprecher, aber auch in diesem Kontinuum im Prozess zu einer integrierten Entlehnung befinden. Die Unterscheidung von integrierten und spontanen Entlehnungen führt deshalb oft zu methodischen Schwierigkeiten. In der Forschung wurden verschiedene Kriterien diskutiert, die etwas über den Integrationsgrad eines Elements aus einer Sprache in einer anderen Sprache aussagen können. Dazu gehören a) die Regularität, d. h. die Frequenz und Verteilung in einem Sprachsystem und die Verteilung der Äquivalente (Poplack 2018), b) die Einbettung im Satz (Muysken 2000), c) die Wortklasse oder die Funktion im Satz (Myers-Scotton 1993), d) das lexikalische Feld (Adamou 2016), e) Zeichen morphologischer Integration in die Sprache, z. B. eigensprachliche Pluralendungen (Riehl 2014), f) die Markiertheit des Switches (z. B. Myers-Scotton 1993) und g) soziolinguistische Faktoren (Matras 2009). Die erwähnten Studien haben, trotz der diskutierten methodischen Schwierigkeiten, das gemeinsame Ziel, einem einzelnen Wort in einer konkreten Situation jeweils eine Sprache zuordnen zu können. Für die Analyse der Sprachkontaktphänomene in den erwähnten Sprachkorpora war es ursprünglich deshalb auch das Ziel, im Rahmen eines Sprachtaggings jedem Wort eine Sprache zuzuordnen, um danach die Verteilung der Sprachen in den Daten untersuchen zu können.

Durch den intensiven Kontakt mit dem Deutschen finden sich im Rätoromanischen viele Elemente, die ursprünglich aus dem Standarddeutschen oder aus dem Schweizerdeutschen stammen, die jedoch oft bereits im rätoromanischen Wortschatz integriert sind. Wie Beispiel (2) zeigte, situiert sich der Einfluss des Deutschen in der rätoromanischen Alltagssprache im Kontinuum zwischen integrierten (wie die Gesprächspartikel *also*) und spontanen Entlehnungen (*Stinktier*) sowie Zwischenformen, bei denen der Grad an Integration ins Rätoromanische nicht klar definierbar scheint (*kippers*). Eine Annotierung der Wörter nach ihrer etymologischen Herkunft würde deshalb den aktuellen Stand der Integration in die Sprache nicht wiedergeben. Eine Differenzierung der Sprachen gemäss dem Wörterbuch birgt ebenfalls ihre methodischen Schwierigkeiten. Zum einen sind Wörterbücher nicht immer aktualisiert und nicht immer aufnahmefreudig, was fremdsprachige Wörter angeht. Zum anderen ist die Integration oder die Wahrnehmung eines lexikalischen Elements als Fremdwort immer auch individuell und kann vom Sprechenden selbst,

aber auch von der Kommunikationssituation abhängen. Für Varietäten wie das Tuatschin fehlt ein Wörterbuch schliesslich ganz.

Eine weitere Schwierigkeit ergibt sich in der Differenzierung von Sprachvarietäten der gleichen Sprache, wie die Unterscheidung von Tuatschin und Sursilvan. Das Tuatschin ist eine dialektale Varietät des Surselvischen. Sehr viele lexikalische Elemente des Tuatschin sind homophon zum surselvischen Äquivalent und lassen sich deshalb nicht von der Standardvarietät Sursilvan unterscheiden. Es kann jedoch nicht davon ausgegangen werden, dass es sich bei diesen Homophonen immer auch um Tuatschin-Wörter handelt, nur weil sie von Sprechenden des Tuatschin produziert wurden. Dies liegt daran, dass Tuatschin-Sprecher ihre Muttersprache manchmal mit dem Surselvischen mischen, wie Beispiel (3) zeigt:

(3) ***ju*** *sa begia, quai era cartaivel cun sis onns sund'****jeu*** *ius a scoletta*
('Ich weiss nicht, das war vermutlich, mit sechs Jahren bin ich in den Kindergarten gegangen')

Das Personalpronomen der ersten Person (Tuatschin *ju*, Sursilvan *jeu*, 'ich') wird hier von einem Sprecher im gleichen Satz in beiden Varianten verwendet. Die Wörter *sa* ('weiss'; 'wissen' 1SG), *begia* ('nicht'), *quai* ('das') und *cartaivel* ('vermutlich') können gemäss der Aussprache klar dem Tuatschin zugewiesen werden, die anderen könnten jedoch sowohl Tuatschin als auch Sursilvan sein. Eine klare Sprachzuweisung ist bei solchen Wörtern in manchen Fällen deshalb gar nicht möglich.

Beim geplanten Sprachtagging war aufgrund der beschriebenen Schwierigkeiten ein Perspektivenwechsel nötig. Statt davon auszugehen, dass jedes Wort nur einer Sprache angehört, wird der Ansatz verfolgt, dass jeweils mehrere Sprachen infrage kommen. In einem Abgleich verschiedener Vollformenlexika sollen deshalb jedem Element im Korpus eine oder mehrere Sprachen zugewiesen werden, deren Gewichtung in einem weiteren Schritt durch verschiedene Faktoren (z. B. soziolinguistische Variablen wie das Alter oder situationsgebundene Faktoren) bestimmt werden. So könnten individuelle Sprecherprofile erstellt werden und eine Entlehnung je nach Situation und Gesprächsbeteiligten als integriert oder als spontan angesehen werden. Den Fragen nach der Gewichtung der Faktoren für die Bestimmung des Integrationsgrads eines Wortes und nach geeigneten statistischen Methoden für die Analyse von grossen Datenmengen muss in weiteren Schritten vertieft nachgegangen werden.

4. Fazit

Der kleine Überblick zu den im Rahmen von zwei Spracherwerbsprojekten erhobenen Sprachdaten hat gezeigt, dass die bisher eher auf qualitativen Studien basierende Forschung zu Sprachkontaktphänomenen im Rätoromanischen durch interessante Daten bereichert wurde, die aufgrund ihrer Menge und der Qualität der Aufbereitung auch für quantitative Studien verwendet werden könnten. Die Einzigartigkeit der Daten in der Rätoromanistik verdeutlicht, dass die Korpora nicht nur für die bereits erwähnten Studien, sondern auch für andere Studien, z. B. zur Phonologie oder zu morphosyntaktischen Aspekten, relevant sind.

Das klassische, in anderen Sprachkontaktstudien praktizierte Sprachtagging hat sich in rätoromanischen Korpora als schwierig erwiesen. Aufgrund der erwähnten methodischen Herausforderungen in der Abgrenzung der verschiedenen Sprachen und Sprachvarietäten wird ein Perspektivenwechsel vorgeschlagen. Statt des binären Entscheids zwischen eigensprachlichen und fremdsprachlichen Elementen sollen der Sprachkontext und die variable Heterogenität eines (individuellen) Sprachsystems miteinbezogen werden und die Individualität der Sprechenden, der Gesprächssituation und anderer Faktoren stärker berücksichtigt werden. Für diesen Ansatz werden jedoch automatische Tools, die spezifisch für rätoromanische Varietäten entwickelt werden, benötigt. Dazu gehört die Entwicklung von automatischen Taggern für die effiziente Annotation von linguistischen Informationen und die automatische Spracherkennung für verschiedene Varietäten, welche die Arbeit mit grossen Sprachdaten erleichtern und die statistische Analyse erst ermöglichen würde.

Bibliografie

Adamou, Evangelia (2016), *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*, Language Contact and Bilingualism 12, Berlin, De Gruyter.

Cathomas, Claudia / Morel, Etienne / Ferretti, Nicola / Bucher, Claudia (2015), *Same same but different: Code-Switching in Schweizer SMS – ein Vergleich zwischen vier Sprachen*, in: *Travaux neuchâtelois de linguistique* 63, 171–189.

Cathomas, Claudia (2015), *Von «I dont Know» zu «Kei problem chara!!» – Eine korpuslinguistische Untersuchung zu rätoromanischen SMS unter besonderer Berücksichtigung verschiedener Formen und Funktionen von Code-Switching*. Inauguraldissertation der Philosophisch-historischen Fakultät der Universität Bern. Bern: Universität Bern, URL: https://boristheses.unibe.ch/925/ [04-01-2023].

Cathomas, Claudia (2021a), *Ju a detg jeu fetsch bu il verstellen mei – Davart fuormas e funcziuns da code-switching el tuatschin*, in: *AnSR* 134, 103–123.

Cathomas, Claudia (2021b), *Projektbericht: Pilotstudie zum Spracherwerb des Rätoromanischen in der Diaspora*, Chur, Institut für Kulturforschung Graubünden.

Cathomas, Bernard (1977), *Erkundungen zur Zweisprachigkeit der Rätoromanen: eine soziolinguistische und pragmatische Leitstudie*, Bern, Lang.

Cathomas, Marietta (2011), *«Mirar vinavon sch'ella dat la natelnumra, ni aschia». L'influenza dal tudestg sin il sursilvan dals giuvenils*, lavur da master, Universitad da Friburg.

De Houwer, Annick (2017), *Minority Language Parenting in Europe and Children's Well-Being*, in: Cabrera, Natasha / Leyendecker, Birgit (ed.), *Handbook on Positive Development of Minority Children and Youth*, Cham, Springer, 231–246.

Ebneter, Theodor / Toth, Alfred (1998), *Die romanisch-deutsche Sprachlandschaft am unteren Hinterrhein*, Aarau / Frankfurt am Main / Salzburg, Sauerländer.

Fishman, Joshua A. (1991), *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*, Clevedon, Multilingual Matters.

Fishman, Joshua A. (2012), *Language Maintenance, Language Shift, and Reversing Language Shift*, in: Bathia, Tej K./ Ritchie, William C. (ed.), *The Handbook of Bilingualism and Multilingualism*, Malden, Wiley Blackwell, 466–494.

Gardner-Chloros, Penelope (2009), *Code-switching*, Cambridge, Cambridge University Press.

Matras, Yaron (2009), *Language contact*, Cambridge, Cambridge University Press.

Maurer-Cecchini, Philippe (2017), *Igl object indirect el tuatschin*, in: *AnSR* 130, 81–87.

Maurer-Cecchini, Philippe (2021), *A Grammar of Tuatschin*, Comprehensive Grammar Library 3, Studies in Diversity Linguistics 99, Berlin, Language Science Press.

Muysken, Pieter (2000), *Bilingual speech: a typology of code-mixing*, Cambridge, Cambridge University Press.

Myers-Scotton, Carol (1993), *Duelling languages: grammatical structure in codeswitching*, Oxford, Clarendon Press.

Poplack, Shana (2018), *Borrowing: Loanwords in the Speech Community and in the Grammar*, Oxford, Oxford University Press.

Riehl, Claudia Maria (2014), *Sprachkontaktforschung. Eine Einführung*, 3. Auflage, Tübingen, Narr Francke Attempto.

Schwartz, Mila (2020), *Strategies and practices of home language maintenance*, in: Schalley, Andrea C. / Eisenchlas, Susana A. (ed.), *Handbook of Home Language Maintenance and Development: Social and Affective Factors*, Berlin/ Boston, De Gruyter Mouton, 2020, 194–217.

Solèr, Clau (1983), *Sprachgebrauch und Sprachwandel. Eine theoretische Faktorenanalyse und die Pragmatik der Sprachbehandlung bei den Rätoromanen von Lumbrein. Mit einem Vergleich der Germanisierung in Präz und Sarn*, Dissertation, Universität Zürich.

Solèr, Clau (1999), *Sprachwandel als Zeichen der Vitalität – oder als Vorzeichen des Sprachwechsels?* in: Kattenbusch, Dieter (ed.), *Studis romontschs. Beiträge des Rätoromanischen Kolloquiums, Giessen/Rauischholzhausen, 21.–24. März 1996*. Wilhelmsfeld, Gottfried Egert, 95–108.

Walther, Géraldine / Sagot, Benoît (2017), *Speeding up corpus development for linguistic research: language documentation and acquisition in Romansh Tuatschin*, in: Degaetano-Ortlieb, Stefania / Kazantseva, Anna / Reiter, Nils / Szpakowicz, Stan (ed.), *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Vancouver, Canada, Association for Computational Linguistics, 89–94.

Wang, Ye / Hartman, Maria / Abdul Aziz, Nurul Akmar / Arora, Sonia / Shi, Lingyun / Tunison, Ellie (2017), *A systematic review of the use of LENA technology*, in: *American Annals of the Deaf* 162/3, 295–311.

Weinreich, Uriel (2011), *Languages in contact. French, German and Romansh in twentieth-century Switzerland. With an introduction and notes by Ronald I. Kim and William Labov*, Amsterdam/Philadelphia, Benjamins ([1]1951).

Weinreich, Uriel (1953), *Languages in contact. Findings and problems*, New York, Linguistic circle of New York.

Dr. Claudia Cathomas, Institut für Kulturforschung Graubünden, Chur, claudia.cathomas@kulturforschung.ch, ORCID 0009-0000-1622-3784