

IST DER WORTSCHATZ DES "LEBENSMÜDEN"
GRÖßER ALS DER DES "SINUHE"?

In seiner Bearbeitung des "Lebensmüden" hat W. BARTA [1] Wortschatzstatistiken über den "Lebensmüden" vorgelegt, die denen W. WESTENDORFS [2] für den "Sinuhe" im Aufbau entsprechen und somit einen Vergleich des Wortschatzes der beiden Literaturwerke ermöglichen.

Unter anderem stellt W. BARTA fest, daß im "Lebensmüden" 63 %, im "Sinuhe" dagegen nur 54 % der verschiedenen Wörter ein einziges Mal auftreten. Aus dem wesentlich größeren Anteil der nur einmal belegten Wörter im "Lebensmüden" zieht er den Schluß, daß der Wortschatz des "Lebensmüden" "bedeutend umfangreicher" ist als der des "Sinuhe".

Dieser Schluß ist unzulässig. Die verglichenen Prozentsätze sind nämlich, wie sich aus den Beobachtungen der linguistischen Statistik unschwer ableiten läßt, von der Länge des ausgezählten Textes abhängig: der Prozentsatz nimmt mit der Länge des Textes ab. Da nun der "Sinuhe" mit einer Textlänge von 2699 Wörtern fast dreimal so groß ist wie der "Lebensmüde" mit 958 Wörtern, ist die Richtung der Divergenz in den verglichenen Prozentsätzen zu erwarten. Ob die festgestellten Prozentsätze in der Höhe ihrer Abweichung allein mit der unterschiedlichen Textlänge zusammenhängen oder ob die Abweichung größer oder kleiner ist als von der unterschiedlichen

Textlänge her zu erwarten, läßt sich überhaupt nur dann entscheiden, wenn man Charakteristika heranzieht, die unabhängig sind von der zufällig verschiedenen Länge der verglichenen Texte.

Nach G. HERDAN [3, 4] verhalten sich Textlänge und Wortschatz folgendermaßen zueinander: Der Wortschatz nimmt mit der Textlänge zu, jedoch mit sich vermindernder Geschwindigkeit. Mathematische Überlegungen und empirische Erfahrungen lassen eine Präzisierung zu: Die Zuwachsrate zu einem beliebigen Zeitpunkt bei der Auszählung des Textes ist umgekehrt proportional zur jeweils erreichten Textlänge. Die von G. HERDAN als von der Textlänge unabhängig angegebene Konstante zur Charakterisierung der Größe eines Wortschatzes ist

$$C = \frac{\log V}{\log N}$$

wobei V die Anzahl der verschiedenen Wörter des Textes, N die Gesamtzahl der Wörter des Textes ist.

Für den "Lebensmüden" und den "Sinuhe" ergeben sich folgende Konstanten:

$$\text{"Lebensmüden": } C_L = \frac{\log 334}{\log 958} = 0,8465$$

$$\text{"Sinuhe": } C_S = \frac{\log 727}{\log 2699} = 0,8340$$

Ob die Differenz zwischen den für den "Lebensmüden" und den "Sinuhe" berechneten Konstanten als signifikanter Unterschied zwischen dem Wortschatz der beiden Literaturwerke zu interpretieren ist, läßt sich nicht mit Sicherheit entscheiden, weil über die Streuung der Größe des Wortschatzes altägyptischer literarischer Texte bis jetzt keine Beobachtungen vorliegen. Zieht man zur Beurteilung der Differenz die Verhältnisse nicht-ägyptischer Literaturwerke, z. B. die des Neuen Testaments [3] heran, so erscheint es als höchst unwahrscheinlich, daß die geringe Differenz zwischen dem "Lebensmüden" und dem "Sinuhe" signifikant ist. Mit anderen Worten: Es läßt sich nicht nachweisen, daß "Lebensmüder" und "Sinuhe" über einen unterschiedlich umfangreichen Wortschatz verfügen.

Eine andere Frage, auf die ich keine Antwort geben kann, bleibt die, in welchem Ausmaß bei doch relativ kurzen Texten wie etwa dem "Lebensmüden" statistische Unschärfen sich bemerkbar machen.

Es ist denkbar, daß die Stilometrie beim Vergleich altägyptischer Literaturwerke gute Dienste leisten kann, z. B. auch bei der Überprüfung der oft versuchten Zuschreibung verschiedener Literaturwerke an ein und denselben Autor. Nur wird man sich dabei schon an den Methoden der linguistischen Statistik orientieren müssen. Die Grenzen der Trivialstatistik sind, wie das Beispiel des Wortschatzes zeigt, eng gezogen.

Literatur:

- [1] W. BARTA, Das Gespräch eines Mannes mit seinem BA, MÄS 18, Berlin 1969, 122 - 125.
- [2] H. GRAPOW, Untersuchungen zur ägyptischen Stilistik I, Der stilistische Bau der Geschichte des Sinuhe, Deutsche Akademie der Wissenschaften zu Berlin, Institut für Orientforschung, Veröffentlichung Nr. 10, Berlin 1952, 122 - 124.
- [3] G. HERDAN, Type-token mathematics, Janua linguarum, Series maior 4, 's-Gravenhage 1960, 26 - 33.
- [4] G. HERDAN, The advanced theory of language als choice and chance, Kommunikation und Kybernetik in Einzeldarstellungen 4, Berlin - Heidelberg - New York 1966, 75 - 77.

Wolfgang Schenkel