HEARSAY AND REPORTED SPEECH:
EVIDENTIALITY IN ROMANCE

Silvio Cruschina, Eva-Maria Remberger

(L. Sciascia, Occhio di Capra, 1984: 53)

1. Introduction

The grammatical category of evidentiality is crosslinguistically realized by different strategies. The use of forms derived from *verba dicendi* constitutes one possible device. For instance, evidentiality can be marked by functional elements like *dizque* in some American Spanish and Portuguese varieties (cf. Aikhenvald 2004). In this paper, we show that this strategy is not restricted to these languages, but is also widespread in European Romance varieties such as Galician (Gal.), Romanian (Ro.), Sardinian (Srd.), and Sicilian ( Sic.). In all varieties these functional elements derive from the verb ‘say’ + complementiser (SAYC), with a meaning similar to English ‘apparently’ or ‘allegedly’:

(1) Sp.   Esto *dizque* va a ser pantano.   (Kany 1944:172)
          this SAYC go.PRES.3S to be.INF swamp
          ‘This is going to be swamp, they say.’

(2) Sic.  *Dicica* ci avivanu finutu i grana.  (AdS)
          SAYC to-him.CL have.IMPF.3P finish.PP the money
          ‘Apparently they had run out of money.’

(3) Srd.  Custas columbas *nachi* s’abbaidana e an cominzadu a faeddare.  (AdS)
          these pigeons SAYC REFL look.PRES.3P and have.PRES.3P start.PP to talk.INF
          ‘These pigeons look at each other and started talking, they say.’
(4) Ro. Cică banul n-adece fericirea.  
SYAC the money not bring.PRES.3S the happiness
‘Money doesn’t give you happiness, they say.’

(5) Gal Disque a filla da Antonia marchou á Coruña vivir co mozo.
SYAC the daughter of-the A. go.PAST.3S to C. live.INF with-the boyfriend.
‘Apparently Antonia’s daughter went to C. to live with her boyfriend.’

In reference to Colombian Spanish, Travis (2006:1276-1277) shows that the SAYC element, even if “not fully grammaticalized”,1 has undergone changes typical of grammaticalisation processes (decategorialisation, inseparability, phonological reduction, semantic generalisation). If we apply the tests given in Travis and several others to the data of the varieties under investigation here, the grammatical status of the SAYC element becomes clear and incontrovertible. Nevertheless, we also must assume different stages in the grammaticalisation process in order to account for the grammatical and functional differences found across the languages in question, in which the roles of the SAYC elements vary from a reported speech marker to an evidential marker proper. Following a cartographic approach to syntactic structures, we show that these markers occupy the functional projection dedicated to evidentiality, namely, the Evidential Mood projection, as identified by Cinque (1999). Moreover, our analysis will provide evidence for a clear-cut distinction between epistemicity and evidentiality, a distinction that is also represented in Cinque’s work in terms of two distinct and independent designated functional projections.

1.1 Types of evidentiality

In recent years, the phenomenon of linguistic evidentiality has aroused increasing interest among researchers in the field of semantics and pragmatics (cf. Dendale et al. 2001, Aikhenvald 2004). The definition of evidentiality given by Aikhenvald is “a linguistic category whose primary meaning is source of information” (2004:3). The name of this relatively new linguistic category comes from the “kind of evidence a person has for making factual claims” (Anderson 1982:273), which is expressed by grammatical means.

There has been a long debate on the relationship between evidentiality and modality, especially epistemic modality. Whereas recent approaches, such as that of Aikhenvald (2004) given above, strictly distinguish between the two and therefore recognize evidentiality as a linguistic category proper, classical works on modality, e.g. Palmer (2001), consider evidential marking to be simply a particular instance of

---

1 We use the term grammaticalisation in a broad sense, as the general process consisting of those phenomena involving change and reanalysis of linguistic forms.
Hearsay and reported speech: evidentiality in Romance

evidentiality: Evidentiality in the broad sense (e.g. Palmer 2001) where evidentiality is part of the system of epistemic modality (or vice versa), since both characterize the speaker’s “attitudes about the epistemic status of information” (Dendale et al. 2001:340) towards the proposition encoded in an utterance; Evidentiality in a narrow sense (e.g. Anderson 1982, Willett 1988, Aikhenvald 2004) where the focus lies on the expression of the “information or sources of knowledge behind assertions” (Dendale et al. 2001:340); if evidential marking results in an interpretation that reveals the speakers’ (e.g. positive or negative) attitudes towards the evidentially-marked proposition, this is purely a consequence of pragmatics and has nothing to do with grammatical epistemicity; Overlapping of evidentiality and epistemicity: this view is adopted particularly for the subfield of inferential evidentiality, which might be easily interpreted as a type of epistemic modality (cf. van der Auwera & Plungian 1998:86 following Dendale et al. 2001:242, Plungian 2001:354). Although we admit that evidentiality and other grammatical phenomena are interacting, in this paper we assume the first position and treat evidentiality as a grammatical category standing on its own.

Given evidentiality as a grammatical category, several subsystems have been proposed. The most elaborate can be probably found in Aikhenvald (2004). Yet, since we concentrate on a subfield of evidentiality, i.e. hearsay and reported speech, for our aims the system delineated by Willett (1988) will suffice. Willett distinguishes between three types of evidentiality, namely direct, indirect and inferential evidentiality (Willett 1988:57). Direct evidentiality is given when the speaker has (and explicitly marks) first-hand sensory evidence for the assertion he makes. Indirect evidentiality covers second- and third-hand evidence and is equivalent to reportative evidence. Inferential evidence is given whenever the speaker marks his assertion as having been deduced. This paper, of course, is concerned with indirect, i.e. reportative evidence. For reportative evidentiality, there is a further distinction, to which we will refer throughout this paper:

6. Types of reportative evidence (cf. Willett 1988:96)

a. second-hand evidence (hearsay): “the speaker claims to have heard of the situation described from someone who was a direct witness”, ‘he says’

b. third-hand evidence (hearsay): the speaker claims to have heard about the situation described, but not from a direct witness”, ‘I heard’, ‘I hear tell’, ‘it is said’, ‘they say’ (impersonal)

c. folklore: the speaker claims that the situation described is part of established oral history (fairy tales, mythology, oral literature, proverbs and sayings)
All types of reportative evidence are found in the evidentials appearing in the Romance data at issue in this paper, as will be shown below.  

1.2 Evidentiality in Romance

Romance languages lack a dedicated paradigm or system of grammatical evidentials. This does not mean that there are no ways to express evidentiality in Romance. The Romance languages express evidentiality through various modal and temporal forms, following a pattern that is typologically characteristic of the languages of Western Europe, where the development of evidentials from modal morphemes is a typical feature (cf. da Haan 2005). Evidentiality as an actual grammatical phenomenon within the Romance verbal system has been recognized and argued for in recent work (cf. Squartini 2001, 2004, 2005, Calaresu 2004).

The conditional, the future, the imperfect, and periphrases with modal verbs (e.g. dovere/devoir/deber + infinitive), are traditionally assumed to take on evidentiality meanings in specific contexts. Although it is undoubtedly true that these verb forms convey evidentiality meanings, their primary role is not to mark the source of information. Evidentiality is just one of the modal and aspectual meanings they can express. Following Aikhenvald (2004), we call these forms “evidential strategies”, as opposed to evidentiality and evidentials proper. The conditional, for instance, can be used with a quotative meaning in French:

(7) Fr. Il y aurait de nombreuses victimes. (Dendale and Tasmowski 2001:345)
   EXPL there have.COND.3S of numerous victims
   ‘It is said/it seems that there are many victims’

However, “[t]his does not mean that this conditional has ‘become’ an evidential” (Aikhenvald 2004:105). The conditional as well as the other Romance evidential strategies resorting to tense and modal specifications of the verbal system typically convey epistemic values, in particular uncertainty and non-responsibility for the information expressed. The interpretation associated with the non-firsthand source of information, therefore, must be considered a “semantic extension” of these forms which frequently overlaps and coexists with a range of modal notions including probability, possibility, factuality, subjectivity, and responsibility. Indeed the

---

2 There is a further confusion concerning a more specialised meaning for “reported” vs. “quotative” evidentiality. Aikhenvald (2004:177ff) e.g. states that “if a language has two reported type evidentials, the most common distinction is that between reported (stating what someone else has said without specifying the exact authorship) and quotative (introducing the exact author of the quoted report).” Thus, “reported” in this sense just means third-hand or folklore, i.e. (6)b and (6)c above, whereas “quotative” is synonym of second-hand evidence, i.e. (6)a.
boundaries between evidentiality and epistemic modality in these strategies are often blurred and dependent on the context.

In Romance, in addition to these strategies realized through the verbal system, various lexical strategies, such as adverbs (Fr. apparemment, visiblement), ‘parenthetic’ expressions (Fr. il paraît, il semble), and verbs of saying (Fr. ils disent que), are used to express evidentiality or evidential-like meanings. These seem to be universal strategies. The impersonal use of verbs of saying, i.e. a matrix clause containing the verb of saying and followed by a complement clause introduced by the complementiser (Fr. on dit que, It. si dice che, Sp. se dice que), is also a very common and widespread means of marking third-hand reported speech.

As pointed out in Willett (1988:79) and Aikhenvald (2004:271ff), the development of evidential strategies into evidential markers is not rare, and the grammaticalisation of verbs of saying + complementiser represent a typical case, as confirmed by our data. The phenomenon has been observed in reference to some varieties of Spanish, particularly Old Spanish (Kany 1994)3 and American Spanish (cf. Travis 2006). We show that it is more wide-spread than it has been assumed so far, since it is found in several modern European Romance varieties, such as Galician, Romanian, Sardinian, and Sicilian. SAYC results from the reanalysis of a biclausal construction comprising a verb of saying with a complement clause. This construction is reanalysed as monoclausal and the matrix verb together with its complementiser is reinterpreted as a single and functional unit, that is a grammatical marker stripped of its lexical features (cf. Aikhenvald 2004:272). The primary meaning of the SAYC element develops into one directly connected to the marking of the source of information. The SAYC elements therefore represent fully-fledged grammatical evidentials, although the path towards grammaticalisation seems not to be complete in all the varieties under investigation here.

2. Tests of Grammaticalisation

In the Romance varieties at issue here, the evidential marker SAYC stems from a verbal form of the verb ‘to say’ plus the morphophonologically incorporated complementiser ‘that’. The original verbal form is normally the third person

3 Indeed, in the database CORDE (cf. Real Academia Española, CORDE), we find several examples of dizque, especially in texts from the end of the fifteenth century, but with the first example dating back to 1293, cf. (i):

(i) E tal postura dizque auien que luego que nascie y el ninno que luegol ponien en quitaçión. (Gran Conquista de Ultramar, anonymous)

And such position SAYC have.PAST.3P that after that be-born.PAST.3S the child that then-it.CL.put.PAST.3P in.PAST.3P in dismissal

‘And they were said to have such a position that, after the child was born, they were dismissed.’
singular, as in Sicilian, Spanish and Romanian. However, in Galician, it is an impersonal version of the third person singular, namely that constructed with the originally reflexive pronoun se. In Sardinian, there is diatopic variation as far as SAVC is concerned and the corresponding verbal form might stem either from a third person singular or from a third person plural. In any case, as has also been observed by Travis (2007:1276), in all varieties at issue the SAVC element is at an advanced, even if perhaps not completed, stage of a grammaticalisation process. That means that on all systematic levels we find the typical phenomena commonly used as diagnostics (cf. e.g. Heine 1993), i.e. phonological erosion (cf. 2.1), morphological decategorisation (cf. 2.2.), change in syntactic distribution and properties (cf. 2.3), as well as desemanticisation (cf. 2.4). In the Romance varieties under discussion, these phenomena manifest themselves to different degrees (cf. the table in (36)).

2.1 Phonology

In all varieties under examination here the SAVC elements show phonological reduction (phonological erosion) or, as in Sicilian and Sardinian, some other phonological peculiarities. In Colombian Spanish, the form dizque was originally dice que, i.e. ‘he/she says that’; the final vowel of the verbal form was apocopated and the complementiser fused with the verbal form to result in the SAVC-element. Following Travis (2007:1276) and Kany (1944) dizque can be further reduced to ‘izque or also ihke and ike (e.g. in Venezuela, cf. Kany 1944:169). In Sicilian, the high vowels /i/ are different in quality in dicica, where they are [+ATR], from the full verb use of the same verb, where they are [-ATR]. This evidence for dicica as a single unit comes from the dialect of Mussomeli (CL):

(8) a. dicica [ˈðɪtʃɪ-κɒ]  
   b. dici ca ['dɪtʃɪ ˌkɒ]

Furthermore, the complementiser ca does not bear a secondary stress in the simplex grammaticalised form, as it typically does when part of the complex form.

We find several forms of SAVC in the Sardinian varieties as well, namely the forms nachi, nanchi (anchi), naca, nanca (anca), etc. These forms come from the third person plural and the third person singular of narrere ‘to say’ plus the

4 The different nature of the high vowels results from a vowel harmony process involving the feature [+ATR] that spreads regressively from the end of the word. This process is triggered by the presence of a high vowel at the end of the word: the final high vowels and all the preceding high vowels become [+ATR]. The harmony process does not take place with word-final a. High vowels within the word therefore stay [+ATR] (cf. Cruschina 2007).

5 Also the use of the compound perfect at nadu is quite frequent e.g. in fairy tales; however, there is no third-hand, but only second-hand use of this form, as far as we know. Thus, we will not treat this form, which also would contradict our claims made in 2.2, as part of the evidential SAVC-system of Sardinian.
complementiser *chi* (< Lat. *quid*) and *ca* (< Lat. *quia*) respectively, i.e. *narrant* *chi/ca* and *narat* *chi/ca* (cf. Wagner 1951:397-398; Pittau 1972:90; Puddu 2000: s.v. *nachi*, *anca*; Blasco Ferrer 1984:258; Jones 1993:126-127). The verbal form has been contracted (*narat* > *nat*, *narrant* > *nant*), phonologically reduced (the *t* ending has been elided) and fused with the complementiser to a single form. Sometimes we find two of the above-mentioned variants, one with the second *n* and one without, in one and the same speaker (AdS; Irgoli):

(9)  
Nanchi it una emina chi nachi teniata trer fizos…

(AdS)

'St told that there was a woman who allegedly had three sons…'

As a consequence of this phonological reduction and fusion no paragogical vowel, i.e. a word final copy of the vowel in the immediately preceding syllable (cf. e.g. Bolognesi 1999), can be inserted between the verbal form and the complementiser (e.g. *nangchi*), in contrast to the use of the verb ‘say’ as a full verb, where the paragogical vowel appears in phrase final or clause final position:

(10)  
Náranta gi bénidi zu zíndiku

(Blasco Ferrer 1984:258)

'They say that the mayor is going to come.'

In Galician, the full verbal form underlying *disque* is not the third person singular *di* plus *que* alone, but must be *dise que*, that is the third person singular form plus an enclitic impersonal pronoun (*se*) ‘one says that’. The clitic in the SAYC use shows final vowel elision.

In Romanian, the SAYC-element *cică* is a contraction of (*se zi) ce că* ‘one says that’ (cf. DEX: s.v. *cică*) or *zice că* ‘he/she says that’ (Tiktin: s.v. *cică*). Here too, we find other regional forms and variants like *şi că* (< *a zis că* / *o zis că*, ‘he/she has said that’ in the compound perfect) (Moţei n.y.). Our data, however, will concentrate on the form *cică* alone.

2.2 Morphology

Travis points out two major morphological characteristics of *disque*, which differ from the full verb. First, only the full verb *decir* can be separated from the complementiser *que*, whereas *disque* functions as a single and indivisible unit.

---

6 However, these short forms of *narrere* like *nat* ‘he says’ can also be used in poetic language, cf. Puddu (2000: s.v. *narrere*).

7 Note that in some contexts (e.g. subordinate clauses), in Galician, the enclitic position of the impersonal pronoun would not be possible. Nevertheless *disque* stays invariant in these contexts, a fact that confirms that it is an amalgamate form.
Second, *dizque* is always morphologically invariable and in no case can it be inflected:

(11) Sp. *dizque* → *diceque* (PRES), *decíaque* (IMPF), *dijoque* (PAST) …

The reduced form *dizque* in Colombian Spanish does not encode any morphological information concerning person, number, tense and mood. Any attempt to assign a morphological specification to the crystallised form *dizque* results in ungrammaticality, even in a context with a past reference-time. Equally, the phonologically reduced forms of SAYC in the other languages also prove to be morphologically fossilised, inseparable and not compatible with tense and mood inflection. For example, Sicilian *dicica* can replace the full-length verbal form, irrespective of its temporal and modal specification, but it always conveys an evidential value and has to appear in its invariable form:

(12) *Dicivanu ca / Dicica l *Dicivaca a mugghieri ci fa / faciva i corna c’u nutaru.
    say.IMPF.3P that/SAYC/SAYC+IMPF the wife to-him.CL do.PRES.3S/do.IMPF.3S
    the horns with the notary
    ‘They said / it is said that his wife is / was cheating on him with the notary.’

Sicilian SAYC, thus, cannot encode any temporal or modal specification and is incompatible with any type of inflection (cf. Menza in press). The same is true of Sardinian *nachi*, Galician *disque* and Romanian *cică*:

(13) a. Sic. *dicica* → *dicivaca* (IMPF), *dissica* (PAST), *dicissica* (SUBJ) …
b. Srd. *nachi* → *naratchi* (PRES), *naraitchi* (IMPF), *naduchi* (PP) …
c. Gal. *disque* → *diseque* (PRES), *digaseque* (SUBJ), *dicíaseque* (PAST)
d. Ro. *cică* → *zicecă* (PRES), *zicăcă* (SUBJ), *ziceacă* (IMPF), *ziscă* (PP) …

The inseparability of SAYC, which behaves as a single unit not divisible into the two original elements, and the morphological invariance, resulting from the impossibility of inflecting for any verbal feature, are the main morphological properties offering primary evidence for the grammatical status of SAYC.

2.3 Syntax

We have seen that SAYC has been reanalysed as a single unit. Not only has the verb of saying lost all its morphological and syntactic traits, but furthermore the complementiser no longer functions as such and is merely part of the unit SAYC. The syntactic consequences of this reanalysis are the following:

(14) a. SAYC can be used in isolation, for instance in an answer to a question.
b. It is incompatible with negation and other lower adverbs.
c. It can occur with other complementiser(-like) elements.

The main function of a complementiser is to introduce an embedded clause. Complementisers are therefore incompatible with the final position of the sentence
Hearsay and reported speech: evidentiality in Romance

(at least in SVO languages). Due to its reanalysis and grammaticalisation, SAYC is not subject to this restriction. It can appear at the end of a sentence when used as parenthetical (cf. also 4.1) and it can also appear in isolation, especially in an answer to a question:

(15) Sic. Chi jè veru ca Maria av’a partiri pi l’America? – *Dicica! beads.PRES.3S that Maria have.PRES.3S to leave.INF for the America SAYC
   ‘Is it true that Mary is going to leave for America? – Apparently!’

   and SAYC wife your REFLECTIVE.PRES.3S liberate.P.PPP SAYC
   ‘Your wife is said to have given birth? – ‘They say so!’

In Romanian, cică can be used in isolation, e.g. as an answer to a question, but it is normally reinforced by ‘yes’ or ‘no’:

   be.PRES.3S true that Ion leave.PRES.3S to New York SAYC yes SAYC no
   ‘Is it true that John goes to New York? – Apparently yes / no.’

Verbs can typically be negated and modified by lower verbal adverbs (e.g. adverbs of time, manner, etc.). By contrast, the grammatical nature of SAYC leads to its incompatibility with negation (e.g. Sic. *un dicica) as well as with any modifier (such as ‘always’, cf. Sic. *dicica sempri, or ‘often’, cf. Sic. *dicica spissu), unless they modify the verb of the clause. Unavailability within the scope of negation is a typical property of evidential markers across languages (cf. Willett 1988, de Haan 1999), although it is not a universal property (cf. Aikhenvald 2004:96-97). Thus, negation can take scope over the speaker’s statement, but not over the source of information. The same happens with adverbs which cannot modify the evidential marker SAYC, but can, on the other hand, modify the stated information (cf. the ungrammaticality of the English expressions *always apparently, *often apparently or *already apparently).

Travis (2006:1276) provides another syntactic test suggesting that SAYC has undergone a grammaticalisation process. Given that the merged complementiser is no longer analysable as an actual complementiser (except in Sardinian and Romanian, as will be clarified below), SAYC can co-occur with a complementiser. There is variation, however, as to the function of the second occurrence of the complementiser (which in many cases is not a real, i.e. complement-introducing complementiser). SAYC can co-occur with a second instance of the complementiser in all languages considered here, performing various functions. This second (or
sometimes third) complementiser often introduces a subordinate causal or also consecutive clause. It can either precede or follow SAYC:

(19) Sic. Vippi troppu assà, (ca) dicica (ca) jera fatti stari na pezza.
    drink.PAST.1S too much that SAYC that be.IMPF.1S make.PP stay.INF a rag
    ‘I drank so much that apparently I was completely out of it.’

Like in Spanish, in these languages too the complementiser is never dependent on SAYC since this element is no longer able to take a complement clause. By contrast, Sardinian and Romanian present a different scenario. In these varieties the complementiser is in fact selected by SAYC, showing that it has not been fully grammaticalised and has not lost its verbal features entirely, insofar as it is able to take an object complement clause. In (20) the complementiser introduces an embedded clause that represents the object of SAYC:

(20) E nachi chi issa no b’andiaada nudda.
    and SAYC that she not there.CL go.IMPF.3S nothing
    ‘And it is said that she didn’t go there at all.’

The example above proves that SAYC in Sardinian is able to take an argument clause. It can also take a subject, as in (21) (from Puddu 2000: s.v. nachi), showing that, as far as theta-role assignment is concerned, it can still behave as a normal verb with its own thematic-grid. It must be noted, however, that even when taking an argument, it is morphologically invariant, failing to agree with the subject:

(21) millu, cojadu, cudhu, mih: e nois nachi si faghiat preíderu!
    look.IMP-him.CL married that look.IMP and we SAYC REFLEX.do.IMPF.3S a priest
    ‘Look at him, married as he is, that one, look: and we always said he would become a priest!’

Since the complementiser integrated in Romanian cică is no longer analysable as a complementiser, there too it frequently occurs that că is repeated, either before or after cică, to introduce a complement clause (cf. (22) – the că in front of cică, again, has a causal meaning):

(22) ... şi nu vine acasă de la serviciu că cică că are şedinţe.
    and not come.PRES.3S home from the job because SAYC that have.PRES.3S meetings
    ‘... and he doesn’t come home from work because he says that he has meetings – allegedly.’

If we analyse SAYC as an evidential marker, the syntactic features discussed in this section are straightforwardly accounted for. Moreover, these features prove to

---

8 The subjunctive particle să can also co-occur with cică to introduce an embedded complement clause.
be shared by other evidential adverbs, which leads us to the conclusion that SAYC has grammaticalised as an adverbial element. However, a clear distinction has to be made between languages like Sicilian and Galician, in which SAYC only exists as a grammatical marker of evidentiality, and languages in which SAYC, irrespective of its phonological reduction and its morphological invariance, can still function as a verb that assigns theta-roles. We will return to this distinction in section 3.

2.4 Semantics

The semantic bleaching (or desemanticisation) typical of elements which are on their way on a grammaticalisation path can be recognized in the following examples where SAYC is used together with the full verb ‘to say’:

(23) Sp. Y dicen que diz que […] no más trabajan en el campo. (Kany 1944:172) and say.PRES.3P that SAYC not more work.PRES.3P in the field ‘And they say that apparently they don’t work on the field anymore.’

(24) Sic. Maria mi dissí ca dicica arrubbaru a machina au dutturi. Maria to-me.CL tell.PAST.3S that SAYC steal.PAST.3P the car to-the doctor ‘Maria told me that apparently the doctor’s car has been stolen.’

(25) Srd … e an cominzadu a faeddare, e nachi ana nadu: … (AdS) and have.PRES.3P start.INF and SAYC have.PRES.3P say.PP ‘… and they started to talk, and allegedly they said: …’

(26) Ro. Se zice că cícă ziua se cunoaşte de dimineaţă. (I. Creangă) REFL say.PRES.3S that SAYC the-day REFL know.PRES.3S from-the-morning ‘It is said that apparently the day can be judged by its morning.’

(27) Gal. Un estado aconfesional como parece que disque dí a Constitución non é lóxico. (www) a state nonconfessional like seem.PRES.3S that SAYC say.PRES.3S the Constitution not be.PRES.3S logical ‘A nonconfessional state as, apparently, the Constitution seems to say, is not logical.’

Since the lexical meaning of the SAYC-element is partially lost and now extended to contexts of more functional usage, the lexical verb ‘to say’ can be inserted again, as a form of “semantic strengthening” (cf. Travis 2007:1277) or “lexical reinforcement” (cf. Aikhenvald 2004). In (23) and (24), a form of lexical ‘say’ plus C has been inserted, in (25) and (27) without C, and in (26) an impersonal verb form is used to lexically reinforce the meaning of SAYC. As for its interpretation, SAYC in its prototypical use can no longer function as a verb, but functions instead as an adverbial with an evidential meaning, cf. the following
examples from Sardinian, which clearly represent a minimal pair and show that naki has lost its original meaning as a declarative verb:

(28) a. Juanne naki si k’est mortu.9
    Juanne SAYC REFL there.CL be.PRES.3S die.PP
    ‘Juanne is said to have died.’

(28)b. Juanne narat ki si k’est mortu. (Sa-Limba 1999-2007)
    Juanne say.PRES.3S that REFL there.CL be.PRES.3S die.PP
    ‘Juanne says that he has died.’

(28)a clearly has an interpretable evidential meaning, whereas (28)b is semantically odd, because of the impossibility of a dead person to speak.

3. Functions of SAYC and degrees of grammaticalisation

As has been shown in 1.1 reportative evidentiality markers can be used to express either second-, or third-hand evidence, or knowledge based on a common ground of traditions, mythology or folklore. As far as the evidential marker at issue here, SAYC, is concerned, a more finegrained differentiation might be needed in some cases. Travis, for example, identifies four functions of dizque in Colombian Spanish (cf. Travis 2006:1278), namely a) reported speech (= second-hand, in Willett’s terms), b) hearsay (= third-hand and folklore, in Willett’s terms), c) labelling, and d) the dubitative use.

As explained in (6), based on Willett (1988), reported speech exemplifies the source of the information (second-hand), whereas hearsay does not (third-hand). Travis’ further distinction between reportative dizque with direct and indirect speech will be important for the Romance varieties under investigation. Following Travis, dizque with direct speech serves also to mark authenticity, whereas it might imply a flavour of doubt when used with indirect speech (Travis 2006:1280-1281). The hearsay function of dizque, where there is no identifiable source of information, is the most common. The labelling function is found when SAYC does not represent an adverbial but a modifier of a noun phrase (or even other phrases); this can sometimes be translated by ‘so-called’, ‘supposed’, or perhaps ‘so to speak’ (Travis 2006:1287):

(29) Se presentò como, dizque narcotraficante. (Travis 2006:1279)
    REFL present.PAST.3S as SAYC drug-dealer
    ‘He presented himself as a so-called drug dealer.’

9 In Sardinian, morrere, ‘to die’, is inherently reflexive (cf. Spanish morirse); furthermore, it is important to distinguish the complementiser ki from the second ki (in this example k’), which is a locative clitic (cf. Italian ci).
Hearsay and reported speech: evidentiality in Romance

The dubitative use of *disque* can be “marking falsity”, but also “unintentional consequences” involving an element of surprise (Travis 2006:1291-2), cf. (30). This is certainly the most difficult function of *disque* to identify, since it remains quite vague in definition. However, the common feature here is that there is no longer a recognizable speech act (Travis 2006:1293).

(30)  Yo, por Dios, *dizque* a limpiar baños.  (Travis 2006:1292)
I by God SAYC to clean:INF toilets
‘And me, by God, I was there cleaning toilets!’

Following Travis (2006), all these functions also reflect different aspects of progressive grammaticalisation. Consequently, the range of possible functions varies from language to language, and not all of them can be found in every language at issue here: For example, Sicilian *dicica* can only mark reported speech in indirect discourse whereas Sardinian *nachi* can also mark direct reported speech; both languages can, of course, encode hearsay, and Sardinian in particular has a very extended use of SAYC in folklore; by contrast, the labelling function is less easy to find – perhaps even impossible – in the varieties under discussion. It is also not yet clear whether each variety has a genuine dubitative use of SAYC or if the dubitative interpretation is due to pragmatic implicatures.

Since the use of SAYC in folklore also seems to exist in Latin American Spanish (cf. Kany 1944:140), in addition to being very common in the Romance varieties at issue, we will add this function (observed by Willett 1988) to Travis’ classification. Thus we will use the following refined typology for SAYC (integrating Willett 1988 and Travis 2006) as a point of reference:

(31)  I. second-hand
e. direct reported speech
   Ib. indirect reported speech
   II. third-hand
      IIA. non specified
      IIB. folklore
      III. Labelling
      IV. dubitative

In what follows, we examine the functions given in (31) with respect to the Romance varieties at issue.

In Sardinian, *nachi* is typically used as a reportative evidential, which does not give the specific source of information, thus encoding third-hand evidence (function

---

10 One could make a further distinction here with respect to the explicit or anaphoric expression of the subject.
It is also extremely common in fairy tales, where it may appear after nearly every sentence, sometimes every constituent (thus function IIb). Besides this third-hand use, nachi can be used in direct (Ia) and indirect speech (Ib). As a direct discourse marker it often occurs with an explicit subject (not necessarily in the third person singular):

(32) “Eh!” nachi Gesù Crilthu “tando bi benzo abberu, azzetto s’invitu.” (AdS)
    “Eh!” said Jesus Christ “in this case I will indeed come with you, I accept the invitation.”

(33) Gesù Crilthu nachi “Pedru” nachi “abbaida: pro te chelzo faghure una cosa totta noa…” (AdS)
    ‘Jesus Christ said “Peter” he said “look: For you I will do something completely new…”’

As far as the dubitative function (IV) is concerned, nachi may express surprise or wonder, but this use does not coincide with the dubitative function described in Travis (2006). The labelling function (III) is not found. Sardinian nachi has certainly reached a certain stage of grammaticalisation, as witnessed by its main function (II) to express third-hand evidence. However, in its second-hand use, especially when used as a direct speech marker with an explicit subject, it seems to be still analysable as a verb (which has strangely enough morphologically incorporated the complementiser).

The primary function of Sicilian dicica is that of a reported speech marker. It can refer to an explicit subject, identifying second-hand evidence (I). In this case, the source of information never occurs in the same sentence containing dicica, which is not able to take a syntactic subject. Rather, it coincides with the subject ‘under discussion’, established as such in the previous discourse. In fact, unless inferable from the context, any specific reference to or indication of the actual author of information is impossible:

11 See the characterization of naki given by Jones (1993:168-9): “Also within this class of modal adverbs we may include the ‘hearsay’ adverb naki […] which is frequently interspersed in reported speech […] but can also have a modal value roughly equivalent to ‘apparently’. It is also frequently used to prefix sentences with which broach a new topic of conversation; e.g. Naki ses cojuande. ‘I hear that you are getting married’.”
Hearsay and reported speech: evidentiality in Romance

(34) (*Maria) dicica (*Maria) iddu un jiva d’accordu cu a soggira.
Maria SAYC Maria he not went of agreement with the mother-in-law
Intended meaning: ‘Mary says he didn’t get on well with his mother-in-law.’

If no subject is present in the discourse or salient in the context, dicica indicates that the speaker is reporting third-hand evidence (IIa). Dicica is sometimes used in sentences expressing surprise or false beliefs, but it is always related to the notion of speech. Therefore, it does not perform the dubitative function (IV). Dicica only rarely reports information as folklore (proverbs, stories, riddles, and songs, IIb). Another property of dicica, confirming its adverbial status, is that it can only take a full clause within its scope, never a noun or a prepositional phrase in the labelling function (III).

Galician disque performs more or less the same functions as Sicilian dicica. It is mainly used in second- (Ib) and third-hand (IIa) indirect reportative speech. As in Sicilian, Galician SAYC cannot take a subject and the exact source of second-hand information can only be gathered from the context. Consequently, it does not function as a direct speech marker, and cannot attribute the utterance to a specific author of the speech act. Unlike Sicilian, disque is also common as a folklore evidence marker (IIb) in fairy tales, stories, and proverbs. Moreover, the meaning of disque has not generalised to cover functions completely unrelated with the notion of speech. Hence, it cannot perform the conventionalised dubitative function (IV). Finally, the labelling function (III) is not found in Galician.

Also Romanian cică can be used in a range of functions. First, it can be used impersonally as an adverb, corresponding to “reportedly”, “allegedly”, as a third-hand evidence marker (IIa). Cică, then, is extremely frequent in fairy tales and folklore (IIb). It can also be used with an explicit indication of the source of information given by the context; in this case, cică does not appear in the same clause with the given source. Thus, it can still be interpreted as a third-hand evidence marker (IIa). Unlike Sicilian dicica and Galician disque, Romanian cică can also appear with an explicit subject in the third person singular in indirect discourse (Ib) (ex. from Delavrancea, following Macrea 1955-1957: s.v. cică):

(35) Mos popa, cînd spune de evanghelie, cică să rabzi şi iar să rabzi.
old priest-the when speak.PRES.3S of gospel SAYC that be-patient.PRES.2S and again that be-patient.PRES.2S
‘The old priest when he preaches always says to be patient and to be patient again.’

However, in Romanian, in contrast to Sardinian, cică cannot appear as a direct discourse marker (Ia), neither with an overt nor with an implicit subject. The labelling and the dubitative functions seem not to be available with cică in Romanian. Thus, Romanian cică does not have all the functions that Sardinian nachi has. In particular, it cannot appear in direct discourse. Yet, it also has no labelling
function. Thus, as far as grammaticalisation is concerned, it can be said to lie somewhere in between Latin American Spanish (more grammaticalised) and Sardinian (less grammaticalised).

The functions of SAYC in Romance are summarised in the following table:

(36) The functions of SAYC in Romance

<table>
<thead>
<tr>
<th></th>
<th>INDIRECT/REPORTATIVE</th>
<th>EVIDENTIALITY</th>
<th>LABELLING</th>
<th>DUBITATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>second-hand (I)</td>
<td>third-hand (II)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Ia)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>indirect speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Ib)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>hearsay</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(IIa)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>folklore</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(IIb)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(III)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(IV)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4. SAYC and the Evidential Mood Phrase

We have already hypothesised that SAYC should show syntactic properties typical of evidential adverbs (cf. 2.3). This section sets out to identify the position of this use of SAYC within the sentence. Adopting a cartographic approach and in particular the hierarchy of functional projections proposed by Cinque (1999), we show that SAYC is generated in the specifier position of the Evidential Mood projection. This position is typically occupied by evidential adverbs, such as, in English, allegedly, reportedly, apparently, obviously, clearly, and evidently (Cinque 1999:86). Evidence for our claim is given with respect to the unmarked position of SAYC within the sentence, adopting some ordering tests within the higher adverbial field.

4.1 The position of SAYC within the clause

In compliance with the syntax of adverbs, SAYC can appear in different positions of the sentence, especially in parenthetical uses, including the final position. SAYC, however, can be identified as a higher sentence adverb which takes the whole sentence within its scope, and its natural position is thus sentence initial. In the unmarked word order, it precedes the subject:

(37) Sic. *Dicica Maria jè malata.*

(38) Gal. *Disque María estaba enferma.*
Hearsay and reported speech: evidentiality in Romance

(39) Srd. Nachi Juanne est malàidu.

(40) Ro. Cică Ion e bolnav.
    SAYC Mary/John be.PRES.3S ill
    ‘Apparently Mary/John is ill.’

On the other hand, it can be preceded by a dislocated topic of the sentence, typically (but not exclusively) coinciding with the subject. In this case, the topicalised constituent does not fall under the scope of the evidential marker:

(41) Sic. Maria dicica jè malata.

(42) Sp. María disque estaba enferma.

(43) Srd. Juanne nachi est malàidu.

(44) Ro. Ion cică e bolnav.
    Mary/John SAYC be.IMPF.3S ill
    ‘As for Mary/John, apparently s/he is ill.’

Examples (37)-(40) offer minimal pairs with the sentences in (41)-(44). The propositional meaning of the sentence pairs is the same. In both word orders SAYC simply means that someone said that Mary/John is ill, but it does not specify the source of this information. We know that the constituent preceding SAYC is the subject of the sentence, which has been left-dislocated in association with a marked information structure of the utterance. The order SAYC-Subject therefore represents the unmarked word order, which can be altered for discourse-related purposes. The topic usually coincides with the subject, but it does not have to be necessarily the case, as shown in the following examples containing a dislocated direct object and a dislocated locative complement, respectively:

(45) Sic. I robbi dicica si l’acetta ’ntra na putìa fina di Palermu.
    the clothes SAYC refl them.CL buy.PRES.3S at a shop posh of Palermo
    ‘As for her clothes, apparently she buys them in a posh shop in Palermo.’

(46) Srd. ... idende unu castello. In custu castello nachi bi istaiada su fizu ‘e su re.
    see.GER a castle in this castle SAYC there.CL stay.IMPF.3S the son of the
    king
    ‘and there they saw a castle. Apparently the son of the king was living in
    this castle.’ (AdS)

The order SAYC-Subj and Top-SAYC are the most common orders. This syntactic distribution reflects the status of the SAYC elements as sentence modifiers, which occur in sentence-peripheral positions, usually at the beginning of the sentence. Other orders, however, are possible as a consequence of the fact that adverbs typically allow for parenthetical uses. It has been observed that higher adverbs generally cannot occur after a finite verb (cf. Cinque 1999:31). This
constraint, however, is overridden whenever the adverb is used parenthetically. As a parenthetical, an adverb can occur at major constituent boundaries (e.g. between the subject and the verb or, more rarely, between the verb and peripheral elements within the VP such as PPs or clauses), and also in final position. Like adverbs that appear at the end of the sentence, SAYC is typically deaccented and prosodically separated from the preceding sentence by a pause (indicated by a comma in the following examples):

(47) Sic. Maria jera malata, dicica  
(48) Gal. María estaba enferma, disque.  
     Mary was ill SAYC  
     ‘Mary was ill, apparently’

As pointed out by Belletti (1990:130, n. 29) and Cinque (1999), a higher adverb can occur after a finite verb also when another constituent follows the adverb:

(49) Sic. Pinu si sciarià cu Turiddu, dicica pi sordi  
     Pinu REFf fight.PAST.3S with Turiddu, SAYC for money  
     ‘Pinu had an argument with Turiddu, apparently for money’

In these examples, SAYC does not have scope over the preceding sentence, but just over the following constituent. However, the frequent placement of SayC at the beginning of the sentence, together with its positional mobility, confirms the claim that SayC behaves as modal adverb.

4.2 The adverbial hierarchy and ordering constraints

Following Cinque (1999), we assume that adverbs occupy the specifier positions of a rich and articulated series of functional projections obeying a rigid hierarchical order. If SAYC is merged into the specifier of the Evidential Mood Projection then some ordering constraints with respect to other verbs are expected, giving rise to the following predictions:

(50)  a. SAYC should not be compatible with other evidential adverbs competing for the same position;  
     b. it should be followed, but not be preceded, by epistemic and possibility adverbs;  
     c. it should be preceded, but not followed, by speech act and evaluative adverbs.

Now, corresponding to these predictions, Cinque’s hierarchical articulation of the (higher) adverbs indeed implies the following ordering constraints:

(51)  a. *apparently + allegedly/clearly/evidently/obviously  
     b. apparently + probably/possibly; *probably/possibly + apparently
Hearsay and reported speech: evidentiality in Romance

c. frankly/honestly/unfortunately + apparently, *apparently + frankly/ honestly/unfortunately

In Sicilian, SAYC is not the only modal adverb derived from the merging of a verb and a complementiser. Other adverbs show the same historical development: The adverb parica, ‘seemingly’, ‘apparently’, also functions as an evidential marker reporting third-hand evidence, and derives from the third person singular of the verb pariri ‘to seem’ plus the complementiser ca. The epistemic adverb pènzica expressing probability comes from a form of the verb pinzari ‘to think’, to which the complementiser has been incorporated. This grammaticalisation process peculiar of Sicilian modal adverbs has also caused the reanalysis of the sequence copula + adjective + complementiser: the locution jè capaci ca (lit. ‘it is capable/possible/ likely that’) has been grammaticalised to the form capacica, which now represents a single unit. The hierarchical system of the Sicilian higher adverbs, equivalent to Cinque’s (1999) system, is illustrated in (52) (cf. also Ledgeway and Lombardi 2005:81 for Calabrian):

(52) Sic. [onestamenti Moodspeech act [pi furtuna Moodevaluative [dicica, parica Mood evidential [pènzica, probabilmenti Moodepisteme] na vota, tannu T(past) [pua T(future) [forsi Modirrealis [pi forza Modnecessity [capacica Modpossibility [...Asp...

The predictions in (50), obeying the ordering constraints in (51), are indeed born out by our data. Dicica proves incompatible with other evidential adverbs (e.g. pàrica), but it can co-occur with epistemic adverbs (pènzica, probabilmenti), as shown by the following examples:

(53) Sic. *Dicica parica / *Parica dicica ora si senti bonu.
SAYC SEEMC / SEEMC SAYC now REFL feel.PRES.3S good
‘Apparently s/he is feeling better.’

(54) Sic. a. Dicica pènzica/probabilmenti s’accattà na casa nova.
SAYC THINKC probably REFL buy.PAST.3S a house new
‘Apparently s/he probably bought a new house.’

b. *Pènzica/probabilmenti dicica...

The Evidential Mood Projection is higher than the position of possibility adverbs, such as capacica. Dicica can thus precede, but not follow capacica:

(55) Sic. a. Dicica capacica ci avivanu finitu i sordi.
SAYC possibly to-him/her.CL have.IMPF.3P finish.PP the money
‘Apparently s/he possibly ran out of money.’

b. *Capacica dicica...

The opposite ordering applies with speech act adverbs, like onestamenti, which can only precede dicica:
(56) Sic. a. **Onestamenti dicica iddu si scantava.**
   honestly SAYC he REF be-scared.IMPF.3S
   ‘Frankly/honestly they say that he was scared.’

   b. *Dicica onestamenti...*

   The same ordering constraints occur in Galician, where epistemic
   (**probablemente**) and possibility (**posiblemente**) adverbs cannot precede **disque**, while the same sentences prove grammatical if they follow it. Along the same lines, evaluative adverbs (such as **desafortunadamente**) need to be placed before **disque**. Romanian cică cannot be combined with parcă 'seemingly', another evidential marker stemming from **se pare că** 'it seems that' (cf. Macrea 1955-1957: s.v. parcă; Berceanu 1971:55-56; DEX s.v. parcă) in the same clause, etc. This impossibility is readily accounted for under the analysis that the two evidential adverbs compete for the same position, namely the specifier of the Evidential Mood Projection.

5. Conclusions

Based on comparative evidence, our paper offers a systematic analysis of the different functions and the varying degrees of grammaticalisation of SAYC in Romance. Firstly, we show that SAYC is found in many Romance varieties, and not just in Latin American Spanish. Secondly, we prove that SAYC has different functions and is characterised by varying degrees of grammaticalisation according to each language. In their prototypical, unmarked use, the SAYC-elements sit in the specifier position of the Evidencial Mood Projection. Like adverbs, they can also appear in all kinds of parenthetical positions in the sentence. Sardinian and Romanian also allow for a (defective) verbal use of SAYC. In Latin American Romance, when performing the **labelling** function, SAYC can also occur within the DP.

Acknowledgements

We would like to thank Beatriz Vizcaíno Pena and Doina Vacaruic for their grammatical judgements on Galician and Romanian, respectively. For Sardinian, we would like to thank the members of the mailing list Sa-Limba (cf. Sa-Limba 1999-2007). Further help with the extraction of the data came from Florian Scheib (Romanian), Heidrun Bohnet (Spanish) and Christian Ferraro (Sardinian). Part of the research of Eva-Maria Remberger concerning this paper has been supported by the DFG with the project “Auxiliary verbs – status, development, variation” within the SFB 471 “Variation and development in the lexicon” at the University of Konstanz as well as by a Feodor-Lynen-Fellowship granted by the Alexander von Humboldt-Foundation for a one year stay in Cambridge, which made this joint paper possible. We would also like to thank Adam Ledgeway for insightful comments.
References


Silvio Cruschina, Eva-Maria Remberger