



# HAFEZ IV. LEXICAL STRUCTURE OF HAFEZ'S GHAZALS

---

## HAFEZ

### iv. LEXICAL STRUCTURE OF HAFEZ'S GHAZALS

In describing the lexical structure of Hafez's ghazals, we must consider three main problems. First, the quantitative valuation may vary, depending on the edition of the ghazals used or of the manuscript(s) chosen for the scrutiny. Second, the data resulting from lexical processing are strongly conditioned by the lexicological choices in singling out tokens, types and lemmas. (By "lemma" we mean the lexical item corresponding to the headword found in Dehḵodā's *Loḡat-nāma*; by "token" we mean any occurrence of a form of a lemma; by "type" we mean any different form a lemma may take according to inflectional or phono-morphological variants.) Third, at present there is no general description of the classical Persian poetic language, and no statistical studies enabling us to analyze deviations in the language of Hafez's ghazals with reference to average data.

Despite these limitations, it is nevertheless necessary to base textual criticism on complete and reliable lexico-statistical inventories of Hafez's ghazals. In this perspective, a simple list of types or lemmas—even if complete with the



relative frequencies—is not enough (see Şadi-qiān and Mir 'Ābedini). A computerized processing of the texts, which will guarantee a greater richness of information, coherence in lexicological choices, and precision of data, thus becomes indispensable (the only such work is Meneghini Correale, 1988, [pp. 21-35], based on the 1983 *Ḳānlarī* edition of the ghazals; the following data were extracted and processed on the basis of that study, and obviously reflect the scientific criteria assumed therein).

The general data pertaining to the lexicon of Hafez's ghazals are as follows: N (number of tokens) = 77,779; V (number of types) = 7,215, of which 3,605 are *hapax legomena* (single occurrences); VI (number of lemmas) = 4,787, of which 2,037 are *hapax legomena*. As there are 486 ghazals, with a total of 4,092 lines, we can extrapolate the following average quantities: 8.42 lines per ghazal; 160 tokens per ghazal; 19 tokens per line; 14.84 different types per ghazal; 1.76 different types per line; 9.85 different lemmas per ghazal; 1.17 different lemmas per line.

The lexicon distribution presents a structure which, on the basis of parameters tested on other linguistic systems, can be considered as regular. The total amount of frequencies of the first 100 lemmas covers 64.39 percent of the lexicon, that of the first 997 covers 89.79 percent, that of the first 2046 covers 95.5 percent. As for lexicon concentration, Hafez's ghazals show values that are typical of lyric poetry; in other words, the occurrences of the first 50 most frequent lemmas account for more than 55 percent of the total number of occurrences (N).

With respect to lexicon subdivision in full words (such as nouns, verbs, adjectives, adverbs) and empty words (such as articles, pronouns, prepositions, conjunctions), we have found that 67 percent of the vocabulary is composed of full words, 33 percent of empty words; occurrences with nominal or adjectival function cover over 46 percent of the vocabulary of Hafez's ghazals. Another important feature is the number of compound words (for criteria see Meneghini Correale, 1988, pp. 33-34): Hafez's ghazals present 1,440 different compound words (types) which account for 4 percent of the occurrences (N) and 41 percent of the total number of types (V).

As to the relationship between the quantity of types and lemmas, the high average frequency of types with reference to the extension of the lexicon (10.6) points to a tendency to introduce new words usually through employing the same lemmas. This feature is confirmed at the consolidation level (9.3



different types per 100 tokens). Both these data are affected by the great number of types occurring just once (*hapax legomena*). This characteristic is further confirmed by the regular and constant introduction of new types: each ghazal presents an average of 14 new types, as has been shown by Zipoli (1990). However, we must keep in mind that the choice of poetic lexicon was strongly influenced by predetermined events (rhyme, *radif*, figures of speech, etc.), which condition the structure of the poems (the lexical elements of the *radif* may, for example, account for up to 15 percent of the lexicon of a single ghazal). The poetic constraints and the strict coherence of a poetry with set themes are therefore particularly important in the lexical universe of Hafez's ghazals (see Meneghini Correale, 1991).

## BIBLIOGRAPHY

---

Ḥāfez, *Divān*, ed. Abu'l-Qāsem Enjavi Širāzi, Tehran, 1361 Š./1982 (includes an alphabetical list of full words with indication of the corresponding lines).

Susan M. Hockey, "A Concordance to the Poems of Hafiz With Output in Persian Characters," in A. J. Aitken et al., eds., *The Computer and Literary Studies*, Edinburgh, 1973, pp. 291-306.

Alan Jones, "Producing a Concordance of the Divan of Hafez by Computer," in *Convegno Internazionale sulla Poesia di Hafez*, Rome, 1978, pp. 99-110.

Daniela Meneghini Correale, *The Ghazals of Hafez: Concordance and Vocabulary*, Rome, 1988.

Idem, G. Urbani, and R. Zipoli, *Handbook of Lirica Persica*, Venice, 1989 (describes a computer-assisted method for textual analysis of Persian poetry uses in the "Lir-ica Persica" project).

Idem, *Hafez: Concordance and Lexical Repertories of 1,000 Lines*, Venice, 1989.

Idem, "Quelques observations sur la structure lexicale des ghazals de Ḥāfiz,"



in Michael Glünz and Johann-Christoph Bürgel, eds., *Intoxication, Heavenly and Earthly: Seven Studies on the Poet Ḥāfiẓ of Shiraz*, Bern, 1991, pp. 105-36.

Idem, "Farroxi, Hafez, Taleb: dati per un'analisi comparativa del lessico," Ph.D. diss., Istituto Universitario Orientale, Naples, 1992.

Abu'l-Faẓl Moṣaffā, *Farhang-e dah hezār vāža az divān-e Ḥāfezá*, Tehran, 1369 Š./1990.

Mahindoḡt Ṣadiqiān with Abu Ṭāleb Mir-Ābedini, *Farhang-e vāža-nāma-ye Ḥāfezá*, Tehran, 1366 Š./1987.

Riccardo Zipoli, "Tecniche informatiche e lirica neopersiana: dalle concordanze di Ḥāfez a Lirica Persica," *Annali di Ca' Foscari* 29, 1990, serie orientale 21, pp. 169-91.

Idem, "Textual Solidarity in the Ghazals of Hafez," in Iraj Afšār and Hans R. Roemer, eds., *Soḡanvāra*, Tehran, 1988, pp. 153-69.

Idem, *Statistics and Lirica Persica*, Venice, 1992.