

# 20 Jahre digitaler Sprachatlas VIVaio Acustico delle Lingue e dei Dialetti d'Italia (VIVALDI)

*Fabio Tosques & Michele Castellarin, Humboldt-Universität zu Berlin*

Zu den zahlreichen schützenswerten kulturellen Gütern in Italien gehören zweifelslos auch die vielen Dialekte und Minderheitensprachen. Während sich Literatur in Texten, Architektur in Bauwerken und Musik in der klanglichen Darbietung ausdrückt, leben Sprachen besonders von der akustischen Realisierung ihrer Sprecher. Bedauerlicherweise liegen im Bereich der romanischen Sprachgeographie allgemein zugängliche Daten bisher fast ausschließlich in gedruckter Form vor. Wenn man einmal absieht von einer Reihe von Schallplatten (v.a. zu italienischen Dialekten), die jedoch erhebliche Nachteile aufweisen (teilweise schlechte Aufnahmequalität, Abnutzung, besonders aber geringe Kapazität), gibt es bisher für Italien kein systematisch erhobenes authentisches auditives Sprachmaterial, das für Forschungszwecke zur Verfügung stünde. Eine Ausnahme bildet der 1998 erschienene Atlas des Dolomitenladinischen (ALD), zu dessen Lieferumfang auch 3 CD-ROM bzw. 1 DVD gehören (vgl. Bauer/Goebel/Haimerl 2005 und <http://ald1.sbg.ac.at/>). Die Dokumentation der sprachlichen Einzigartigkeit Italiens hat sich das Projekt VIVALDI (VIVaio Acustico delle Lingue e dei Dialetti d'Italia / Akustischer Sprachatlas der Dialekte und Minderheitensprachen Italiens) zur Aufgabe gestellt. Das Hauptziel besteht darin, authentisches und aktuelles phonetisches Material zu sammeln und der Wissenschaft und Forschung zur Verfügung zu stellen, das besonders die kontemporären Einflüsse des Standarditalienischen auf die untersuchten Dialekte und Minderheitensprachen sichtbar und hörbar machen soll, den Vergleich der vielen verschiedenen Dialekte untereinander erlauben soll und auch den Vergleich auf phonetischer Ebene mit älteren Sprachstufen zulässt (z.B. AIS, Atlas Italiens und der Südschweiz).

## 1 VIVALDI - Von der Idee bis zur Realisierung der digitalen Versionen

Sprachatlanen zählen im Bereich der Sprachgeographie und Geolinguistik zu den wichtigsten Arbeitsmitteln in der Forschung. Als Forschungsinstrument hat auch in Italien der Sprachatlas eine lange Tradition, wobei die bisherigen ausschließlich in gedruckter Form erschienen sind. Im Umfeld des ALD I (*Atlante linguistico del ladino dolomitico e dei dialetti limitrofi*)<sup>1</sup> entstand die Idee eines akustischen Sprachatlases, der alle Regionen Italiens abdeckt.

Die ersten Sprachatlanten, die über die romanische Sprachgeographie hinaus methodologische Maßstäbe setzten, sind der ALF (*Atlas linguistique de la France*) aus den Jahren 1902 bis 1910 (Gillieron & Edmont 1902-1910) und der AIS (*Sprach- und Sachatlas Italiens und der Südschweiz*) aus den Jahren 1928 bis 1940 (Jaberg/Jud 1928-40), die in Form von gedruckten, zu mehreren großformatigen Bänden gebundenen Sprachkarten vorliegen, jedoch kein authentisches Tonmaterial enthalten.

Neuere Initiativen wie beispielsweise das in den 80er Jahren von Sobrero geleitete Projekt NADIR (*Nuovo Atlante del Dialetto e dell'Italiano per Regioni*, vgl. Sobrero 1991) oder das von Ruffino initiierte Projekt des sizilianischen Sprachatlases (ALS, vgl. Ruffino 1986) stellen bis heute der Öffentlichkeit vom gesammelten erhobenen Material nichts bzw. lediglich gedruckte Kostproben zur Verfügung.

Mit der Schallplatte steht zwar seit den zwanziger Jahren des 20. Jahrhunderts ein Medium zur Verfügung, das eine Konservierung und größere Verbreitung sprachlichen Materials ermöglicht, aber die Aufnahmepraxis war umständlich und kostspielig, so dass

<sup>1</sup> [http://www.sbg.ac.at/rom/people/proj/ald/ald\\_home.htm](http://www.sbg.ac.at/rom/people/proj/ald/ald_home.htm)

sie kaum in der Feldforschung eingesetzt wurde. Es gibt eine Reihe von Schallplatten – z.B. zu italienischen (Cortelazzo 1974-88) oder zu schweizerdeutschen (Hotzenköcherle/Brunner 1972-76) Dialekten –, die jedoch erhebliche Nachteile aufweisen: schlechte Aufnahmequalität, Abnutzung, besonders aber geringe Kapazität sowie keine direkte Verbindung des visuellen Elements (Karte) mit dem auditiven (Ton).

Mit der Einführung der CD und später der DVD wurde zwar das Problem der Kapazität geringer und die Einführung von Aufnahmegegeräten wie MiniDisc und DAT-Recorder erleichterte die Aufnahmepraxis. Dennoch sind die inzwischen digital aufgezeichneten und digitalisierten Tondokumente, die seit den 1990er Jahren auf CD erschienen sind, in der Praxis nur schwer erhältlich.

Heute findet man natürlich zahlreiche dialektale Kostproben in Portalen wie Youtube usw., die jedoch völlig unsystematisch dargeboten werden und daher für wissenschaftliche Zwecke weitgehend unbrauchbar sind. Neuere Projekte wie das von Google unterstützte Projekt *endangered languages*<sup>2</sup>, welches die weltweit bedrohten Minderheitensprachen katalogisieren möchte, sind ebenfalls unsystematische Sammlungen von Video- und Audioaufzeichnungen der jeweiligen Sprachen und für die wissenschaftliche Arbeit schwer zu verwenden.

Die Einführung moderner, computerunterstützter Verfahren der Tonaufnahme, -konservierung, -verarbeitung und -wiedergabe erlaubte eine qualitativ hochwertige, verlustfreie und kostengünstige Verbreitung der Sprachdokumente. Der Sprachatlas ALD I erschließt der Sprachgeographie diese Neuerungen erstmals in größerem Rahmen (Goebel 1998): Die abgefragten Wörter und Sätze (Stimuli) werden im Computer in einer Datenbank systematisch nach genau definierten Kriterien erfasst, welche die Transkription der einzelnen Dialektantworten unterstützt, die Verwaltung der Sonderzeichensätze übernimmt, über ein Koordinatensystem eine Zuordnung zur geographischen Lokalisation herstellt und die Daten über das Postscript-Format (heute PDF-Format) direkt in die Karte exportiert. Darüber hinaus werden die Tondokumente auf DVD mit direktem Zugriff auf die einzelnen Stimuli bereitgestellt. Damit wurden die technischen Möglichkeiten umfassend genutzt und mit Hilfe des Computers eine direkte Verknüpfung von Ton, Karte und Transkription realisiert.

Den Einsatz gängiger Software zur Erstellung von Sprachkarten kleinräumiger Sprachatlanten sowie die Bearbeitung von Tondokumenten beschreiben Harder und Boller (Harder & Boller 1996), wobei von ihnen die Verbindung von Transkription und Ton noch nicht realisiert, aber als in der Zukunft realisierbar gesehen wird (vgl. Müller & Köhler & Kattenbusch 2001).

Seit Ende der 1990er Jahre werden mit der immer schnelleren Verbreitung des Internets gehäuft digitale Sprachatlanten als Online-Version realisiert. Als eines der ersten funktionierenden online-Projekte ist hier in jedem Fall VIVALDI zu nennen, welches seit 1999 die erhobenen Daten kontinuierlich im Netz zur Verfügung stellt. Die Internetversion des ALD-I – mit dem das Projekt VIVALDI im ständigen Wissens- und Informationsaustausch steht – basiert beispielsweise komplett auf der im Projekt VIVALDI entwickelten Softwareumgebung. Zahlreiche andere Projekte, die das Ziel verfolgen, Sprachkarten mit authentischen Audiodaten zu verknüpfen, haben seit kurzem mit der Entwicklung von eigenen Lösungen begonnen und zeigen auch verschiedene Beispielkarten. Zu nen-

---

<sup>2</sup> <http://www.endangeredlanguages.com>

nen wäre hier beispielsweise das Projekt ALMURA (*L'Atlas linguistique multimédia de la région Rhône-Alpes*)<sup>3</sup>, welches sich an VIVALDI (und in der Konsequenz am ALD-I) orientiert:

En général, ces atlas [ALMURA, Anm. d. Verf.] disposent d'un module phonétique qui sur interrogation de l'internaute reproduit la prononciation d'un item donné. Cela nous paraît particulièrement réussi pour les projets [ALD-I](#) et [Vivaldi](#) qui complètent la reproduction de l'enregistrement du témoin par la transcription phonétique de l'item (Reisdoerfer 2009).

Neueste Publikationen, wie die beiden in der Reihe *Handbuch der Sprach- und Kommunikationswissenschaft* erschienenen Bände *Language and Space* verweisen auf die Bedeutung und die nicht von der Hand zu weisenden Vorteile von multimedialen Sprachatlanten im Internet:

The advantages of internet publication are obvious. Multimedia and interactivity offer map users new ways of perceiving areal variation. Internet publications are readily accessible to the public. They allow for the integration of different research projects and access to a large quantity of data. In contrast to finalized print publications, internet publications have a dynamic character since, in principle, they remain open for corrections and additions (Girnth 2010, S. 117).

Seit 1999 existieren von VIVALDI eine online-Version und eine CD-ROM bzw. seit 2007 eine DVD, die stets weiterentwickelt werden. Stand 1999 nur die Region Sizilien zur Verfügung, ist seither kontinuierlich fast jedes Jahr eine weitere Region veröffentlicht worden. Aktuell sind elf Regionen abgeschlossen und weitere fünf in Bearbeitung, in 230 Orten wurden Interviews mit Sprechern durchgeführt (Stand: 08/2013). VIVALDI bietet inzwischen gut 100.000 Datensätze online (Audio-Dateien und Transkriptionen).

Im Wesentlichen hat sich VIVALDI folgende Ziele gesetzt:

1. Sammlung aktueller Dialektdaten in allen 20 Regionen Italiens.
2. Nutzung elektronischer Datenträger (CD-ROM/DVD) und moderner Medien (Internet) zur schnellen Bereitstellung der Daten in Form von Audio-Dateien.
3. Transkriptionsvorschläge, somit Möglichkeit des Vergleichs der aktuellen Daten mit den Daten des AIS (*Sprach- und Sachatlas Italiens und der Südschweiz*, vgl. Jaberg/Jud 1928-40) und des ALI (*Atlante Linguistico Italiano*, vgl. Bartoli 1995-2008).
4. Flächendeckende Dokumentation der italienischen Dialektlandschaft.
5. Wissenschaftliche Dokumentation der Basisdialekte.
6. Verwendungsmöglichkeit in der linguistischen Forschung.

Mit der L.N. (nationales Gesetz) 482 von 1999 zum Schutz der historischen Minderheitensprachen in Italien war Italien eines der ersten Länder, die die Minderheitensprachen tatsächlich gesetzlich schützen. Das liegt zum einen daran, dass in Italien nach wie vor eine Vielzahl an Minderheitensprachen aktiv gesprochen werden und dass zum anderen erkannt wurde, dass jene, die weniger als 10.000 Sprecher haben, in naher Zukunft vom Aussterben bedroht sein könnten. In Artikel 2 der L.N. 482/1999 werden die zu

---

<sup>3</sup> <http://w3.u-grenoble3.fr/almura/index.php>.

schützenden Minderheitensprachen explizit genannt: „[...] la Repubblica tutela la lingua e la cultura delle popolazioni albanesi, catalane, germaniche, greche, slovene e croate e di quelle parlanti il francese, il franco-provenzale, il friulano, il ladino, l'occitano e il sardo.“ Von dieser Gesetzesinitiative ausgenommen sind aber sämtliche Dialekte, die in Italien aktuell gesprochen werden (vgl. Toso 2006, 67). Jede noch so gut gemeinte Gesetzesinitiative kann aber alleine das Aussterben einer Sprache nicht verhindern. Davon betroffen sind besonders die Minderheitensprachen mit wenigen 100 oder 1.000 Sprechern wie beispielsweise das Okzitanische in Guardia Piemontese (ca. 340 Sprecher, vgl. Toso 2006, 89) oder die kroatischen Kommunen in Molise mit weniger als 2.000 Sprechern (vgl. ebd. 127).

Komplizierter stellt sich die Situation bei den Dialekten dar, die – abgesehen von vereinzelten regionalen Initiativen – keinerlei staatlichen Schutz genießen. Das betrifft zum einen dialektale Sprachinseln wie das Galloitalische in der Basilikata (vgl. Toso 2006, 87) oder auf Sizilien (vgl. ebd. 159), aber auch viele Kommunen auf der Halbinsel, die von starker Abwanderung betroffen sind. Eine Vielzahl italienischer Dörfer besteht nur noch aus wenigen hundert Einwohnern und deren Bevölkerungsprognose ist mehr als pessimistisch. So werden wir in den nächsten Jahrzehnten vermutlich erleben, dass mit dem Aussterben der Dörfer auch die jeweiligen Dialekte dem Untergang geweiht sind.

Besser ist die Situation für das Tabarchino auf Sardinien, einen ligurischen Dialekt, der dort noch von ca. 10.000 Sprechern in den beiden Kommunen Calasetta und Carloforte gesprochen wird. Problematisch ist hier, dass die kleine Minderheit zwar durch regionale Gesetze<sup>4</sup> geschützt wird, nicht hingegen von dem wichtigen nationalen Gesetz 482. Das könnte u. U. zum Aussterben der Sprache führen, auch wenn aktuell davon nicht die Rede sein kann, da die Sprecher ein starkes traditionelles Sprachbewusstsein pflegen. Das zeigt, dass die dialektale Situation zwar nicht hoffnungslos ist, aber eine Vielzahl an Dialekten möglichst schnell systematisch dokumentiert werden sollten. VIVALDI ist hier bereits sehr gut vorangekommen und hat die technischen und wissenschaftlichen Möglichkeiten und Voraussetzungen, dies zu realisieren, und zwar in einem sehr kurzen Zeitplan.

Seit 1992 wurde in 230 Orten der jeweilige Dialekt bzw. die Minderheitssprache dokumentiert und publiziert. Daraus resultiert ein Korpus von ca. 100.000 Tondateien und Transkriptionen mit einer Gesamtlänge von ca. 100 Stunden, das Dank der für VIVALDI entwickelten Software dem Nutzer bequem und übersichtlich in Form von Karten oder Listen zur Verfügung gestellt wird. Es bildet so die Basis für zahlreiche Publikationen, Abschlussarbeiten und Forschungsvorhaben, deren Zahl in den letzten Jahren stetig steigt. Mit dem Abschluss der Erhebungen würde erstmalig das gesamte Dialektkontinuum Italiens mit authentischem Tonmaterial abgebildet werden und könnte so die Grundlage für weitere zukünftige Forschungsarbeiten bilden.

Dass das Projekt zusehends an Bekanntheit und Bedeutung gewinnt, wird zum einen durch die Nutzerzahlen der online-Version deutlich (ca. 1.000 Zugriffe/Monat – Google Analytics, vgl. Abbildung 1) und zum anderen wird in der einschlägigen Fachliteratur VIVALDI häufig zitiert und als funktionierendes Beispiel für einen sprechenden interaktiven Sprachatlas angeführt. Es handelt sich schließlich um eine systematische Sammlung einer nicht-zufälligen Stichprobe von italienischen Dialekten mit einem definierten

---

<sup>4</sup> L.R. 26/1997.

Fragebuch und ist zudem auf den HU-Seiten<sup>5</sup> weltweit hörbar. Das ist ein absolutes Novum, da ansonsten i.d.R. „nur“ die gedruckten Transkriptionen veröffentlicht werden und wurden.

Während Forscher das gesammelte Material als Grundlage für weiterführende Untersuchungen und Analysen der Dialekte und Minderheitensprachen Italiens nutzen können, bietet VIVALDI Laien, Schülern und Studenten, Einheimischen und Ausgewanderten, Angehörigen der Minderheitensprachen und Einwohnern aus den Ursprungsländern die Möglichkeit, sich die Originalstimmen anzuhören und so eine sprachliche Zeitreise zu unternehmen, die dank neuester Technologien heute tatsächlich im Bereich des Möglichen liegt. Wir freuen uns, dass wir Teil dieser dokumentarischen und wissenschaftlichen Erhebungen sein dürfen, deren Wert von Jahr zu Jahr steigt. Und vielleicht gelingt es uns, einen Teil des immensen Kulturschatzes Italiens zu bewahren, zu dokumentieren, verfügbar zu machen und zu analysieren, da einige der von uns aktuell erhobenen Sprachschätze vom Sprachentod bedroht sind. Wohl wissend, dass das Aussterben von Sprachen nicht verhindert werden kann, möchten wir doch unseren Teil dazu beitragen, dass die Stimmen der Vergangenheit (und Gegenwart) nicht in Vergessenheit geraten und so die Grundlage für aktuelle und zukünftige Forschungsvorhaben bilden.

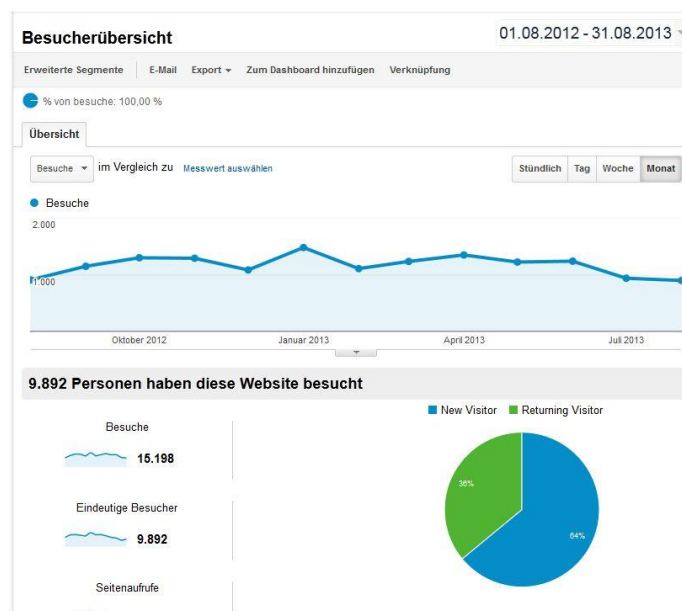


Abbildung 1: Screenshot von Google Analytics im Zeitraum vom 01.08.2012 bis 31.08.2013: Zu sehen ist, dass monatlich mindestens 1000 Besucher die VIVALDI-Seiten aufrufen.

Bisher wurden die folgenden Regionen aufgenommen, bearbeitet und veröffentlicht:

- 1992: Sizilien (erste Probeaufnahmen, die teilweise seit 1999 wiederholt wurden)
- 1999-2000: Sardinien
- 2001-2002: Ligurien
- 2003: Aostatal
- 2003-2004: Umbrien

<sup>5</sup> <http://www2.hu-berlin.de/vivaldi>

- 2005-2007: Trentino-Südtirol
- 2005: Molise
- 2007-2008: Piemont
- 2009-2010: Friaul-Julisch Venetien
- 2009-2012: Venetien
- 2012: Basilikata

Seit 2009 sind Toskana, Apulien, Kalabrien, Lombardei und Kampanien in Bearbeitung. Die in diesen Regionen aufgenommenen Dialekte sind bereits online verfügbar. Es fehlen noch: Emilia-Romagna, Marken, Latium und Abruzzen.

Seit dem Jahr 2011 können die Minderheitensprachen mit dem gesonderten Visualisierungsmodul PALMI (Panorama Acustico delle lingue minoritarie d'Italia) in ihrer Gesamtheit dargestellt werden (s. Abbildung 2). Zum ersten Mal stehen Tondokumente von den aktuell in Italien gesprochenen Minderheitensprachen der Allgemeinheit zur Verfügung.

Es ist uns natürlich bewusst, dass mit dem Fragebuch, das speziell die Entwicklungen in den italienischen Dialekten untersuchen soll, insbesondere die nicht-romanischen Sprachen nicht vertieft analysiert werden können. Andererseits können mit den gut 350 Antworten in jedem Fall unidirektionale Interferenzen vom Italienischen bzw. der lokalen Dialekte auf die Minderheitensprachen herausgearbeitet werden.

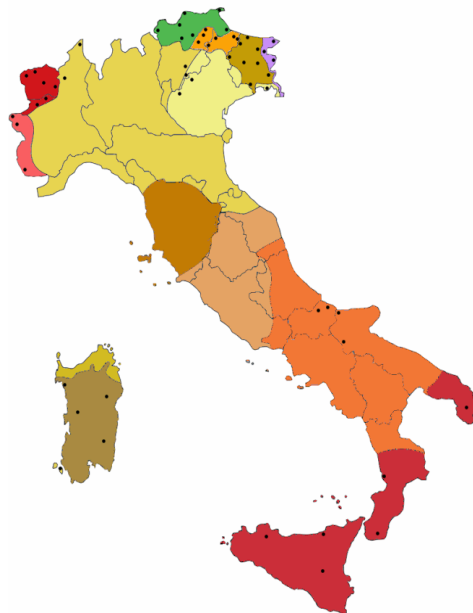


Abbildung 2: Startseite von PALMI, eine Übersicht von den verzeichneten Minderheitensprachen

Die Wahl der bisherigen Aufnahmeorte berücksichtigt hauptsächlich die grobe dialektale Gliederung Italiens (Nord-, Mittel-, Süditalien). Die Auswahl der bisher aufgenommenen Regionen erfolgte auch immer mit dem Ziel, eine romanische oder nicht-romanische Minderheitensprache zu dokumentieren.



Nach der Erhebung der Daten im Feld und der digitalen Aufarbeitung und Nachbereitung dauert es in der Regel höchstens sechs Monate bis zur Veröffentlichung im Internet. Neben der zeitaufwendigen Erhebung der Daten im Feld, der Konzeption und Erstellung des Fragebuchs sowie der Veröffentlichung der Daten im Netz und auf DVD wurden die folgenden Vorarbeiten geleistet: Entwurf der Datenbank, Konzeption der Software zur Präsentation der Daten und Erstellung der Webseiten.

Die relationale Datenbank von VIVALDI bildet das Herzstück für die Präsentation und Publikation des im Feld gewonnenen Materials. Nachdem die Daten eingegeben wurden, werden diese mit Hilfe verschiedener Softwaremodule ausgelesen. Die Eingabe betrifft einerseits die Transkription und andererseits die geschnittenen Tondateien aus dem Aufnahmekontinuum. Dafür wird im Projekt die Software Goldwave benutzt, mit deren Hilfe die einzelnen Antworten aus der Aufnahme, die in der Regel zwei bis drei Stunden dauert, geschnitten werden. Goldwave hilft auch beim Segmentieren der einzelnen Antworten, um die Transkriptionen zu verbessern und/oder zu kontrollieren.

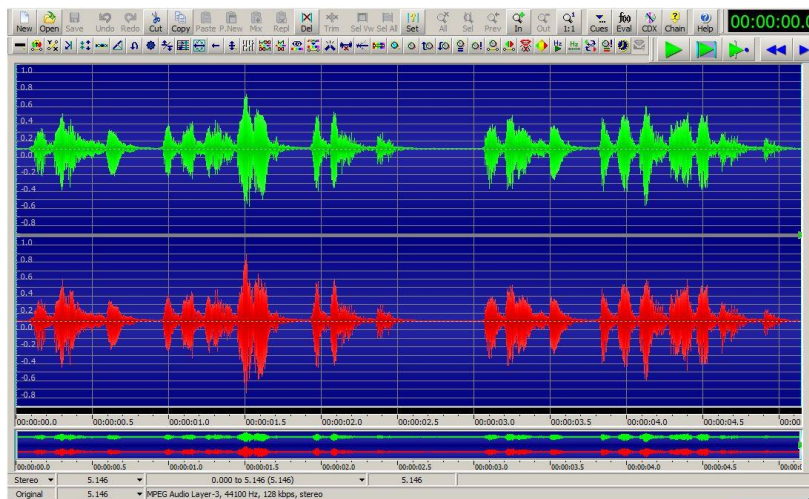


Abbildung 3: Screenshot Goldwave vom Stimulus Nr. 83 „Adesso devi andare a destra“ von Ortsnr. V08 (Verona)

Bei schwer zu identifizierenden Lauten wird zur Unterstützung auf das Programm Praat zurückgegriffen, da hier alle drei Dimensionen (Dauer, Frequenz und Intensität) sichtbar werden und dem Phonetiker beim Erkennen der Laute behilflich sind. So erleichtern die bei Praat dargestellten Formanten die Identifizierung von Vokalen und Diphthongen.

Die Transkriptionen werden zunächst in eine Excel-Tabelle, dem im Projekt entwickelten VIVALDI-Transkriptionsassistenten (VivTKA) eingetragen, der die Codes in die jeweiligen Transkriptionszeichen umsetzt. Dafür mussten spezielle Module für Excel entwickelt werden, die die Umsetzung vornehmen.

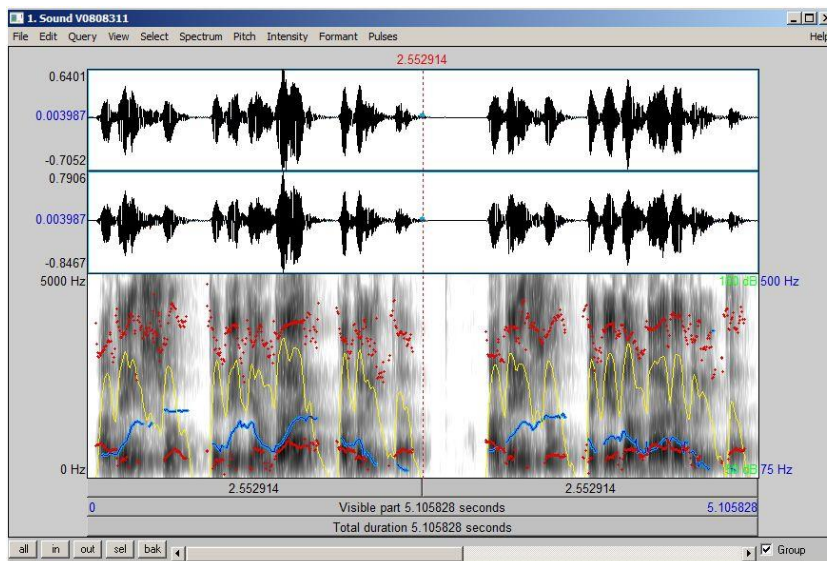


Abbildung 4: Screenshot Praat vom Stimulus Nr. 83 „Adesso devi andare a destra“ von Ortsnr. V08 (Verona)

Die Aufbereitung des Tonmaterials beruht darauf, dem linguistisch forschenden Benutzer keine subjektive Transkription anzubieten, damit er nicht auf die eines anderen Wissenschaftlers angewiesen ist. So entfällt z.B. auch die fehleranfällige Übertragung von einem in ein anderes Transkriptionssystem, mit dem der romanistische Linguist ständig konfrontiert ist. Benutzer können das Tonmaterial z.B. auch für eigene phonetisch/akustische Analysen verwenden.

A	B	X	Y	Z	AA	AB	AC
Ort Nr.		V07			V08		
SprecherIn Nr.		1			1		
Version		1			1		
Ort Name		Pozzale			Verona		
Bemerkung		fertig			fertig		
Sprachatlas Ort							
WWW Ort							
StNr	Stimulus	Code	Zeichen	Notiz	Code	Zeichen	Notiz
1	l'acqua	al\3D\1ga	al'ɛ ga		l'ákwa	l'ákwa	
2	l'acqua è calda	al\3D\1ga l(r) é	al'ɛ ga i' é cáwða		l'ákwa l é kálda	l'ákwa l é kálda	
3	l'agnello	al k\3e\1tín	al kɛ tín		l'al6n\3D\1l	l'a ñ ɛ l	
4	l'aglio	l'ái	l'ái		l'áy\2O\1	l'áyɔ	
5	agosto	ag\2d\l6s\1t\2o	agó s tɔ		agó\6s\1t\2O\1	agó s tɔ	
6	l'ala	l'áa	l'áa		l'ál\6a\1	l'ála	
7	alto	l'á\1á\6w\1t\2o\1	á u tɔ		ált\2O\1	áltɔ	
8	altro	á\6w\1tro	á u tro		á\6t\2O\1	á\6tɔ	
9	dammi un altro pezzo	dá m in un ál\6	dá m in un á u tro tɔ		dá m\3E\1 n á\6	dá m ɛ n á\6 tɔ <sup>2</sup> tɔ kɔ	
10	l'anca	l\7n\1n\6c\1a	l'á n ca		l'á\6N\1k\6a\1	l'á y ka	
11	l'angelo	l\7n\1nd\2O\1l	l'á n dɔ l		l'án\6g\1e\l\5U\1	l'án ɛ l ɛ l ɛ	
12	l'anno	l'á\6N\1	l'á y		l'án\2O\1	l'án ɔ	
13	aprile	ap\6r\l\3e\1	ap\6r\l ɛ		ap\6r\4n\1l	ap\6r\l	
14	l'argento	l'ar\6d\3D\1nto	l'ar d ɛ nto		l'arg\3G\1nt\2C	l'arg ɛ ntɔ	

Abbildung 5: Eingaben im VIVALDI Transkriptions Assistenten (VIVTka)

Nach der Endkontrolle der eingetragenen Transkriptionen mit den Tondateien wird aus den Excel-Daten eine Exportdatei für die Datenbank erzeugt. Die Daten werden dann in die Datenbank eingespielt, worauf die im Projekt entwickelte Software den Ort mit der Tondatei und der Transkription verknüpft. Entweder kann die Antwort auf die gestellte



Frage durch einen Klick auf den jeweiligen Ort in der Karte angehört werden, wobei die Transkription dann über der Karte erscheint, oder die Antworten stehen mit Ortsname, Transkription und Tondatei als Liste unter der Karte zur Verfügung.



Abbildung 6: Die Region Piemont im Projekt VIVALDI

Für die Präsentation wurden mit Hilfe von LAMP (Linux, Apache, MySQL, PHP) die Internetseiten entwickelt. Die Karte selbst wurde als Java-Applet realisiert. Dabei wurde stets darauf geachtet, dass standardisierte und freie Software zum Einsatz kommen. Die Seiten werden regelmäßig gepflegt, zum Einen werden sie mit neuen Daten aus neuen Regionen bzw. Orten aktualisiert und zum anderen werden sie durch Korrekturen der Transkriptionsvorschläge und das Ersetzen von toten Links durch funktionierende auf dem neuesten Stand gehalten. Diese stetigen Verbesserungen und Aktualisierungen sind auch ein wesentlicher Vorteil von online-Sprachatlantem im Vergleich zu gedruckten Atlanten.

Seit gut sieben Jahren steht dem Nutzer auch die „Einzelauswahl von Orten und Stimuli“ zur Verfügung, die besonders für wissenschaftliche Arbeiten von größtem Nutzen ist. Der Nutzer kann sich bei der Einzelauswahl eine Treffermenge bestimmter Orte und Stimuli selbst zusammenstellen. Dies erleichtert besonders die Vergleichsmöglichkeiten von Antworten in verschiedenen Orten, wie beispielsweise das Frankoprovenzalische in Piemont und jenes in Apulien, oder den Vergleich der galloitalischen Dialekte untereinander usw.

VIVALDI Maps ist seit gut fünf Jahren online. Hier werden die Daten aus der Datenbank von VIVALDI mit den Geodaten von Google gemischt. Dadurch entsteht eine andere Form der Präsentation: Der Nutzer hat die Möglichkeit sich einen Gesamtüberblick über die aufgenommenen Orte im gesamten Raum zu verschaffen. Auch bei VIVALDI Maps werden die Transkriptionen angezeigt und die Tondateien abgespielt.

### Vivaldi Maps



Abbildung 7: Startbildschirm von VIVALDI Maps

VIVALDI nutzt die von Girnth (2011, S. 117) genannten Vorteile von Sprachatlanten im Internet in vollem Umfang. Alle aufgenommenen Daten werden schnell bearbeitet, transkribiert und zeitnah der Öffentlichkeit zur interaktiven Nutzung zur Verfügung gestellt.

PALMI ermöglicht eine schnelle und unmittelbare Visualisierung aller Sprachinseln und sprachlichen Minderheiten in Italien, die mit den Transkriptionen und Audiodateien von VIVALDI verknüpft werden und so die wissenschaftliche Beschäftigung mit diesen Minderheitensprachen vereinfacht.

Mit der Vervollständigung VIVALDI wird so das gesamte Kontinuum der italienischen Dialektlandschaft hör- und sichtbar. Untersuchungen zu aktuellen phonetischen Phänomenen werden damit ermöglicht. Daher sind die Dokumentation und die systematische Archivierung von authentischem basilektalen Tonmaterial für den Gesamtraum Italien, das nicht nur für kurzfristige, sondern auch für langfristige Untersuchungen erhoben worden ist, das primäre Ziel von VIVALDI.

## 2 Zur Methodik

### 2.1 Erstellung des Fragebuchs

Das Fragebuch ist so konzipiert, dass alle phonetischen Besonderheiten im Bereich Vokalismus und Konsonantismus, die bei der Entstehung der Dialekte von Bedeutung sind, berücksichtigt sind. Das heißt: Ausgehend von den lateinischen Ursprungsformen wird für alle potentiellen Entwicklungen mindestens ein Vertreter (Etymon) abgefragt.<sup>6</sup> Im

<sup>6</sup> Eine komplette Etymaliste mit Hinweisen auf die AIS-Karten und VIVALDI-Stimuli befindet sich im An-

Bereich Vokalismus bedeutet dies beispielsweise, dass für die zehn lateinischen Vokalphoneme und Diphthonge die heutigen Ausprägungen in den jeweiligen Dialekten überprüft werden. Dabei berücksichtigt wird natürlich auch die konsonantische Umgebung, in der die Vokalphoneme im Lateinischen stehen. Gleiches gilt für die Konsonanten und Konsonantenverbindungen. Auch hier befindet sich im Fragebuch mindestens ein Vertreter für das Auftreten im Anlaut, im Inlaut und in bestimmten Umgebungen (z.B. *muta cum liquida* uvm.).

Losgelöst vom Kontext könnten einige Stimuli auf den ersten Blick als fragwürdig erscheinen, z.B. (88) „io dico“ und (89) „lui dice“. In diesem wie in ähnlich gelagerten Fällen geht es aber nicht um die konkrete lexikalische Realisierung, sondern um die phonetischen Variationen wie z.B. die Realisierung des intervokalischen stimmlosen velaren Verschlusslautes. So kann im norditalienischen Raum aus dem velaren Verschlusslaut auch bei der ersten Person Singular eine Veränderung zum Frikativ stattfinden.



Abbildung 8: "io dico" in verschiedenen Dialekten (Ausschnitt: VIVALDI)

So entstehen aus den gut 300 Fragen (Originalkarten) zur Phonetik bei einer genauen Taxierung der phonetischen Phänomene mindestens 1.000 Arbeitskarten. Beispielsweise beim Stimulus Nr. 60 „la chiave“ (< CLAVE(M): Hier stellen sich die Fragen:

- Bleibt der konsonantische Nexus *cl-* im Anlaut erhalten?
- Wie entwickelt sich der Haupttonvokal *a*?
- Bleibt der intervokalische stimmhafte Frikativ erhalten?
- Wie entwickelt sich der Auslautvokal?

D.h. aus der einen Grundkarte „la chiave“ können so vier Arbeitskarten entstehen. Die folgende Tabelle zeigt exemplarisch einige Entwicklungsmöglichkeiten auf, die aus dem lateinischen Etymon CLAVE(M) entstehen können.

Ortsname	San Severo (Apulien)	Moggio (Friaul)	Claut (Friaul)	Verona (Venetien)	Cencenighe (Venetien)	Reba (Venetien)
la chiave (< CLAVE(M))	<i>a čěvə</i>	<i>la klāf</i>	<i>la tšá</i>	<i>la cávę</i>	<i>la tšáf</i>	<i>la klě</i>

## 2.2 Auswahl der Orte

Die Auswahl der Orte erfolgt nach festen Kriterien. Höchste Priorität haben die Aufnahmeorte des AIS-Punktenetzes. Grundsätzlich wird versucht, sämtliche im AIS verzeichneten Orte neu aufzunehmen. Sind in einer Region jedoch nur wenige AIS-Punkte dokumentiert – besonders im süditalienischen Raum ist das AIS-Punktenetz relativ dünn –

hang.

werden die ALI Punkte herangezogen, um einen Vergleich mit älteren Daten zu ermöglichen. Darüber hinaus wird immer versucht, in den einzelnen Regionen die dort vorhandenen sprachlichen Minderheiten bzw. Sprachinseln zu katalogisieren.

### 2.3 Sprecherwahl und Aufnahmesituation

Die Informanten müssen über eine hohe dialektale Kompetenz verfügen, sollten im Ort aufgewachsen sein und möglichst wenig Zeit außerhalb der Region verbracht haben. Das bedeutet: Es muss ein Sprecher<sup>7</sup> gefunden werden, dessen erste Muttersprache der Dialekt ist und der diesen in seiner alltäglichen Kommunikation verwendet.

Bei der Wahl der Sprecher ist weniger die Schulbildung oäd. von Bedeutung<sup>8</sup>, sondern vielmehr dass sie die Pragmatik verstehen, d.h. die Sprecher eine basilektale Entsprechung des italienischen Stimulus korrekt und klar aussprechen und bereit sind, den Fragebogen vor der Befragung zu bearbeiten<sup>9</sup>, damit alle gefragten Teile vollständig beantwortet werden.

VIVALDI geht auch der Frage nach, inwieweit sich im Gesamttraum Italien die Dialekte italianisiert haben. Besonders in den 1970er und 1980er Jahren wurden die Dialekte geradezu verteufelt und beispielsweise aus den Schulen verbannt bzw. verboten. Auch in den Familien wurde immer weniger im Dialekt kommuniziert. Daher interessiert sich VIVALDI besonders für jene Generation, die in den 1960er und 1980er Jahren aufgewachsen ist und zu dieser Zeit die Schule besucht hat. Das sind Personen, die heute im Alter zwischen 45 und 65 Jahren sind.

Die Informanten dürfen natürlich keine Sprach- und Artikulationsfehler (wie z.B. Lispeln) haben. Von größter Wichtigkeit ist ein komplettes, möglichst natürliches, Gebiss, damit beispielsweise die Sibilanten korrekt artikuliert werden. Ältere Informanten (> 70 Jahre) haben oft kein natürliches Gebiss und keine stabile und klare Stimme, um die Laute deutlich auszusprechen. Bei jüngeren Sprechern ist in der Regel das Kriterium „erste Muttersprache = Dialekt“ nicht mehr gegeben.

Meistens reicht die Befragung eines Vertreters des Ortsdialektes aus. Zeigen sich bei diesem jedoch Unsicherheiten und Diskrepanzen, so wird ein zweiter Vertreter – wenn nötig auch weitere – befragt. Die Befragung mehrerer Sprecher erfolgt aber nicht aus soziolinguistischen Motiven. Bei VIVALDI geht es weniger darum, die verschiedenen Varietäten auf kleinem Raum zu untersuchen, wie es beispielsweise Ziel soziolinguistischer Studien ist, sondern darum, den Basilekt der jeweiligen Ortspunkte im Untersuchungsnetz zu dokumentieren. Die Erforschung des Varietätengefüges kann und will VIVALDI nicht leisten. Das wäre schon wegen der Größe des Untersuchungsgebiets – ganz Italien – mit den derzeit vorhandenen Mitteln nicht erreichbar. Ziel ist und bleibt, die phonetischen Besonderheiten im gesamten Raum Italien zu dokumentieren.

---

<sup>7</sup> Mit „Sprecher“ sind natürlich sowohl die weiblichen als auch die männlichen Informanten gemeint. Aus Gründen der Lesbarkeit verzichten wir auf die umständlichen Formulierungen, wie „Sprecher/Sprecherin“ oder „SprecherIn“ usw.

<sup>8</sup> Darauf weisen schon Jaberg/Jud (1928, S. 190): „Bildung gefährdet den Bestand der Mundart, weil ihr Vehikel die Schriftsprache ist“. Sujets von natürlicher Intelligenz stellen unabhängig vom Bildungsstand für den Dialektforscher eine gute Quelle für die Datenerhebung dar.

<sup>9</sup> Der Informant bekommt den Fragebogen mindestens einen Tag vor dem Interviewtermin und hat somit Zeit, die Fragen, Sätze und das Gleichnis vorzubereiten und durchzuarbeiten.

Feste Regeln, die zum Auffinden der Gewährsleute führen, gibt es nicht. Die meisten Aufnahmeorte sind bezüglich der Einwohnerzahl so überschaubar, dass dort von der Gemeinschaft gute Hinweise auf den optimalen Informanten gegeben werden. Solche Hinweise können in der Gemeindeverwaltung, Bibliotheken, Kirchen, Bars und (falls vorhanden) Hotels gewonnen werden. In größeren Städten gestaltet sich die Informantensuche meistens schwieriger. Hier gibt es aber Kultur-Organisationen, die einem weiterhelfen. Ansonsten ist der *Assessore alla cultura* von Amts wegen verpflichtet, den Dialektforscher zu unterstützen. Wer frühzeitig weiß, wo und wann er eine Aufnahme macht, kann auch per e-Mail zuvor sein Kommen bei der Gemeinde ankündigen und u.U. den Fragebogen vorab schicken.

Viel schwerer als das Auffinden eines geeigneten Informanten gestaltet sich häufig die Suche nach einem geeigneten Aufnahmeort. Da es sich, wie erwähnt, um einen phonetischen Sprachatlas handelt, sind nach Möglichkeit alle störenden Geräusche abzuschirmen. Zu diesem Zweck werden stets ein Richtmikrofon und eine „Schallbox“ verwendet.



Abbildung 9: Typische Aufnahmesituation bei Aufnahmen für das Projekt VIVALDI

## 2.4 Transkriptionssystem und Transkriptionen

Bei der Auswahl des Transkriptionssystems lagen besonders pragmatische Gesichtspunkte zugrunde. Drei Anforderungen wurden an das System gestellt:

- a) direkte Vergleichsmöglichkeit mit dem Transkriptionssystem des AIS, da dieser nach wie vor das Referenzwerk der italienischen Sprachatlanten ist.
- b) Einfache Schreib- und Lesbarkeit sowohl für Projektmitarbeiter als auch für wissenschaftlich Interessierte.
- c) Da der Atlas über seinen wissenschaftlichen Zweck hinaus auch dokumentarischen und landeskonservatorischen Wert hat, soll der Zugang zu den Transkriptionen auch für nicht-Linguisten ermöglicht werden.

Nur das AIS-System scheint alle drei Anforderungen zu erfüllen. Aus diesem Grund wurde dem Internationalen Phonetischen Alphabet (IPA) das AIS-System vorgezogen, welches nach und nach an die eigenen Bedürfnisse angepasst wurde. Darüber hinaus wird das System in vielen Standardwerken zur italienischen und romanischen Sprachge-

schichte verwendet (z.B. Tagliavini 1968), sowie in renommierten älteren Einführungen wie Walther von Wartburg (1970) und Gerhard Rohlfs (1966) und sollte daher bekannt sein. Sämtliche Transkriptionszeichen werden auf den VIVALDI Seiten artikulatorisch beschrieben und sind daher einfach konsultierbar.

Transkribiert werden aus dem Fragebuch alle Stimuli des phonetischen Teils. Nicht transkribiert wird dagegen das „Gleichnis vom verlorenen Sohn“, da aktuell die dafür benötigten Arbeitskräfte fehlen. Das Transkriptionsverfahren erfolgt in drei Etappen:

1. Vor Ort, direkt nach der Aufnahme, um Unstimmigkeiten beseitigen und eventuelle Zweifel sofort mit dem Informanten klären zu können.
2. Eingabe in die Datenbank mit Hilfe des VIVALDI Transkriptionsassistenten.
3. Endkontrolle vor der Publikation im Internet und auf DVD. Hier wird noch einmal jede Frage von wenigstens zwei Mitarbeitern genau geprüft, Flüchtigkeitsfehler, die bei der Eingabe in die Datenbank auftreten können, werden eliminiert und die Verknüpfung von Ton und Transkription kontrolliert.

Trotz aller Sorgfalt (Kontrolle mit Goldwave und Praat) möchten wir darauf hinweisen, dass es sich dabei um einen Transkriptionsvorschlag handelt, da uns durchaus bewusst ist, dass der individuelle Höreindruck eine gewisse Toleranz bezüglich des Gehörten zulässt. Auch die technische Umgebung (Lautsprecher, Kopfhörer usw.) können u. U. zu leichten Variationen der auditiven Wahrnehmung führen. Um dieses Problem zu beseitigen, bieten wir Interessierten sämtliche Tondateien auf Nachfrage auf DVD an, die dann von diesen kontrolliert werden können.

Bezüglich der Präsentation der Daten folgen wir festen Prinzipien, die auch u.a. von Girth (vgl. 2010, S. 117) postuliert wurden:

- *Principle of quality* – Daten sollen transparent sein und die sprachliche Realität möglichst genau abbilden.
- *Principle of quantity* – es soll darauf geachtet werden, dass keine Datenlücken entstehen und die Karten den Raum vollständig abbilden.
- *Principle of thematic relevance* – Daten müssen systematisch erhoben werden und die Karten dürfen nur für das Untersuchungsobjekt relevante Daten enthalten.
- *Principle of modality* – Erhebung und Präsentation sollen für die Nutzer nachvollziehbar sein.

Die schnelle Bearbeitung der gesammelten Daten ermöglicht ihre zeitnahe Veröffentlichung im Netz – ein weiteres Prinzip, das sich VIVALDI zum Ziel gesetzt hat und seit Jahren strikt verfolgt. Damit wird vermieden, dass die Daten bei der Publikation schon veraltet sind bzw. Datenfriedhöfe entstehen.

Aus Zeit- und Kostengründen ist es zwar derzeit nicht möglich, das Gleichnis durchgehend zu transkribieren, dennoch dient es Untersuchungen als Kontrollfunktion für vorher abgefragte Stimuli oder um spezielle syntaktische Phänomene im Kontext eines durchgängig gesprochenen Textes zu analysieren (vgl. Mensching 2012). Dann sollte nicht vergessen werden, dass das Gleichnis vom verlorenen Sohn in den vergangenen hundert Jahren immer wieder abgefragt wurde, u. a. auch im ALI, und daher eine ausgezeichnete Quelle für den diachronen Vergleich unterschiedlichster Phänomene bietet.



Die Erfahrung zeigt, dass die Darbietung der gesprochenen Sprache entspricht, obschon wir Wert darauf legen, dass die Informanten das Gleichnis mindestens einen Tag vorher bekommen und vorbereiten können. Ein unabdingbares Kriterium für die systematische Arbeit mit einem längeren Text, damit die Vergleichbarkeit gewährleistet wird. Aus technischen Gründen wird das Gleichnis nach der Aufnahme in 24 Teile segmentiert, was den spontanen Sprechfluss während der Aufnahme jedoch keineswegs behindert.

### 3 Umgang mit den im Projekt erzielten Forschungsdaten

Die im Projekt systematisch erhobenen und gesammelten Forschungsdaten werden für einen Zeitraum von mindestens zehn Jahren auf Servern der Humboldt-Universität zur Verfügung stehen. Darüber hinaus werden im Projekt regelmäßig, d.h. alle 6 Monate DVDs von allen Daten erzeugt. Diese werden, falls gewünscht, an interessierte Forschungseinrichtungen und Forscher geschickt, die die im Projekt VIVALDI erhobenen Daten nutzen möchten. Natürlich haben diese auch die Möglichkeit, die aktuellsten Daten vom Server der HU kostenlos zu nutzen. Bis auf Weiteres werden wir immer eine aktualisierte DVD (Lebensdauer ca. 30 Jahre) der Institutsbibliothek zur Verfügung stellen, die für alle interessierten Nutzer im Bibliothekskatalog verzeichnet ist und ausgeliehen werden kann.

Es ist uns durchaus bewusst, dass 30 Jahre, besonders im Bereich Sprache und Sprachgeographie, einen relativ kurzen Zeitraum darstellen. Daher sind wir an zwei Archive herangetreten, die es sich zur Aufgabe gemacht haben, Sprachdaten zu sammeln, zur Verfügung zu stellen, zu dokumentieren und zu archivieren. Dies ist zum einen das Projekt *Documentation of Endangered Languages (DOBES)*<sup>10</sup>, das besonders für die bedrohten Minderheitensprachen in Italien von Interesse ist. Zum anderen sind wir in engem Kontakt mit der Leitung des Projekts *Gra.fo* (Scuola Normale Superiore di Pisa/Università degli Studi di Siena), das sich darauf spezialisiert hat, sämtliche Dialektdaten der Toskana zu sammeln und zu archivieren. Es wäre natürlich für eine langfristige Archivierung wünschenswert, wenn in den nächsten Jahren Archive entstehen würden, die es sich zur Aufgabe machen, analoge und digitale Sprachdaten zu sammeln, zu archivieren und wie oben beschrieben zur Verfügung zu stellen.

### Literaturverzeichnis

- Bartoli, Matteo Giulio (Hrsg.) (1995-2008): *Atlante linguistico italiano*. 7 Bände. Roma.
- Bauer, Roland/Goebel, Hans/Haimerl, Edgar (2005): *Der „sprechende“ Ladinienatlas*, San Martin de Tor.
- Calamai, Silvia (2012): „Ordinare archivi sonori: il progetto Gra.fo“, in: *Rivista Italiana di Dialettologia*, 35, 135-164.
- Cortelazzo, Manlio (Hrsg.) (1974–88): *Profilo dei dialetti italiani*. 17 Bände, z.T. mit Schallplatten. Pisa.
- Gilliéron, Jules/Edmont, Edmond (1902–1910): *Atlas linguistique de la France (ALF)*. 10 Bände. Paris.

---

<sup>10</sup> <http://dobes.mpi.nl/>

- Girnth, Heiko. (2010): „Mapping language data“, in: Lameli, Alfred/Kehrein, Roland/Rabanus, Stefan (Hrsg.): *Language and Space. Volume 2 Language Mapping*, Berlin, Boston, 98-145.
- Goebel, Hans (Hrsg.) (1998): *Sprachatlas des Dolomitenladinischen und angrenzender Dialekte*, 1. Teil (ALD I). 4 Kartenbände, 3 Indexbände, 3 CD-ROM. Wiesbaden.
- Harder, Andreas/Boller, Fred (1996): *Sprachgeographie und PC. Sprachkarten, Datenorganisation, Tonproben mit Mikrorechnern*. Kiel.
- Hotzenköcherle, Rudolf/Brunner, Rudolf (Hrsg.) (1972–76): *SDS-Phonogramme. Tonaufnahmen für den Sprachatlas der deutschen Schweiz*. 16 Langspielplatten. Zürich.
- Jaberg, Karl/Jud, Jakob (1928): *Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz*, Halle (Saale).
- Jaberg, Karl/Jud, Jakob (1928–40): *Sprach- und Sachatlas Italiens und der Südschweiz*. 8 Bände. Zofingen.
- Mensching, Guido (2012): „Anmerkungen zur sardischen Syntax anhand des Vivaio Acustico delle Lingue e dei Dialetti d’Italia (VIVALDI)“, in: (Das) diskrete Tatenbuch – Festschrift für Dieter Kattenbusch zum 60. Geburtstag, Berlin. [online: <http://www2.hu-berlin.de/festschrift-kattenbusch/mensching-sardisch-syntax.html>].
- Müller, Marcel Lukas/Köhler, Carola/Kattenbusch, Dieter: „VIVALDI - ein sprechender Sprachatlas im Internet als Beispiel für die automatisierte, computergestützte Sprachatlasgenerierung und –präsentation“, in: *Dialectologia et Geolinguistica*, 9/2001, 55-68. [online: <http://www2.hu-berlin.de/vivaldi/publikationen/vivaldi-sprechender-sprachatlas-2001.pdf>].
- Reisdoerfer, Joseph (2009): „Géolinguistique et Informatique“, in: *Le Monde Blog*, <http://laurette.blog.lemonde.fr/2009/08/29/geolinguistique-et-informatique/> (Zugriff am: 03.07.2012)
- Rohlf, Gerhard (1966): *Einführung in das Studium der romanischen Philologie*, Heidelberg.
- Ruffino, Giovanni (Hrsg.) (1986): „Prospettive di lavoro per un atlante linguistico-etnografico della Sicilia“, in: *Atti della tavola rotonda*, Palermo, 11 ottobre 1985. Palermo.
- Sobrero, Alberto A./Romanello, Mariateresa/Tempesta, Immacolata (1991): *Lavorando al NADIR: un’idea per un atlante linguistico*. Galatina.
- Tagliavini, Carlo (1968): *Le origini delle lingue neolatine*, Bologna.
- Toso, Fiorenzo (2006): *Lingue d’Europa. La pluralità linguistica dei Paesi europei fra passato e presente*. Milano.
- Wartburg, Walther von (1970): *Einführung in die Problematik und Methodik der Sprachwissenschaft*, Tübingen.
- Wieland, Katharina/Plikat, Jochen/Küster, Lutz: „VIVALDI – eine reiche Fundgrube für Lehramtsstudierende? Fremdsprachen- und hochschuldidaktische Überlegungen“, in: Carola Köhler/Fabio Tosques (Hg.), (Das) diskrete Tatenbuch. *Digitale Festschrift für DIETER KATTENBUSCH zum 60. Geburtstag*, Berlin 2012. [online: <http://www2.hu-berlin.de/festschrift-kattenbusch/kwp-vivaldi-didaktik.html>].