

In Defense of *Romancero* Geography

Suzanne Petersen

Folklore materials derive their defining characteristics from their behavior within the space-time coordinates as they are subjected to certain forces and influences through environment and the very mechanisms of transmission (Foster 1968:247).

Ramón Menéndez Pidal's pioneering essay "Sobre geografía folklórica. Ensayo de un método" (1920) constituted the first full-scale implementation of geographic methods in *romancero* studies.¹ At a time when the very concept of traditionalism was hotly debated and the centrality of the transmission process to a definition of traditional poetry was largely ignored, this study provided Menéndez Pidal with several valuable insights into the mechanisms of variation governing orally transmitted Hispanic ballads. By meticulously plotting the geographic dispersion of key motifs present in two widely disseminated modern *romances*, each represented by some 160 traditional versions, he found that each independently conceived motif, each element that appears in a traditional *romance*, attains its own continuous and compact, but ever-evolving area of diffusion and has a history distinct from that of every other motif or element in the text.² In addition to documenting the important role played by the independent propagation of individual motifs in the continuous transformation of oral *romances*, his geographic study of these two sizeable bodies of evidence also revealed the existence of clearly discernible "local types" or groups of versions in geographic proximity which manifest considerable thematic uniformity and a high proportion of common

variants. In the face of the “prodigiously multiform” variation to which his two corpora attested, this evidence further supported his recently formulated theory of collective re-creation, exemplifying the way and extent to which individual creative freedom is restricted and partially neutralized, by the pressure of communal sensibilities and attitudes (artistic, moral, and ideological).

When some thirty years later Diego Catalán and Álvaro Galmés again studied the same two ballads (1954), then represented by more than a thousand traditional versions, the addition of a temporal factor to Menéndez Pidal’s geographic method allowed them to reconstruct with far greater precision the paths of dissemination and the complex history of the two ballads.³ With heightened awareness of the diachrony implicit in folklore geography and less preoccupied than their predecessor with hypothetical prototypes, Catalán and Galmés were able to identify three successive, but overlapping stages in the evolutionary development of these two ballads: an initial and most truly “traditional” stage characterized by the independent propagation of isolated motifs, resulting in numerous highly diversified local forms; a subsequent phase in which a particular version or versions, having gained exceptional prestige, were for a time propagated *en bloc*, leading to the implantation throughout entire regions of highly uniform “types” that competed with and modified pre-existing local forms; and, in more modern times, a third stage brought on by the recent invasion of modern civilization and the rapid erosion of traditional rural society, and characterized by the tendency toward peninsular uniformity in conjunction with the ever more accentuated expansive force of Andalusian ballad types.

While the particular history of each ballad is, of course, unique and the specific details of the Catalán-Galmés study with regard to the propagation of isolated motifs, the formation of compact as opposed to diffuse geographic areas, the creation of fixed “types,” the southern invasion of the north and its reaction, as well as the chronology and rhythms of change are valid only for the two ballads studied, subsequent research has shown that the major evolutionary tendencies and patterns identified in this historical-geographical study are, for the most part, characteristic of the modern oral *romancero* tradition.

To be sure, the importance of geography and the utility of geographic methods have frequently been questioned in subsequent years and, on occasion, flatly denied in Hispanic ballad criticism.

In this regard *romancero* studies have not been remiss in reflecting major trends in contemporary Western ballad scholarship. Yet, contrary to prevailing opinion in many of its sister disciplines, and despite repeated attacks, geographic considerations play a very significant role in much of *romancero* criticism to this day.

To explain and justify the continued reliance on geographic data in this branch of balladry, I will first examine the explicit and implicit objections of those who have proposed alternative approaches to the study of the modern oral *romancero*. In so doing, my aim is not to deny that subsequent criticism has accurately identified certain shortcomings in one or another geographic study, or that it has legitimately taken issue with certain procedures and even with portions of the results of *romancero* geography studies undertaken to date. My objective is rather to demonstrate with supporting evidence from recent scholarship that despite the difficulties inherent in the use of both space and time factors, these two variables are nonetheless essential to any theory which purports to define the processes of transformation and reproductive mechanisms of the *romancero* and similar bodies of traditional folk literature in which the processes of creation and transmission are inseparable. It is my contention that rather than seek to discredit or supplant folklore geography, what we require are more appropriate tools for incorporating it in our analyses. To this end I will conclude my review and defense of historical-geographical methods with sample analyses which illustrate how time and space as well as numerous interdependent, internal variables were integrated in a pilot program of computer-aided literary analysis specifically designed to facilitate exploration of the processes of variation in traditional ballad narratives.

As I hope to make clear in the following pages, the modern oral Hispanic ballad tradition differs significantly in a number of ways from many of the models of oral poetic discourse with which contemporary ballad research concerns itself. One very distinctive feature whose importance should not be overlooked or underestimated in seeking to justify the prevalence of geographic and historical-geographic studies in *romancero* scholarship is the unparalleled spatial and temporal span of the Hispanic *romancero* tradition.⁴ Documentary evidence of this tradition not only encompasses six centuries of uninterrupted traditional poetic activity, but exists, often in abundance, for nearly every region

inhabited by speakers of one of the four Hispanic languages.⁵

As we would expect, the earliest and most unconditional denials of the value of geographic considerations issue from those *romancero* scholars most closely allied with the individualist position. Thus the French critic Jules Horrent, in affirming the essential similarity of the creative processes in learned and traditional poetry, voices strenuous objection to the aforementioned studies by Menéndez Pidal and Catalán-Galmés primarily on the grounds that their focus on development and transformation tends to obscure the fact that “the innumerable variants that give the ballad its continual and unique movement, they too are individual in essence” (1957:392).

From a similar position Daniel Devoto launches a scathing attack on the same two studies in folklore geography, insisting that “the accent of the investigation ought to be put, above all, on the individual folkloric act, the conduct of *that* individual, of the *individual* alone, capable of varying the ballad each time he brings it to life” (1955:253).⁶ When Devoto asserts that geographic methods are invalidated by the very fact that variations inevitably arise in successive repetitions of a ballad by the same subject (253), he, like Horrent, attributes to the individual singer a degree of freedom and independence from the tradition far in excess of that which documentary evidence bears out. No one would deny that each and every variation that arises in the course of a ballad’s oral life originates with an individual (or specific individual situation), but, as the evidence of folklore geography has demonstrated time and again, once the variant has been accepted by the tradition, and in fact becomes the tradition, that element conditions subsequent individual folklore behavior to a far greater degree than Devoto is willing to acknowledge.⁷

At the root of Devoto’s dissatisfaction with these geographic studies is his conviction that such methods are virtually useless for the study of what really matters, namely, the underlying reasons for the behavior of folklore materials (1955:250). His contention that the study of this behavior leads to the discovery of the “sentido profundo” of the recurrent, invariant motifs of the tradition has considerable merit, but in rejecting the traditionalist conception of collective re-creation and with it the relevance of geographic data, Devoto takes his argument too far, insisting that this invariant latent symbolic content informs all versions of a ballad, old and modern alike. When he counsels that our attention

should be directed to the question of how the words can vary while the essential content remains unchanged, he not only overlooks the significance of the ways in which, at the lexical and morphological level, local geographic traditions determine and explain individual behaviors, but more importantly, he fails to perceive how, at the semantic level, these essential contents (the meanings of myth) can and frequently do undergo profound alteration. In seeking to identify and explain that which is constant, Devoto largely ignores the equally important, opposing (but complementary) forces responsible for innovation. As a result, he underestimates and unduly trivializes the creativity of the modern *romancero*, whose narratives have retained their relevance and thus survived only by adapting to the diverse and ever-evolving social and historical contexts in which they are reproduced. To perceive how the latent content of inherited narrative structures is subject to profound and often radical revision as it is reinterpreted in ways consistent with the prevailing attitudes and values of culturally bound collectivities, we need only compare the dominant regional types within a well-documented corpus of versions of a single modern ballad—one that adequately represents the geographic and cultural diversity of the Hispanic world.

Exemplary in this sense are the very disparate reactions manifested in the different areas of the modern tradition to the problem of sibling incest posed in *Amnon y Tamar*. Within Spain alone the tradition proffers a wide variety of potential “solutions” to the family crisis created when the father/king unwittingly hands over his daughter to be brutally ravaged by her brother. In the various northern traditions these solutions range from the victim’s conventual reclusion, her suicide, or the vengeful murder of the rapist, to conjectured papal sanction of the marriage of brother and sister—all but the last in keeping with traditional morality, but denounced as insensitive, useless, or unjust and regarded as equally unsatisfactory. By contrast, in the vulgate versions originating in the southern half of the Peninsula, rather than explore ways to deny, cover up, atone for, or legitimize sinful behavior, the reaction to the incestuous act is one of open consent and even defiant maternal defense of the singularly beautiful offspring proudly baptized “son of brother and sister.”⁸

As numerous recent studies devoted to the analysis of variation have demonstrated, this ongoing process of reactualization

is possible by virtue of the essential, unique property of oral narrative structures: their openness to unlimited and potentially irreversible transformation at all levels of articulation.⁹ However, it is the *romancero*'s re-creative process, its particular brand of orality, that determines the *modus operandi* of this characteristic feature of all oral discourse.

Like the individualists (and Devoto) who emphasize the importance of the individual creative act in the folkloric performance, the oralists who attempt to apply to the modern *romancero* the principles of oral-formulaic composition established by Parry and Lord fail to make the necessary distinction between an oral tradition carried on by a small core of trained semi-professionals and one that has survived for centuries as a communal activity in which any and all can participate. In the oral epic tradition of the Balkans an individual singer selects elements from the body of stock epithets, phrases, and forms the tradition offers, to create a text at the moment of each performance. By contrast, in the *romancero* those traditional elements form part of a text, an organic whole, endowed with its own structure, which the individual subject has memorized. In the "performance" what the singer actualizes is the text she has learned.¹⁰ While motifs, formulas, and verses can and often do travel from one text to another, their normal existence is as part of a larger unit, which is the text, the *romance*.¹¹ Precisely because Hispanic ballads are not individual creations that come into being on the occasion of a public performance, but re-creations of a text previously committed to memory, they are characterized by a high degree of fidelity to local tradition in both of the complementary aspects of traditionality: the retention of inherited narrative structures and their renovation in response to communal attitudes, values, and aesthetic sensibilities.

A verse by verse comparison of two versions of *Muerte del príncipe don Juan* collected in the same town (Uña de Quintana, Zamora, Spain), one in 1912 and the other in 1981, illustrates the "permanence" of a text in a given geographic location.

Muy malo estaba don Juan,	muy malo estaba en la cama.
Muy malo estaba don Juan,	muy malo estaba en la cama,
(Don Juan was gravely ill, gravely ill in his bed;)	

siete doctores lo curan	de los mejores de España;
siete doctores lo curan	de los mejores de España; (seven
doctors attend him, among the best in Spain.)	

aún faltaba por venir aquel dotor de la Parra,
 y aún faltaba por venir aquel dotor de la Parra,
 (There was yet to arrive that doctor La Parra.)

que dicen que es gran dotore, gran dotor que adivinaba,
 que dicen que es gran dotor, gran dotor que adivinaba.
 (a wise doctor they say he is, a wise doctor who reads signs.)

Trae solimán en el dedo y en la lengua se lo echara.
 Trae el veneno en el dedo y en la lengua se lo echara. (He
 brings poison on his finger and places it on his tongue.)

—Tres horas tienes de vida, la media ya va pasada:
 —Tres horas tienes de vida, la una ya va pasada:
 (“Three hours of life you have, a half has already gone by:)

una para confesarte y otra pa enmendar tu alma
 una para confesarte, otra para enmendar tu alma
 (one to make your confession, another to attend to your soul)

y otra para despedirte de la tu esposa doña Ana. —
 y otra para despedirte de la tu esposa doña Ana. —
 (and another to bid farewell to your wife doña Ana.”)

Estando en estas razones entró doña Ana en la sala.
 Y estando en estas razones entró doña Ana por la Sala. (At
 this very moment dona Ana entered the room.)

—¿Tu qué tienes, el don Juan, tú qué tiés en esa cama?
 —
 (“What is ailing you, don Juan, what ailment keeps you abed?”)

—Tengo mal de calentura, que otro mal no se me hallaba
 —
 (“I have a very high fever, no other ill has been found.)

¿Dónde vienes, la mi esposa, dónde vienes, doña Ana?
 —¿Dónde vienes, la mi esposa, dónde vienes, doña Ana?
 (“Where have you been, my wife, where have you been, doña Ana?”)

—Vengo del templo de Dios, del convento 'e Santa Clara; —
 Vengo de rezar por ti del convento 'e Santa Clara;
 (“I come from the temple of God, from the Convent of Saint Clare:)

vengo de rezar por ti te levantes de esa cama.
 vengo de rezar por ti te levantes de esa cama.
 (I have been praying for you, that you get up from that bed.”)

—Yo sí me levantaré, antes de por la mañana,
 —Yo sí me levantaré, antes de por la mañana,
 (“Yes, I will surely get up before a new day comes.)

para unas andas de pino, para una iglesia sagrada;
 —
 ([headed] for a pine bier, for a holy church:)

of the ballad—identical save for minimal verbal variation and minor gaps in the later version—clearly points to two acts of memorization of a common model.¹²

Proponents of ethnographic and anthropological approaches who contend that in order to understand the behavior of this body of folklore material the primary focus of critical attention must be the specific social context of each individual performance similarly overlook the larger implications of the distinctive character of *romancero* orality. It being the case that the process of transmission routinely begins with an act of memorization, we must bear in mind that this activity is, at one and the same time, a creative, interpretive act. What is learned is not what was emitted, but what the receiver heard and/or was able (wanted) to understand. Consequently, the most significant adaptation or change with respect to the model occurs in the acquisition phase, rather than in subsequent executions of the “text” that has been committed to memory. Thus, as a rule, it is not primarily the immediate, external circumstances surrounding a given performance, but the broader sociocultural context and the ballad’s coexistence and contact with other traditional texts that together condition change.

Moreover, for the vast majority of traditional *romances*, execution is not normally tied to specific social events as is frequently the case in neighboring European oral traditions. As long as traditional Hispanic cultures survive, open-structured *romancero* narratives can and do maintain their cultural relevance by adapting to changes in the referent, but the possibility for profound change is lost when the narrative itself ceases to function as a reflection on life. This is just what has occurred with the relatively small number of ballads that owe their continued existence largely to their having become ritualized (e.g., ballads sung in chorus as part of children’s games or those sung only on certain holidays or as part of specific social events). As J. Antonio Cid (1979:329) so aptly points out, in these exceptions to the norm the context has come to prevail over the text with the result that “variation approaches zero.”¹³ For this reason the study of the contexts in which these particular ballads are reproduced, however interesting, can shed little light on the mechanisms governing variation in truly traditional re-creative activity.

Finally, although in the past the *romancero* enjoyed an active life in rural areas, forming part of everyday traditional communal

activities, today it only rarely surfaces spontaneously even among those who participated most actively in former times. Many of the best of today's informants admit to not having sung the ballads in twenty or thirty years (and particularly not in public). Thus today very few opportunities arise to witness the spontaneous reproduction of ballad themes in their natural, immediate contexts. In the face of this reality, folklorists and ballad specialists anxious to document what most have assumed to be the final stages of a dying plurisecular tradition have been forced in recent decades to seek out prospective Informants and induce them to reactivate their fund of traditional knowledge. The intense collecting activity of the past ten years has extracted in this manner an unprecedented wealth of ephemeral documentary evidence from the "archive" of popular memory.¹⁴ In these circumstances, however, considerable temporal distance often separates the elicited execution and the primary interpretive activity which configured the text committed to memory. Consequently, much of the variation that arises in this process is the result of vacillation as the singer struggles to retrieve her "text" from the depths of her memory. Simple memory lapses, together with the interference of variants recalled from competing local versions and the intertextual contamination of motifs and sequences (or even entire *fabulae*) among several ballads in the singer's repertoire are increasingly characteristic of today's recitations. These phenomena are by no means anti-traditional in nature, but given the dormant state of the tradition, the transformation they occasion is less likely a reflection of current social values and collectively (or individually) held opinions on present-day life than was the case in former times when traditional re-creative activity was vastly more intense and flourished naturally.¹⁵

Although Menéndez Pidal's concept of traditionalism had gained general acceptance by the sixties, the methods of folklore geography continued to be viewed with skepticism by several of the neo-traditionalists most actively involved in the nascent critical re-evaluation of the *romancero*. In seeking to counter early traditionalism's overriding concern with the retentive aspects of the *romancero* while avoiding Menéndez Pidal's occasionally excessive reliance on the fragmented evidence of independent variants, both Paul Bénichou (1968a) and Guiseppe Di Stefano (among others) wisely advocated the straightforward synchronic comparison of texts or groups of texts as a more appropriate method of evaluating the

creative aspects of orally transmitted poetic structures.

Bénichou, the most outspoken critic of the earlier geographic studies, echoes Devoto in arguing that in view of the genealogical and chronological uncertainty with regard to the derivation of the texts, the history of a ballad is an inaccessible goal with or without geography (1968b:328). While he concedes its occasional utility as a tool with which to investigate the processes of traditional re-elaboration, he maintains that the labor-intensive methods of folklore geography generally provide no better nor more conclusive evidence than can be obtained through simple comparison of the texts. Despite the implicit skepticism of such remarks as “. . . but this does not keep us from asking ourselves just what end this method serves and what difference knowing the geographic origin of each text makes when considering a group of versions of a ballad” (327), much of Bénichou’s own extremely insightful work makes readily apparent the relevance of geography.

When Bénichou sets out to evaluate the creative activity and poetic value of the modern tradition in *Creación poética en el romancero tradicional* (1968a), he focuses primarily on the “variety of content and structure of the modern versions, *in their different families*” (95, italics mine) to demonstrate how this activity in modern times is neither substantially different from, nor inferior in its results to that which produced the venerated *romances tradicionales* of the sixteenth century.

In each of his three comparative studies of modern ballad texts Bénichou establishes these different families or groups of versions on the basis of common motifs and similar thematic structure—the two criteria used by proponents of *romancero* geography to identify the various local traditions of a given ballad.¹⁶ Not surprisingly, in all three studies his families of versions coincide with recognized traditional geographic areas and even highly localized traditions. To cite but one example, in this same essay on *Muerte del príncipe don Juan* the author observes “this group [*La Montaña*], geographically very near the first, has a totally distinct physiognomy” (1968a:103). In his study of forty-one modern versions of this ballad, Bénichou dramatically demonstrates how in four distinct areas of the modern tradition the ballad’s transmitters have variously developed the latent poetic content of the primitive historical *romance*, transforming it in both form and spirit. After meticulously examining the narrative function of the motifs that characterize each group of versions,

Bénichou affirms:

... we find innovations that change the spirit and the construction of the ballad: in the Castilian versions of the first group, the anticipated description the sick man makes of his funeral and of the helplessness of the young widow; in some versions from *La Montaña*, the final scene of tenderness between husband and wife and their death; in the Portuguese versions, the scruples of the dying protagonist and his parents regarding the dishonored mistress; in the Jewish versions, the family lament (118).

In his eloquent defense of the modern tradition Bénichou's case for the freedom with which modern transmitters of the ballad have exercised their poetic creativity is well made, but the limitations within which that freedom is seen to operate clearly indicate the relevance of geography to the processes of traditional re-elaboration. He concludes his study of the ballad with the astute observation that the real difference between the old and the modern traditions is that they reflect different worlds (123). Bénichou's analysis admirably demonstrates traditional poetry's ability to adapt to the changing environments in which it is reproduced, but these "different worlds" are not merely those of the sixteenth and twentieth centuries, as his introductory comments make abundantly clear:

The ballad has been transforming itself in environments so distinct and distant from one another as Old Castile, Portugal, Morocco, and the Eastern Sephardic communities and therefore, in the main, has experienced a very high degree of differentiation (97, italics mine).

By respecting the integrity of the poetic structure and shifting the focus of attention to the creative reinterpretation and restructuring of poetic content, Bénichou as well as Di Stefano, Catalán, and others avoid some of the undeniable limitations inherent in the earlier geographic and historical-geographic studies. Nevertheless, their findings invariably confirm the significant role of geography in the transformation of collectively re-created poetic narratives.¹⁷ Despite methodological improvements, the more recent investigations that stress synchronic comparison of ballad texts also impose severe restrictions on the analysis of transformation processes. If the early studies in *romancero* geography could

analyze an exhaustive corpus only by virtually limiting analysis to a single variable (thereby ignoring numerous other interdependent factors), more comprehensive analysis based on the “simple comparison of texts” has proven feasible only by limiting the textual evidence to a relatively small, but manageable sample. This restriction inevitably tends to undermine the possibility of identifying and comparing genetically related structures and thus jeopardizes our ability to determine exactly how orally transmitted poems are reproduced. The Hispanic *romancero* constitutes an ideal corpus for exploring the mechanisms of transmission in a dynamic open-structured model precisely because it offers bodies of evidence of unprecedented numerical, spatial, and temporal amplitude.

With the advent of modern electronic data processing, rigorous, systematic analysis of large, highly complex corpora such as that constituted by the oral *romancero* became a real possibility for the first time. Computers capable of storing, sorting, tabulating, and cross-correlating vast quantities of data can perform exhaustive synchronic and diachronic comparisons of hundreds (or thousands) of traditional poetic texts. By exploiting this potential in conjunction with appropriate samples of the *romancero*, we can explore the processes of variation at all levels of organization of the narrative(s) and, in the case of multiple manifestations of a single theme, examine and measure the dynamic interaction among these levels.

The first attempt to devise a computer-based system of literary analysis designed to facilitate a comprehensive description of the reproductive patterns and system of poetic communication of the *romancero* was undertaken in the early seventies. The data base for this pilot project consisted of 612 versions of *La condesita*, a modern ballad first documented in 1820 and widely diffused in all areas of Spain.¹⁸ To maximize the potential for subsequent comparative operations, this entire corpus was first segmented into various smaller units of analysis. These included the three most obvious breaks in the chain of discourse—the version, the hemistich (the basic unit of composition), and the word—as well as three other units of analysis related to the narrative structure and its dramatic presentation: the dramatic scene, the thematic segment, and the element of narrative information. In order to generate these and other paradigmatic groupings related to the poetic structure or to the spatial and temporal classification of the

material, eleven different sets of codes were manually or mechanically assigned to each of the 34,233 hemistichs of the corpus.

Each version was defined with a code identifying 1) its place of origin (province and region), 2) its date of collection (year and period), 3) the nature of its testimony (distinguishing between fragmentary and more or less complete versions) and 4) its ballad type (distinguishing among oral versions of the autonomous ballad *La condesita*, oral versions of the double ballad *Gerineldo + La condesita*, non-traditional versions propagated by virtue of the popularity of certain published texts, and versions derived from Menéndez Pidal's influential composite text published in *Flor nueva de romances viejos*).

The text of each hemistich was defined with a code to:

- a) insert it into its immediate syntagmatic context;
- b) identify it as a 1st (non-rhyming) or 2nd (rhyming) hemistich;
- c) attribute it to a particular narrative sequence, specifying the dramatic scene and thematic segment in which it appears;
- d) classify it by its mode of narrative discourse, identifying 1st or 3rd person narration or direct discourse;
- e) identify the speaker, if part of a direct discourse;
- f) identify the type of narration, if part of a narrated sequence, distinguishing among inter- and intra-scenic narration, narration denoting a time/space transition, narration announcing the entrance or exit of a character in the drama, or narrative introduction to direct discourse;
- g) specify the element of narrative content it carries, identifying it with one of 250 minimal elements of narrative information;
- h) specify the invariant or "archetypal" hemistich of which it is but one manifestation, identifying it with one of 1,937 archetypal hemistichs;¹⁹

Finally, each of the 170,978 words in the corpus was assigned a code to 1) identify its position in the hemistich, 2) identify its grammatical function and part of speech, and 3) define its function in the metric scheme by identifying the words that carried the assonance.

Although much of the above information is logically associated with a particular level or unit of analysis, by uniformly encoding each and every hemistich with all the information it was all retained and available at all levels of analysis or subdivisions

thereof. The capacity to correlate any number or any combination of variables allows taking into account simultaneously all the simple and complex variable factors that must be weighed in order to determine which are responsible for a given phenomenon, a particular aspect of variation.

In testing the applicability of a system with almost unlimited potential for generating and tabulating information on significant subsets of the data base, we focused primarily on problems that had consistently defied analysis with more traditional methodologies, that is, those that required either a multilevel analysis or the interpretation of large batches of sorted and tabulated data. Geographic and temporal variables figured prominently in a number of our computer-aided analyses of lexical variation and lexical-syntactic, semantic, and structural transformation. In view of the fact that a serious limitation of the earlier geographic studies was their inability to adequately assess structural transformation in the versions whose variants they studied, the examples selected all utilize geographic and/or temporal variables to analyze an important aspect of this process.

As illustration of even the simplest of the analyses requires some familiarity with the ballad studied, the following description of the narrative content of *La condeaita* will serve to introduce both the ballad and two of the units of analysis that figure in the subsequent examples.

ACT #	SEGMENT #
1 Separation of the spouses	10 Semi—autonomous introduction
	11 The husband is called to war
	12 Grief of the spouses in the face of forced separation
	13 Estimated duration of the separation
2 The wife decides to go in search of her husband	14 The husband's departure
	21 The estimated time limit on the separation elapses and the husband has not returned
	22 The wife's family pressures her to remarry
	23 The wife discusses going in search of her husband with a member of her family
2 Encounter with an informant	24 The wife, disguised in a pilgrim's garb, leaves in search of her husband
	31 After long travels the "pilgrim"

	(wife) encounters an informant
32	First news of the husband
33	The “pilgrim” is appraised of how and where to gain an audience with her husband
4 Recognition	41 The “pilgrim” manages to obtain an interview with her newly affianced husband
	42 Dialogue culminating in the revelation of the “pilgrim’s” true identity
	43 The husband’s reaction
	44 The rival fiancée’s reaction on learning what has transpired
	45 The husband decides which of the women he wishes to claim as his wife
	50 Final comment, marginal to the story line

The first example, in which geographic considerations are brought into play only in a second stage of the analysis, illustrates how the coding system enables us to analyze the transformation of individual invariant hemistichs—the underlying base structures or “archetypes” that generate the 34,233 hemistichs (19,403 variant hemistichs) of the *La condesita* corpus.

Segment 44, in which the count’s new fiancée [N] or a member of her household m comes upon her betrothed [C] lying in a dead faint before a pilgrim whom he has just discovered to be his long-abandoned wife [R], frequently begins with one of two hemistichs that carry the information coded as element 505 (Curses cast on the pilgrim): **Damn the pilgrim* [44505111] and **You, Pilgrim (~ Madame), are the Devil* [44505511]. A sample listing of 17 of the 87 variant hemistichs that manifest archetype 44 505 511 reveals how, in the course of its reproduction, this archetype has acquired new semantic values and taken on different functions in the final scene of the ballad.

<i>Hemistich</i>	<i>Speaker</i>	<i>Segment</i>	<i>Element</i>	<i>Archetype</i>
12939A Vaya al diablo la romera	N	44	505	44 505
23330A Eres el diablo romera	N	44	506	44 505
19929A Eres mujer o demonio	N	44	506	44 505
22522A Mujer es usted el demonio	N	44	506	44 505
21627A Que mujer o que demonio	T	44	506	44 505
28733A Que demonio de princesa	N	44	506	44 505
51117A Que diablo de romea	C	42	462	44 505
56219A Eres er diablo niña	C	42	462	44 505

52718A	Romera eres el demonio	C	42	462	44 505
61949A	Tu eres el mismo demonio	C	42	462	44 505
25116A	Vete romera del diablo	C	42	438	44 505
19540A	Mujer eres el demonio	N	45	472	44 505
20428A	Hombre tu eres el demonio	N	45	471	44 505
50367A	Que romera ni demonios	N	42	410	44 505
25116A	Vete romera del diablo	C	42	438	44 505
54053A	Vete demonio e romera	C	41	433	44 505
65954A	Sera mujer o demonio	N	41	433	44 505

The allusion to the devil that appears frequently in conjunction with archetype 44505111, **Damn the pilgrim, / (who in) the devil brought her here*, is here formulated as an insinuation that the pilgrim might herself be the devil: **You, Pilgrim (~, Madame), are the devil / who has come to tempt him* (element 506). Once the association surfaces as an insinuation or expression of doubt, it can readily be transferred from the irate fiancée [N] to the count himself [C], whereby the same archetype comes to express his last moment of vacillation between committing himself to his present life or accepting the superior claims of his first love [C 42 462]: **You, Pilgrim, are the devil / who has come to tempt me*.

As revealed by other discrepancies between the archetype code and the codes identifying the segment in which the hemistich actually appears and the information it actually carries, this archetype occasionally takes on other meanings as well. It travels intact to segment 45 where, as an accusation lodged either against the count (element 471) or the countess (element 472), it functions to express the rival's protestations in the face of her defeat: **You, Madame (~ Sir), are the devil / who came to insult me (~ for you have deceived me)*. With only minor lexical variation ("Qué mujer o qué demonio" ~ "Qué romera ni demonios"), the same base structure can introduce the pilgrim's revelation of her true identity (element 410): **What pilgrim or devil?, / I'm your lawful wife!*. With similar ease the curse is transformed into an expression of disgust at the pilgrim's haughty refusal of alms (element 438) and finally, reformulated yet again in segment 41, it functions as an order for the alms-seeking pilgrim to withdraw from the palace door, thus replacing the more usual expression **Get you back, Madame*, issued either by the fiancée or the count himself.

By providing data that allows us to analyze each archetype's

capacity for acquiring new semantic values and functioning in different contexts, the codes allow us to detect all instances of syntagmatic reorganization of the narrative, including those with repercussions on the paradigmatic plane. The relative importance of the syntagmatic reorganization occasioned by the mobility of archetypes can be measured for parts or all of the narrative simply by tabulating occurrences of discordance between the two sets of codes. Not surprisingly, these statistics reveal that structural transformation in *La condesita* is most radical in the final scene, where discordance between the ideal and the actual position of the hemistichs averages 43% (as compared to 30% for the corpus as a whole). In other words, only 57% of hemistichs that reproduce an invariant ideally belonging to a particular segment in the fourth act actually appear in that segment. Mobility is most extreme precisely in segment 44, where 54% of the hemistichs we expect to find there have travelled to another segment (this, compared to the minimal variation attested to in segment 13 of act 1, for example, where discordance drops to 10%).²⁰

Table I

Region	Number of Versions	% of Versions with 44505511	Occurrences 44505511 by BT	Breakdown by Function	Breakdown of Function by BT
1 (Northwest)	84	4.70%	75% BT I		
2 (Leonese area)	79	3.80%	100% BT II		
3 (Old Castile)	83	4.80%	75% BT I		
4 (Extremadura)	78	16.70%	62% BT I	N44 = 54% C42 = 46%	BT I: N44 88% C42 12% BT II: C42 100%
5 (New Castile)	72	40.30%	62% BT I	N44 = 72% C42 = 17% C45 = 7% N41 = 3%	BT I: N44 82% C42 5% N45 11% BT II: N44 54% C42 36% N45 9%
6 (Andalusia)	86	23.30%	92% BT II		
7 (Levant)	66	16.70%	82% BT I		

(BT = Ballad type)

Regional Classification of Provinces



Key to Regions and Provinces

<i>Regions</i>	<i>Provinces</i>		
1 Northwest	LC = La Coruña PO = Pontevedra LU = Lugo OV = Oviedo ST = Santander		
2 Leonese Area	LN = León ZM = Zamora OR = Orense Z = Unspecified within region		
3 Old Castile	BU = Burgos PA = Palencia VD = Valladolid SG = Segovia MD = Madrid SO = Soria LO = Logroño VT = Álava BI = Bilboa [Vizcaya] SS = San Sebastián [Guipuzcoa] PP = Pamplona		
4 Extremadura	SM = Salamanca TM = Trás-os-Montes BE = Beira AV = Ávila CC = Cáceres BD = Badajoz	7 Levant	BA = Barcelona LE = Lérida TA = Tarragona CS = Castellón VA = Valencia AL = Alicante PM = Palma de Mallorca [Balears]
5 New Castile	TO = Toledo CR = Ciudad Real AB = Albacete MU = Murcia CU = Cuenca GJ = Guadalajara		
6 Andalusia	CO = Córdoba SV = Sevilla HL = Huelva CD = Cádiz MG = Málaga GR = Granada JN = Jaén AM = Almería A = Unspecified within region HS = Huesca ZZ = Zaragoza TE = Teruel GE = Gerona	8	M = Morocco

In order to perceive any significant pattern in the reproduction of any of the 1,937 invariant hemistichs of the corpus across time and space, we need only sort the data already obtained on the reformulations of archetypes by the appropriate temporal and/or geographic variables.²¹ Internal sorting of the 87 variant hemistichs of archetype 44505511, for example, reveals very definite trends in its implantation in the tradition. The following table summarizing the output of several cross-tabulations on this archetype allows us to identify these trends as well as the relationship between external factors (time and space) and structural characteristics of the ballad.²²

It is immediately apparent (col. 3) that while archetype 44505511 appears in all regions of the Peninsula, it can only be said to characterize a significant number of versions in four of the seven regions (4-7), with notable prevalence in region 5 (where a full 35% of all the manifestations of this invariant hemistich occur). The next column clearly suggests that the archetype, where used, is found primarily in versions of the autonomous ballad (type I), the descendants of the older structure. (In the Leonese area, where the archetype is unknown to type I versions, its three occurrences in the more modern double ballad are statistically irrelevant and in Andalusia, where the double ballad constitutes 95% of all known versions, the archetype nevertheless occurs in 50% of the four documented type I versions.) The most significant patterns in the reproduction of this archetype emerge in the last two columns, which register, within each region, the proportionate use of each of the reformulations of the archetype (col. 5) and within that context, their proportionate use in the type I and type H versions (col. 6). In Extremadura, although the archetype functions with approximately the same overall frequency, now as a curse on the pilgrim uttered by the irate fiancée [N44], now as an expression of the count's vacillation [C42], the latter reinterpretation corresponds exclusively to versions of the more modern ballad type, while 88% of the versions in which the invariant functions as an expression of the rival's outrage correspond to the older autonomous ballad (type I). Manifestations of the archetype in Andalusia fully corroborate the correlation between type II versions and the archetype's reformulation in segment 42. In region 5, where re-creative activity has produced maximum diversification in the reinterpretation of this archetype, its function in the context of the rival's reaction clearly dominates

72%) over all other reformulations. So tenacious is this interpretation in New Castile that it has seriously compromised the normal syntagmatic organization of type H versions, prevailing in 54% of the area's double ballad texts that retain the archetype.²³

Syntagmatic reorganization of a higher order can be analyzed in a similar fashion simply by substituting or incorporating additional variables in the cross-tabulation program and changing the internal sorting priority. This procedure allows us to explore three key processes of transformation intimately related to the poetic restructuring of the narrative for dramatic effect: the distortion of the natural order of the narrative (Petersen 1976a:190-92; Catalán 1976:72-74); modifications in the absolute and relative length of the ballad and of its dramatic scenes and segments; and variation in the proportionate use and distribution of direct discourse and narration.

A brief look at some statistics on the average length and proportionate use of direct discourse in the second act of the ballad will serve to illustrate how electronic processing of secondary data can identify important evolutionary trends in the structural transformation of the ballad across time and space.

Table II lists the average length of act 2 of all non-fragmentary versions in each of the seven peninsular regions (and in parentheses, the percentage of the overall narrative those figures represent), followed by a breakdown of the same information by ballad type.

Table II. Length (in hemistichs) and Relative Weight of Act 2

<i>Region</i>	<i>avg. length</i>		<i>BT I</i>		<i>BT II</i>	
	<i>length</i>	<i>%</i>	<i>avg. length</i>	<i>%</i>	<i>avg. length</i>	<i>%</i>
1	9	(13%)	11	(14%)	6	(11%)
2	9	(14%)	11	(15%)	5	(10%)
3	13	(18%)	14	(20%)	9	(14%)
4	11	(17%)	13	(18%)	7	(17%)
5	14	(23%)	14	(23%)	12	(24%)
6	6	(14%)	7	(14%)	6	(14%)
7	13	(22%)	13	(22%)	12	(23%)
8	8	(18%)	15	(25%)	5	(13%)

The average length of the second act is seen to vary considerably from region to region, from maximum elaboration in New Castile (14 hem.) to minimum elaboration in Andalusia (6 hem.). In all areas, however, the average length of the act is

greater in the autonomous ballad (type I) than in its more modern counterpart, and in several regions the contrast is extreme. In regions 1, 2, and 4 the second act is approximately twice as long in type I versions. The two radically different conceptions of the narrative suggested by the dramatic contrast in Morocco (region 8) between versions descended from the older ballad structure and those which reproduce the more modern double ballad theme is entirely consistent with the well-documented duality of the Moroccan tradition wherein, even today, the poetic legacy of the Sephardic communities coexists with recent importations from the Peninsula. Equally significant are the figures for regions 5, 6, and 7, where we find no appreciable differences between type I and type II versions in the average length of the second act. This similarity suggests an approximation between the two ballad types in these areas, particularly in regions 5 and 7 where the importance accorded to act 2 is maintained even in type II versions (24% and 23%).

As the next table reveals, the proportionate use of direct discourse in the ballad's second act also varies radically from region to region.

Table III. Proportionate Use of Direct Discourse in Act 2

<i>Region</i>	<i>% dir. Disc.</i>	<i>BT I % dir. Disc.</i>	<i>BT II % dir. Disc.</i>
1	59.6%	73.9%	30.8%
2	49.4%	59.3%	16.7%
3	62.2%	64.3%	47.0%
4	48.9%	53.7%	19.8%
5	62.3%	68.0%	67.0%
6	25.0%	36.4%	24.7%
7	59.9%	60.9%	52.7%
8	42.8%	56.8%	25.6%

The breakdown of direct discourse percentages by ballad type confirms and reinforces the structural tendencies seen in the previous table. Just as the average length of the ballad is greater in all type I versions, so too is the proportionate use of dialogue always higher in these versions than in the double ballad. Moreover, the greatest contrasts between the two ballad types occur in the same regions that manifested the greatest contrast in average length (regions 1, 2, 4, 8), as do the regions where the contrast is minimal (regions 5, 6, 7).

Setting aside other important regional differences that point to

a need for further investigation, the numerous parallels we have identified between length and proportionate use of dialogue are also born out in a global comparison of the autonomous and double ballad structures, both for the second act alone and for the entire narrative. The following table summarizes these statistics for the non-fragmentary oral versions of *La condesita*.

Table IV

	<i>Act 2</i>		<i>Acts 1-4</i>	
	<i>BT I</i>	<i>BT II</i>	<i>BT I</i>	<i>BT II</i>
Avg. length	13	7	69	47
% dir. Disc.	63%	35%	67%	63%

The most interesting statistic here is the percentage of direct discourse for all type H versions of the ballad taken as a whole (63%). The dramatic contrast seen between the two ballad types in act 2 largely disappears, indicating that the use of direct discourse for dramatic effect has not been replaced by narration in the double ballad, but rather that it has been used, very extensively, at other points in the narrative. The redistribution of discursive modes signalling important shifts in the focus of interest is a clear indication of profound reinterpetive activity.

The sample analyses discussed thus far have demonstrated that certain structural characteristics of the ballad are manifested, not in isolated versions, but in whole groups of versions, and that the poem adjusts itself to a number of very different poetic molds, with certain structural characteristics tending to dominate in certain regions and at certain stages in the ballad's oral life. The evidence clearly suggests that the study of any structural aspect of a traditional ballad requires an analysis of its transformation across time and space. While the study of the effects of these two factors on individual, simple variables provides valuable insights into the ballad's evolution, the most interesting comparative operations are those that allow correlating the data on all relevant variable factors so as to determine the effect of each one and the relationship or degree of interdependence among them. The final example illustrates how a first-generation computer cartography system was used to display the results of global comparative operations in such a way as to make the effects of time and space visually discernible.²⁴

Whereas the earlier manual geographic studies could only aspire to cartographically illustrate affinity among certain versions

of a ballad based on a number of subjectively selected hemistichs (i.e., one variable), computer cartography and geographic information systems allow us to take into account the geographic distribution of all the hemistichs and consider concurrently any number of variable factors (length of the poem, length of its scenes and segments, sequences of narrative events, discursive modes employed throughout, etc.). In this particular application, the geographic processing package was used to determine the degree of narrative affinity among all strictly oral, non-fragmentary versions of *La condesita* and to measure the extent to which that affinity is conditioned by the time and space factors, while the mapping program (primitive by today's standards) provided visual display of those relationships.

The first step in determining narrative affinity among versions was to create a base structure to which individual versions or groups of versions could be compared. To do this we chose to generate not one but several model or archetypal versions, each representative of the poem's dominant narrative structure in one of the temporally and spatially restricted subsets of the total corpus. Using the *ballad type* and *region* variables for the temporal and geographic subclassification of the texts, a representative version was mechanically generated for each of the two traditional ballad types (I and II) in each of the seven peninsular regions.

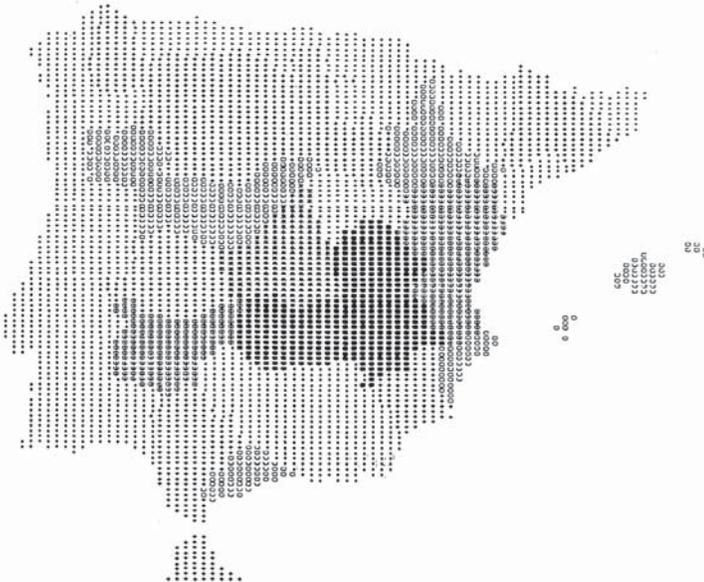
To assure that the fourteen archetypal versions faithfully reproduced the narrative structure of the ballad characteristic of the group of versions used to generate it, the computer first tabulated for each group of versions the total number of versions in the subset, the average length (in hemistichs) of the versions, the number of versions in which each of the poem's thematic segments appears, and the average length of each segment. With these statistics it created a "skeleton" defining, in outline form, both the narrative content and the syntagmatic structure of the narrative in each group of versions.

To give concrete form or "texture" to the narrative skeleton of each model version, a second program, run once for each subset, listed and tabulated occurrences of all archetypal hemistichs present in each of the poem's thematic segments. Based on their number of occurrences in these lists, the most representative archetypal hemistichs (that is, those of maximum incidence in each segment) were then selected to fill in the corresponding, predetermined "slots" in each narrative skeleton.

**Map 1 Archetypal Version of *Region 1, BT I*
against all *BT I* versions**



**Map 2 Archetypal Version of *Region 5, BT I*
against all *BT I* versions**



To evaluate affinitive relationships among the versions of the corpus, we elected to determine the precise degree of narrative affinity between each of the archetypal poems and all the real versions of each *ballad type* in each *province*, using as a measurement the average proportion of a model version's archetypal hemistichs preserved in the real versions of each temporally and geographically defined subset. Once the computer had determined, for a given archetypal version, the average proportion of affinity of all versions of a given *ballad type* in each *province*, it merely converted those proportions into a number on a graduated scale of ten values and then printed those values, translated into one of ten corresponding graduated over-print symbols.

The first map produced in this way constitutes a province by province comparison of the model versions of *ballad type I - region 1* with all 288 non-fragmentary versions of the autonomous (type I) ballad. As Map 1 reveals, this archetypal version represents a highly regionalized ballad "type" with its center in the Asturian provinces of Oviedo and Santillana (today, Santander). Lugo, the third of the three provinces in region 1 whose versions served to generate the model, is shown to be no more closely related to the nuclear group than the neighboring province of León (region 2). A considerable part of Old Castile, excluding the eastern portion (Logroño, Soria), maintains an attenuated second and third degree relationship of affinity with the model. The difference between the provinces of Palencia and Valladolid (degree 4), more akin to the model, and Burgos, Segovia, and Ávila (degree 3), bears out the progressive debilitation in relative affinity as one moves from west to east within the Castilian area. Aside from this zone adjacent to the nuclear area, the affinity between the archetypal version of *region 1 - type I* and the remaining type I versions is minimal. In reconfirming the existence of a regional Cantabrian "type" previously identified by Catalán and Galmés (1954) on the basis of nine invariants, this computer map confirms the validity of both this and the earlier geographic studies. In addition, the computer generated map visually displays not only the implantation of this regional "type" and the greater or lesser conformity of the versions included in the Cantabrian region to the details of the "type," but also the degree of affinity between it and the contiguous versions not belonging to the regional type.

To further test the validity of this regional archetypal version, the computer generated another map of narrative affinity using as

the basis of comparison the oldest real type I version from the northwest, rather than a composite, artificial model. As expected, the numerous inevitable singularities of a real version resulted in a general debilitation of the degrees of affinity, but in all other respects this comparison reproduced the same image as Map 1.

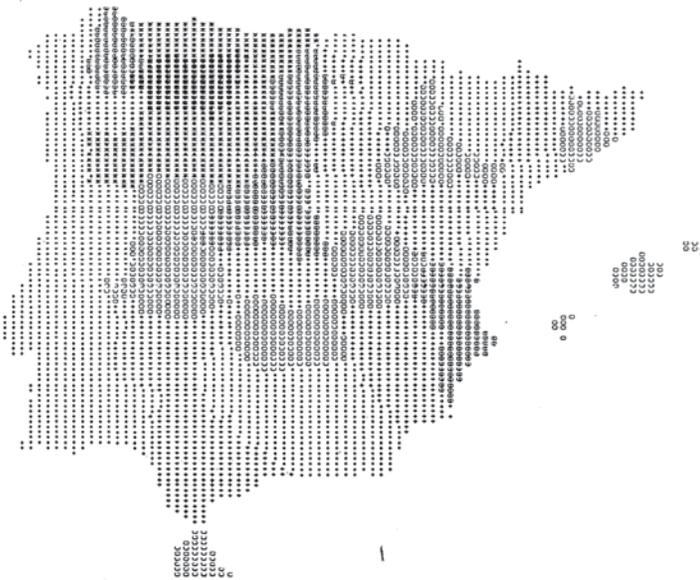
Map 2 compares the archetypal version generated from the thirty-four ballad type I versions of region 5 (New Castile and Murcia, excluding the province of Madrid) with all 288 type I versions. This map, like the first one, reveals the importance of the spatial factor in the transformation of a ballad narrative. Again we find a nucleus of marked affinity (degree 9) in the Manchegan provinces of Ciudad Real, Albacete, and Cuenca in the center of the region, and a secondary zone (degree 6) that includes not only Toledo, but Madrid, therein belying our conviction (influenced by the maps in *Cómo vive un romance*) that the versions from Madrid ought to be grouped with those of Old Castile. Conversely, the affinity of the versions from Guadalajara (degree 1) and Murcia (degree 2), used to generate the model, is considerably less than that of other provinces outside the region. As in the case of the first map, here too there exists in the periphery of the region a third zone of somewhat less pronounced affinity that includes, to the west, the province of Badajoz (degree 4) and, to the east, the kingdoms of Valencia and Aragón (degrees 3-4). No less significant and logical is the higher proportion of affinity with respect to the model in the versions from the southern band of the northern plateau than in those of the northern part of this plateau.

Not all the ballad type I maps are similar to the two de-scribed thus far because not all regions have a dominant "type" with marked regional characteristics. In contrast to regions 1 and 5, regions 2, 3, and 4 (Zamora-León, Old Castile and Extremadura) generate maps in which there is no nucleus of maximum affinity clearly identifiable with the center of the region that has served to generate the model. Although geography is still seen to be relevant in that the focus of affinity shifts (from the Leonese area in region 2), now to the east (region 3), now to the southwest (region 4), in all three maps we find that the versions which contributed to their respective regional models show no greater affinity to those models than the versions from surrounding provinces that did not contribute to the model. Moreover, in all three cases the patches of attenuated affinity extend to provinces

Map 3 Archetypal Version of *Region 6, BT II*
against all *BT II* versions



Map 4 Archetypal Version of *Region 2, BT I*
against all *BT I* versions



at considerable distance from the region whose model serves as the basis of comparison. Both these characteristics demonstrate that the model versions representative of these regions contain numerous archetypal hemistichs of non-regional character, that is, hemistichs that very nearly approach a universal ballad type I model. It is in these three areas where we find the vast majority of what Catalán and Galmés call the “independent versions” representative of the oldest surviving tradition of the ballad.

In contrast to the maps generated from versions of the autonomous ballad, Map 3 confronts the *ballad type II - region 6* model generated from the seventy-six Andalusian versions with all 208 ballad type II versions. The most striking feature of this map is the absolute homogeneity of the Andalusian region that has served to generate the model: all seven of the provinces with type II versions reflect the maximum degree of affinity. (Jaén, which is virtually unexplored, as well as Madrid and Toledo, have no versions of this type.) Equally evident is the fact that the fourteen Moroccan type II versions are intimately related to their Andalusian counterparts (unlike the five Moroccan type I versions). The affinity of the periphery of Andalusia with the Andalusian model encompasses Murcia as well as all of Extremadura (degree 6) and the Manchegan provinces of Albacete and Ciudad Real (degree 5). However, the greatest novelty of this map with respect to all those that compared type I versions is that here the highest degrees of affinity (8 and 9) with the archetypal version occur no less in the Balearic Islands than in Soria, in Ávila, and, to a significant degree (6), even in La Coruña. It must be noted that in all these provinces ballad type II is represented by no more than one or two versions—a circumstance which indicates that the double ballad *Gerineldo + La condesita* is a rarity of recent importation in these areas. Of greater importance is the fact that in a province such as León, with as many as eighteen type II versions, the affinitive relationship is still quite pronounced (degree 5). The fact that in areas far removed from Andalusia rather dense populations of ballad type H versions manifest a high degree of affinity with the Andalusian model seems to confirm the hypothesis that, in the last analysis, ballad type II depends on the success achieved by some individual initiative to fuse into a single romance the two Andalusian ballads *Gerineldo* and *La condesita*.

The remaining maps generated from ballad type II models all reflect a certain affinity with the Andalusian double ballad, but at

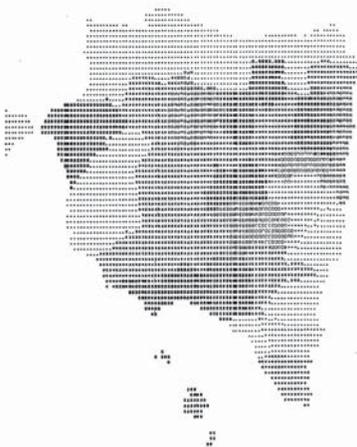
the same time each manifests strong regional characteristics that contrast sharply with one another. This tendency is particularly in evidence in the entire northwestern part of the Peninsula (regions 1 and 2), where an exceedingly compact area of maximum degrees of affinity denounces the existence of a very pronounced and firmly established regional “type” whose unique personality renders it utterly distinct from the type II versions in the eastern half of the Peninsula.

The first series of fourteen maps treats the ballad narrative as a whole, but the coding system also allows us to explore independently the different behavior of the various parts of the overall narration. Thus, in a second series of maps, we elected to use as contrastive units the ballad’s four dramatic scenes or acts. The earlier analyses suggested that the archetypal versions of *region 2 - ballad type I* and *region 6 - ballad type II* were ideally suited to this purpose, as these two geographically and temporally defined subsets illustrate extremes in the total corpus. As Map 4 demonstrates, in region 2 no pronounced regional “type” prevails; the majority of the versions that have contributed to the generation of the archetypal version are representative of local models devoid of expansive force. This model therefore constitutes a good example of the more traditional, conservative geography. In contrast, *region 6 - type II* (Map 3) generates a model of unified regional character that enjoys tremendous expansive force in modern times. (The four period maps that display the relative proportion of type I and type II versions collected in each province in each of the four *periods of collection* established for this corpus confirm the expansive power of the double ballad and its decided triumph over the autonomous *La condesita*.)

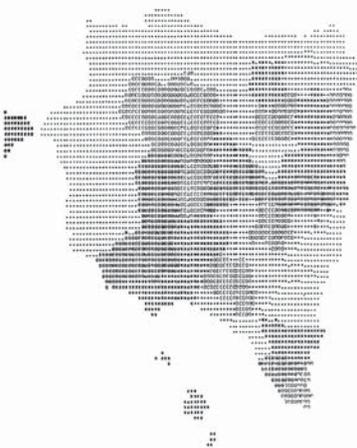
Maps 5-8 constitute an act by act comparison of the *region 2 - type I* model with all 288 type I versions (subclassified by act).

Even the most casual sequential “reading” of these four maps exposes the contrast between act 4 and the other three. In spite of their differences, Maps 5, 6, and 7 have in common an essential feature: in all three the distribution of the diverse degrees of affinity is not organized around a focus or geographic nucleus more or less identifiable with the region that has served to generate the model. None of the three reproduces the configuration depicted in Map 4, the map that displays an image of the global comparison

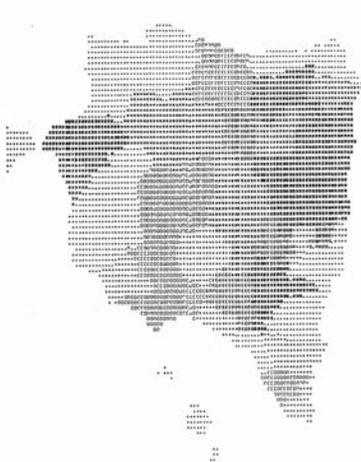
Map 5 Archetypal Act 1 from Region 2, BT I
against all BT I, act 1



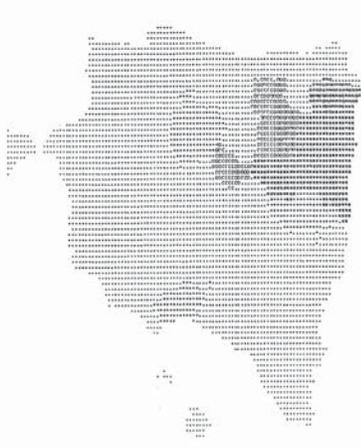
Map 6 Archetypal Act 2 from Region 2, BT I
against all BT I, act 2



Map 7 Archetypal Act 3 from Region 2, BT I
against all BT I, act 3



Map 8 Archetypal Act 4 from Region 4, BT I
against all BT I, Act 4



of the *region 2 - type I* archetypal version with the 288 type I versions. It is obvious that acts 1, 2, and 3 of the model versions generated from the thirty-nine type I versions of León, Zamora, and Orense are not regionally defined. In Maps 5, 6, and 7 geography is shown to be relevant in some cases, but only in a negative sense: for example, in act 1, Catalonia, with proportions of affinity oscillating between degrees 2 and 9, maintains a certain personality in the face of the model; in act 2, Andalusia is the region to separate itself most ostensibly from the norm (as opposed to what happens in acts 1 and 3); and in act 3 regions 5 (New Castile) and 7 (Levant), *en bloc*, demonstrate the relative strength of their respective “types” in differentiating themselves more than the others from a model that is seen to be quite universal.

In sharp contrast to the first three acts of this model version, act 4 is seen to be eminently regional. Map 8 clearly depicts a focus of greater affinity, comprised of the nineteen Leonese versions, surrounded by a very limited area of moderate affinity (degrees 5 and 4) which includes, to the north, three provinces of region 1, and to the east, Palencia and Valladolid (region 3). A second band with weak proportions of affinity (Zamora, Trás-os-Montes, and Ávila) serves only to emphasize the regional character of the last act. As in all maps that manifest a well-defined geographic nucleus, we once more observe here that the nuclear region is redefined, or rather redefines itself, thereby correcting our regional classification of provinces. In this case two of the provinces within region 2, although they contribute 46% of the versions used to generate the model, have a lower degree of affinity than other adjacent provinces in regions 1 and 3. It is apparent that act 4 and only act 4 is responsible for the clearly delineated geographic focus displayed in Map 4. The possibility that a single act can account for the general outline of a map generated on the basis of an entire version (Map 4) is not so surprising if we keep in mind that in the model version under consideration the archetypal hemistichs of act 4 constitute 51% of the whole poem. What does merit emphasizing is the fact that the individual geographic study of the four acts demonstrates that the only part of the ballad to contain a high proportion of distinctive creations is the final scene.

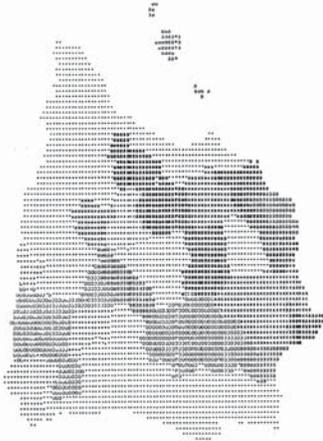
The obvious contrast between the configuration of Map 8 and that of Maps 5, 6, and 7 must not obscure the similarity, in another important aspect, between Maps 6 and 8 as opposed to Maps 5 and 7. The first and third act maps both display large

patches of elevated degrees of affinity: in the first, these are dispersed throughout the most diverse regions of the Peninsula; in the third, they are basically concentrated in two separate areas in the northwest. Conversely, acts 2 and 3 have in common the absence of all degrees of affinity superior to 6 (except in the second act of the five Moroccan type I versions). This indicates that these are the two acts which are subject to the greatest amount of revisionary activity, the two which most attract the creative attention of the artisans of the traditional ballad. With respect to act 4 we already knew this to be the case, but the act by act comparison of type I versions also offers objective proof that act 2, the act in which the wife decides to assert her rights and abandon her family and native land, represents a secondary focus of creative activity. In this second act the personality of the various type I versions undoubtedly achieves maximum definition within region 2 itself, since the three provinces whose versions have generated the model of the archetypal act offer even lower proportions of affinity to that model than provinces outside the region.

The four acts of the *region 6 - type II* model version also produce highly contrasting maps.

The two most outstanding features of Map 9 are the almost absolute uniformity of the degrees of affinity in all provinces with type II populations and the fact that the maximum degree (9) is maintained both without and within the region that has generated the model. Clearly, the Andalusian act 1 model serves to define the first act of the double ballad *Gerineldo + La condesita* in the Spanish oral tradition such as it is known in modern times. This uniformity is not so surprising in type II versions where the first act functions as a transition between the two fused ballads, retaining in its six hemistichs only the essential information: the proclamation of war, mobilization of the count, and estimated duration of the protagonists' forced separation. Only in region 5 do we find a more limited affinity with the Andalusian model (Murcia, degree 8; Albacete, 6; Cuenca, 7; and Guadalajara, with only one version, degree 4). In the preliminary statistical analyses of different aspects of the ballad's narrative structure, both the type I and type II versions from region 5 were seen to be often at variance with the general tendencies manifested by the rest of the

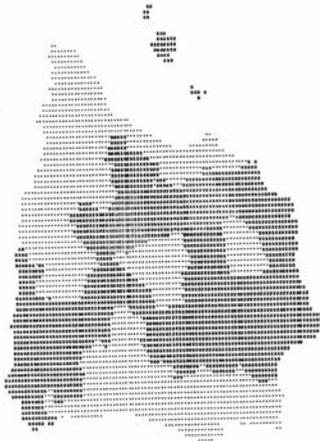
Map 10 Archetypal Act 2 from Region 6, BT II
against all BT II, act 2



Map 12 Archetypal Act 4 from Region 6, BT II
against all BT II, act 4



Map 9 Archetypal Act 1 from Region 6, BT II
against all BT II, act 1



Map 11 Archetypal Act 3 from Region 6, BT III
against all BT II, act 3



corpus. In the map of act 1, the versions from this region once again demonstrate their independence from the other type II versions of *La condesita*.

The map of act 3 is more significant. First, unlike the other three, it comes quite close to duplicating the map generated from the global comparison of all type II versions with the *region 6 - type II* model version (Map 3). The strong affinity between the Andalusian model of act 3 and the third act of the poem in the vast majority of provinces where the double ballad is known confirms the high degree of uniformity among ballad type H versions, for, unlike the abbreviated first act, the third act of the Andalusian archetypal poem constitutes 36% of the entire model and is proportionately far more important than in the Leonese type I model. On the other hand, the map clearly reflects a contrast between the south, where the highest degrees (7-9) prevail, and the central and northwestern portions of the Peninsula, where the affinity ranges between degrees 4 and 6.

The regional character of act 4 (Map 12) is readily apparent. The affinitive relationship between the versions from Andalusia and those from the north virtually ceases to exist. The affinity between the model version and the whole of region 5 also diminishes greatly. Far more surprising is the fact that even Extremenian versions have little similarity to the Andalusian model of act 4 despite the extraordinary similarity between the two maps that compare the Andalusian archetypal version and the Extremenian model to all type II versions. These two maps reveal that while the vast majority of hemistichs characteristic of the Extremenian type II model survive in the Andalusian model, a lower proportion of the latter's hemistichs are retained in the Extremenian versions. Map 11 proves that this attenuation is largely the result of the singularity of the final act of the ballad in Andalusia. This observation is further confirmed by the Moroccan versions, which, on the whole, are also closely akin to those of Andalusia (see Map 3) and yet maintain only slight affinity (degree 3) with the Andalusian model of the last act. Unlike the situation in the previous series of maps for region 2, the regional character of the fourth act in the Andalusian model is not sufficient to determine the map image produced from the global comparison of the Andalusian model version with all type II versions (Map 3). This is due to the fact that in this model the final resolution, although distinctive, is proportionately less important: act 4 here

represents only 31% of the total narration as opposed to 51% in the *region 2 - type I* model (Map 4). The variability of the final act is, nonetheless, apparent within Andalusia itself. Whereas Cádiz (degree 9), Granada (8), Huelva, Sevilla, and Córdoba (7) offer quite high proportions of affinity, Málaga (6) and Almería (5) denounce the existence of very real divergencies in the denouement of their versions.

Map 12 is the most surprising of the series. The second act of the seventy-six Andalusian versions that have served to generate the archetypal act of the region do not seem to be particularly homogeneous since, in general, they have less affinity with the model than the versions from region 5 (New Castile and Murcia) and its prolongation as far as Teruel. Even two provinces as densely populated with type II versions as Huelva and Almería (which together make up 41% of the Andalusian type II versions) have a very limited affinitive relationship (degree 4) with the act 2 model they have largely defined. This can undoubtedly be explained by the fact that nearly half of the Andalusian versions lack two of the three segments that survive in the archetypal second act of the region.

Considered together, these fragmented act by act maps provide objective proof that, as Menéndez Pidal argued, the ballad narrative does not undergo parallel transformation throughout. The opening up of the system is manifested above all in the last act of the ballad, whether this act is accorded maximum development and elaboration or is reduced to a bare minimum. The singers of each region, and even within each region, inevitably exercise the greatest amount of creative activity after they have received the better part of the ballad's actantial message. While this fact merely confirms an observation made on other occasions, its verification with this corpus is extremely important, for it proves that even a final resolution so readily anticipated and highly formulaic as that of our ballad provokes variable reactions in those who carry on the modern oral ballad tradition. The initial act, on the other hand, is less susceptible to regionalization, undoubtedly because its functions, in addition to being highly formulaic, must lay the groundwork for the narration. Acts 2 and 3 are more open to regionalization. Act 3 (the encounter between the wife and the informant), more episodic and formulaic in nature, tends to be expressed via more universal archetypal verses (maps 7 and 11). Nevertheless, we can perceive in it, more than in act 1, certain regional tendencies

(namely, the divergent personality of New Castile and Morocco in type I versions and the differentiation between type II versions in the south and those of New Castile and the north of the Peninsula). The most outstanding characteristic of act 2 (Maps 6 and 10), which relates the wife's departure, is its very great instability, even among versions of the same region or province. This instability is an indication of the singers' very diverse reactions to the act's functions which (as was demonstrated in other analyses) are of far greater transcendence for the underlying message of this ballad than a casual reading leads us to believe.

Unquestionably, the 612 ballad texts studied in this pilot project represent a mere fraction of those elaborated at one time and place or another in the oral life of *La condesita*. Nonetheless, as the preceding examples demonstrate, with the aid of modern technology statistical evidence can be obtained from the diachronic and synchronic analysis of this sample which clearly identifies a number of important, interrelated evolutionary trends—trends that cannot be perceived by simple manual comparison of the texts. Computer-based geographic and temporal analysis of the material, while only one method and one focus among many, can substantially enhance our knowledge of a ballad's history (a history less inaccessible than many have claimed). In this undertaking we stand to gain valuable insights into the mechanisms of reproduction and transformation of orally transmitted poetic narratives.

University of Washington

Notes

¹For his comments on the affinities and differences between his methodology and that of the early Finnish folklorists, see Menéndez Pidal (1973:304, n.2).

²Following the generally accepted convention in *romancero* studies, I retain Menéndez Pidal's distinction between *version* ("complete or fragmentary text of a ballad, taken as a whole") and *variant* ("each of the details of which the version is comprised, in so far as this detail differs from the analogous contents of the other versions") (1973:299). Thus I use interchangeably the terms *motif*, *element* (of content), and *variant*.

³Interest in synchronic and diachronic consideration ought not to obscure the fact that in living folklore traditions spatial projection is in no small measure a reflection of diachronic movement. Consequently, time and space factors should not be conceived as two independent dimensions of

transformation. Transmission theories often tend to suggest just such a dichotomy. For example, J. W. Foster states “though the spatial dimension of transmission may in certain cases be more important than the temporal, most transmissionists think of transmission as a process in time” (1968:239).

⁴See, for example, S. G. Armistead’s comments on the chronological and geographic continuity of the Hispanic ballad tradition (1985:109-10).

⁵Classification on the basis of linguistic areas traditionally distinguishes (minimally) among the following branches of the oral *romancero* tradition: 1) Peninsular Castilian linguistic domain, 2) Canary Islands, 3) Spanish-speaking America (South and Central America, Mexico, and Southwestern United States), 4) Moroccan Sephardic, 5) Eastern Sephardic, 6) Galicia, 7) Portugal, 8) Portuguese islands and overseas communities, 9) Brazil, and 10) Catalán linguistic domain. A simple list of the Judeo-Spanish communities represented in but one major ballad collection (of 2150 texts) will suffice to underscore the geographic and cultural diversity implicit in this broad classification. *Eastern Sephardic*: Sarajevo, Belgrade, Vienna, Istib, Sofía, Stanke Dimitrov, Tatar-Pazardjik, Plóvdiv, Ruse, Bucarest, Rósióri, Salonika, Véroia, Kastoría, Lárisa, Edirne, Istanbul, Bursa, Tekirdag, Çanakkale, Izmir, Rhodes, Beirut, Damascus, and Jerusalem; *Northern Africa*: Tangier, Tetuán, Arcila, Larache, Alcazarquivir, Casablanca, Morocco, Gibraltar, and Oran. See Armistead (1978:1:34.) For information on ballad-collecting activity in the various branches of the modern tradition, see A. Sánchez Romeralo (1979a:15-51), S. G. Armistead (1979c:53-60), and D. Catalán (1979:217-56), all in *The Hispanic Ballad Today: New Frontiers*. For recent, major bibliographic sources for the various branches of the tradition, see S. G. Armistead (1985:111).

⁶See D. Catalán’s reply (1959b:149-82) and Devoto’s counter-reply (1969:11-44).

⁷This phenomenon will be discussed and documented at several points in the ensuing pages. For further discussion of the extent to which the individual’s possibilities of selection and invention are governed by local tradition, see D. Catalán, (1970-71:18-19). Catalán’s observations are substantiated in his reply to Devoto (1959b:155ff). Even casual perusal of the geographically, temporally, and thematically organized ballad corpora published in the *Romancero tradicional del las lenguas hispánicas* series makes apparent the pressure of local tradition.

⁸See D. Catalán’s detailed analysis of the final sequences of the ballad (1986:94-97). Catalán’s study further documents the profound alteration of latent semantic content in response to social changes in the referent with partial analyses of *Una fatal ocasión*, *Muerte del duque de Gandía*, *Fatricida por amor*, *El moro que reta a Valencia*, and *Muerte del príncipe don Juan*. For other studies that discuss the relationship between prevailing social structures and traditional Hispanic narratives, see D. Catalán (1978:245-70), Manuel Gutiérrez Esteve (1978:551-72), J. Antonio Cid (1979:281-359), and Beatriz Mariscal de Rhett (1984-85:19-56 and 281-333).

⁹Since the early 1970’s the majority of publications by D. Catalán and members of the Instituto Seminario Menéndez Pidal which he directs have dealt with multi-level analysis of variation and include theoretical discussion of the *romancero* as a model of open-structured narratives. The basic tenets of this conception of the *romancero* first appear in D. Catalán (1976:55-62). My doctoral dissertation (S. H. Petersen 1976a) constitutes the most exhaustive

analysis to date of poetic, verbal, and semantic transformation of a single Hispanic ballad.

¹⁰I refer to individual informants as “she” in recognition of the dominant role women have played as singers of *romances* in the last two centuries.

¹¹Any field researcher who has recorded *romances* sung or recited by a group of informants has witnessed the lively debates that often ensue among neighbors or family members who argue over the “correct” text (“No, mama, that’s not the way it goes. Let’s see if I can get it right . . .”). Equally indicative of the textual memorization process are the comments made by informants who, upon faltering in their recitation, almost invariably either lament the passing of a friend or relative whose memory was vastly superior to their own or recommend consulting other family members or friends from whom they learned the ballad or with whom they used to sing (e.g., Felisa Fernández Naranjo, age 44, from Hinojosa de Calatrava, who, having forgotten portions of *La loba parda*, suggested the collectors seek out her father who learned the ballad as a boy and used to sing it to her and her brothers when they were children to put them to sleep. As promised, Eusebio Fernández Martínez, a 72-year-old shepherd, later sang a more complete, but otherwise essentially identical text with the help of his wife. See A. Sánchez Romeralo 1977:155-57.)

¹²The texts (cited from *CGR* 1A:199-200) also illustrate one of the difficulties of using the date of collection as a basis for measuring diachronic movement. As the older version was sung by a woman of 31 and the more recent version, by an 80-year-old woman who learned the ballad in her youth, the temporal “distance” separating the two is, in fact, minimal. On the other hand, as some areas of the tradition are extremely conservative and others highly innovative, the rhythm of change varies greatly from area to area, with the result that a considerable temporal “distance” often exists between contemporary versions originating in geographically distant areas. For other examples of the enormous fidelity to local tradition as well as additional commentary on the rhythms of transformation in the *romancero*, see *CGR* 1A:195-204.

¹³See Cid’s lucid exposition of the larger implications of ritualization, contamination, and other factors that restrict traditional variation (325-35).

¹⁴For example, in 1985 the Seminario Menéndez Pidal’s three-week field trip to León yielded some 2700 ballad texts on 110 traditional themes. Approximately 545 informants from 255 Leonese towns and villages contributed to this sizeable corpus. Naturally, not all the evidence recorded during these recent campaigns is of equal value, but superb versions of rare ballads—traditional *romances* scarcely documented in this century—continue to surface with surprising frequency. A list of some of the most invaluable testimony unearthed in the last fifteen years appears in D. Catalán’s article in this issue.

¹⁵On the other hand, recent critical interest in traditional Hispanic cultural activities, and specifically the field research on the *romancero*, has served to rekindle popular interest in traditional cultural products and has thereby contributed to affirming and prolonging the life of this body of oral poetic discourse.

¹⁶In his study of *Muerte del príncipe don Juan* (1968a), for example, Bénichou explains “versions PV and RC contain the dying protagonist’s

declaration that the wife he leaves behind is pregnant; therefore I include them in this group” (98-99), and further on, “As for the details or poetic elements that are scarcely found outside this group . . .” (101), and again, “They [the Portuguese texts] total sixteen versions, which must be divided into two subgroups on the basis of their divergent endings, that is, first, the ending with the projected marriage, and second, the ending with the young woman’s complaint” (106, n.18).

¹⁷Although from a different perspective, few studies provide such conclusive evidence of the importance of geography as A Sánchez Romeralo’s article in this volume on *La loba parda*, a modern shepherd’s ballad propagated exclusively along the six migration routes that traverse the Peninsula from north to south. Both in the relative homogeneity of the ballad (owing to the crisscrossing and convergent patterns of the sheep tracks and the common winter pastures) and in the few distinctive characteristics of one or another of the six groups of versions, geography is seen to function with singular clarity.

¹⁸The test ballad was selected on purely practical grounds (its availability in published form), but our prior knowledge of it (owing to the two earlier major geographic studies) proved extremely useful in anticipating system design requirements and in providing some means of verifying the validity of computer-generated analyses. The system, designed to accommodate many different types of ballad corpora, is fully described in my dissertation (Petersen 1976a). Numerous publications that include sample analyses have appeared since the mid-seventies: Catalán et al. (1975), Catalán (1976), and Petersen (1976b, 1978, 1979).

¹⁹For the present discussion only the first five digits of this eleven digit code require some explanation. They define each hemistich with reference to the thematic segment to which it “ideally” belongs (2 digits) and the element of information it “ideally” carries (3 digits). Any discrepancy between these five digits and the five that identify the thematic segment in which a given hemistich actually appears (c above) and the element of information it actually carries (g above) automatically denotes some type of reorganization or modification of the narrative structure and content of the poem. Both these interrelated phenomena—the archetypes’ potential for opening up and acquiring new semantic values and their mobility (or capacity to function in different contexts within the overall narrative) —constitute an important mechanism of variation and are central to the whole question of structural transformation in oral poetic discourse.

²⁰The figures become more significant when cross-tabulated to indicate precisely where the hemistichs that reproduce the archetypes of a given segment have travelled to. This information is even more useful when further broken down by geographic areas, period, or any other relevant internal variable.

²¹Given certain characteristics of our data base, the particular history of this ballad, and our coding system, the classification of versions by ballad type and region was considered the most valid and useful procedure for establishing diachronic and synchronic breaks in the corpus. The two oral ballad types (BT I and II), which largely eliminate the interference of diachrony from geography, constitute a more valid means of temporally classifying the versions of *La condessita* than the time periods that can be established for this corpus (see Petersen 1979:175-76, n.11). With respect to geography, although all

versions were identified by both province and region, only the latter classification guaranteed comparison of roughly equal populations. (With a larger data base, smaller geographic units—province, or provincial districts, or even townships—would be possible and clearly preferable.)

²²To avoid unnecessary confusion in reading the chart, I have suppressed statistically Irrelevant data.

²³In as much as only 15% of the Levantine type I versions make use of archetype 44505511, little significance can be attributed to the apparent contradiction in Levant to the trends identified above.

²⁴For a more detailed description of this system and the results of its application, see Petersen (1979:176-228 and 1976a:212-82). Future applications of this sort will benefit enormously from recent technological advances in computer cartography. Graphics-oriented microcomputers (with appropriate peripherals and software) provide a sophisticated tool for geographic data processing and mapmaking and largely free the individual researcher from costly and cumbersome mainframe computing.