

# Dialectologia et Geolinguistica

DiG

Journal of the International Society  
for Dialectology and Geolinguistics

9 / 2001

Editor: *Wolfgang Viereck*  
Assistant to the editor: *Hedwig Gwosdek*

Edizioni dell'Orso

## **VIVALDI – ein sprechender Sprachatlas im Internet als Beispiel für die automatisierte, computergestützte Sprachatlasgenerierung und -präsentation**

*Marcel Lucas Müller, Carola Köhler und Dieter Kattenbusch*

### 1. Einleitung

Das wohl wichtigste Instrument der auf eine nunmehr hundertjährige Tradition zurückblickenden Sprachgeographie ist der Sprachatlas, der die dialektalen Realisierungen bestimmter hochsprachlicher Stimuli an genau definierten geographischen Punkten darstellt, indem die dialektalen Ergebnisse als Transkription oder mittels kategorisierender Symbole den entsprechenden Koordinaten auf einer einheitlichen Grundkarte zugeordnet werden. Obwohl das Tondokument das eigentliche Forschungsobjekt und die primäre Informationsquelle ist, gibt es bisher so gut wie kein authentisches Sprachmaterial, das für Forschungszwecke oder zur Verwendung im universitären Unterricht der Allgemeinheit zur Verfügung stünde.

Das Projekt "Vivaldi" (*Vivaio Acustico delle Lingue e dei Dialetti d'Italia*) hat sich zum Ziel gesetzt, einen akustischen Sprachatlas Italiens zu entwickeln, welcher die dialektalen Besonderheiten und Minderheitensprachen der 20 Regionen Italiens repräsentativ abbildet. Durch die Nutzung moderner Computertechnologie für die Erfassung und Verarbeitung der Tondokumente und deren Transkriptionen einerseits sowie für eine internetbasierte Präsentation mit direkter Verbindung von Karte, Tondokument und Transkription andererseits sollen neue Möglichkeiten für Sprachatlasprojekte entwickelt werden.

### 2. Hintergrund

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 (Gilliéron und Edmont 1902-1910) und der AIS (*Sprach- und Sachatlas Italiens und der Südschweiz*) aus den Jahren 1928 bis 1940 (Jaberg und Jud 1928-1940), die in Form von gedruckten, zu mehreren großformatigen Bänden gebundenen Sprachkarten vorliegen.

Die Aufzeichnung von Sprache war in den Anfängen nach der Erfindung des Phonographen 1877 durch T.A. Edison aufwendig und schlecht in der Wiedergabequalität sowie unpraktikabel in der Anwendung. Mit der Schallplatte stand zwar ab dem Ende der zwanziger Jahre des 20. Jahrhunderts ein Medium zur Verfügung, das eine größere Verbreitung von Tonaufnahmen ermöglichte, aber die Aufnahmepraxis war weiterhin umständlich und kostspielig, so daß sie kaum in der Feldforschung eingesetzt wurde. Es gibt eine Reihe von Schallplatten – z.B. zu italienischen (Cortellazzo 1974-1988) oder zu schweizerdeutschen (Hotzenköcherle und Brunner 1972-1976) Dialekten –, die jedoch erhebliche Nachteile aufweisen: teilweise schlechte Aufnahmequalität, Abnutzung, besonders aber geringe Kapazität sowie keine direkte Verbindung des visuellen Elements (Karte) mit dem auditiven (Ton).

Erst die Einführung moderner, computerunterstützter Verfahren der Tonaufnahme, -konservierung, -verarbeitung und -wiedergabe erlauben eine qualitativ hochwertige, verlustfreie und kostengünstige Verbreitung der Sprachdokumente. Der Sprachatlas ALD I (*Atlante linguistico del ladino dolomitico e dei dialetti limitrofi*) erschloß der Sprachgeographie diese Neuerungen erstmals in größerem Rahmen (Goebel 1998): Die Stimuli werden im Computer in einer Datenbank 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 direkt in die Karte exportiert. Darüber hinaus werden die Tondokumente auf CD-ROM 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 ([http://www.sbg.ac.at/rom/people/proj/ald/ald\\_home.htm](http://www.sbg.ac.at/rom/people/proj/ald/ald_home.htm) – hier und im folgenden: zuletzt besucht am 05.01.2001).

Den Einsatz gängiger Software zur Erstellung von Sprachkarten kleinräumiger Sprachatlanten sowie die Bearbeitung von Tondokumenten beschreiben Harder und Boller (1996), wobei von ihnen die Verbindung von Transkription und Ton noch nicht realisiert, sondern für in Zukunft möglich gehalten wird.

Mit dem Internet hat sich in den letzten Jahren ein Medium etabliert, welches eine neue Dimension für die multimediale Darstellung von Inhalten aufzeigt. Die relativ starre Präsentation mittels Buch und CD-ROM könnte durch die Nutzung von Internet-Technologie sinnvoll ergänzt und erweitert werden. Die Präsentation des Datenmaterials kann zeitnah aktualisiert und nahezu ubiquitär verbreitet werden; der Zunahme des Datenmaterials im Verlaufe eines Sprachatlasprojektes sind keine engen Grenzen mehr gesetzt; technische Hürden aufgrund verschiedener Computerhardware oder Be-

triebssysteme können überbrückt werden; durch die direkte Verknüpfung von Karte und Tondokument, d.h. die direkte akustische Impression aus der Kartendarstellung heraus, wird die unmittelbare Interaktion der Benutzenden mit dem Gesamtsystem ermöglicht.

Ein Beispiel für eine internetbasierte Sprachatlaspräsentation ist LAMSAS (*Linguistic Atlas of the Middle and South Atlantic States*, <http://us.english.uga.edu/lamsas/>). Hier werden dialektale Gegebenheiten von 11 mittel- und südatlantischen Staaten der USA auf der Basis von Erhebungen, die zwischen 1933 und 1974 gemacht wurden, dargestellt (Kretzschmar et al. 1993). Die phonetische Umsetzung erfolgt jedoch nicht direkt auf der Karte, sondern dient als Ausgangspunkt für eine tabellenartige Zusammenstellung von Informationen ohne Präsentation der entsprechenden Tondokumente. Der Einsatz proprietärer Windows-Techniken bei der Präsentation der Transkriptionen verhindert deren plattformunabhängige Nutzung.

Für den romanischen Sprachraum wären zwei Projekte zu nennen: Der ALAVAL (*Atlas linguistique audiovisuel du Valais romand*, <http://www.dialecto.unine.ch/AtlasPresent.html>) präsentiert neben den Tondokumenten auch kleinere Videosequenzen, um die dialektalen Äußerungen durch non-verbale Aspekte zu ergänzen. Trotz der angewendeten Videokompressionstechniken ist die Menge der pro Dokument zu übertragenden Daten immer noch sehr hoch, so daß eine umfassende Präsentation des Datenmaterials derzeit nur auf lokalen Datenträgern praktikabel erscheint.

Der *Atlas linguistique parlant d'une région alpine* (<http://www-ellug.ugrenoble3.fr/ellug/livres/atlas/couverture/index.html>) verbindet eine Atlaskarte mit einer Stimuliliste. Der Zugriff auf die Daten ist durch die verwendete, webbrowsersabhängige Skriptsprache problematisch beziehungsweise im schlimmsten Fall sogar unmöglich.

Ein weiteres Projekt der multimedialen Darstellung in der Geolinguistik ist ANAE (*Atlas of North American English*), welches für das Jahr 2001 eine CD-ROM angekündigt hat ([http://www.ling.upenn.edu/phono\\_atlas/home.html](http://www.ling.upenn.edu/phono_atlas/home.html)), über die sich aber zum jetzigen Zeitpunkt (Jan. 2001) noch keine weiteren Aussagen treffen lassen.

### 3. Das Projekt Vivaldi

Im Umfeld des AID entstand die Idee, einen akustischen Sprachatlas Italiens nach Regionen namens "Vivaldi" zu entwickeln. "Vivaldi" ist das Akronym für "Vivaio Acustico delle Lingue e dei Dialetti d'Italia", das italienische Wort *vivaio* (< lat. VIVARIUM) wird hier verwendet im Sinne eines Lebensraumes unterschiedlicher Gattungen.

Ziel ist die Sammlung aktueller Dialektdaten in allen Regionen Italiens und deren Bereitstellung unter Nutzung elektronischer Medien (CD-ROM beziehungsweise Internet). Da keine erschöpfende Erfassung aller Dialekte und Minderheitensprachen Italiens angestrebt wird, sollen pro Region nur etwa 10 bis 12 Orte exploriert werden, nicht zuletzt, um mit geringem Personalaufwand zügig Ergebnisse präsentieren zu können. Das Fragebuch selbst setzt sich aus verschiedenen Komponenten zusammen, den extra-, sozio- und metalinguistischen Gegebenheiten, dem phonetischen, lexikalischen, morphologischen und syntaktischen Teil sowie als zusammenhängendem Text der Übersetzung des *Gleichnisses vom verlorenen Sohn*. Die Entwicklung des Fragebuches und die Ortsauswahl orientieren sich eng am AIS, um die aktuellen Daten mit den Daten des AIS diachron vergleichen zu können (Bauer 1995; Kattenbusch 1995). Insgesamt sollen etwa 350 Stimuli pro SprecherIn erfaßt und verarbeitet werden, bisher wurden die Regionen Sizilien und Sardinien exploriert und umgesetzt.

Das prinzipielle Vorgehen orientiert sich an folgenden Maßgaben (siehe Abbildung 1): Die Interviews werden während der Exploration vor Ort auf DAT-Bändern digital verlustfrei aufgezeichnet. Die Auswertung und Aufarbeitung der Sprachaufnahmen beginnt mit dem digitalen Überspielen der DAT-Kassetten in den Computer – jedes Interview liegt jetzt jeweils in Form einer einzigen, oft mehrere hundert Megabytes großen, digitalen Wave-Datei vor. Diese wird mit Hilfe geeigneter Software in Fragmente segmentiert. Jedes Fragment repräsentiert eine dialektale Antwort, über den Dateinamen lassen sich eindeutig Ort, SprecherIn und Stimulus zuordnen. Jede dialektale Antwort muß auf ihre Qualität (Aussteuerung, Lautstärke, Schnitt, Störgeräusche) kontrolliert und gegebenenfalls modifiziert werden. Mit Hilfe eines vom ALD übernommen Codierungssystems werden die phonetischen Transkriptionen der einzelnen Tondokumente erfaßt. Nach der Endredaktion mit Datenkorrektur sollen die Ton-, Karten- und Transkriptionsdaten für ihre Präsentation auf CD-ROM bzw. für das World Wide Web aufbereitet und zusammengebracht werden. Die Transkription als solche kann vom Computer bisher nicht geleistet werden, sondern muß in traditioneller Weise durch den Dialektologen oder die Dialektologin erfolgen. Jedoch kann dieser Arbeitsschritt – also das Abhören des Tondokuments und die schriftliche Fixierung der Transkriptionszeichen – nun mit Hilfe ein und desselben Geräts erfolgen.

Im folgenden wird näher erläutert, an welchen Punkten im Prozeß der Sprachdatenverarbeitung und -präsentation sich moderne Computertechnologien sinnvoll einsetzen lassen.

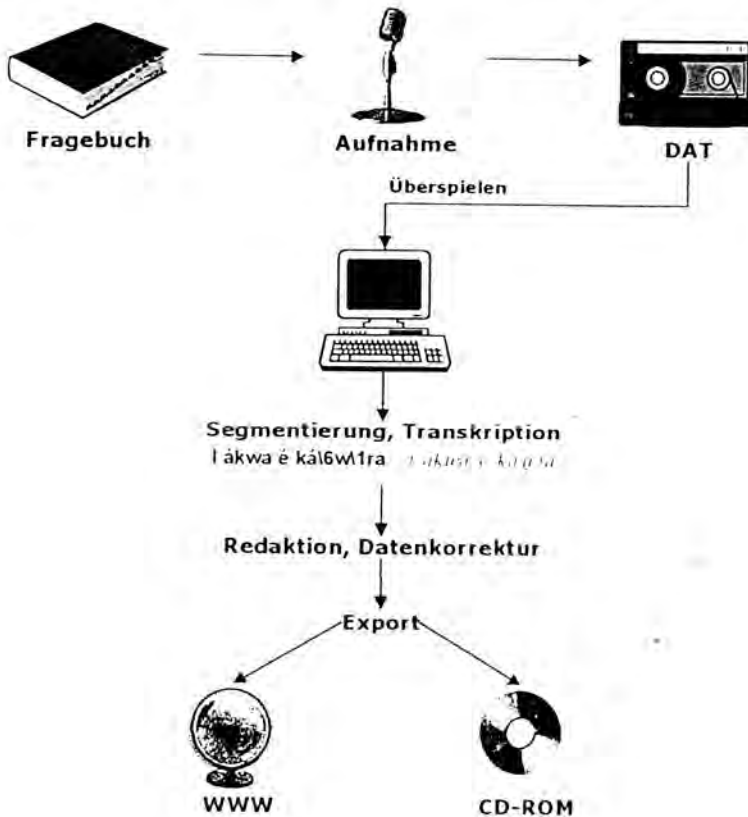


Abbildung 1: Grundsätzliches Vorgehen

### 3.1 Die “Vivaldi Transkriptionsassistentin (TKA)”

Die Entwicklung der “Vivaldi Transkriptionsassistentin (TKA)” ist ein Resultat aus der Anforderung, große Datenmengen verarbeiten und verwalten zu müssen und dies so benutzerInnenfreundlich wie möglich zu gestalten. Daraus ergeben sich zwei Schwerpunkte: der eine liegt auf der Seite der BenutzerInnen, die nach kurzer Einarbeitungszeit ein übersichtlich strukturiertes und leicht zu bedienendes Arbeitsinstrument zur Verfügung haben, der andere betrifft die Realisierung dieser Möglichkeiten durch die darauf abgestimmte Programmierung.

Ein generelles Problem bei der Verarbeitung von Transkriptionen mittels des Computers stellt die Darstellung der phonetischen Zeichen dar, für die

in der Regel keine standardisierten Zeichensätze vorhanden sind. Besonders kritisch ist die Lage, wenn man bei der Transkription nicht das System der IPA (*International Phonetic Association*) einsetzt, sondern sich an einem anderen orientiert – bei VIVALDI z.B. aus Gründen der Vergleichbarkeit an dem von Jaberg und Jud für den AIS entwickelten (vgl. Jaberg und Jud 1928: 24-36). Da gleiches auch für den ALD gilt, hatte man dort schon eine Lösung gefunden, die wir für die TKA übernehmen konnten. Jedem Laut wird ein Sonderzeichen zugeordnet, welches aus einer Matrix von Zeichensätzen entstannt, so daß jedes Zeichen über die Ordnungsnummer seines Zeichensatzes und seiner Position im Zeichensatzalphabet definiert ist (Abbildung 2). Dieses Vorgehen hat den Vorteil, daß sich beim Auftreten bisher noch nicht codierter phonetischer Phänomene problemlos ein neues Zeichen hinzufügen läßt.

Stimulus	Code	Zeichen
l'agnello	l a\6n\1é\6FF\5U\1	l a ñ é ð ð u

Abbildung 2: Stimulus mit Transkriptionscode und transkribierten Zeichen

Ursprünglich war vorgesehen gewesen, die ebenfalls im Rahmen des ALD-Projekts entwickelte, auf MS-DOS und Clipper/dBase basierende Erfassungssoftware CARD zur Verwaltung der Daten zu nutzen. Diese unterstützt unter anderem die Eingabe und Codierung der Transkriptionen sowie den Export und die Druckausgabe auf Plottern, enthält darüber hinaus jedoch eine Vielzahl von Funktionsmöglichkeiten, die für das primäre Ziel der Erstellung eines sprechenden Sprachatlas nicht benötigt wurden ([http://www.sbg.ac.at/rom/people/proj/ald/card/card\\_home.htm](http://www.sbg.ac.at/rom/people/proj/ald/card/card_home.htm)).

Nachteilig war dabei, daß CARD eine Umsetzung der Codes in Zeichensätze erst zum Zeitpunkt der Druckausgabe beziehungsweise nach dem Export in ein Windows™-basierendes Anzeigemodul realisiert, eine direkte Kontrolle der Richtigkeit der eingegebenen Codes somit nicht möglich ist. Zudem ist das Abspielen des zu einer Transkription gehörigen Tondokuments zur Kontrolle mit einer zeitraubenden Suche im Dateisystem verbunden.

Wir entschieden uns für die Neuprogrammierung der Eingabesoftware, da der dazu nötige Mehraufwand die Zeitersparnis beim Transkribieren zu rechtfertigen versprach. Die Programmierung sollte folgenden Anforderungen genügen:

- Direkte audiovisuelle Kontrolle der eingegebenen Codes (Verknüpfung mit der Tondatei, Anzeige der transkribierten Zeichen)
- Gleichzeitige Darstellung mehrerer Stimuli in einer Matrix aus Stimulus und Ort beziehungsweise Sprecher, leichte Erweiterbarkeit
- Schnelle Programmierbarkeit und Benutzungsfreundlichkeit durch Standard-Softwarekomponenten auf der Basis von Microsoft Windows™
- Exportfunktionalitäten

Als Programmierumgebung für die “Vivaldi Transkriptionsassistentin (TKA)” entschieden wir uns für das Tabellenkalkulationsprogramm Microsoft Excel™ – eine Wahl, die zunächst überraschen mag, da ein Datenbankmanagementsystem sicherlich näher gelegen hätte. Excel bietet einige Vorteile: neben den relativ geringen Anschaffungskosten erleichtert die vertraute Umgebung die Benutzung. Die Excel-Oberfläche bringt bereits eine Menge erwünschter Funktionen der Benutzungsschnittstelle mit sich. Ergänzt durch die mächtige, objektorientierte Programmiersprache *Visual Basic for Applications* bereitet so die Programmierung der Basisfunktionalitäten zur Eingabe und Kontrolle der Transkriptionen keine großen Probleme. Damit war die Programmierung der Benutzungsschnittstelle relativ schnell erledigt, der Schwerpunkt konnte mehr auf erweiterte Import- und Exportfunktionalitäten und das Datenmanagement gelegt werden – die Vivaldi Transkriptionsassistentin übernimmt also die Rolle eines Regiezentrum, welches die Ton-, Karten-, Transkriptions- und Metainformationen unter einer gemeinsamen, intuitiven Benutzungsoberfläche integriert und diese in diverse Standardformate (z.B. dBase, PDF, Postscript, RTF, HTML) exportiert.

Abbildung 3 zeigt die Programmoberfläche: In der Excel-üblichen Tabellendarstellung sind die Stimuli zeilenweise und die SprecherInnen und Orte (gelegentlich wurden mehrere SprecherInnen pro Ort erfaßt) spaltenweise angeordnet, die verschiedenen Orte einer Region werden in einem Tabellenblatt zusammengefaßt, jedes Tabellenblatt repräsentiert also eine der 20 verschiedenen Regionen. Diese Anordnung der Stimuli, Sprecher beziehungsweise Orte und Regionen in Zeilen, Spalten und Tabellenblättern ermöglicht eine übersichtliche Darstellung und leichte Navigation mit kurzen Zugriffszeiten.

An den Schnittpunkten der Zeilen und Spalten finden sich jeweils drei Felder für den Transkriptionscode, dessen Umsetzung in die entsprechenden Zeichen und ein zusätzliches Feld für Metainformationen. Nach der Eingabe des Transkriptionscodes wandelt ein Programmmodul diesen in die entsprechende Zeichenkette um und färbt sie grün, wenn die korrespondierende Tondatei vorhanden ist, anderenfalls rot ein.



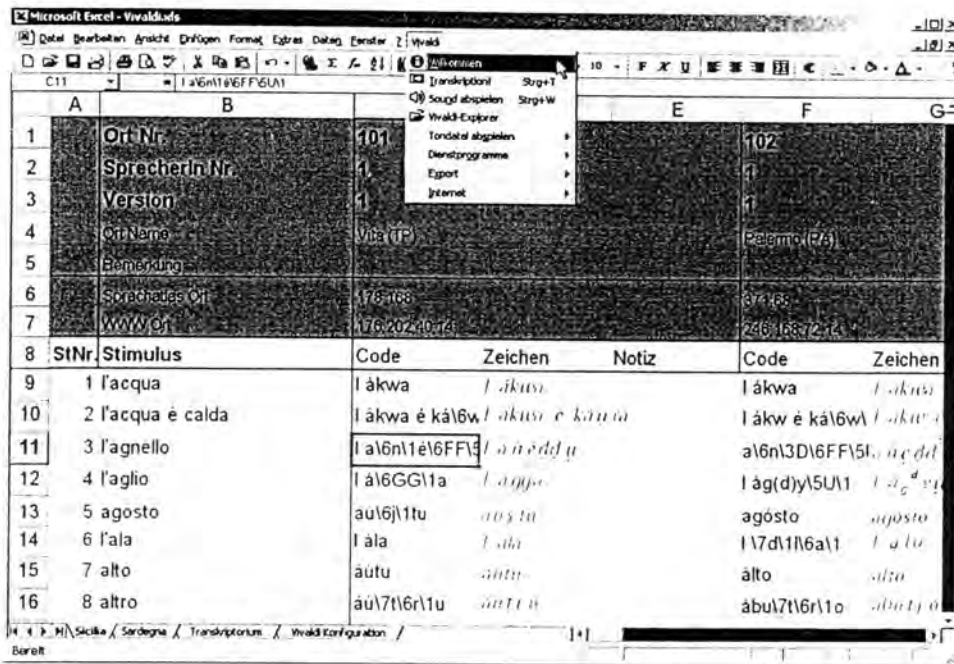


Abbildung 3 Vivaldi Transkriptionsassistentin

Ein Shortcut erlaubt das direkte Abspielen der Tondatei (Abbildung 4), hier greift das Programm auf eine Windows-Standardkomponente zur Ausgabe von Tondateien zurück. Über einen anderen Shortcut kann die Tonbearbeitung mit der entsprechenden Spezialsoftware direkt aktiviert werden.

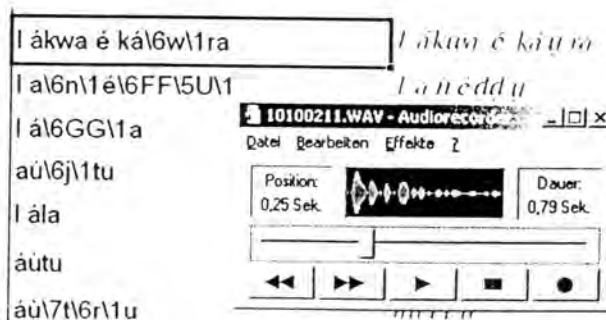


Abbildung 4 Abspielen des Tondokuments

Wie bereits erwähnt, konnten wir dank der raschen Umsetzung der Benutzungsschnittstelle mit den Grundfunktionalitäten zur Bearbeitung der Transkriptionen einen Schwerpunkt auf die Exportschnittstellen legen: Die Vivaldi Transkriptionsassistentin unterstützt, um die Kompatibilität zum Datenbankformat des ALD zu erhalten, einen Datenexport im CARD-dBase-Format, so daß theoretisch alle Daten direkt in CARD weiterverarbeitet werden könnten beziehungsweise mit dem ALD-Sprachatlasprogramm angezeigt werden können. Ein Export im dBase-Format wahrt die Kompatibilität mit Datenbankmanagementsystemen.

Weitaus wichtiger für das Vivaldi-Projekt ist der Export im Rich-Text-Format (RTF), einem weit verbreiteten und relativ plattformunabhängigen Textformat. Damit lassen sich ganze Tabellenblätter oder nur Teilbereiche der Matrix originalgetreu in eine Textverarbeitung, etwa zur Anfertigung von Korrekturlisten, transportieren (Abbildung 5).

Mit der Generierung von Postscript-Code, einer weit verbreiteten und gut standardisierten Programmiersprache zur Ansteuerung von Druckern und Plottern, eröffnet sich wie im ALD die Möglichkeit, direkt hochwertige Sprachatlaskarten zu generieren. Das Portable-Document-Format (PDF) der Firma Adobe hat sich als Webstandard zur originalgetreuen Darstellung von Texten etabliert, die TKA unterstützt dieses ebenfalls unter Nutzung einer geeigneten Zusatzsoftware.

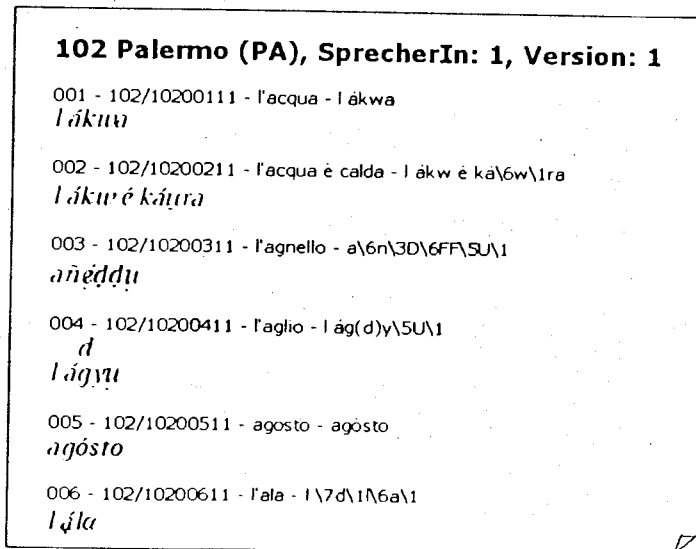


Abbildung 5: Beispiel für RTF-Export

Zusätzlich wird das Transkriptorium, die phonetische Beschreibung der verwendeten Zeichen, innerhalb der TKA verwaltet. Das Programm verbindet außerdem automatisch die Zeichenbeschreibungen mit Beispielen aus den entsprechenden Tondokumenten.

Zuletzt kann die Transkriptionsassistentin den gesamten Inhalt in eine HTML-Präsentation exportieren, die – auf CD-ROM oder über das Internet verbreitet – mit jedem beliebigen Webbrowser dargestellt werden kann.

### 3.2 Multimediale Präsentation

Ein weiterer wesentlicher Aspekt des Vivaldi-Projektes ist die Präsentation der Daten in direkter Verbindung von Karte, Tondokument und Transkription. Der traditionelle, statische Sprachatlas in Buchform, gegebenenfalls ergänzt durch das Zuordnen der Tondokumente auf einer Schallplatte oder CD, muß naturgemäß auf eine direkte Integrierung dieser Informationsquellen verzichten. Erst die Nutzung des Computers erlaubt eine dynamische, interaktive Präsentation, hier bietet sich die Internet-Technologie mit der *Hypertext Markup Language* (HTML) an. Dabei spielt es im allgemeinen keine Rolle, ob die Informationen lokal (beispielsweise von CD-ROM) oder über ein Netzwerk von einem fremden Rechner eingelesen werden.

Eine Internetpräsentation kann im einfachsten Falle aus einer Reihe durch Hyperlinks verbundener, statischer HTML-Dokumente bestehen. Ein interaktiver Sprachatlas dagegen muß dynamisch die Text-, Bild- und Toninformationen verbinden – z.B. soll die Auswahl eines Ortes auf einer Karte die Darstellung der korrespondierenden Transkription und das Abspielen des zugehörigen Tondokumentes auslösen. Für eine dynamische, webbasierte Präsentation von multimedialen Inhalten haben sich zahlreiche Standards etabliert. Die serverbasierten Programmiersprachen (z.B. php, ASP, Perl) kommen nicht in Betracht, da der hierfür notwendige Einsatz eines Webserver der Forderung einer hybriden Präsentation auf CD-ROM und im Internet widerspricht. Die aktive Komponente muß also im Webbrowser liegen, hierfür kommt eine Reihe von Standards in Frage (Java™, Java Script, Macromedia Flash™). Java (<http://java.sun.com>) hat sich als Internet-Programmiersprache etabliert und eine weite Verbreitung gefunden. Wir entschieden uns vor allem deshalb für Java, weil die meisten Webbrowser diese Umgebung unterstützen, ohne daß proprietäre Zusatzprogramme (*plugins*) vorhanden sein müssen.

Die Internetpräsentation des Vivaldi-Projektes (<http://www2.hu-berlin.de/Vivaldi>), die im Herbst 1998 in ihrer ersten Version gestartet ist, besteht zunächst aus mehreren statischen HTML-Dokumenten zur Anzeige allgemeiner Informationen und Hintergründe zum Projekt (Abbildung 6).

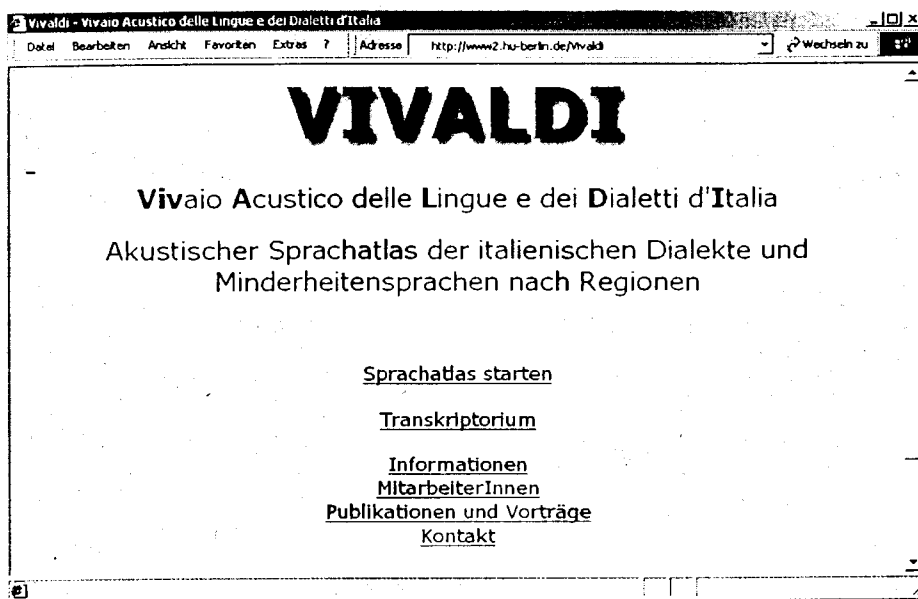


Abbildung 6: Vivaldi-Startseite

Abbildung 7 zeigt die eigentliche Darstellung der interaktiven Atlaskarten beispielhaft für Sizilien: das Klicken auf einen der Stimuli im Navigationsframe, dem linken Teil des Bildschirmausschnitts, lädt im Hauptframe die entsprechende Kartendarstellung mit einer Liste aller verfügbaren Transkriptionen. Bewegt man den Mauszeiger über einen Ortsnamen auf der Karte, wird am oberen Rand sofort die Transkription angezeigt und das entsprechende Tondokument abgespielt. Unter der Karte findet sich eine statische Liste aller verfügbaren Transkriptionen mit Links auf die Tondateien. Der Link "Transkriptorium" öffnet die aus der TKA exportierte Liste der Zeichenbeschreibungen mit den darin automatisch generierten Verweisen zu den entsprechenden Beispielen aus den Tondokumenten.

Technisch betrachtet versteckt sich hinter der Kartendarstellung die aktive Komponente, das Java-Applet, welches die Informationen über die auszugebenden Daten aus der HTML-Datei, in die es eingebettet ist, bezieht. Zunächst lädt es die Karte der gewünschten Region, anschließend werden die Hotspots definiert und angezeigt. Die Aktivierung eines Hotspots löst jetzt das Abspielen der in der HTML-Datei definierten Sounddatei und die Anzeige der Transkriptionsgrafik aus. Damit kann das gleiche Applet für die Darstellung aller Karten und Stimuli eingesetzt werden, nur

der Export der Anzeigeinformationen aus der Vivaldi Transkriptionsassistentin in die zu einem Stimulus gehörige HTML-Datei bestimmt die individuelle Erscheinungsform.

Den Prinzipien guter Webprogrammierung folgend haben wir reines Java (100% pure Java) eingesetzt und auf alle anderen Erweiterungen (Java Script, Cookies, Plugins oder ähnliche) verzichtet. Damit können die Sprachatlaskarten mit jedem Webbrowser, der Java unterstützt, dargestellt werden.

Navigationsframe    Hauptframe    Transkription    Java-Applet    Hotspot

VIVALDI - Vivaldi Acustico della Capota e dei Dialetti d'Italia

Sicilia: in maggio fa già molto caldo  
*na māyu fà gā assā: kāuru*

**VIVALDI**

101 Vico (TP)  
*na māyu fà gā kāuru assā*

102 Palermo (PA)  
*a mādyu fà gā u oppu kaldu*

103 Piana degli Albanesi (PA)  
*te māi bān ān sum vāu*

Abbildung 7 Interaktive Sprachkarte

#### 4. Ergebnisse und Ausblick

Die Erfahrungen beim Einsatz der Vivaldi Transkriptionsassistentin zeigen, daß damit die Generierung eines interaktiven Sprachatlas erheblich beschleunigt wird. Vorteilhaft ist insbesondere die Unterstützung der Transkription, die Minimierung von Fehlern, die Erhöhung der Transparenz bei der Verwaltung einer nicht unerheblichen Datenmenge und der rasche Datenexport in Fremdformate bei geringen Zusatzkosten.

Die Internetseiten dokumentieren zeitnah den aktuellen Projektstand, es dauert so nicht Jahre, bis neue Ergebnisse publiziert werden. Damit kann kostengünstig einem breiteren Zielpublikum nahezu ubiquitär die Geolinguistik erschlossen werden: eine Rezeption der Daten ist nicht auf kostspielige Atlanten, die meist nur in einigen Bibliotheken verfügbar sind, beschränkt. Für die intensive Auseinandersetzung mit dem gesamten Datenmaterial unabhängig von Übertragungsraten und –kosten ist eine CD-ROM weiterhin das Mittel der Wahl. Diese läßt sich ohne Mehraufwand erstellen, die Plattformunabhängigkeit ermöglicht hier ebenfalls eine breitere Anwendung. Durch die vorhandene Exportfunktion der TKA wäre sogar die Herstellung eines traditionellen Kartenwerks unproblematisch zu realisieren. Die Offenheit des Modells und der Softwarekomponenten setzen der Erweiterung des Datenmaterials bei fortschreitender Exploration kaum Grenzen, auch eine Nutzung des Modells für ein anderes Sprachatlasprojekt wäre durchaus denkbar.

Der Einsatz eines leistungsfähigen Notebooks bei den Explorationen wird es in Zukunft ermöglichen, die Tondokumente direkt, also ohne zwischengeschaltete DAT-Bänder, aufzunehmen und zu verarbeiten. Dies verringert den Geräte- und Materialbedarf, spart dadurch Kosten und minimiert eventuelle Übertragungsverluste und –fehler.

Für die nähere Zukunft planen wir, die Kartendarstellung in Richtung eines interaktiven Lehr- und Lernsystems mit der Möglichkeit zur individuellen Datenauswahl (z.B. Playlists) zu erweitern.

Weiterhin ist an die Realisierung einer verteilten, internetbasierenden Dateneingabe mit einem zentralen Datenbankserver auf der Basis der frei verfügbaren Webserver-Programmiersprache php gedacht. Mehrere ExploratorInnen könnten dann unabhängig voneinander ihre Transkriptionsdaten und Tondateien zum Server übertragen, so daß diese rasch für die Internetpräsentation bereitstünden.

Im Vordergrund steht aber vor allem die Intensivierung der Exploration, sofern die entsprechenden finanziellen Mittel zur Verfügung stehen.

## 5. Bibliographie

- Bauer, Roland. 1995. "Vivaldi-Sicilia. Documentazione sonora dei dialetti siciliani." In: Giovanni Ruffino, Hrsg. 1995. *Percorsi di geografia linguistica. Idee per un atlante siciliano della cultura dialettale e dell'italiano regionale*. Palermo, 543-50.
- Cortelazzo, Manlio, Hrsg. 1974-1988. *Profilo dei dialetti italiani*. Pisa. 17 Bde., einige mit Schallplatten.
- Gilliéron, Jules, und Edmond Edmont, Hrsg. 1902-1910. *Atlas linguistique de la France (ALF)*. 10 Bde. Paris.
- Goebel, Hans, Hrsg. 1998. *Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 1. Teil (ALD I)*. Wiesbaden. 4 Kartenbände, 3 Indexbände, 3 CD-ROM.
- Harder, Andreas, und Fred Boller. 1996. *Sprachgeographie und PC. Sprachkarten, Datenorganisation, Tonproben mit Mikrorechnern*. Kiel.
- Hotzenköcherle, Rudolf, und Rudolf Brunner, Hrsg. 1972-1976. *SDS-Phonogramme. Tonaufnahmen für den Sprachatlas der deutschen Schweiz*. Zürich. 16 Langspielplatten.
- Jaberg, Karl, und Jakob Jud. 1928. *Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz*. Halle (Saale).
- , Hrsg. 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen, 8 Bde.
- Kattenbusch, Dieter. 1995. "Atlas parlant de l'Italie par régions: VIVALDI." In: *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*. Barcelona, 443-55.
- Kretzschmar, W.A., V. McDavid, T. Lerud und E. Johnson, Hrsg. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago.

Marcel Lucas Müller,  
 Institut für Medizinische Informatik und  
 Biomathematik,  
 Westfälische Wilhelms-Universität Münster,  
 Domagkstr. 9, D-48129 Münster  
 E-mail: mamulle@uni-muenster.de

Carola Köhler,  
 Institut für Romanistik,  
 Humboldt-Universität Berlin,  
 Dorotheenstr. 65, D-10099 Berlin  
 E-mail: carola.koehler@rz.hu-berlin.de

Dieter Kattenbusch,  
 Institut für Romanistik,  
 Humboldt-Universität Berlin,  
 Dorotheenstr. 65, D-10099 Berlin  
 E-mail: dieter.kattenbusch@rz.hu-berlin.de