The Linguistic Atlas of the Iberian Peninsula (ALPI): A geolinguistic treasure ‘lost’ and found*

David Heap
University of Western Ontario

One of Jack’s Dialectology assignments required us to examine data from a linguistic atlas: in my case, this project put me on a trail which would eventually lead me to the *Atlas Lingüístico de la Península Ibérica* (Linguistic Atlas of the Iberian Peninsula or *ALPI*, by its Spanish acronym). I was not able to work on the *ALPI* data at that time (it was hard to find even the one volume published in 1962), and it took me more than a decade to eventually track down the unpublished data from this major dialect survey. Nonetheless, it was that project in Jack’s course which eventually led me to what became both a voyage of discovery and a major research publication: the online *ALPI* will soon provide more than 36,000 pages of data from the original notebooks, exactly as gathered by the *ALPI* fieldworkers between 1930 and 1956.

In 1990 I took Jack Chambers’ Dialectology course as part of my doctoral programme in Romance Linguistics at the University of Toronto. The first term assignment required us to examine data from a linguistic atlas—any linguistic atlas. Since we had already looked at Gilliéron and Edmont’s (1902–1910) *Atlas linguistique de la France*, my interest in Romance dialectology sent me in the direction of the Iberian Peninsula. Little did I suspect then that Jack’s course assignment was to start me on a journey that would lead to the continuation of a dialectological project begun almost a century earlier and which had lain dormant for decades: the *Atlas Lingüístico de la Península Ibérica* (Linguistic Atlas of the Iberian Peninsula or *ALPI*, by its Spanish acronym). I was not able to work on the *ALPI* data at that time (it was hard to find even the one published volume, *ALPI* 1962), and it took me more than a decade to track down the unpublished data from this major dialect survey. Nonetheless, that assignment in Jack’s course eventually led to what became both a voyage of discovery and a major research undertaking: the online *ALPI* (www.alpi.ca), an electronic dialect database which will provide more than 36,000 pages of transcriptions from original survey notebooks, exactly as gathered by the *ALPI* fieldworkers between 1930 and 1956.

* Many thanks to Inés Fernández-Ordóñez, Stephanie Kelly and Enrique Pato for their help and support in preparing this paper, to Sarah Cummins for her gentle but firm efforts to rein in my prose, especially my proliferation of full colons, and to Bridget Jankowski for her patient and precise copy-editing; all remaining errors, omissions and infelicities are of course the fault of this author alone. This work supported in part by the Social Science and Humanities Research Council of Canada (Heap 410-2003-0461).

Copyright © 2008 David Heap
There was a fair amount written by the 1950s about the linguistic cartography of Spain and Portugal (e.g. Sanchis Guarner 1953), and this was sufficient material available for me to put together a sketch of Iberian dialectology for Jack’s course. The ALPI was originally conceived of by Ramón Menéndez Pidal, the eminent Hispanic philologist to whom we owe (among other things) the rediscovery of medieval epic poems or romances as part of living oral literature throughout the Spanish-speaking world. Already in the early 1900s, when he took up his Chair in Hispanic philology at the University of Madrid, Menéndez Pidal considered that the creation of a dialect atlas of Iberian Romance (along with a history of the Spanish language and a Romancero or collection of traditional oral literature) was among the scholarly needs which had to be addressed in the field. The publication of Gilliéron & Edmont’s Atlas Linguistique de la France (1902–1910) had already begun and the potential of linguistic atlas surveys as fundamental materials for research on vernacular language was very much in the air of the times.

It would be decades before Menéndez Pidal’s project of an Iberian linguistic atlas would become a reality, but in the meantime he made the crucial decision to involve one of his most promising disciples, Tomás Navarro Tomás, in the atlas project, and to send him abroad to study dialectology and phonetics at French, Swiss and German universities (Pérez Pascual 1999). By the early 1920s Menéndez Pidal, along with Navarro Tomás and other students, had carried out initial fieldwork on the Leonese-Asturian border in northern Spain, covering the rugged terrain on horseback. 1923 also saw the first published mention of the ALPI, in the Revista de Filología Española. In 1928 Navarro Tomás, by this time a renowned specialist in Spanish phonetics and phonology, spent a year as visiting professor in Puerto Rico, where he conducted dialect surveys with a protocol which was later to form the core of the ALPI questionnaire; the data from these surveys would eventually become the first linguistic atlas of a Spanish-speaking territory (Navarro Tomás 1948). Even while on vacation in Caracas, Navarro Tomás spent part of his trip completing a few dialect surveys; he wrote to Menéndez Pidal that while interviewing some subjects in the street he had nearly been caught in round-up of political demonstrators by Venezuelan police, but that this was not a problem since if imprisoned he could always interview the jailguards or the other prisoners (a remark which eerily foreshadows the fate of some of the ALPI fieldworkers in Spain a decade later).

In 1930, with the backing of the Centro de Estudios Históricos (Centre for Historical Studies) in Madrid, the ALPI surveys finally got under way. Navarro Tomás gave intensive training seminars in phonetic transcription and dialect survey techniques to a team of six fieldworkers, two for each of the major Ibero-Romance linguistic zones: Manuel Sanchis Guarner and Francesc de Borja Moll for the Catalan-Valencian area, Aurelio M. Espinosa (Jr.) and Lorenzo Rodríguez-Castellano for the Castilian-Andalusian zone, and Aníbal Otero with Rodrigo de Sá Nogueira for the Galician-Portuguese zone.

Navarro Tomás, with input from Menéndez Pidal and other scholars, prepared a questionnaire made up of two notebooks, Cuaderno I (411 questions on Phonetics and Grammar) and Cuaderno II (Lexicon), the latter in two versions (IIG or General, with 413 questions, and IIE or Extended, with 833 questions), although the Extended version ended up being used more frequently than the General one. Normally, a pair of
fieldworkers would work together in a selected community, each of them completing either Cuaderno I or Cuaderno II with a different subject. The initial teams of two ended up varying due to different circumstances: Sá Nogueira withdrew from the project early on, for health reasons, and was replaced by Armando Nobre de Gusmão for some of the survey points in Portugal; meanwhile, Otero conducted all of the Galician surveys (both Cuadernos I and II) on his own, travelling on foot or by public transport. Since the fieldworkers had to fit ALPI surveys in between other duties and commitments, they would hit the road in different combinations according to their availability, especially in Castilian-speaking areas where they could all conduct surveys (for example, Otero joined Espinosa for some survey points in the north, while Rodríguez-Castellano covered some points in Teruel with Sanchis Guarner; Navarro Tomás himself took part in at least one survey, point 544, Cadalso de los Vidrios in the province of Madrid).

The ALPI fieldwork continued during the years 1931–1936, as described by Moll in his memoirs (1970), and by Rodríguez-Castellano and Sanchis Guarner in their road diary (reproduced in Casanova 2004), among others (Navarro Tomás 1975). The overwhelming majority of the points had been surveyed by June 1936, when the fieldwork was interrupted by the military uprising against the Republican government which was to become the Spanish Civil War. The Phalangist-Fascist rebellion not only led to the suspension of fieldwork, it also caught some of the ALPI fieldworkers behind enemy lines. Espinosa, a native of New Mexico, made his way rapidly from the Burgos area (where he was collecting oral literature; see Espinosa 1988, de Prada Samper 2004) to Portugal, where he eventually got a ship back to the U.S. Otero was less fortunate: he was doing ALPI fieldwork in northern Portugal at the time, and although he attempted to return to Republican territory via the Extremadura border (in west-central Spain), the Portuguese authorities instead turned him over to the Spanish military at Tuy, in his native Galicia, which had fallen under rebel military control at the start of the uprising. On the strength of ‘evidence’ that he was a spy—his avowed sympathy for the Republican cause, the fact that his ALPI fieldwork was funded by the Republican government in Madrid, and especially those suspicious notebooks he had with him, which were apparently full of incomprehensible writing in ‘code’—a summary military tribunal convicted Otero of high treason and sentenced him to death by firing squad.

Tipped off by family members of another condemned inmate, Menéndez Pidal (by then self-exiled in Havana) wrote an urgent message in support of Otero, arguing that his work on the ALPI was of great scientific importance for the prestige of Spain. Other evidence (including testimony from an ideologically reliable cleric with philological training) was produced to show that his notebooks of ‘spy code’ in fact contained phonetic transcriptions, and Otero’s appeal was partially successful: his death sentence was commuted to life imprisonment. He was eventually pardoned and released in 1941 (two years after the end of the Civil War), having spent more than five years in prison for the crime of linguistic fieldwork and Republican sympathies (Catalán 2001). He had however continued his research during his years in various military prisons: following the ALPI methodology, he surveyed different fellow prisoners’ speech, while also collecting traditional oral literature (romances) which he transcribed from memory after his release and sent to Menéndez Pidal, who had by then returned to Madrid.
The rest of the *ALPI* team was dispersed as a result of the Civil War and the ensuing upheaval: Moll returned to his native Mallorca, Navarro Tomás became Director of the National Library, among other positions, in the Republican government, while Sanchis Guarner became an officer in the Republican army, was imprisoned following the war and released in 1942. The nearly-complete *ALPI* field notebooks were evacuated from the Centro de Estudios Históricos in Madrid (along with other national treasures, including Menéndez Pidal’s own papers and paintings from the Prado) to the successive seats of the Republican government in Valencia and Barcelona, and eventually sent into exile. Navarro Tomás was one of the thousands of Republican refugees who fled Franco’s military victory (he helped Antonio Machado cross the Pyrenees to Collioure in France, which was to be the great Spanish poet’s final resting place). Navarro Tomás’ exile took him to France and then to the U.S., where he obtained a teaching position at Columbia University in New York.

Like many exiled Spanish Republicans, Navarro Tomás remained hopeful about a return of legitimate democratic government to Spain, and he kept the *ALPI* field materials safe with him, looking forward to the day when he could return to Madrid and complete the project. He published the Puerto Rico atlas materials in 1948, and proposed and published a questionnaire for a Hispanico-American atlas (Navarro Tomás 1943). When it became painfully clear that the Allied victory in Europe would not extend to removing fascism from Spain as well (on the contrary: Franco’s regime became a U.S. ally), Navarro Tomás, realising that there was now little hope of returning to a free Spain to continue his work, began negotiating the return of the *ALPI* materials to Madrid, in order for the project to be completed. His one condition for the return of the *ALPI* materials was that the remaining fieldwork and the preparation of the atlas for publication be carried out by members of the original fieldwork team trained by him. In the late 1940s Sanchis Guarner, Otero and Rodríguez-Castellano were granted researcher status by the Consejo Superior de Investigaciones Científicas (CSIC, or Higher Council of Scientific Research), the Francoist body which has replaced the former Centro de Estudios Históricos. In 1947 they began work on the remaining fieldwork points: contemporary documents include their government permits to circulate and interview people in various provinces, at a time when movements of any sort were closely scrutinized by Francoist authorities. In 1951, Sanchis Guarner and Rodríguez-Castellano were allowed to travel to New York to collect the *ALPI* fieldwork materials from Navarro Tomás, along with his instructions for how to proceed with editing and mapping the data. From 1952 to 1956 Otero, joined by the Portuguese scholar Lindley Cintra (and occasionally by Sanchis Guarner as well) completed the remaining surveys in Portugal, bringing the final total to 527 points surveyed.

The trail of the *ALPI* data goes quite cold by the mid 1960s: after the publication of the first and only volume in 1962 there are a few reviews (one of them quite negative, by Manuel Alvar (1964), who by this time was promoting his own series of regional linguistic atlases in various parts of Spain) and references by scholars such as Diego Catalán (Menéndez Pidal’s grandson), who used data from the first *ALPI* volume in a couple of studies (1964, 1974, 1975). Near the end of his life, Navarro Tomás published his *Noticia histórica del ALPI* (1975) in which he notes that the publication had been
suspended for lack of funds and makes a plea for the preservation of the ‘now historical’ ALPI materials by the CSIC until such a time as they can be properly published.

The interest in the ALPI and its fate (which Jack’s assignment had sparked) stayed with me, especially since my later work with Yves Roberge in his Groupe de recherches en dialectologie comparative at the University of Toronto led me to look at linguistic atlases and dialect studies in general as underutilised and neglected sources of data for studying morphosyntactic variation. I worked on pronominal variation in other Romance varieties but never forgot the goal of tracking down the ALPI materials to see what interesting data they might include. In 1993 I posted a query to LinguistList (http://linguistlist.org/issues/4/4-566.html) seeking information about the ALPI data but received no useful leads.

When I lived in Spain in 1994–1995, I again tried to find out about the ALPI materials but hit a brick wall: although Navarro Tomás (1975) suggests that the field notebooks be preserved at the Consejo Superior de Investigaciones Científicas, nobody at the CSIC in Madrid could tell me anything about their whereabouts. It was however suggested that Sanchis Guarner, the coordinator of cartography for the publication of the first volume, had had all of the ALPI notebooks in his possession at the time of his death in the early 1980s, and had willed the material (along with the rest of his papers) to a private cultural foundation in his native Valencia, the Acció Cultural del País Valencià. I was lucky in that another one of my University of Toronto professors, renowned Catalan philologist Joseph Gulsoy, was able to put me in touch with colleagues at the Universitat de València who had worked with Sanchis Guarner, and who confirmed that the ALPI materials had indeed been in the Acció Cultural del País Valencià archives up until the early 1990s. At that time the archives had closed and all of Sanchis Guarner’s papers had been sent to Barcelona for safekeeping until another archive in Valencia could accommodate them. Through the Institut d’Estudis Catalans in Barcelona, I was eventually able to track down what had now become known as the ‘Fons Manuel Sanchis Guarner’ i.e. the collection of all of his scholarly papers, at a library run by the Universitat de Barcelona which specialised in materials from the Republican years. The staff at the Pabelló de la República made it very clear to me that they were merely taking care of the Fons Manuel Sanchis Guarner until it could be returned to Valencia, but they nonetheless allowed me to examine and photocopy as much of the materials as I wished.

When I examined the materials left by Sanchis I found that they only included about one third of the ALPI field notebooks; his correspondence indicated that Otero and Rodríguez-Castellano had the remaining two thirds. That is, for the purposes of editing and mapping the data, the three had divided the materials up, along the same lines as the original fieldwork teams: from about 1957 onward, Otero got the Portuguese-Galician materials, Rodríguez-Castellano had most of the notebooks from the Castillian-Andalusian area, while Sanchis kept the notebooks from the Catalan-Valencian zones as well as some bordering areas.

So now the task became one of tracking down the notebooks which had been in the possession of the other two fieldworkers. Through Inés Fernández-Ordóñez (from the Universidad Autónoma de Madrid), I was able to make contact with Diego Catalán, who was very helpful in allowing me access to correspondence between his grandfather...
and the various ALPI fieldworkers held in the vast collection of papers at the Fundación Ramón Menéndez Pidal in Madrid. With encouragement and help from Inés and Diego, I eventually tracked down Lorenzo Rodríguez-Castellano’s daughters Aurora-Isabel and Matilde, living in Madrid and Gijón, respectively. They confirmed to me that the papers their father had left included the ALPI notebooks. The two sisters (both secondary school teachers) were quite aware of the importance of these papers for scholarly research, but they were in no hurry to turn them over to any Spanish institution; in fact, they had at one time considered donating them to a university in the U.S. where one of their uncles worked. They seemed quite happy that I had come along—“a North American, like in a film”—and was interested in their father's research papers. In May of 2001 Inés and I travelled with Aurora-Isabel Rodríguez-Castellano to her parents’ now-empty flat in Oviedo, where we found the ‘missing’ ALPI field notebooks, perfectly conserved and neatly bundled by province, as one might expect a professional librarian such as Lorenzo to have left his materials.

The remaining third of the ALPI notebooks (the Galician-Portuguese materials) turned out to be in the collection of the Instituto da Língua Galega (Galician Language Institute or ILG) in Santiago de Compostela, where Aníbal Otero’s family had deposited them after his death in 1973. Professor Antón Santamarina, who was in charge of the materials at the ILG, had been a personal friend of Otero’s, and considers the ALPI notebooks to be held ‘in trust’ for Otero’s family. As a consequence, he would only allow me to examine the materials, not copy them, unless I could get an authorisation from Horocel Otero, the fieldworker’s son and only surviving heir. This presented something of a challenge, as Horocel had no fixed address at that time: he lived in a series of residential hotels and spent much of his time in different cafés. Luckily, I was able to track him down in the nearby port city of Vigo, where I arranged to meet him at one of his regular spots, the Cafeteria Arrondo. He was very pleased that someone was finally taking an interest in his father’s work and signed a letter on the spot authorising me to copy and publish any of the ALPI notebooks his father had left. With this authorisation in hand, I was able to obtain a complete copy of the remaining ALPI materials from the ILG.

So the most complete collection of ALPI fieldwork materials in existence is now held (in photocopy form) at the Theoretical and Applied Linguistics Laboratory (or TALL, at the University of Western Ontario: http://ling.uwo.ca), where the copies fill two large filing cabinets. At the time of writing, the last missing pages and ‘stray’ notebooks have been replaced, largely from the Fons Manuel Sanchis Guarner, which has since been returned to València, where it is now held at the Arxiu Històric Municipal d’Elx, near Alicante.

Now that we had the material, the question became: what to do with it? The original project—producing a fully mapped, traditional dialect atlas—would require enormous resources: specially trained transcribers and cartographers, as well as expensive printing in such a large format. Since the high costs of producing the 1962 volume were the main reason for halting the original publication, it seemed unlikely that any funding source could be found which would be willing to finance such a vast editorial and publishing undertaking. Furthermore, as had been pointed out at least since reviews
of the first volume (Catalán 1964), linguistic display maps are not always the most useful medium for researchers to use when searching for data:

...it would even be appropriate to abandon cartographic representation whenever searching on a map gets in the way of consulting data, rather than facilitating it: simple lists of results would in many cases save editors’ time and money while sparing interested linguists’ eyesight. (Catalán 1975, our translation)

However, even transcribing and editing lists of data for publication (as was done, for example, with the Survey of English Dialects, Orton et al. 1962–1971) would entail printing and labour expenses well beyond what was (and is currently) available for this project, quite apart from years of painstaking work. In order to make the ALPI data available to the international scholarly community as quickly as possible, the ideal solution turns out to be the internet: scanned images of the original fieldwork notebook pages can be posted to webpages at a fraction of the costs of printing, and with minimal delay. The only requirement is an internet server and a database interface which allows users to find the pages of transcriptions they are looking for. This is exactly what can now be found at www.alpi.ca: a database (at the University of Western Ontario’s Theoretical and Applied Linguistics Laboratory) which allows researchers to select survey points either from lists (by province and name) or from an interactive map, and then select the page(s) of the ALPI field notebooks which interest them. The data—original phonetic transcriptions, exactly as the fieldworkers initially recorded them—can be downloaded and/or printed, and have already begun to be used as raw data for a number of studies. Researchers in Spain and around the world have begun to realise that this “lost” geolinguistic treasure is in fact alive and well and living in southwestern Ontario where it is accessible to them by internet, and the site now gets over 1200 hits a month. The scanning of the notebook images is currently more than half finished, and work is under way on a number of digital transcription tools which will be used to create further databases from the notebook data, eventually leading to electronic dialect database for on-the-fly dynamic mapping of linguistic variables (cf. Kretzschmar 1999, Ruiz Tinoco 2002).

Although some have criticised its questionnaire as well as its elapsed time to completion, the ALPI compares rather well with later projects on both grounds. For example Alvar (1964), a strong proponent of regional rather than national atlases, criticised the fact that the ALPI surveys began in 1930 and were not completed until 1954. In 1974 Alvar proposed that a coordinated series of regional atlas surveys could cover the Iberian Peninsula better and more efficiently than a single large-scale atlas could hope to do: since that proposal, more than 30 years have elapsed and fieldwork is still not complete for Alvar’s projected Atlas Lingüístico de España y Portugal (the Linguistic Atlas of Spain and Portugal or ALEP) by regions. The 24 years spanned by the ALPI surveys now seem fairly reasonable in contrast, especially considering that the overwhelming bulk of points surveyed (including all of the Castilian and Andalusian areas) were completed within just six years, making these pre-Civil War notebooks truly a synchronic ‘snapshot’ from a period for which we have no other comparable sources of linguistic data. Furthermore, the ALPI’s questionnaire seems in some ways refreshingly modern when compared with many
traditional dialect surveys (including Alvar’s series of regional atlases). Morphosyntactic variables are particularly well-represented, with the second part of the first notebook eliciting full sentences rather than isolated forms, including data on issues which are of continuing importance to contemporary theoretical linguistics: reflexives, clitic sequences, conditional structures, verb forms, variable agreement and pronoun paradigms, to name but a few. Navarro Tomás and his collaborators who put together the questionnaire have left us a precious legacy of variation data which remains unparalleled both in breadth and depth of coverage (Heap 2002, in preparation). As more and more of the original notebook data are made available online in facsimile format and being used for new analyses (see for example Fernández-Ordóñez 2001, Heap 2005, 2006, Pato 2003ab, 2005, Heap & Pato 2006), we are also developing tools which will soon make it possible to search for specific data using criteria such as the year, the fieldworker, and the socio-demographic characteristics of the informants for each survey (Heap & Bayona 2004).

A fitting tribute to the farsighted vision of Menéndez Pidal, Navarro Tomás and their dedicated team of fieldworkers, the ALPI surveys contain more than enough data for dozens of doctoral theses and other studies (the growing number of visits the site receives each month is testimony to how researchers are now taking up some of these many possibilities). Contemporary linguists (whether ‘theoretical’, ‘descriptive’ or ‘variationist’ by inclination) are far too often unaware of the extraordinary wealth of data that lie waiting to be discovered in linguistic atlases—some of them published but often under-utilised, while a surprising number of others have field data collected but still waiting to be edited, published and analysed. Just how many such nearly-lost linguistic corpora are out there in need of recovering and preserving is far from obvious, but one thing is very clear: part of Jack Chamber’s legacy to dialectology has been his constant insistence on pointing out the importance of paying serious attention to such traditional linguistic atlas scholarship.

References


