Nancy Joe Dyer

René Pellen. *"Poema de Mio Cid": Dictionnaire lemmatisé des formes et des références.* Annexes des Cahiers de Linguistique Hispanique Médiévale 1. Paris: Séminaire d'Études Médiévales Hispaniques de l'Université de Paris XIII, 1979. Pp. 307.

The work includes an "Introduction" (1-10), a "Dictionnaire lemmatisé des références" (13-242), an "Index des formes classées par ordre de fréquence" (244-68), an "Index alphabétique des formes" (271-306), and a Table of Contents (307).

Pellen's conceptual and theoretical framework situates lexicography in its "véritable contexte" (2) and explores the possibilities that information processing offers for a more rational

knowledge of texts and languages. He posits a symbiotic relationship between the examination of concordances, linguistic and critical studies of the poem, and of its historical period, leading to improved comprehension of the text and resultant modification of information contained in it. His experiment, based on Ramón Menéndez Pidal's paléographie edition of the *Cantar de Mio Cid,* involves a "nouvel enregistrement du texte" based on a cohesive system of conventions. Throughout the discussion of the methodology, Pellen exhibits a profound respect for the integrity of the paléographie text. "Les descriminations qu'il concrétise relèvent, non plus de la présentation matérielle, mais de la lecture, de l'interprétation que le chercheur retient par chaque détail du texte" (2).

"Codifying" *(coder)* often can involve assigning a diacritical mark to one of a pair of homographs; it contrasts with "codification" *(codage),* which already begins when one divides words. Pellen's traditionalist orientation underlies his approach to this project: "Au premier niveau de l'établissement d'un texte en vue de sa mémorisation, personne n'échappe donc au codage—et de toute façon l'orthographe est aussi un codage" (4). Since the transcription of an oral text entails orthography, a codification process previously has taken place and the forms presented in the dictionary are already "reconstituées" (4).

The indexing program for the dictionary, "Programme d'établissement de concordances et d'index," was written in Fortran by Pellen with the help of an engineer, M. C. Siredey. It features the capacity to produce (1) indices ordered either alphabetically, by frequency, by length, and in reverse; (2) concordances with several contexts and partial possibility of editing; and (3) a reference index citing each form with all its uses in the text. The lemmatized dictionary of forms, which had been numbered and ordered alphabetically, was merged with the non-lemmatized index of references; the resulting list was analyzed from an aesthetic point of view to eliminate useless marks and to retain the utilitarian forms which are deciphered by a code.

The introduction explains that homographs for the first and third person forms ORTOGO 1 and ORTOGO 3 are differentiated; proper nouns are marked with an asterisk (CASTEION*), and truncated forms marked ≠ (ESPADO≠ 'espadado'). Unfortunately,

the symbols #, $, and + following some words remain to be explained anywhere in the dictionary: CASTEION#, D#, DE#, DESTELLADO#, DIA#, DO#, DON#, E#, ENTRAMOS#, FUERES#, FUERO#, HY#, L#, LA#, ONDREDES#, OUIESTES#, PIDIMOS#, SEN=AS#, SSEN#, SI#, SOMO#, SUELTA#, TOD#, TOMA#, TORNA#, VENGA#, VESTID#, VN#; DECIR+, Y+, YA+, MAS+, NOMBRE+, O+, ESCAPARA $ 3, FIZIERADES $, JUGARA $ 3, LIDIARAN $ [person not indicated]. To complicate this apparent oversight, the copy examined for this review was faultily printed on pages 276-77, partially truncating word endings, with the result that the letters appearing after the lacunae easily could have been construed as additional unexplained symbols.

Following other readers' suggestions, homographs with different grammatical functions (e.g., TANTO adj./adv.) are indicated and listed separately. For example, the use of UNO as numeral is separated from the article and the indefinite. A short list of words with variable function (ALLENDE, DELANTE, MUCHO, POCO, TANTO, UNO) has been published in *Cahiers de linguistique hispanique médiévale 2.*

In the "Dictionnaire lemmatisé des références" each entry centers on the page and is preceded by its global frequency and followed in parentheses by its frequency in each *cantar.* The "Index des formes classées par ordre de fréquence" lists all words in descending order of frequency, alphabetically. The "Index alphabétique des formes" orders reflexes of a root alphabetically and, to its right, the main entry under which the form is listed in the "Dictionnaire lemmatisé des références." For example, this index would direct a user who cannot find TECA*, TECHA* in the "Dictionnaire lemmatisé des references" to ATECA*. In this index the page format (four equally spaced columns) fails to reflect the all-important relationship between the form and the prime entry.

Pellen suggests, and this reviewer concurs, that this research tool provides "un outil de travail d'un genre nouveau" (9), applicable in studies of linguistics, text and discourse analysis, semiology, and ideology of the epic poem of the Cid.

Nancy Joe Dyer
Texas A&M University

-o-oOo-o-