

the political work and their collaborators in the case of more serious disorders, demonstrations or public incidents were arrested on the spot, as it happened in December 1942. The conservative elements in the Iranian government mostly represented by people „of old-age“ did their best to remove persons of active agents of parties standing in their way and to destroy the parties themselves as the potential and real danger for the traditional system.

However there was no longer any means to stop the impelled wheel of history once set to movement. The most dynamic and aware of its ends People's Party of Iran imposed a rate of the political and social development on other parties and forced them to a more or less fruitful rivalry.

All the parties of Iran in the years 1941—1946 had then fulfilled their historical task: they were the awakeners of a new generation of Iranians and an excellent school educating phalanges of future fighters for public freedom within the constitutional-parliamentary system.

WOJCIECH SKALMOWSKI

EIN BEITRAG ZUR STATISTIK DER ARABISCHEN LEHNWÖRTER IM NEUPERSICHEN

Der vorliegende Beitrag befasst sich mit quantitativer Untersuchung des arabischen Lehnwortschatzes in den persischen Ghaselen von Sa'di und Hāfiz. Zahlenmässige Angaben über den Umfang und Verwendungsbereich des arabischen Wortschatzes bei einzelnen persischen Autoren sind aus zwei Gründen interessant: 1^o sie charakterisieren die Sprache (parole) eines gegebenen Autors und gehören somit in den Bereich der Stylometrie; 2^o als vergleichbare Grössen können sie Auskunft über die Rolle des Arabischen in einzelnen Perioden der Geschichte des Neupersischen geben und dürfen deshalb zu den Hilfsmitteln der Sprachsystemanalyse eingerechnet werden.

Bis unlängst fehlten derartige Untersuchungen fast vollständig; wenn man von den statistisch unvollständig ausgearbeiteten Untersuchungen über die arabischen LW bei Firdausi von Horn¹ und Humbert² absieht, liegt erst seit 1960 eine Pionierarbeit auf diesem Gebiet, ein Aufsatz von R. Koppe u. d. T. *Statistik und Semantik der arabischen Lehnwörter in der Sprache 'Alawī's'*³ vor. Der vorliegende Beitrag wurde als eine Ergänzung zu diesem Aufsatz und eine Erweiterung der Untersuchung auf ältere Schichten des Neupersischen beabsichtigt.

MATERIAL

Die Grundsätze der Zusammenstellung des Materials wurden dem Aufsatz Koppe's entnommen (s. da, S. 588 ff.). Die Unter-

¹ *Grundriss der iranischen Philologie*, 1895, Bd. I, 2, S. 2ff.

² P. Humbert, *Observations sur le vocabulaire arabe du Châch-nâmeh*, 1953.

³ „Wissenschaftliche Zeitschrift der Humboldt-Univ. zu Berlin“, Ges.-Sprachwiss. R. IX, 1959/60, S. 585—619.

suchung wurde auf Textproben durchgeführt. Für jeden der beiden Dichter — mit S. und H. bezeichnet — wurden je zwei Proben genommen: Probe I bestand jeweils aus einzelnen Versen (jeweils wurde nur der zweite *beit* eines Ghasels herausgenommen); Probe II bestand jeweils aus ganzen Ghaselen. Die Anzahl der Wortvorkommen in den Proben — mit N bezeichnet — schwankte von 1185 bis 2772. Für jede Probe wurden folgende Grössen gesucht:

1. die Gesamtanzahl der persischen Wortvorkommen — N_p
2. die Gesamtanzahl der arabischen Wortvorkommen — N_a
3. die Anzahl der lexikalischen Einheiten (Wortschatz) — V
4. der persische Wortschatz — V_p
5. der arabische Wortschatz — V_a
6. die Anzahl der lexikalischen Einheiten, die nur einmal auftraten (hapax legomena) und zwar für:
 - a) den persischen Wortschatz — V'_p
 - b) den arabischen Wortschatz — V'_a

Die Ergebnisse der Zählung werden in der Tabelle I dargestellt. Nach freundlicher Erlaubnis von Herrn Koppe werden auch die entsprechenden Zahlen für die Textproben aus 'Alawī's Werken (Pr. I — *Čašmhāyaš*, Pr. II — *Nāmehā*) aus seiner Arbeit (op. cit. S. 590 ff) abgedruckt.

Tab. I

	S. I	S. II	H. I	H. II	'A. I	'A. II
N	2240	1925	2772	1185	17370	9691
N_p	1854	1532	2203	937	13969	8572
N_a	386	393	569	248	3401	1119
V	646	663	881	476	2177	1821
V_p	392	393	503	282	1149	990
V_a	254	270	378	194	1028	831
V'_p	208	235	264	177	500	496
V'_a	192	205	277	155	514	496

AUSWERTUNG

Zur Bestimmung des Verhältnisses einzelner Proben zueinander wurden sie in bezug auf die folgenden sie charakterisierenden Grössen: V'_p , ($V_p - V'_p$), V'_a , ($V_a - V'_a$) mit Hilfe des χ^2 — Testes miteinander verglichen. Der Test erlaubt die Wahrscheinlichkeit P festzustellen, mit welcher die beobachteten Differenzen als zufällige Abweichungen betrachtet werden können; im Allgemeinen nimmt man $P = 0,05$ als Grenze zwischen wesentlichen und unwesentlichen Differenzen an¹.

Proben I und II für S. und H. können als einheitlich bezeichnet werden: für S. ist $P > 0,5$, für H. $P > 0,05$; ebenso die Proben S. II, H. I und S. II, H. II: $P > 0,3$; dagegen die beiden Proben von 'Alawī weichen stark voneinander ab: $P \ll 0,01$.

Das Verhältnis des arabischen Wortschatzes zum gesamten Wortschatz $V_a : V$ und der Anteil der arabischen Wortvorkommen am Gesamtwortvorkommen $N_a : N$ werden in Prozenten ausgedrückt, s. Tab. II:

Tab. II

	S. I	S. II	H. I	H. II	'A. I	'A. II
$V_a : V$	39,16	40,72	42,90	40,75	47,23	45,58
$N_a : N$	17,23	20,41	20,53	20,41	19,57	11,56

Um eine gründlichere Einsicht in die Rolle des arabischen Wortschatzes in der Sprache der untersuchten Autoren zu gewinnen ist die Kenntnis der Häufigkeitsverteilung einzelner Wörter notwendig. Zur Beschreibung der Verteilungen arabischer und persischer Wörter in den untersuchten Proben dient hier die Grösse ϑ , welche die sogen. Texttemperatur ausdrückt. Der Terminus ist von B. Mandelbrot eingeführt worden und bezeichnet $1/B$, d. h. den reziproken Wert einer der Konstanten in der Gleichung: $f_r = P \cdot (r + \varrho)^B$, welche eine modifizierte Formulierung des sogen. Zipfschen Gesetzes darstellt². Die zahlenmässigen Werte

¹ Vgl. G. Herdan, *Language as Choice and Chance*, 1956, S. 88 ff.

² S. u. a. *Structure formelle des textes et communication*, Word, 10, 1954, S. 1—27.

der Variablen dieser Gleichung: f_r — der Häufigkeit des r -ten Wortes, und r — der Stellung in der Reihenordnung fallender Häufigkeiten, können für einen gegebenen Text empirisch bestimmt werden. Obwohl die Theorie, die das Zipfsche Gesetz erklären soll noch diskutabel ist¹, stellt die Formel selbst ein bequemes Mittel zur Beschreibung der Häufigkeitsverteilungen der Wörter dar.

Die Texttemperatur wurde für den persischen und arabischen Teil des Wortschatzes einzelner Proben errechnet (ϑ_p und ϑ_a bezeichnet). Für die Proben 'A. I und 'A. II wurden die Häufigkeitstabellen im Aufsatz Koppe's (op. cit. S. 592) benutzt; die Probe H. II wurde wegen kleinen N nicht berücksichtigt. Die gewonnenen Werte wurden in der Tab. III zusammengestellt.

Tab. III

	S. I	S. II	H. I	A. I	A. II
ϑ_p	0,90	0,86	0,96	0,72	0,80
ϑ_a	1,73	1,80	1,73	1,73	1,73

ERGEBNISSE

Das Verhältnis des Arabischen zum Persischen in den Ghaselen von Sa'di und Hāfiz kann als identisch bezeichnet werden. Die Wörter arabischer Herkunft bilden ca. 40% ihres Wortschatzes und ca. 20% ihrer Wortvorkommen.

In bezug auf die arabischen LW weist der Wortschatz 'Alawī's keine wesentlichen Unterschiede mit dem der klassischen Ghaselen auf. Das Verhältnis $V_a : V$ in den Proben S. I, 'A. II und H. II, 'A. II gibt χ^2 — Werte, denen $P > 0,05$ und $P > 0,1$ entsprechen. Der ein wenig grössere Anteil der LW im 'Alawī's Wortschatz kann als stil- und themabedingt angesehen werden — dafür spricht die Tatsache, dass die Texttemperatur der arabischen LW in allen Proben beinahe gleich ist. Dieselbe Tatsache zeugt auch davon, dass das arabische Wortschatz als Grundwortschatz verwendet

wird — sowohl in den Ghaselen, als auch im modernen Neupersischen. Die stilistischen Unterschiede zwischen den untersuchten Autoren drücken sich viel deutlicher in Verwendung des rein persischen Wortschatzes aus, wie es die ϑ_p — Werte zeigen.

Es kann zusammenfassend gesagt werden, dass das moderne literarische Neupersische, dessen 'Alawī einer der angesehensten Vertreter ist, in bezug auf die Anwendung der arabischen LW der Sprache der Denkmäler klassischer Literatur folgt. Dies bekräftigt die These Koppe's (op. cit. S. 593), ein absolut „reines Persisch“ (پارسی سره) sei unmöglich. Auf der anderen Seite die Untersuchung zeigt, dass die Auswüchse eines übermässigen, modebedingten Arabisierens, die von Zeit zu Zeit in der persischen Literatur auftauchten, keinen dauernden Einfluss hinterlassen haben. Das Eindringen des Arabischen scheint sich einmalig und zwar noch in der Zeit der Herausbildung des Neupersischen vollzogen zu haben; seitdem bilden die arabischen LW einen organischen Teil dieser Sprache.

¹ Vgl. G. Herdan, *Type/Token Mathematics*, 1960, S. 35 ff.