Intramolecular basis set superposition error effects on the planarity of benzene and other aromatic molecules: A solution to the problem

David Asturiol, Miquel Duran, and Pedro Salvador

Additional information on J. Chem. Phys.

Intramolecular basis set superposition error effects on the planarity of benzene and other aromatic molecules: A solution to the problem

David Asturiol,[1] Miquel Duran,[2] and Pedro Salvador[2,a)]

[1]*Institut de Química Computacional, Parc Científic i Tecnològic de la Universitat de Girona, Edifici Jaume Casademont, Pic de Peguera 15 (la Creueta), 17003 Girona, Spain*
[2]*Institut de Química Computacional and Departament de Química Universitat de Girona, Campus de Montilivi, 17071 Girona, Spain*

Recently, the surprising result that *ab initio* calculations on benzene and other planar arenes at correlated MP2, MP3, configuration interaction with singles and doubles (CISD), and coupled cluster with singles and doubles levels of theory using standard Pople's basis sets yield nonplanar minima has been reported. The planar optimized structures turn out to be transition states presenting one or more large imaginary frequencies, whereas single-determinant-based methods lead to the expected planar minima and no imaginary frequencies. It has been suggested that such anomalous behavior can be originated by two-electron basis set incompleteness error. In this work, we show that the reported pitfalls can be interpreted in terms of intramolecular basis set superposition error (BSSE) effects, mostly between the C–H moieties constituting the arenes. We have carried out counterpoise-corrected optimizations and frequency calculations at the Hartree–Fock, B3LYP, MP2, and CISD levels of theory with several basis sets for a number of arenes. In all cases, correcting for intramolecular BSSE fixes the anomalous behavior of the correlated methods, whereas no significant differences are observed in the single-determinant case. Consequently, all systems studied are planar at all levels of theory. The effect of different intramolecular fragment definitions and the particular case of charged species, namely, cyclopentadienyl and indenyl anions, respectively, are also discussed. © *2008 American Institute of Physics.* [DOI: 10.1063/1.2902974]

INTRODUCTION

In a recent Communication, Moran *et al.*[1] reported a vast number of *ab initio* calculations on benzene and other planar arenes at different correlated levels of theory and Pople's basis sets that yielded nonplanar minima. The planar optimized structures turned out to be transition states exhibiting one (or more) large imaginary frequency. Single-determinant-based methods such as Hartree–Fock (HF) and density functional theory (DFT) methods (BLYP and B3LYP) lead to the expected planar minima and hence no imaginary frequencies.

These results are in line with the earlier reports on failures of electronic structure methods for the correct description of low-lying out-of-plane vibrational frequencies of benzene,[2,3] several planar arenes,[4–8] and other nonrigid molecules.[9] Similar pitfalls have been recently found by Shabahzian[10] for the challenging $(B_6C)^{2-}$ anion, particularly at the MP2 level of theory. Rather to any deficiency of the post-HF electronic structure methods including electron correlation, the origin of the problem has been suggested to be rooted on (atom-centered) basis set deficiencies. Simandiras *et al.*[9] found that the use of *f*-type basis functions was necessary to obtain accurate bending frequencies. Martin *et al.*[2] dealt in detail with the benzene case and concluded that out-of-plane bending modes are pathologically basis set dependent, putting forward also a basis set superposition origin, based on the work of Sellers and Almlöf.[11] Jensen[12] found that imaginary frequencies can also appear at the single-determinant level for an uncontracted double-zeta quality basis set including *d*-type diffuse function (aug-pc-1) for a rather narrow range of diffuse function's exponent values.

Perhaps the most important error introduced by the use of truncated atom-centered basis sets is the so-called basis set superposition error (BSSE).[13] It has been recognized for years[14] that BSSE introduces some spurious extra binding in the *ab initio* calculations and that its correction is essential to properly describe intermolecular interactions. In the last decade, it has been also shown that BSSE does not merely affect the interaction energy but also the topology of the potential energy surface (PES) of these systems, which translates into geometrical and vibrational effects. There are a few examples in the literature where BSSE accounts for large geometric effects, namely, the hydrogen fluoride dimer[15,16] or weak C–H⋯O (Ref. 17) interactions.

There are different ways to deal with BSSE. One option is to increase the basis set quality, since in the complete basis set limit BSSE should vanish. Whereas this could be accomplished at the HF or DFT levels of theory, for post-HF methods, BSSE has been shown to converge to zero very slowly with basis set improvement.[18,19] On the other hand, there are several strategies[20–23] to correct for BSSE, one is the counterpoise (CP) method[20] which is the simplest and is most widely used. Given a system formed by N fragments, the CP correction to the energy, in its simplest form,[24] reads as
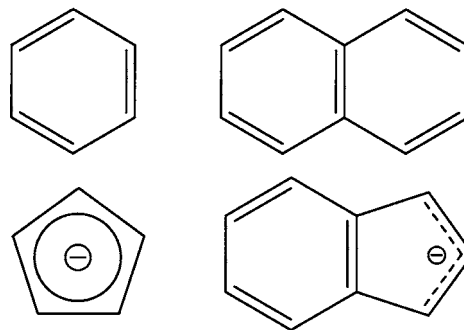
a)Author to whom correspondence should be addressed. Electronic mail: pedro.salvador@udg.edu.

$$\delta^{CP}(\vec{R}) = \sum_i^N \varepsilon_i^i(\vec{R}) - \varepsilon_i^{full}(\vec{R}), \tag{1}$$

where $\varepsilon_i^i(\vec{R})$ and $\varepsilon_i^{full}(\vec{R})$ represent the energy of the ith fragment of the system calculated with its own basis set and with the full basis set of the system (ghost-orbital calculation), respectively. Explicit dependence on the position of the atoms of the systems has been included to point out that the CP correction is geometry dependent. Note also that the charge, multiplicity, and electron state of the fragment's calculations must be specified, and that the value of $\delta^{CP}(\vec{R})$ will actually depend on the choice. This does not represent practical difficulties for closed-shell interacting molecules. However, the case of radical or charged molecular complexes, where the spin density or the charge is spread over the whole system, is not so trivial. We will go back to this point when dealing with the cyclopentadienyl anion case.

When the CP correction is applied as correction term to the total energy of the system,[25] it is possible to obtain BSSE-corrected energies and any molecular property that may be stated as a derivative of the energy, which includes optimized geometries or vibrational frequencies, as implemented in GAUSSIAN 03.[26] Experience gathered on the BSSE correction of intermolecular interactions over the last years shows that the results obtained with BSSE-corrected calculations using moderate basis sets are generally close to those that one can obtain with much larger basis sets with or without correction, provided that (a) the basis set is flexible enough to describe the true physical interactions in the system, and (b) BSSE effects on the geometry are also taken into account by optimization on the BSSE-corrected PES.

Much less attention has been paid to BSSE effects in single molecules, what one can refer to as intramolecular BSSE.[11,27–30] Noteworthy exceptions are recent studies on aromatic-backbone intramolecular interactions[31–34] on peptide models, which have put forward that the accurate determination of relative stabilities of conformers is heavily affected by BSSE. In fact, the first CP calculation by Jansen and Ros[13] was applied for a single cation (HCO$^+$). In single molecules, there is not *a priori* problem with the fact that one atom or fragment can use the basis functions centered on other atoms or parts of the molecule. This helps to naturally describe polarization or charge transfer effects. The problem arises when the use of external basis functions is merely due to a lack of flexibility of the fragment's basis set; i.e., when the basis set incompleteness error (BSIE) is strongly geometry dependent, spurious geometry changes can be induced by intramolecular BSSE effects. Unbalanced descriptions of PES can also emerge. This directly translates into poor vibrational frequencies and likely changes in the topology of the stationary points, as found by Moran *et al.*[1] In that revealing report, the authors showed that in those cases where a nonplanar optimized structure was found for benzene, the two-electron BSIE tends to dramatically increase for geometries away from planarity. In fact, even for the levels of theory where no imaginary frequencies are observed, BSIE still shows moderate dependence on the geometry. Thus, the two-electron BSIE diagnostic seems indeed a valuable tool

SCHEME 1. Arenes considered in this study.

to detect possible spurious geometries induced by BSSE effects, but it is not definitive yet. A recent study[35] based on the use of chemical hardness profiles is also able to detect such spurious structures, but again, it does not offer a solution to the problem.

In this work, we wish to show that the reported pitfalls can be interpreted in terms of intramolecular BSSE effects and, as such, the calculations can be corrected using BSSE correction techniques typically applied to intermolecular interactions, particularly the CP method. For this purpose, we have carried out CP-corrected optimizations and subsequent CP-corrected frequency calculations for a number of the arenes (see Scheme 1) at the HF, B3LYP, MP2, and CISD levels of theory with several basis sets. We will show that, with no exception, correcting for intramolecular BSSE fixes the anomalous behavior of the correlated methods, whereas no significant differences are observed in the single-determinant case.

FRAGMENT'S DEFINITION

A key point when estimating intramolecular BSSE effects is the proper definition of the fragments forming each molecule in order to apply Eq. (1). A natural and unambiguous way to proceed could be to take atomic fragments. The problem with this approach is that the number of necessary extra calculations to obtain the CP correction would be $2N$, N being the number of atoms. It would be desirable to be able to define larger fragments while maintaining the molecular symmetry of the system. In this sense, it is essential to note that in all the reported cases the problems are associated with out-of-plane bending low-lying modes. This tells us that in these cases the intramolecular BSSE does not seem to appreciably affect bonds or bond angles (otherwise stretching and other bending modes would have been affected). Taking all this into account, we can infer that proper moieties that can be used as fragments are the C–H units constituting the arenes.

In order to assess the effect of different fragment definitions, we have first determined the single-point CP-corrected energy profile along the b_{2g} vibrational mode of benzene at the MP2/6-311G level of theory, for which an imaginary frequency of $722i$ cm^{-1} is obtained. The uncorrected and CP-corrected results using twelve atomic fragments, six C–H fragments, as well as three $(C–H)_2$ fragments are displayed in Fig. 1.
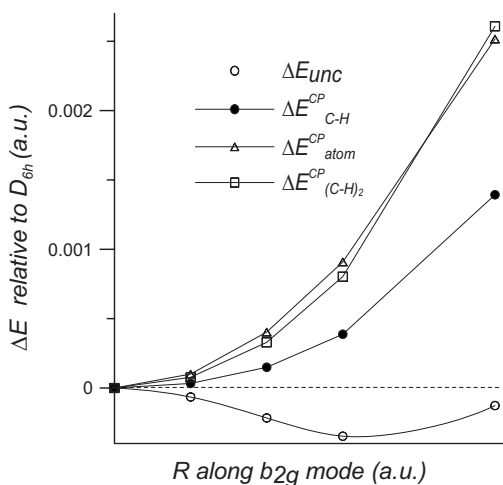
FIG. 1. CP-corrected energies along the b_{2g} vibrational mode in benzene.

It can be readily seen that, even though the results obviously depend on the fragments' definition, all CP-corrected energy profiles properly describe the distortion as a vibration, contrary to the uncorrected profile.

COMPUTATIONAL DETAILS

As detailed in the Supporting Information,[36] we have used both GAUSSIAN 03 and our own code that allows us to exploit the symmetry of the molecules by minimizing the number of fragment calculations in both the CP-corrected optimization and frequency calculations. This is particularly important for the case of CISD, as in any ghost-orbital calculation symmetry must be turned off (fragments usually do not have the same symmetry as the supermolecular system).

The case of cyclopentadienyl anion is a special one because, even though the 6π electrons are fully delocalized in the ring, the extra electron must be assigned, in principle, to one of the five C–H fragments, thus breaking the symmetry of the CP-corrected calculation. Since our program allows us for a flexible definition of the CP-correction term, we have carried out the following approach in order to equally share the extra π electron among five identical CH fragments: We

have determined the CP correction using both neutral, $\delta_{\text{CH}}^{\text{CP}}(\vec{R})$, and negatively charged, $\delta_{\text{CH}-}^{\text{CP}}(\vec{R})$ C–H moieties and combined them as

$$\delta^{\text{CP}}(\vec{R}) = \tfrac{1}{5}\delta_{\text{CH}-}^{\text{CP}}(\vec{R}) + \tfrac{4}{5}\delta_{\text{CH}}^{\text{CP}}(\vec{R}),\qquad(2)$$

thus preserving the symmetry of the calculation and also taking into account the effect of the extra electron. This approximate CP-correction term has been used for both geometry optimization and frequency calculations.

The case of indenyl anion is also problematic. Note that up to five imaginary frequencies are obtained at the MP2/6-311G level of theory, the largest one exceeding $1000i$ cm^{-1}. For this system, a CP function, which keeps the C_{2v} symmetry of the molecule and equally shares the electron charge among the three C–H moieties of the five-member ring, has been used, in the spirit of Eq. (2).

RESULTS AND DISCUSSION

The lowest vibrational frequencies for benzene along with the symmetry of the vibrational mode, both uncorrected and CP corrected, are gathered in Table I. We have used three basis sets ranging from qualitatively good behavior (6-31+G*) to dramatic BSSE effects (6-311++G) and applied the same levels of theory as in Ref. 1 for comparison.

Noticeably, the CP correction fixes the anomalous behavior of the correlated methods in all cases. Consequently, benzene is found to be planar for all levels of theory. No significant differences are observed for higher frequencies or for the already qualitatively correct single-determinant calculations such as HF or B3LYP. Also, the CP-corrected optimized frequencies do not appreciably differ from the uncorrected ones (tables with all CP-corrected geometries and frequencies for all levels of theory are given in the Supporting Information[36]). The CP-corrected frequencies are also in better agreement with both experimental[37] and best theoretical estimates[2] than the uncorrected ones, even ignoring the lowest five vibrational modes (see Table II).

Noticeably, Jensen[12] has detected similar problems even for a single-determinant case, (BLYP, in particular) in combination with a double-zeta quality uncontracted basis set

TABLE I. Lowest uncorrected and CP-corrected vibrational frequencies of benzene for different levels of theory and basis sets (spurious frequency values in italics).

		MP2		CP-corrected		CISD		CP-corrected
a	b_{2g}	*182*	e_{2u}	384	e_{2u}	418	e_{2u}	425
	e_{2u}	379	b_{2g}	473	b_{2g}	591	e_{2g}	639
	e_{2g}	618	e_{2g}	613	e_{2g}	643	a_{2u}	713
b	b_{2g}	*722i*	e_{2u}	382	b_{2g}	*189i*	e_{2u}	382
	e_{2u}	333	b_{2g}	442	e_{2u}	397	b_{2g}	442
	a_{2u}	620	e_{2g}	625	e_{2g}	652	e_{2g}	625
c	b_{2g}	*1852i*	e_{2u}	409	b_{2g}	*1384i*	e_{2u}	455
	e_{2u}	*468i*	e_{2g}	631	e_{2u}	281	e_{2g}	653
	e_{2g}	470	a_{2u}	703	e_{2g}	650	a_{2u}	760

[a]6-31+G*.
[b]6-311G.
[c]6-311++G.

TABLE II. Average error (in %) for the computed harmonic frequencies of benzene with respect to the experiment (Ref. 37). In parenthesis the error computed without the lowest five frequencies.

	CP-HF		HF		CP-B3LYP		B3LYP	
a	7.9	(7.4)	8.6	(7.9)	1.0	(1.0)	1.1	(1.1)
b	9.0	(8.1)	9.6	(8.4)	2.0	(1.9)	2.4	(1.9)
c	8.2	(7.6)	9.8	(8.6)	2.4	(2.3)	2.4	(1.9)
	CP-MP2		MP2		CP-CISD		CISD	
a	4.5	(3.0)	7.2	(3.9)	4.4	(4.4)	5.4	(4.7)
b	5.1	(3.2)	16.1	(5.2)	5.1	(3.2)	10.6	(4.9)
c	3.7	(3.9)	31.1	(13.2)	9.7	(7.4)	18.7	(20.7)

[a]6-31+G*.
[b]6-311G.
[c]6-311++G.

(aug-pc-1), which includes a set of diffuse d functions with exponent $\alpha_d = 0.10$. We have noticed that this basis set suffers from strong linear dependencies at the benzene geometry. In fact, the overlap matrix is not positive definite and exhibits up to 24 eigenvalues below Gaussian's default threshold of 10^{-6} (being numerically zero). We have observed that the ghost-orbital calculations of the CH moieties yield unphysical delocalizations of the charge density in the neighboring atoms, which indicates that this basis set is very unbalanced upon removal of the dependent basis sets. Using the basis set in its contracted form leads to the expected planar structures, the CP correction not changing this situation.

Qualitative wrong behavior is observed also in the case of naphthalene, as seen in Table III. At the modest MP2/6-31G level of theory, the out-of-plane b_{2g} vibrational mode is poorly described, exhibiting an imaginary frequency of $400i$ cm^{-1}. The correction of the intramolecular BSSE leads to a value of 570 cm^{-1}, whereas no meaningful differences are observed for the rest of low frequencies.

Finally, the results for the charged species are gathered in Table IV. It is remarkable that the CP correction fixes even the pathological case of the indenyl anion at the MP2/6-311G level. The intramolecular BSSE effects are so large in this system that the matching between uncorrected and corrected out-of-plane normal modes is dubious. In the case of cyclopentadienyl anion, the imaginary frequency associated with a degenerate e_2'' normal mode also disappears upon CP correction, whereas large effects on other low-frequency normal modes is also observed. Hence, the rather involved CP func-

tions used for these systems to cope with the extra electron and keep overall molecular symmetry have proved to be successful.

For testing purposes, we have also applied to these systems a simpler CP function ignoring the extra charge (i.e., neutral C–H and C–C moieties). The frequencies obtained (not reported) are again real in all cases, while the differences with the frequencies reported in Table IV do not exceed 25%. Hence, the charge (and electronic state) of the fragments does not seem to be critical to correct for BSSE for these systems. This had been observed before also in the case of charged intermolecular complexes[30] and it is at the heart of the general success of the CP method. The BSSE effects are noticeable when the BSIE is strongly geometry dependent. The latter affects the description of the occupied molecular orbitals and also the virtual ones, which is the reason why post-HF methods are more prone to suffering from strong BSSE effects. Therefore, the BSSE is not directly related to the number of electrons of the system. This also explains the fact that hardness profiles, which have been successfully applied to the present problem,[35] are much more robust to level of theory and basis set effects than energy profiles (the chemical hardness is computed as the difference between the vertical ionization potential and electron affinity). Indeed, invoking the CP philosophy, an energy difference-based measure will perform better than a single energy value.

CONCLUSIONS

We have provided evidences that intramolecular BSSE accounts for the existence of nonplanar optimized minima structures predicted by typical electronic structure methods at the correlated level of theory. We have shown that by taking as fragments the C–H and C–C moieties, the CP-corrected optimized structures correspond to planar minima with no imaginary frequencies. We have also addressed charged systems, wherein in order to maintain the molecular symmetry more involved CP function have been used. For this purpose, we have employed a code[38] that allows for flexible definitions of the CP function and deals with molecular symmetry.

TABLE III. Lowest MP2 uncorrected and CP-corrected vibrational frequencies of naphthalene for different basis sets (spurious frequency values in italics).

Symmetry	MP2/6-31G	CP-corrected
b_{3u}	159	159
a_u	177	181
b_{1u}	369	362
b_{1g}	359	368
b_{3u}	419	446
b_{2g}	426	450
b_{2g}	*402i*	570

TABLE IV. Lowest MP2 uncorrected and CP-corrected vibrational frequencies of indenyl and cyclopentadienyl anions (spurious frequency values in italics).

Indenyl anion			Cyclopentadienyl anion		
Symmetry	MP2/6-311G	CP-corrected	Symmetry	MP2/6-311G	CP-corrected
a_2	*1072i*	331	$e_{2''}$	*401i*	614
b_1	*555i*	1058	$e_{1''}$	284	640
a_2	*458i*	604	$a_{2''}$	523	680
b_1	*200i*	823	$e_{2''}$	632	848
a_2	*125i*	848	$e_{2'}$	860	854
b_1	199	197			
a_2	305	221			
b_2	391	390			

ACKNOWLEDGMENTS

[1] D. Moran, A. C. Simmonett, F. E. Leach, W. D. Allen, P. V. Schleyer, and H. F. Schaefer, J. Am. Chem. Soc. **128**, 9342 (2006).
[2] J. M. L. Martin, P. R. Taylor, and T. J. Lee, Chem. Phys. Lett. **275**, 414 (1997).
[3] L. Goodman, A. G. Ozkabak, and S. N. Thakur, J. Phys. Chem. **95**, 9044 (1991).
[4] A. Dkhissi, L. Adamowicz, and G. Maes, J. Phys. Chem. A **104**, 2112 (2000).
[5] H. Lampert, W. Mikenda, and A. Karpfen, J. Phys. Chem. A **101**, 2254 (1997).
[6] D. Michalska, W. Zierkiewicz, D. C. Bienko, W. Wojciechowski, and T. Zeegers-Huyskens, J. Phys. Chem. A **105**, 8734 (2001).
[7] M. Saeki, H. Akagi, and M. Fujii, J. Chem. Theory Comput. **2**, 1176 (2006).
[8] H. Torii, A. Ishikawa, R. Takashima, and M. Tasumi, J. Mol. Struct.: THEOCHEM **500**, 311 (2000).
[9] E. D. Simandiras, J. E. Rice, T. J. Lee, R. D. Amos, and N. C. Handy, J. Chem. Phys. **88**, 3187 (1988).
[10] S. Shahbazian, Chem. Phys. Lett. **443**, 147 (2007).
[11] H. Sellers and J. Almlof, J. Phys. Chem. **93**, 5136 (1989).
[12] F. Jensen, J. Chem. Phys. **118**, 2459 (2003).
[13] H. B. Jansen and P. Ros, Chem. Phys. Lett. **3**, 140 (1969).
[14] F. B. Van Duijneveldt, J. G. C. M. Van Duijneveldt-van de Rijdt, and J. H. Van Lenthe, Chem. Rev. (Washington, D.C.) **94**, 1873 (1994).
[15] P. Hobza and Z. Havlas, Theor. Chem. Acc. **99**, 372 (1998).
[16] P. Salvador, B. Paizs, M. Duran, and S. Suhai, J. Comput. Chem. **22**, 765 (2001).
[17] P. Salvador, S. Simon, M. Duran, and J. J. Dannenberg, J. Chem. Phys. **113**, 5666 (2000).
[18] M. Masamura, Theor. Chem. Acc. **106**, 301 (2001).
[19] T. Van Mourik, A. K. Wilson, K. A. Peterson, D. E. Woon, and T. H. Dunning, Adv. Quantum Chem. **31**, 105 (1999).
[20] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[21] I. Mayer, Int. J. Quantum Chem. **23**, 341 (1983).
[22] T. Nagata and S. Iwata, J. Chem. Phys. **120**, 3555 (2004).
[23] S. Iwata and T. Nagata, Theor. Chem. Acc. **117**, 137 (2007).
[24] L. Turi and J. J. Dannenberg, J. Phys. Chem. **97**, 2488 (1993).
[25] S. Simon, M. Duran, and J. J. Dannenberg, J. Chem. Phys. **105**, 11024 (1996).
[26] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., GAUSSIAN 0.3, Revision B.02, Gaussian, Inc., Pittsburgh, PA, 2003.
[27] G. J. Halasz, A. Vibok, S. Suhai, and I. Mayer, Int. J. Quantum Chem. **89**, 190 (2002).
[28] F. Jensen, Chem. Phys. Lett. **261**, 633 (1996).
[29] N. Kobko and J. J. Dannenberg, J. Phys. Chem. A **105**, 1944 (2001).
[30] P. Salvador, M. Duran, and J. J. Dannenberg, J. Phys. Chem. A **106**, 6883 (2002).
[31] T. Van Mourik, J. Comput. Chem. **29**, 1 (2008).
[32] N. Y. Palermo, J. Csontos, M. C. Owen, R. F. Murphy, and S. Lovas, J. Comput. Chem. **28**, 1208 (2007).
[33] L. F. Holroyd and T. van Mourik, Chem. Phys. Lett. **442**, 42 (2007).
[34] J. Csontos, N. Y. Palermo, R. F. Murphy, and S. Lovas, J. Comput. Chem. **29**, 4 (2008).
[35] M. Torrent-Sucarrat, P. Geerlings, and J. M. Luis, ChemPhysChem **8**, 1065 (2007).
[36] See EPAPS Document No. E-JCPSA6-128-023815 for CP-corrected geometries and CP-corrected harmonic frequencies for all systems considered. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[37] N. C. Handy, C. W. Murray, and R. D. Amos, J. Phys. Chem. **97**, 4392 (1993).
[38] P. Salvador, CP-opt driver for GAUSSIAN98 (Girona, 2001).