representation to the generic user DTD.

In this second phase, we can take advantage of all existing back-conversion tools and methods, built around the RAINBOW DTD.

Hence there are numerous advantages in coding the recognition result in compliance with the ODIL DTD :

- you start from a non ambiguous syntax, checked by an SGML parser ;
- for the first up-conversion phase, you can use an SGML-SGML conversion tool ;
- the same conversion tool can be used both for phase 1 and phase 2 ;
- in both phases, you take advantage of synergy between the various studies conducted and products implemented in order to achieve conversions in the SGML world.

For instance, with OMNIMARK, an ODIL-RTF transcoder was developed in three days.

You find on Fig 3 an example of a document processed by our PRASAD prototype, and its corresponding ODIL expression on the following page.

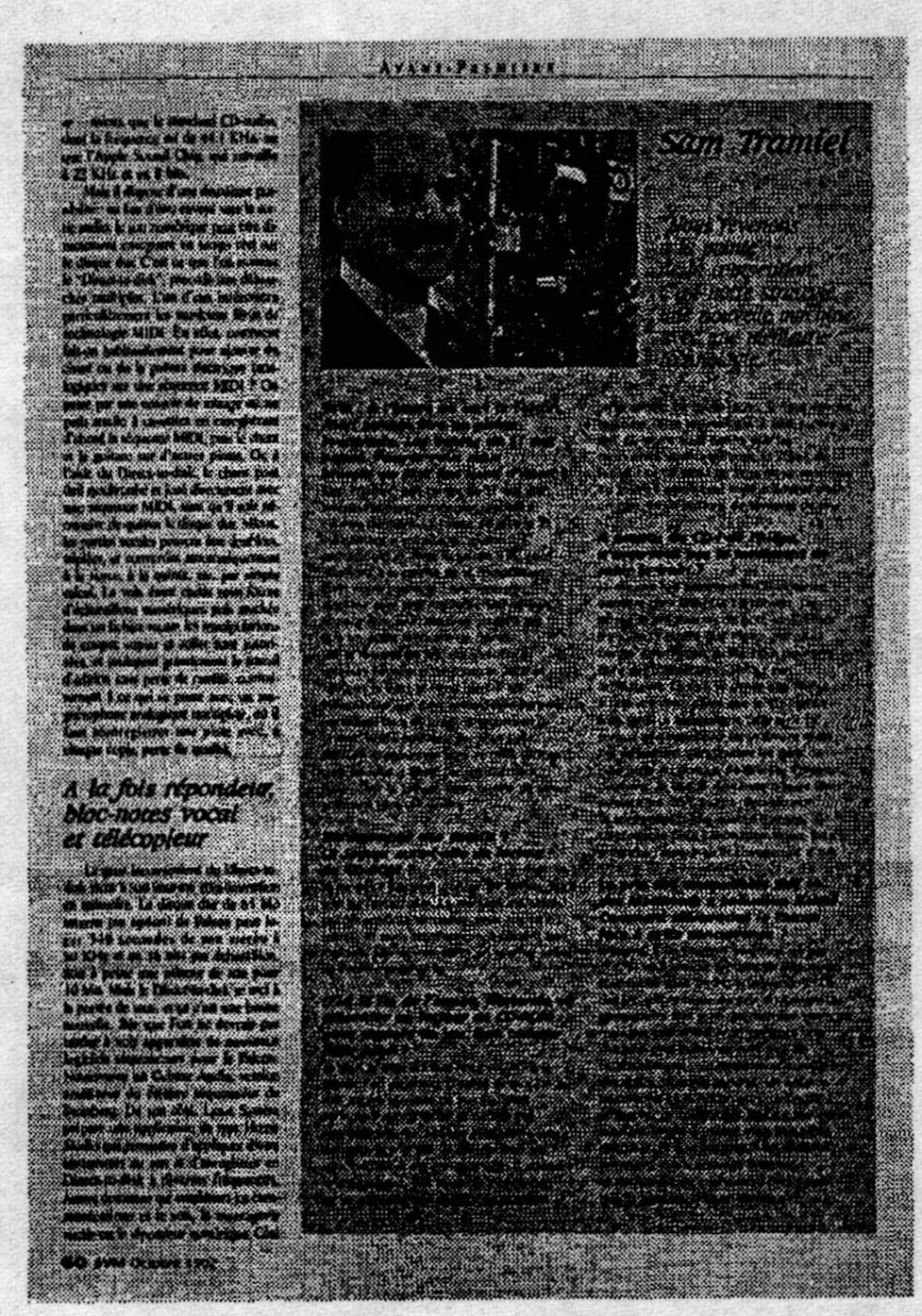

Fig 3 : Example of a processed page and its expression in ODIL

References

[1] E Van Herwijnen, *Practical SGML*, Kluwer Academic Publishers.

[2] ISO 8613-5, *Office Document Architecture (ODA) and Interchange Format*, Part 5, Annex E.

[3] Ph Lefèvre, C Felter, J.Y Christine, *Prototype de reconnaissance de documents PRASAD*, Note EDF-DER HN-46/93/136, (dec 1993).

[4] Ph Lefèvre, F Reynaud, *Choix d'un format dereprésentation normalisé de la structure physique des documents*, EDF-DER HN-46/93/035, (mar 1993).

[5] Ph Lefèvre, F Reynaud, *Spécification d'un format SGML de représentation de la structure physique des documents après segmentation et reconnaissance : ODIL*, Note EDF-DER HN-46/93/036, (jun 1993).

[6] R Ingold, « A document description language to drive document analysis », in *Proceedings of the First International Conference on Document Analysis and Recognition*, 294-301, (oct 1991).

[7] A Belaïd, J.C Anigbogu, Y Chenevoy, *Qualitative analysis of low-level logical structures*, electronic Publishing, Vol 6(4), 435-446, (dec 1993).

[8] O.T Akindele, *Vers un système de construction automatique de modèles génériques de structures de documents*, Thèse de doctorat de l'Université Henri Poincaré Nancy 1, (jan 1995).

[9] D Sklar, *Accelerating Conversion to SGML via the Rainbow Format*, <TAG>, (jan 1994).

[10] D Sklar, *The Annotated Rainbow DTD - Rainbow version 2.2*, Electronic Book Technologies, (feb 1994) e-mail : rainbow@ebt.com.

Proceedings of the Third International Conference on Document Analysis and Recognition (ICDAR '95)
0-8186-7128-9/95 $10.00 © 1995 IEEE