



You are accessing the Digital Archive of the Catalan Review Journal.

By accessing and/or using this Digital Archive, you accept and agree to abide by the Terms and Conditions of Use available at [http://www.nacs-catalanstudies.org/catalan\\_review.html](http://www.nacs-catalanstudies.org/catalan_review.html)

Catalan Review is the premier international scholarly journal devoted to all aspects of Catalan culture. By Catalan culture is understood all manifestations of intellectual and artistic life produced in the Catalan language or in the geographical areas where Catalan is spoken. Catalan Review has been in publication since 1986.

Esteu accedint a l'Arxiu Digital del Catalan Review

A l' accedir i / o utilitzar aquest Arxiu Digital, vostè accepta i es compromet a complir els termes i condicions d'ús disponibles a [http://www.nacs-catalanstudies.org/catalan\\_review.html](http://www.nacs-catalanstudies.org/catalan_review.html)

Catalan Review és la primera revista internacional dedicada a tots els aspectes de la cultura catalana. Per la cultura catalana s'entén totes les manifestacions de la vida intel·lectual i artística produïda en llengua catalana o en les zones geogràfiques on es parla català. Catalan Review es publica des de 1986.

## ***Catalan's Place in Romance Revisited*** **Mathew L. Juge**

**Catalan Review, Vol. XXI, (2007), p. 257- 277**

# CATALAN'S PLACE IN ROMANCE REVISITED\*

MATTHEW L. JUGE

## ABSTRACT

---

Catalan is unique among the Romance languages in having a relatively large number of speakers in a thriving speech community but not being the dominant language of a major nation-state. It is also unusual in that its position within the Romance subfamily is a matter of some debate. I argue that the application of the principle of contact linguistics to data from Catalan dialects, especially the Alguerès variety, support rejecting the traditional treatment of Catalan as Ibero-Romance and Occitan as Gallo-Romance in favor of placing Catalan and Occitan together in a separate subbranch.

---

## INTRODUCTION

Catalan is unique among the Romance languages in having a relatively large, thriving speech community but not being the dominant language of a major nation-state. Also unusual is the ongoing debate over its position within Romance, "one of the oldest and most insoluble problems in Romance linguistics" (Blasco Ferrer, *Grammatica* 132). I argue that historical and contact data favor grouping Catalan and Occitan together over the traditional treatment of Catalan as Ibero-Romance (I-R) and Occitan as Gallo-Romance (G-R). Higher-level groupings affected by reclassifying Catalan lie outside the scope of this paper.

The placement of Catalan in I-R appears to result from geographic distribution rather than linguistic features and history. Likewise, the inclusion of Occitan in G-R seems to stem from its location across the Pyrenees. Though common, this view is not the only grouping scholars have proposed. Paden (3) notes that "Occitan regionalists sometimes considered Catalan a form of Occitan" but also cites Elcock (448), who states that "'from the standpoint of the linguist there can be no doubt concerning its status' as an independent language." Since the European varieties of the Romance languages have remained relatively static, the long-standing contact relations among them may complicate the process of subgrouping.

---

\* I thank my consultants in Alguer, three anonymous reviewers, and the editors of *Catalan Review*, who also provided help in making arrangements for my research in Sardinia, and William F. Weigel. This work was supported by a grant from the Research Enhancement Program at Texas State University-San Marcos. Any shortcomings are my responsibility alone.

The *Alguerès* dialect of Catalan, however, provides a special opportunity to separate contact features from signs of genetic relatedness. Brought to Alguer, Sardinia, in the 14<sup>th</sup> century, this dialect has been in contact with Italian and Sardinian, whose greater distance from Catalan within Romance makes identifying contact-based influence easier than with more closely related varieties.

Here I explore the difficulties involved in applying the comparative method to the Romance languages and adduce lexical, morphological, and phonological evidence as well as specifically *Alguerès* data. I aim to synthesize previous work in a broader context and take the first steps in using the extraordinary situation of Catalan in Alguer to sort out complicated issues in language contact and genetic linguistics.

#### TRADITIONAL SCHOLARSHIP

Posner (24) calls Hall's family tree with Catalan in Northern I-R (northern branch) and Occitan in G-R from his article on Proto-Romance reconstruction "[t]he now most widely accepted version of the Romance Stammbaum" but concludes, "The end-result looks suspiciously like areal grouping." She notes, though, that Trager (no citation given) puts Catalan closer to Occitan than to French despite the use of morphosyntactic criteria found in French and Occitan ("Gallo-Roman") but not in Catalan (Posner 197). Meanwhile, Harris (16) notes the "far closer parallels, both synchronic and diachronic, between Occitan and Catalan than between Occitan and French" and that a protolanguage encompassing Occitan and Catalan is "quite plausible" but a group with French and Occitan but not Catalan "certainly is not."

Posner and Harris do not mention Bec (468), who concludes, "Everything becomes simple, however, if one lists Catalan, as we do, in an *Occitan-Romance* unit more flexibly defined and intermediate, by definition, between Gallo-Romance proper (*langue d'oïl*) and Ibero-Romance (Spanish and Portuguese)." While Bec's statement might seem conclusive, the debate goes on. Baldinger and Badja Margarit ("El catalán"), for example, both cite G-R (Meyer-Lübke, *Das Katalanische*) and I-R (Menéndez-Pidal) claims for Catalan. Additionally, Baldinger cites Rohlfs (among others) favoring a grouping of Catalan, high Aragonese, Gascon, and possibly Provençal, while both Baldinger and Badja Margarit see Catalan as a bridge language.

<sup>4</sup> Blasco Ferrer ("La posizione" 132-33) cites several problems with previous scholarship on the issue: a) the underestimation of Modern Occitan; b) incomplete knowledge of both Old and Modern Catalan;

c) the lack of satisfactory comparative analysis of morphosyntax in favor of phonology and lexicon; and d) undue emphasis on extralinguistic factors, including political, ethnic, and cultural issues.

Points a) and b) stand out because limited understanding of the modern languages impedes the use of dialect data that may provide an important complement to older texts in reconstruction. Additionally, limited knowledge of the modern languages hinders internal reconstruction, and thorough analysis requires a clear understanding of the relationship between the old and modern languages.

Blasco Ferrer does not give a precise definition of the term "position" with respect to Catalan and Romance. Though genetic classification seems a likely interpretation, he also mentions the relative number of "typological traits" shared by Catalan with I-R and G-R ("La posizione" 134).

In this paper, I focus on Catalan's genetic position in Romance, using typological data where appropriate.

#### ESTABLISHING INTERNAL RELATIONSHIPS

Harrison identifies several problems with subgrouping methodology. On sound change he states, "Our notions regarding naturalness are grounded in nothing more than vague intuition and anecdote" (236). Within this "bleak scenario," he emphasizes that "subgrouping arguments must make crucial reference to changes" (233). Further, he states, "Eliminating 'shared retention from an earlier antecedent state' as an account of similarities in outcome is a problem unique to subgrouping," which he seeks to distinguish clearly from the goal of using the comparative method to establish genetic relatedness apart from family-internal relations. He sees subgrouping as the step taken after establishing genetic relatedness and expresses a view compatible with Hall's approach: "one must subgroup in order to reconstruct" (239).

Garrett discusses the interpretation of family-internal relations (in the context of Indo-European branches) and, without rejecting the Stammbaum model, suggests that the apparent unity of a family like Italic may result from diffusion of traits across members of a dialect continuum. Meanwhile, Baldi observes that extreme contact cases can hinder genetic classification, as with Armenian, "first thought to be a radical dialect of Persian" (42).

With languages that have been essentially static for a long time, how to separate genetic and contact factors is not always immediately apparent. In historical phonology, shared innovations are commonly cited as evidence of branching. Some sound changes, however, cannot

be used for subgrouping. The presence of front rounded vowels, for example, is an areal feature stretching from Norway in the north to France in the southwest to Hungary in the east (plus Iceland and Scotland), spanning more than a dozen countries and crossing not only boundaries between family branches but also between language families and thus does not neatly indicate subgroups. So we must identify and separate genetic and contact features in order to firmly establish Catalan's position within Romance.

Hock explores the relationship between genetic relatedness and shared features in his discussion of convergence in Europe. He argues that a sprachbund involving unrelated languages can show traits like those found in dialect continua. He explains, "The similarity between sprachbund and dialect continuum should perhaps not be surprising, for...the use of different languages in bilingual societies is in many ways comparable to the use of different dialects in monolingual societies" (509). Hock (510) points out that "major changes in grammatical structure can diffuse over large geographic areas through a dialect-like chain of bilingualism. And this diffusion can be completely oblivious of the genetic relationship between languages".

This issue surfaces in the Catalan/Occitan context in the use of the definite article, which both languages use with possessives and personal names, unlike both French and Spanish. However, Italian uses the definite article with possessives and (dialectally) with proper names,<sup>1</sup> so this feature does not establish Catalan/Occitan unity.

In a family like Romance, the interpretation of textual evidence is a key issue, as modeling change on the basis of written records requires care. First, as Lass (58) notes, "The lag between phonological and orthographic change is a commonplace." Meanwhile, de Dardel (30) discusses the "risks that recourse to written Latin includes" in reconstructing Proto-Romance, and Herman and Wright (18) caution that interpreting texts "is not a straightforward matter." Nonetheless, de Dardel reminds us that there are advantages, too. In the end, such texts require judicious use.

Haiman and Benincà explore methodological issues in subgrouping in Rhaeto-Romance (R-R). They start with two independent questions: a) "do they share enough features to justify their being grouped together?" and b) "does this unit differ sufficiently from other languages to justify status on a par with them?" (5). With respect to Catalan and Occitan, these questions can be thought of this way: a) do they constitute a subgroup within the family? and b) do they group especially closely to other members of the family?

While Walsh (405) recognizes a "sharp linguistic boundary

<sup>1</sup> Thanks to an anonymous reviewer for this observation.

separating the Eastern from the Western dialect" of Catalan, at the Pan-Romance level, some traditional criteria are subject to criticism. Malkiel, for instance, cites repeated challenges to the East-West division. Maiden echoes Malkiel, pointing out that even the La Spezia-Rimini line, perhaps the most famous of Romance isoglosses, presents complications (273-74):

The importance of this Line as the division between "eastern" and "western" Romance languages (cf. Wartburg (1967)) should not be overstated: it does *not* mark the boundary between "western Romance" weakening of intervocalic voiceless consonants, and "eastern Romance" non-weakening ...nor between "western" inflectional -s and "eastern" inflectional -i or -e...

Addressing the second question I will focus on the relationship to G-R and I-R in terms of membership rather than sister relations. That is, I will address whether Catalan and Occitan belong in either of these groups and, if not, leave higher-level branching for further research.

#### SHARED INNOVATIONS

Cases in which two or more languages share a feature that is demonstrably new relative to the mother language provide the strongest evidence for subgrouping. These shared innovations suggest that the languages derive from an intermediate stage in which certain changes not attested elsewhere in the family took place before these languages diverged (this does not rule out parallel developments elsewhere in the family). However, Hock (579) notes that "it is not always easy to establish whether given common innovations are to be attributed to a period of exclusively shared prehistory or to diffusion between different members of a dialect continuum within the proto-language."

#### SHARED RETENTIONS

Common features that languages have retained wield less diagnostic force. Hock (579) explains, "Common archaisms (or inheritances) can be found between any two members of a larger language family. In fact, sometimes they survive best in quite distantly related members of the family, located on or near the periphery." Thus, for example, Occitan *can* 'dog' and its French cognate *chien* do not furnish evidence of a special Occitan-French connection, merely a lack of evidence of a connection between either of these and, say, Spanish *perro*.

## AREAL FEATURES

As mentioned above, several areal features affect the evaluation of the Catalan-Occitan relationship, the most important being front rounded vowels, the merger of /b/ and /v/, uvular vs. alveolar rhotics, and the so-called spirantization of voiced stops.

The presence of front rounded vowels in Occitan and their absence in Catalan correlates directly with the phonologies of their neighbors — Occitan follows French and, across the Pyrenees (and across the Mediterranean in Algier), Catalan resembles Spanish, Portuguese, and Italian (except Gallo-Italian; see Maiden 235). The ambit of front rounded vowels in Europe, crossing linguistic lines both within families and between families, eliminates this trait as a diagnostic feature for subgroup relationships for Catalan and Occitan. A useful parallel here (courtesy William F. Weigel, personal communication) is the relationship between Yiddish and German, known to be very closely related but differing in this phonological trait.

Wheeler (249) notes that, like Basque, Spanish, Galician, and most of Catalan, southwestern Occitan dialects show no /b~/v/ distinction (but Scala 31 and Veny 107 note that *Alguerès* does). Basque's presence in the first group shows clearly that this trait does not reliably indicate family relations.

Some Occitan dialects feature uvular rhotics, but this difference from Catalan does not diminish the case for Catalan/Occitan unity, as not all dialects share this trait. Furthermore, uvular rhotics are another areal feature of northwest Europe, present in French, Dutch, German, Danish, and some Norwegian and Swedish dialects.

Fricative or approximant realizations of voiced stop phonemes occur not only in standard Catalan (but not *Alguerès*), some Occitan dialects (*Lengadocian* and *Gascon*; cf. Wheeler 248), Spanish, and Basque. The presence of Basque in this group and the absence of Portuguese require mention. The fact that most dialects of Portuguese, so close to Spanish, do not show this trait and the fact that the non-Indo-European language Basque does show this trait indicate that areal features are not reliable signs of genetic relationships.

In sum, certain potential markers of degree of genetic distance cannot always be taken at face value and must be weighed in conjunction with other traits.

## INDEPENDENT PARALLEL DEVELOPMENT

Besides shared innovation and areal features, independent parallel developments may play a role. For example, the change of initial /f/ to

/h/ in Spanish and Gascon need not be seen as a sign of a special affinity, just as the development of phonemic nasal vowels is a widely attested type of change. Thus the presence of nasal vowels in Southeastern Lemosin (Wheeler 254) should not be taken as a sure sign of special affinity with French, since such developments are common across the world's languages.

This point also raises the important issue of individual developments in given dialects. If a trait cannot be reconstructed for the proto-languages, i.e., proto-Catalan or proto-Occitan, then it does not merit the same diagnostic status as features that can be attributed to the proto-languages. In other words, developments after the breakup of proto-Catalan and proto-Occitan into their respective dialects do not bear directly on possible earlier Catalan/Occitan relations.

### MORPHOSYNTAX

Blasco Ferrer ("La posizione" 136) writes that "the decisive role in determining the linguistic position of Catalan will be up to morphosyntax" and emphasizes the importance of typological analysis. He clearly identifies which factors he considers most important. The bulk of the analysis comprises the examination of 21 morphosyntactic traits:<sup>2</sup> 1. Definite articles from IPSE 'self'; 2. Compound 1P and 2P pronouns; 3. Demonstratives; 4. Adverbial pronouns from IBI and INDE; 5. Relative pronouns; 6. The interrogative *quin*; 7. The neuter pronoun (< HOC); 8. Negative reinforcement with *pas*; 9. Quantifiers; 10. The so-called preterit infix /-r-/; 11. The periphrastic preterit; 12. The imperfect subjunctive in the protasis of counterfactual conditionals; 13. -UTU and -ITU second-conjugation participles; 14. TENERE vs. HABERE as the principal verb of possession; 15. Participle agreement in the perfect with HABERE; 16. TORNARE as an auxiliary of repetition; 17. *En* + infinitive; 18. Tmesis in the future and conditional; 19. Temporal and locative adverbs; 20. MAGIS vs. PLUS for 'more'; and 21. The nominal suffix -ATA.

I will typologize these traits according to their diagnostic value for a possible proto-Catalo-Occitan and hence a Catalo-Occitan subgroup within Romance. Some features may be shared retentions (not evidence of genetic unity) or innovations found in other

<sup>2</sup> Grammatical abbreviations used here are the following: 1 = first person, 2 = second person, 3 = third person; S = singular, P = plural; ICONJ = first conjugation, ART = article, IMPF = imperfect, COMP = complementizer, IND = indicative, M = masculine, NEG = negation, POSS = possessive, PRS = present, SBJ = subjunctive.

languages (suggesting independent parallel development). In the latter instance, the case for Catalan/Occitan unity is strengthened only if it can be established that, while the other developments elsewhere in the family are independent but parallel, those in Catalan and Occitan reflect a shared history. Certain traits are areal features and may provide reason to discount differences between Catalan and Occitan because we can identify contact, rather than genetic, factors for divergence between the two languages. Finally, unique shared innovations, new developments found in Catalan and Occitan, will provide the strongest evidence of a special genetic relationship between them.

Before proceeding I should note a certain ambiguity in Blasco Ferrer's use of the terms *Occitan*, *Gascon*, and *Provençal*. His presentation of examples and other data often shows Gascon and Provençal as equals of French, Spanish, and Portuguese, implying that they are all separate languages. He rarely uses the label *Occitan*, though he often uses the Italian term *occitanico*, apparently in contrast with terms like *Gascon*, not as a cover term (e.g., 146). Meanwhile, though other sources commonly use *Provençal* to identify a variety of dialects, "Provençal refers properly only to the local dialect of Provence" (Harris 16). Thus I use *Occitan* as the cover term for the set of dialects that includes Provençal as well as other varieties, such as Gascon and Lemosin.

As Blasco Ferrer notes, the first trait, the change of the intensive pronoun/adjective into a definite article, is found not only in Catalan and Occitan (old Gascon), but also in Sardinian. This distribution suggests that it is either a shared retention or an innovation shared by distantly related languages. Thus, it constitutes neither strong evidence in favor of nor against Catalan-Occitan unity.

Another apparent shared retention is the use of adverbial pro-forms from Latin *IBI* 'there' and *INDE* 'from there' (e.g., Catalan *anem-hi* 'let's go there' and *n'hi ha tres* 'there are three of them') (trait 4). As Blasco Ferrer ("La posizione" 146) notes, these particles occurred generally in the old western Romance varieties, including Spanish, Portuguese, Gascon, and French, but are no longer found in modern Spanish and Portuguese. This says very little about the genetic relations between Catalan and the other Romance languages. The loss of these particles in Spanish and Portuguese offers no information about Catalan's possible membership in I-R, since the old languages all used them.

Other shared retentions that do not inform us about Catalan's place in Romance are relative pronouns (trait 5); the use of *HOC* as a neuter pronoun (trait 7); the agreement of the past participle in perfect constructions with *HABERE* (trait 15); the use of the imperfect

subjunctive in the protasis of conditional sentences (trait 12) — shared by Catalan with Portuguese, Spanish, Gascon, and Provençal; and the use of TORNARE to mark repetition (trait 16) — shared by Catalan with Portuguese, Provençal, and Old Spanish.

Some traits developed after Catalan and Occitan would have separated from a putative proto-Catalo-Occitan and thus are not useful for establishing genetic relationships. For example, Blasco Ferrer's second criterion (trait 2) is the development of compound 1P and 2P tonic pronouns (NOS ALTEROS/VOS ALTEROS) as opposed to the simple forms NOS/VOS still found, for example, in French and Valencian Catalan.

Trait 8, *pas* marking or reinforcing negation, also seems to have developed too late to definitively address Catalan's position in the family. Further, its presence in French shows a broader distribution, possibly independent parallel development. Nor can we rule out, a priori, spread through contact. Blasco Ferrer ("La posizione" 153) points out, however, that such uses are not characteristic of Spanish or Portuguese, which at least is consistent with the view that Catalan does not belong in I-R.

Demonstratives — specifically whether a given language has a two-term (Catalan, Italian, R-R, Sardinian) or a three-term (Spanish, Portuguese, Gascon) system — are the subject of trait 3. Here again we see that Catalan falls into a group of languages not belonging to any Romance subgroup and shows independent parallel development. The interrogative *quin*, trait 6 on Blasco Ferrer's list, falls into the same category of independent parallel development, since, as his data show, it is found not only in Catalan, Provençal, and Gascon but also in southern Italian, Sardinian, and Romanian. Likewise, trait 14, the use of TENERE rather than HABERE as the principal verb of possession, is a case of late development. In this case the contiguity of the languages in question — Portuguese, Spanish, and Catalan — leaves open two possibilities: an independent parallel development or an areal feature. In either event, this change does not shed light on Catalan's genetic affiliation.

Similarly, trait 10, the so-called preterit infix, is a poor candidate for an inherited feature of Catalan and Occitan. Before addressing this feature as a subgrouping criterion, it must be noted that it is not an infix, just part of the perfective past marker (cf. the voiced labial in the imperfect indicative, e.g., *cant-a-ve-s* 'sing-ICONJ-IMPF.IND-2S = you sang'). Blasco Ferrer's argument for the development of forms with this element rests on homophony avoidance and presents problems similar to those presented by Pérez Saldanya ("Gramaticalització") and rejected by Juge ("Morphological"). Perhaps more significant for Catalan's place in the family are the patterns of the spread of /-r-/

within the preterit. As Table 1 shows (cf. Nadal and Prats 288, Bybee 55 [citing Meyer-Lübke, *Grammaire* 352; Ronjat 193]), in Old Catalan and Old Occitan the *-r-* was found in the 3P form. Blasco Ferrer identifies the 14<sup>th</sup> century as the time when these extensions first appear. (In table 1, bolded forms reflect substantial intraparadigmatic influence, i.e., leveling, or so-called analogy, and italicized forms are the base forms for intraparadigmatic influence.)

TABLE I  
INTRAPARADIGMATIC INFLUENCE IN THE PRETERIT  
IN CATALAN AND OCCITAN

	Catalan		Occitan		
	Old	Modern	Old	Charente	Clermont-Ferrand
1S	canté	cantí	cantéi	<b>cantèri</b>	<b>cantéte</b>
2S	cantast	<b>cantares</b>	cantést	<b>cantères</b>	<b>cantétes</b>
3S	cantà	cantà	<i>cantét</i>	cantèt	canté
1P	cantam	<b>cantàrem</b>	cantém	<b>cantèrem</b>	<b>cantétem</b>
2P	cantats	<b>cantàreu</b>	cantétz	<b>cantèretz</b>	<b>cantétez</b>
3P	<i>cantaren</i>	cantaren	<i>cantéren</i>	cantèron	<b>cantéton</b>

This trait (i.e., the so-called preterit infix) presents some difficulties. As Blasco Ferrer ("La posizione" 154) notes, a similar development has occurred in Romanian — certainly not a close relative of either Catalan or Occitan. If the Romanian changes are parallel to but independent of what happened in Catalan and Occitan, then the developments in the latter two languages may also be independent of each other. Indeed, Table 1 shows different patterns within Occitan, with some dialects, such as Gascon, as Blasco Ferrer ("La posizione" 154) notes, showing no such extensions at all. Perhaps the strongest refutation is the fact that Old Catalan texts show the paradigm without the extension.

The sound changes preceding these morphological changes, though, may provide more convincing evidence of Catalan/Occitan unity than the data from the preterit itself. Both languages reflect the loss of final *-US* in 1P forms. This change, however, also characterizes R-R, French, and Romanian and thus is not a definitive sign of Catalan/Occitan unity.

The next item (trait II), the periphrastic preterit, repeats the problem of relative chronology. This construction, whose morphological and

narrative features, including others' analyses, are discussed by Juge ("Morphological," "Narrative"), marks perfective past with an auxiliary plus infinitive. In early Catalan texts, simple preterits predominate in general usage and as forms of *anar* 'go' used as auxiliaries in the periphrasis. The morphological changes that accompanied the shift from preterit auxiliary to present auxiliary and led some scholars to adopt the misnomer 'GO-past' are described in Juge ("Morphological"). Again, the relatively late development does not affect the case for Catalo-Occitan.

For trait 13, second-conjugation participles like Catalan *venut* 'sold' (infinitive *vendre*), Blasco Ferrer ("La posizione" 161) rightly acknowledges that the use of /-i-/ in Spanish and Portuguese (*vendido* 'sold') is not evidence of genetic closeness between Catalan and G-R, since, as discussed above ("Shared retentions"), shared retentions do not indicate special affinity.

Another shared retention is the inchoative suffix typical of fourth conjugation (-IRE) verbs in various Romance languages, including Catalan, Occitan, French, Italian, R-R, and Romanian. The fact that Spanish and Portuguese do not feature this type as robustly — especially in terms of the -Ø-/-SC- intraparadigmatic alternation — does not constitute strong evidence in favor of Catalan/Occitan unity.

The ninth trait in Blasco Ferrer's list is a collection of reflexes of (Late) Latin quantifiers in Provençal/Occitan, Gascon, Catalan, Aragonese, Castilian, and Portuguese (French and R-R are excluded without explanation). At first, some of the items seem to suggest a relationship connecting Provençal/Occitan (in his use of the terms), Gascon, and Catalan. However, some of the items are shared with Old Spanish and Old Portuguese, like RES/REM 'nothing,' while the uncited French equivalents, such as *rien* 'nothing,' *quelconque* 'any, some,' and *chacun* 'each,' show that the others are not exclusive to Occitan and Catalan.

Tmesis in the future and the conditional (trait 18) — the intercalation of one or more pronouns between the infinitive and the person/number markers deriving from HABERE — is, by Blasco Ferrer's own account, present in all the medieval Western Romance languages. Its loss in most Romance varieties, including Catalan and Occitan (but not Gascon; cf. "La posizione" 165), offers no information about subgrouping.

Adjectival and adverbial comparison with MAGIS or PLUS (trait 20) offers little insight into Catalan's relations with the other Romance languages. As Blasco Ferrer notes, both occur early, and now the basic pattern is MAGIS in the Iberian Peninsula, with PLUS predominating elsewhere. This trait seems to have much more to do with geography and contact than with genetic connections.

Blasco Ferrer's final feature (trait 21), the use of the suffix *-ATA* to encode a blow (e.g., Catalan *bastonada* 'a blow with a stick'; cf. *bastó* 'stick'), involves a pattern of use found to some extent in Spanish (now generally favoring *-azo*), Portuguese, and French in addition to Catalan and Occitan. Such a distribution provides little insight into Catalan's place in Romance.

The use of the preposition *en* with an infinitive (rather than the gerund as in other Romance languages) shows some potential as a sign of a special connection between Catalan and Occitan, as this feature (trait 17) appears in Catalan, Provençal, and Gascon. The late date of its first appearance, however, bodes ill. Occurring alongside the structure *en* + gerund, the infinitive construction does not appear in Catalan until the 14<sup>th</sup> century. Using this trait as diagnostic would require positing that it was in use before but earlier texts happened not to include it. This is not implausible, but it is not especially compelling.

Blasco Ferrer's nineteenth trait is reminiscent of the ninth. It concerns a set of 55 temporal and locative adverbs found in Western Romance, though why these particular items were chosen is not explained. A handful are found only in one or another language, the remainder divided into the following groups: pan-Western Romance (21), I-R (15), and Catalan/Provençal (13). The exclusion of French and R-R is not explained. Nonetheless, the 12 items shared only by Catalan and Provençal (the thirteenth item is found in Portuguese as well) suggest a special connection.

Table 2 shows how Blasco Ferrer ("La posizione") summarizes his examination of the 21 traits and identifies how many points show agreement among the various Romance languages (171):

TABLE 2  
SUMMARY OF MORPHOSYNTACTIC FEATURES FROM BLASCO FERRER  
("LA POSIZIONE")

Items	Languages with the features
8	Catalan, Provençal
8	Catalan, Provençal, Gascon
2	Catalan, Gascon
18	"Occitano-Romance"
1	Catalan
1	Catalan, Spanish
1	Portuguese

This summary, however, is somewhat misleading. Consider trait 13, second conjugation participles in -UTU, which is counted as one of the traits aligning Catalan with G-R. He notes that "past participles in -UTU are found in all the western Romance languages during the medieval period" and concludes, "The Gallo-Romance character of Catalan is traceable to the innovative evolution of Spanish and Portuguese, which abandoned participles in -UTU from the end of the 15<sup>th</sup> century on" ("La posizione" 161). This reasoning falls short, however, in two respects. First, the late date of the developments in Spanish and Portuguese means that the change from participles in -UTU to participles in -ITU is not an I-R phenomenon since it happened in already differentiated I-R languages, not in proto-I-R. Thus it shows independent parallel development or contact. More importantly, this is another instance of shared retention and thus says little about the position of Catalan.

Blasco Ferrer ("La posizione" 172) concludes that "the typology of the Catalan language stands in sharp contrast with the structures of Portuguese and Spanish; it shows, instead, a tight connection with the Trans-Pyrenean languages, especially with the Occitano-Romance group (the modern Occitan dialects and Gascon)." The analysis shows greater emphasis on typological patterns (but not the classic kinds of alignment questions commonly covered in typological studies, most of which would in any case reveal few differences among these languages) than on traditional genetic concerns. While I do not deny the value of this kind of typological evaluation, I do, in light of the malleability of such features in the face of language contact, argue that separating such factors out stands to provide greater insight into genetic relations.

Certainly the combination of shared retentions and parallel developments creates the appearance that Catalan is closely related to G-R and especially to Occitan. Unless we accept, however, some inherent quality of a language that persists after the breakup of a proto-language into daughter languages, these features should be viewed not as evidence for a relationship between Catalan and these other languages but rather as a possible impediment to establishing exactly what kind of relationship the languages actually have — hence the aforementioned importance of clearly distinguishing such factors from genetic considerations.

#### PHONOLOGY

One phonological change attested in these languages presents a bit of a puzzle. In the history of Catalan, the -ND- > -n- is regular (e.g., VENDEMUS > *venem* 'we sell,' MANDARE > *manar* 'to rule'). Generally

the corresponding Occitan forms (*vendèm, mandar*) do not show this change, although Gascon does (Wheeler 249). However, in the basic verb of motion, the change appears to be reflected in most if not all Occitan dialects (\*ANDARE > *anar* 'to go,' \*ANDAMUS > *anam* 'we go'). Given the tendency for highly frequent lexemes like those meaning GO to resist analogical pressure (cf. Bybee, among others), the diffusion of the form without the cluster to the other dialects is unexpected. This change initially creates the appearance of a closer relationship between Catalan and Occitan than closer analysis shows. The issue of the spread of the forms with *-n-*, like *anar*, remains for future research.

Verbal morphophonology provides strong evidence of a Catalan-Occitan subgrouping. In these two languages certain verb forms display velar consonants that come from sequences of labials and velars, along with various analogical extensions. Naturally there is variation in these patterns not only between the two languages but also among their respective dialects. Pérez Saldanya (*Del llatí*) devotes an entire chapter to the history of such forms in Catalan. The preterit most clearly shows the development of these velar consonants, as shown in Table 3 with reflexes of HABERE 'have' (Old Catalan and Old Occitan are used here to highlight the antiquity of the pattern; note that <gu>, <c> represent /g/, /k/, respectively). The fact that other Romance languages do not reflect these changes strongly favors Catalan/Occitan unity.

TABLE 3  
VELAR CONSONANTS IN CATALAN AND OCCITAN  
IN THE PRETERIT OF HABERE

Infinitive	Old Catalan haver	Old Occitan aver	Spanish haber	French avoir
1S	hac	aic, aguí	hube	eus
2S	haguist	aguist	hubiste	eus
3S	hac	ac	hubo	eut
1P	haguem	aguem	hubimos	eûmes
2P	haguets	aguetz	hubisteis	eûtes
3P	hagren	agron	hubieron	eurent

#### CATALAN AS A BRIDGE LANGUAGE

The notion of Catalan as a "bridge language" between Ibero- and Gallo-Romance, mentioned above, is a recurring theme in Catalan studies. Among those endorsing the notion are Baldinger, Blasco Ferrer, Badia

Margarit, Francescato (see below, "Quantitative methods"), and Balacciu-Matei. Balacciu-Matei analyzes 1070 Catalan words to find lexemes shared with French, Occitan, Spanish, and Portuguese and establishes twelve sets of correspondences within Romance. She finds that Catalan shares the same number of lexemes with I-R as with G-R. Balacciu-Matei concludes that *Gallo-Romance* and *Ibero-Romance* remain useful terms, with the term *bridge language* serving not only for Catalan but also for Occitan. These are, she says, part of Ibero-Romania and Gallo-Romania, respectively, but each has at the same time close relations with the other group. This finding fits with what we might expect both in terms of linguistic family trees and language contact. Both perspectives are consistent with —and may in fact be considered to predict— such an outcome. Depending on how a tree branches, we should not be at all surprised to find languages that appear to bridge subgroups. Certainly in terms of contact, we can expect gradient features in languages located between others —that is, continuum effects at the level of languages rather than dialects, as discussed above with respect to Hock's analysis of convergence areas.

#### QUANTITATIVE METHODS

Francescato proposes a classification depending in part on a quantitative, rather than a qualitative, analysis of features in the Romance languages. Using 44 unidentified features, he produces figures indicating how close given pairs of Romance varieties are. His results yield numbers ranging from a low of 2 for the close relationship between the R-R varieties of Fassa and Cadore to a high of 61 for more distant pairings like French and Romanian or French and Sardinian. In this context it is no surprise that Spanish and Portuguese have a low score, 9, indicating a close affinity. While Catalan and Provençal yield the relatively low score of 21 (equal to Provençal-Francoprovençal; note that he does not use the name *Occitan*), Spanish and Catalan produce the surprisingly low score of 15 (Francescato 80). He concludes that Catalan belongs to the "bridge" varieties along with Vegliote, Provençal, and Engadine (82).

#### LEXICOSTATISTICS AND GLOTTOCHRONOLOGY

Other scholars have tried to use glottochronology and lexicostatistics to subgroup the Romance languages. This approach has a long history of criticism, however. For example, Rea (who cites some key works in

this area) raises several objections. One — the possibility of establishing a universal word list suitable for comparing a wide variety of languages — primarily addresses cultural concerns that do not materially affect Romance. Noting that Latin and Romance were used in establishing the technique but that it was never checked against all that is known about Romance, Rea (147) writes that “the standard mathematical ‘proof’ consisting of running one’s end product back through the original equation has never been undertaken.”

These techniques mainly aim to establish dates of separation of varieties and degree of relatedness as indicated by vocabulary patterns. (Rea notes that the time span between two samples of a given language may also be covered, but this really falls under the first goal and is not directly relevant here.) Even if the absolute dates of separation provided by these techniques do not coincide with what we know from other sources, they may still shed light on the relationships involved.

In fact, Kroeber, responding to Rea’s criticism of the technique, argues that, although the dates that Rea achieved in his critical examination differ from known data by a factor of more than two, the data reinforce what we know about the internal structure of Romance. In defense of such apparent repetition, he argues, “It may be objected that such a finding is unnecessary because something like it has long since been believed by comparative students of Romance. The basic answer to such an objection is that a conclusion is more certain if it is reached by two methods than by one” (455). He does concede, however, that “the greatest danger in lexicostatistics is its involvement with chronology, which is its weakest point, though also the one of widest appeal” (455). He claims, though, that for well established families whose “internal classification has too often been made haphazardly, impressionistically, without recourse to full knowledge ...or...left to nonlinguists to discuss” like Romance, Uto-Aztecan, Athabaskan, and Salish, “certain fairly reliable conclusions can be reached...[about] internal organization” (455-56).

Meanwhile, additional concerns arise with this technique. Percentages of cognate vocabulary do not in themselves distinguish shared innovations, which reliably indicate shared history, from shared retentions, which provide no information about subgrouping.

Furthermore, languages differ in many ways, not just lexically. Consider, for example, the relationship between French and Catalan on the one hand and French and Sardinian on the other. Scholars rarely if ever place French as close to Sardinian as to Catalan, but by Rea’s lexicostatistical analysis, both pairs show 80% cognacy in the 100-word Swadesh list. Nonetheless, when used judiciously, this method has the potential to provide some insight when other techniques come up short.

In the case of Catalan, Occitan, and Spanish, each pair shows 80% cognacy on the 100-word list, which runs counter not only to the common grouping of Occitan with G-R and Catalan with I-R but also to the view that Catalan and Occitan belong together at some distance from both of the other groups. How effectively we can measure the distance separating these three groups will depend on other criteria.

### CONTACT FEATURES IN ALGUERÈS

Kuen's unfinished treatment of Alguerès presents morphological analysis, detailed phonetic and phonological information, and detailed historical analysis and provides a great deal of information about the relationship of Alguerès to Catalan in general. Caria and Bosch are also key works for Alguerès.

In separating genetic and contact-related factors, the presence of a Catalan dialect in Sardinia offers a perspective not available from continental and Balearic varieties. As advances in reconstruction techniques make recovering older syntactic patterns more practicable (Harris and Campbell), comparison of dialect data becomes increasingly important. Such comparison will facilitate identifying and weighing differing historical factors. In some cases, it will be relatively easy to pick out structures resulting from contact. As shown below, for example, in Alguerès, the use of subjunctive in nominal clauses after verbs of opinion indicates Italian influence (though this is a minority pattern).

<i>Standard Catalan:</i>	Penso	que	és	important
	think-PRS.IND.IS	COMP	be.3S.PRS.IND	important
<i>Alguerès:</i>	Pens	que	sigui	important
	think-PRS.IND.IS	COMP	be.3S.PRS.SBJ	important
<i>Italian:</i>	Credo	che	sia	importante
	think-PRS.IND.IS	COMP	be.3S.PRS.SBJ	important
	'I think that it is important'			

While this particular example presents little difficulty and requires little analytical sophistication, future research may reveal ways in which data from Alguerès can complement other Catalan dialect data to more accurately place the language in the family tree.

Blasco Ferrer (*Grammatica*) analyzes phonological, morpho-syntactic, and lexical "interference" in Alguerès and typologizes cases of influence on the dialect from both Italian and Sardinian. Use of Alguerès data requires attention to a question Blasco Ferrer (*Grammatica*) raises, namely, whether it is innovative or conservative.

Analogical developments like *sés* (cf. *ets* 'you are') and archaizing traits like the /a/ in the 2<sup>nd</sup> plural (*anau* vs. *aneu*) combine to illustrate that this is not a yes-or-no issue. Careful comparison allows for beneficial use of dialect data in evaluating the status of the language as a whole.

## CONCLUSIONS

The exploration of Catalan's place in Romance raises important questions in historical linguistics, contact linguistics, and sociolinguistics, not the least of which concerns the importance of balanced analysis. Taking advantage of well-attested languages furthers our understanding of the principles and processes involved in their histories and facilitates identifying comparable changes in less well-attested languages, ultimately improving our accounts of change in all languages. Hock (510) recognizes the importance of this approach, stating that we can apply discoveries from studies of languages with ample attestation and to studies of those with less documentation.

While we need not belabor minor points, it is important to explore those issues whose direct analysis is possible only in well-attested language families. For example, certain Swadesh items may be learned words in a given group of languages, as with *persona* 'person' and its variants. In such a case the forms may not appear to be borrowings but nonetheless inflate the apparent closeness of the languages in question. Considering such possibilities increases the accuracy of our analyses of less-attested languages.

I have attempted to show how potentially conflicting methodologies can be used jointly to better understand relations within Romance. The evidence from traditional genetic analysis, including phonological, morphosyntactic, and lexical examination, along with insights gained from contact linguistics, shows a certain degree of ambiguity both on the general issue of Catalan's place in Romance as well as the specific topic of whether Catalan and Occitan belong together rather than in I-R and G-R, respectively, as they are commonly categorized. Perhaps this should be no surprise in light of the view that "good subgrouping evidence is very hard to find" (Harrison 238).

I have emphasized the importance of the timing of the development of features like the Catalan periphrastic preterit as well as the value of distinguishing among areal features, shared retentions, and shared innovations, the last of which provides the strongest evidence of relatedness.

The strongest sign of Catalan/Occitan unity is the development of velar consonants in the verb system. The totality of the evidence seems

to slightly favor a special affinity between Catalan and Occitan apart from G-R and I-R. Many of the similarities between Catalan and one or another subset of the Romance languages can be explained by independent parallel development, the spread of areal features, or shared retentions. Exactly how these three groups relate to each other remains a matter for further research.

In considering contact issues, I have suggested that these questions may benefit especially from analysis of the dialect of Catalan spoken in Alguer, Sardinia, where the language has not been in contact with its closest relatives, thus allowing greater separation of contact and genetic factors in the analysis.

The equivocal nature of the available evidence calls for closer analysis of the sociolinguistic circumstances of the times from which the earliest Catalan and Occitan texts date in order to explore the possibility of reconstructing and extrapolating the relations among the relevant speech communities. In this way, perhaps it will be possible to develop greater insight into the forces affecting these languages not only to improve our understanding of how such relations affect dialects spoken in proximity but also to increase the accuracy with which we classify the Romance languages.

MATTHEW L. JUGE

TEXAS STATE UNIVERSITY-SAN MARCOS

#### WORKS CITED

- BADIA I MARGARIT, Antoni M. *Gramàtica històrica catalana*. Barcelona: Noguer, 1951. Trans. cat. Barcelona: Tres i Quatre, 1981.
- . "El catalán: Evolución lingüística interna I. Gramática." *Lexikon der Romanistischen Linguistik*. Ed. Günter Holtus, Michael Mietzeltin, and Christian Schmitt. Vol. 2. Tübingen: Niemeyer, 1991. 127-52.
- BALACCIU-MATEI, J. "Poziția catalanei și a occitanei din perspectiva vocabularului de bază al latinei." *Studii și cercetări lingvistice* 43 (1992): 363-67.
- BALDI, Philip. "Indo-European Languages." *The World's Major Languages*. Ed. Bernard Comrie. Oxford: Oxford UP, 1990. 31-67.
- BALDINGER, Kurt. *La formación de los dominios lingüísticos en la Península Ibérica*. Madrid: Gredos, 1972.
- BEC, Pèire. *Manuel pratique de philologie romane*. Vol. 1. Paris: A & J Picard, 1970.
- BLASCO FERRER, Eduard. *Grammatica storica del catalano e dei suoi*

- dialetti con speciale riguardo all'algherese*. Tübingen: Gunter Narr, 1984.
- . "La posizione linguistica del catalano nella Romania." *Zeitschrift für romanische Philologie* 102 (1986): 132-78.
- BOSCH, Andreu. *El català de l'Alguer*. Barcelona: Abadia de Montserrat, 2002.
- BYBEE, Joan L. *Morphology: A Study of the Meaning of the Relation between Meaning and Form*. Amsterdam: Benjamins, 1985.
- CARIA, Rafael. "L'alguerès des d'una perspectiva històrica." *Revista de l'Alguer* 1 (1990): 33-53.
- DE DARDEL, Robert. *A la recherche du protoroman*. Tübingen: Niemeyer, 1996.
- ELCOCK, W. D. *The Romance Languages*. Rev. John N. Green. London: Faber, 1975.
- FRANCESCATO, Giuseppe. "A Proposal for the Classification of Romance Languages." *Italic and Romance: Linguistic Studies in Honor of Ernst Pulgram*. Ed. Herbert J. Izzo. Amsterdam: Benjamins, 1980. 75-84.
- GARRETT, Andrew. "A New Model of Indo-European Subgrouping and Dispersal." *Proceedings of the Twenty-Fifth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 1999. 146-56.
- HALL, Robert A., Jr. "The Reconstruction of Proto-Romance." *Language* 26 (1950): 6-27.
- HARRIS, Alice C., and Lyle CAMPBELL. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge UP, 1995.
- HARRIS, Martin. "The Romance Languages." *The Romance Languages*. Ed. Martin Harris and Nigel Vincent. London: Croom Helm, 1988. Rpt. London: Routledge, 1997. 1-25.
- HARRISON, S. P. "On the Limits of the Comparative Method." *The Handbook of Historical Linguistics*. Ed. Brian D. Joseph and Richard D. Janda. Malden: Blackwell, 2003. 213-43.
- HOCK, Hans Henrich. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter, 1991.
- JUGE, Matthew L. "Morphological factors in the grammaticalization of the Catalan 'go' past." *Diachronica* 23 (2006): 313-39.
- . "Narrative and the Catalan GO-past." *Folia Linguistica Historica* (forthcoming).
- KROEBER, A. L. "Romance History and Glottochronology." *Language* 34 (1958): 454-57.
- KUEN, Heinrich. "El dialecto de Alguer y su posición en la historia de la lengua catalana." *Anuari de l'Oficina Romànica* 5 (1932): 121-77; 7 (1934): 41-112.

- MAIDEN, Martin. *A Linguistic History of Italian*. London: Longman, 1995.
- MALKIEL, Yakov. "The Classification of the Romance Languages." *Romance Philology* 31 (1978): 467-500.
- MENÉNDEZ-PIDAL, Ramón. *Orígenes del español*. 8<sup>th</sup> ed. Madrid: Espasa, 1976.
- MEYER-LÜBKE, Wilhelm. *Grammaire des langues romanes*. New York: Stechert, 1923.
- . *Das Katalanische*. Heidelberg: Winter, 1925.
- NADAL, Josep, and Modest PRATS. *Història de la llengua catalana*. 4<sup>th</sup> ed. Barcelona: Ed. 62, 1993.
- PADEN, William D. *An Introduction to Old Occitan*. New York: MLA, 1998.
- PÉREZ SALDANYA, Manuel. "Gramaticalització i reanàlisi: el cas del perfet perifràstic en català." *Actes del desè Col·loqui Internacional de Llengua i Literatura Catalanes*. Ed. Axel Schönberger and Tilbert D. Stegmann. Barcelona: Abadia de Montserrat, 1996. 3: 71-107.
- . *Del llatí al català: morfosintaxi verbal històrica*. València: U de València, 1998.
- POSNER, Rebecca. *The Romance Languages*. Cambridge: Cambridge UP, 1996.
- REA, J. A. "Concerning the Validity of Lexicostatistics." *International Journal of American Linguistics* 24 (1958): 145-50.
- ROHLFS, Gerhard. *Le gascon: études de philologie pyrénéenne (avec 2 cartes)*. Halle, Saale: Niemeyer, 1935.
- RONJAT, Jules. *Grammaire historique des parlers provençaux modernes*. Vol. 3. Montpellier: Société des Langues Romanes, 1937.
- SCALA, Luca, ed. *Català de l'Alguer: criteris de llengua escrita*. Barcelona: Abadia de Montserrat, 2003.
- VENY, Joan. *Els parlars catalans*. Mallorca: Moll, 1993.
- WALSH, Thomas J. "Origin of the Catalan Dialects: Pre-Roman Substrates or Medieval History?" *Hispanic Review* 54 (1986): 405-25.
- WARTBURG, Walther von. *La fragmentation linguistique de la Romania*. Paris: Klincksieck, 1967.
- WHEELER, Max W. "Occitan." *The Romance Languages*. Ed. Martin Harris and Nigel Vincent. London: Croom Helm, 1988. Rpt. London: Routledge, 1997. 246-78.