Jermo van Nes*

Hapax Legomena in Disputed Pauline Letters: A Reassessment

Abstract: To argue for the pseudonymity of disputed Pauline letters on the (partial) basis of their disproportional high number of *hapax legomena* continues to be common practice among New Testament scholars. By means of linear regression analysis, it is shown that only 1 and 2 Timothy use significantly more *hapaxes* than the undisputed Pauline letters. If, however, proper noun *hapaxes*, *hapaxes* taken from citations, and *hapaxes* which – upon removal of its preposition – have a common lexeme are not taken into consideration, it appears that none of the disputed Pauline letters use significantly more *hapaxes* than the undisputed Pauline letters.

Zusammenfassung: In der neutestamentlichen Forschung ist es nach wie vor üblich, als Argument für die Pseudonymität umstrittener Paulusbriefe (unter anderem) deren disproportional große Anzahl von *Hapax Legomena* anzuführen. Unter Anwendung der linearen Regressionsanalyse wird im vorliegenden Artikel aufgezeigt, dass im Vergleich mit den unumstrittenen Paulusbriefen lediglich die beiden Timotheusbriefe jeweils signifikant mehr *Hapax Legomena* aufweisen. Sieht man jedoch von *Hapax Legomena* ab, bei denen es sich entweder um Eigennamen handelt, welche sich in Zitaten finden und/oder welche bei gleichem Lexem nur in einer unterschiedlichen Präposition bestehen, dann scheint es sogar so zu sein, dass im Hinblick auf die durchschnittliche Anzahl von *Hapax Legomena* keiner der umstrittenen Paulusbriefe im Vergleich mit den unumstrittenen Briefen des Apostels wesentlich nach oben abweicht.

Keywords: (un)disputed Paulines, *hapax legomena*, linear regression analysis

*Corresponding author: Jermo van Nes, Evangelische Theologische Faculteit, Sint Jansbergsesteenweg 95, B-3001 Leuven, Belgium; jermo.vannes@etf.edu

Introduction

Hapax legomena (or *hapaxes*), usually referring to "words found only once in a particular text,"[1] have recently been called "a neglected field of study"[2] in biblical scholarship. While many *hapaxes* have not been studied thoroughly, they have played a major role in Pauline authorship studies. Serving as one argument among many, scholars since the turn of the nineteenth century have argued that a disproportionate high number of *hapaxes* in the disputed Pauline letters evidences their pseudonymity.[3] This applies in particular to the Pastoral Epistles or Pastorals (1–2 Timothy, and Titus),[4] but it is not uncommon to see the argument being used in support of the pseudonymity of Ephesians, Colossians, and 2 Thessalonians as well.[5]

---

1 David E. Aune, Hapax legomenon, in: idem, The Westminster Dictionary of New Testament and Early Christian Literature and Rhetoric, Louisville (KY)/London 2003, 210–211.

2 Hellen Mardaga, Hapax Legomena: A Neglected Field in Biblical Studies, CBR 10 (2012) 271.

3 While Edward Evanson (The Dissonance of the Four Generally Received Evangelists and the Evidence of Their Respective Authenticity Examined, Ipswich 1792, 261–263.267–279) was probably the first to dispute the Pauline authorship of Ephesians, Colossians, and Titus on the (partial) basis of their peculiar vocabulary, Friedrich D. E. Schleiermacher (Ueber den sogenannten ersten Brief des Paulos an den Timotheos. Ein kritisches Sendschreiben an J. C. Gass, Berlin 1807, 27–74) pioneered the *hapax* argument by pointed to over 60 unique expressions in 1 Timothy in order to prove its pseudonymous character.

4 See, for example, Wolfgang Schenk, Die Briefe an Timotheus I und II und an Titus (Pastoralbriefe) in der neueren Forschung (1945–1985), ANRW II.25.4 (1987) 3409; Jürgen Roloff, Der erste Brief an Timotheus (EKK 15), Düsseldorf/Neukirchen-Vluyn 1988, 28–30; Jerome D. Quinn, Epistles to Timothy and Titus, ABD 6 (1990) 563; Cesare Marcheselli-Casale, Le Lettere Pastorali. Le due lettere a Timoteo e la lettera a Tito (SOC 15), Bologna 1995, 23; James D. G. Dunn, The First and Second Letters to Timothy and the Letter to Titus (NIB 11), Nashville (TN) 2000, 778; Lorenz Oberlinner, Die Pastoralbriefe. Kommentar zum ersten Timotheusbrief (HThKNT 11/2), Fribourg/Basel/Vienna 1994, repr. 2002, xxxvii; Raymond F. Collins, I & II Timothy and Titus. A Commentary (NTL), Louisville (KY) 2002, 2; Alfons Weiser, Der zweite Brief an Timotheus (EKK 16/1), Düsseldorf/Neukirchen-Vluyn 2003, 58; I. Howard Marshall, The Pastoral Epistles (ICC), London/New York 1999, repr. 2004, 60–61; Bart D. Ehrman, The New Testament: A Historical Introduction to the Early Christian Writings, New York ⁶2016, 453–454; Udo Schnelle, Einleitung in das Neue Testament, Göttingen ⁹2017, 407.

5 So, for example, Earl J. Richard, First and Second Thessalonians (SP 11), Collegeville (MN) 1995, 22–23; Pheme Perkins, Ephesians (ANTC), Nashville (TN) 1997, 20; Joachim Gnilka, Der Kolosserbrief (HThKNT 10), Fribourg/Basel/Vienna 1980, repr. 2002, 16–17; B. Heininger, Die Rezeption des Paulus im 1. Jahrhundert, in: Oda Wischmeyer (ed.), Paulus. Leben – Umwelt – Werk – Briefe (UTB 2767), Tübingen ²2012, 353; M. Eugene Boring, An Introduction to the New Testament: History, Literature, Theology, Louisville (KY) 2012, 332–333; Charles B. Puskas and Mark Reasoner, The Letters of Paul: An Introduction, Collegeville (MN) ²2013, 145.158.

In this study, however, it is aimed to reassess the evidence for the commonly held view among New Testament scholars that the disputed Paulines have a disproportionate high number of *hapaxes* in their texts.[6] To this end, the first part ties in with a recent discussion on the definition of *hapaxes*. The second part introduces a neglected statistical method to analyze the number of *hapaxes* in the *Corpus Paulinum*. In the third and final part, some critical factors are discussed which affect the number of *hapaxes* and call into question whether they should have any discriminating force.

1 Terminological Considerations

In light of the scholarly weight ascribed to *hapax legomena* in Pauline authorship studies, it is surprising to see few scholars properly define the term. For example, Percival Neale Harrison in his often-cited and influential studies seems to understand *hapaxes* to refer to all words (excluding proper nouns) found in the Pastorals but not in any of the other Paulines, even if they are used more than once.[7] Such a definition violates the nature of *hapaxes* as it allows them to be found in more than one text, and more than once in those texts. In order to avoid such confusion, Hellen Mardaga argued a few years ago that a *hapax legomenon* in biblical studies is best understood as a word that is used only once in a specific biblical corpus, including the Masoretic Text, LXX, or New Testament.[8]

Thomas Kraus, however, countered that the textual reference frames as set by Mardaga in terms of the Masoretic Text, LXX, or New Testament are too broad, because they include different genres, have different socio-cultural origins, and were written in different time periods.[9] He also thinks Mardaga's definition of *hapaxes* excludes the possibility to look for words distinctive to a single text (e. g.

---

6 The argument is a refined version of the section on *hapax legomena* in Jermo van Nes, Pauline Language and the Pastoral Epistles: A Study of Linguistic Variation in the *Corpus Paulinum* (LBS 16), Leiden/Boston 2018.

7 Percival Neale Harrison, The Problem of the Pastoral Epistles, London 1921, 68–70; idem, Important Hypotheses Reconsidered. III. The Authorship of the Pastoral Epistles, ExpTim 67 (1955) 77–81; idem, Paulines and Pastorals, London 1964, 9–30. For the critical reception of Harrison's work, see Jermo van Nes, The Problem of the Pastoral Epistles: An Important Hypothesis Reconsidered, in: Stanley E. Porter/Gregory P. Fewster (eds.), Paul and Pseudepigraphy (PAST 8), Leiden/Boston 2013, 153–169.

8 Mardaga, *Hapax Legomena* (see n. 2), 264–274.

9 Thomas J. Kraus, *Hapax Legomena*: Definition eines *terminus technicus* und Signifikanz für eine pragmatisch orientierte Sprachanalyse, NTS 59 (2013) 545–564.

2 Peter), homogeneous corpus (e. g. Johannine letters), or heterogeneous texts of the same corpus (e. g. Genesis and Ezekiel). For Kraus, the multiple occurrence of a word in a single text of a well-defined text corpus is equally, if not more, distinctive than single attestations of particular words. Accordingly, he suggests using a more pragmatic definition of *hapaxes* that suits the purpose(s) of one's study:

> Ein *Hap[ax] leg[omenon]* ist somit nur dahingehend ein Einzelwort, als es in einem Text durchaus mehrmals auftreten darf, jedoch dann in einem bestimmten sprachlichen Referenzrahmen sonst nicht vorkommt.[10]

Following Kraus's definition, *hapaxes* for the purpose of this study can be defined in the context of the *Corpus Paulinum*, which is the set frame of textual reference. These "Pauline" *hapaxes* may refer to words found exclusively, whether once or more, in a Pauline letter but not in any of the other twelve.

   Yet the inclusion of words that are used more than once in a Pauline letter alerts to a missing element in Kraus' definition of *hapaxes*. What if some of these words semantically differ, and, despite their multiple occurrences, become semantically unique at some places? According to the standard English-Greek lexicon by Walter Bauer,[11] there are 31 of such words in the *Corpus Paulinum*: 6 in Romans,[12] 12 in 1 Corinthians,[13] 3 in 2 Corinthians,[14] 2 in Gala-

---

10 Kraus, *Hapax Legomena* (see n. 9), 555–556.

11 Walter Bauer, A Greek-English Lexicon of the New Testament and other Early Christian Literature (ed. Frederick William Danker; transl. William F. Arndt, F. Wilbur Gingrich, and Frederick William Danker), Chicago (IL) ³2000.

12 Ἀδύνατος = "impossible" (8,3) and "powerless" (15,1); ἀσύνετος = "foolish" (1,31; 10,19) and "senseless" (1,21); δικαίωμα = "requirement" (1,32; 2,26; 8,4), "justification" (5,16), and "righteous deed" (5,18); ἐκκλίνειν = "turning aside" (3,12) and "to avoid" (16,17); κοίτη = "sexual intercourse" (13,13) and "seminal emission" (9,10); νικᾶν = "winning (a legal case)" (3,4) and "to (be) overcome" (12,21).

13 Ἀγοράζειν = "to acquire something (as property in the sense of securing one's rights)" (6,20; 7,23), and "to buy something (in exchange for money)" (7,30); ἀνά = "between" (6,5) and "in turn" (14,27); ἀνακρίνειν = "to judge" (2,14–15; 14,24), "to examine" (4,3–4; 9,3), and "to ask questions" (10.25.27); ἄτιμος = "in disrepute" (4,10) and "insignificant" (12,23); ἄφωνος = "without sound" (12,2) and "without meaning" (14,10); γάλα = "doctrine (figurative)" (3,2) and "milk (literally)" (9,7); θύειν = "ceremonial killing" (5,7) and "(the practice of) sacrificing" (10,20); μέλει = "source of concern" (7,21) and "(someone's) interest" (9,9); μετέχειν = "to have a share" (9.10.12) and "to partake" (in something in common with someone)" (10.17.21.30); ὅλως = "actually" (5,1; 6,7) and "completely" (15,2); συνήθεια = "accustomed" (8,7) and "habit" (11,16); ψυχικός = "unspiritual" (2,14) and "physical" (15,44.46).

14 Ἐκδημεῖν = "being away" (5.6.9) and "to leave" (5,8); καθαίρεσις = "destruction" (10,4) and "disability" (10,8; 13,10); φωτισμός = "enlightenment" (4,4) and "revealing" (4,6).

tians,[15] 2 in Ephesians,[16] 1 in Colossians,[17] 3 in 1 Timothy,[18] 1 in 2 Timothy,[19] and 1 in Titus.[20] These so-called "semantic *hapaxes*" suggest that Kraus's definition is in need of refinement. *Hapaxes* need not only be defined in relation to a set frame of textual reference and their number of attestation, but also in relation to their semantics.

Given these considerations, *hapaxes* for the purpose of this study are best understood as words that are univocal in meaning and that are used exclusively, whether once or more, in one Pauline letter and not in any of the other twelve. They need not necessarily be similar in form, but must be of the same lexeme. Ἀτιμάζειν, for example, qualifies as one *hapax*, since both ἀτιμάζεσθαι (Rom 1,24) and ἀτιμάζεις (Rom 2,23) are used exclusively in Romans in the sense of "to dishonor"; ἀδύνατος, on the other hand, qualifies as two *hapaxes*, since ἀδύνατον (Rom 8,3) and ἀδυνάτων (Rom 15,1) have the meanings of "impossible" and "powerless" respectively.

2 Statistical Analysis

The comparative analysis of *hapax legomena* in the *Corpus Paulinum* in relation to questions of authorship has proven a real challenge for both New Testament scholars and statisticians over the past decades. The use of statistics in Pauline authorship studies dates back to 1851 at least,[21] but became in vogue in the twentieth century, especially with the advent of the computer in the 1950's.[22] Issues

---

15 Προσανατίθεσθαι = "to consult with" (1,16) and "to contribute" (2,6); ὠδίνειν = "to have birth-pains (literally)" (4,27) and "to suffer (figurative)" (4,19).
16 Ἐκτρέφειν = "to nourish" (5,29) and "to bring up" (6,4); ὕψος = "height" (3,18) and "heaven" (4,8).
17 Ἀπεκδύεσθαι = "to disarm" (2,15) and "to take off" (3,9).
18 Ἐπακολουθεῖν = "devoting oneself" (5,10) and "to follow" (5,24); κόσμιος = "appropriate" (2,9) and "respectable" (3,2); προσμένειν = "to remain" (1,3) and "to continue" (5,5).
19 Κακοπαθεῖν = "suffering misfortune" (2,9) and "bearing hardship patiently" (4,5).
20 Νομικός = "matters relating to law" (3,9) and "being well informed about law" (3,13).
21 In 1851, Augustus de Morgan in one of his private letters written to a clergyman friend (which was later collated by his wife Elisabeth de Morgan in Memoir of Augustus de Morgan, London 1882, 214–216) suggested to proof the (non-)Pauline authorship of Hebrews by dividing its total number of letters by its total number of words in comparison to the results of letters claiming to be written by Paul.
22 For historical overviews and critical assessments, see *inter alia* David L. Mealand, Computers in New Testament Research: An Interim Report, JSNT 33 (1988) 106–110; Kenneth J. Neumann, The Authenticity of the Pauline Epistles in the Light of Stylostatistical Analysis (SBLDS 120),

that have proven critical in the debate concern the letters' differing text-lengths[23] and the (un)grouping of letters.[24] A statistical method that is able to overcome both issues is (simple) linear regression analysis. This type of analysis has not been applied to New Testament studies, but in the sciences probably "is the most widely used statistical technique" for "modeling the relationship between variables."[25] Linear regression analysis basically is the process of making predictions, based on the relationship between dependent and independent factors. It sets itself to the task of finding a model or equation in order to make such predictions. For the purpose of this study, linear regression analysis enables us to see whether or not the observed number of *hapaxes* in the disputed Paulines fit the predicted

Atlanta (GA) 1990, 23–114; A. Dean Forbes, Statistical Research on the Bible, ABD 6 (1992) 187–193; Matthew Brook O'Donnell, Linguistic Fingerprints or Style by Numbers? The Use of Statistics in the Discussion of Authorship of New Testament Documents, in: Stanley E. Porter/D. A. Carson (eds.), Linguistics and the New Testament: Critical Junctures (JSNTSup 168/SNTG 5), Sheffield 1999, 206–262; idem, Corpus Linguistics and the Greek of the New Testament (NTM 6), Sheffield 2005, 87–101; J. José Alviar, Recent Advances in Computational Linguistics and Their Application to Biblical Studies, NTS 54 (2008) 139–159; Michael P. Oakes, Literary Detective Work on the Computer (NLP 12), Amsterdam 2014, 175–187.

23 According to the British statistician Udny Yule (The Statistical Study of Literary Vocabulary, Cambridge 1944, repr. 2014, 2.281), the minimal number of words for literary statistics should be 10,000. Because "the sum total of the words in all three Pastoral Epistles is far less than this figure," Bruce Metzger (A Reconsideration of Certain Arguments Against the Pauline Authorship of the Pastoral Epistles, ExpTim 70 [1958] 94) once noted, "the New Testament scholar who uses the statistical method on them must at the outset resign himself to accepting results that are less than generally trustworthy." Kenneth Neumann (Authenticity [see n. 22], 135–136), however, referred to over 10 studies which used text samples ranging from 85 to 3,500 words giving accurate results. Cf. Douglas Biber, Methodological Issues Regarding Corpus-based Analyses of Linguistic Variation, LLC 5 (1990) 258–261.

24 The importance of this issue was shown by Thomas Robinson (Grayston and Herdan's "C" Quantity Formula and the Authorship of the Pastoral Epistles, NTS 30 [1984] 282–288) in his critique of a study by Kenneth Grayston and Gustav Herdan (The Authorship of the Pastorals in the Light of Statistical Linguistics, NTS 6 [1959] 1–15), arguing that the data of the Pastorals should not be analyzed together. The individual study of the Pastorals is encouraged by a growing group of scholars. Cf. Jens Herzer, Abschied vom Konsens? Die Pseudepigraphie der Pastoralbriefe als Herausforderung an die neutestamentliche Wissenschaft, TLZ 129 (2004) 1267–1282; Michel Gourgues, La recherche sur les Pastorales à un tournant?, ScEs 61 (2009) 73–86; I. Howard Marshall, The Pastoral Epistles in Recent Study, in: Andreas J. Köstenberger/Terry L. Wilder (eds.), Entrusted with the Gospel: Paul's Theology in the Pastoral Epistles, Nashville (TN) 2010, 308–309; Hans-Ulrich Weidemann, Die Pastoralbriefe, TR 81 (2016) 400–403.

25 Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining, Introduction to Linear Regression Analysis, Hoboken (NJ) [5]2012, 1.

number of *hapaxes* as based on the standard set by the undisputed Paulines.[26] The procedure can be explained in three steps.

The first step is to collect all relevant data. Based on Robert Morgenthaler's seminal study of New Testament vocabulary, table 1 lists all Pauline *hapaxes* (i. e. words that are univocal in meaning and that are used exclusively, whether once or more, in one Pauline letter and not in any of the other twelve) for both the undisputed and disputed Paulines according to their total vocabulary tokens (i. e. total number of words):[27]

Tab. 1: *Hapax legomena* in the *Corpus Paulinum*

		Vocabulary tokens (x)	Hapax legomena (y)
Undisputed Paulines	Rom	7111	281
	1Cor	6830	255
	2Cor	4477	177
	Gal	2230	91
	Phil	1629	75
	1Thess	1481	36
	Phlm	335	10
Disputed Paulines	Eph	2422	86
	1Tim	1591	131
	Col	1582	63
	2Tim	1238	102
	2Thess	823	21
	Tit	659	51

Linear regression analysis assumes a linear relationship between the independent variable (x), in this case the number of vocabulary tokens, and the dependent variable (y), in this case the number of *hapax legomena*. By means of the linear

---

26 To accept the seven undisputed Paulines as Paul's genuine writings is an important assumption in the analysis. Its weakness is that the undisputed Paulines can never become "disputed" in this regard (cf. Eduard Verhoef, The Authenticity of the Paulines Should Not Be Assumed, PzB 19 [2010] 129–151), but the genuine status of Romans, 1–2 Corinthians, Galatians, Philippians, 1 Thessalonians, and Philemon is almost universally recognized by contemporary New Testament scholars. Cf. Schnelle, Einleitung (see n. 4), 31–192.

27 Robert Morgenthaler, Statistik des neutestamentlichen Wortschatzes, Zürich ³1982, 67–157. Morgenthaler's numbers are based on the 26th edition of Nestle-Aland's *Novum Testamentum Graece*, but the Greek text of the Pauline letters has not been altered in the latest 28th edition. Table 1 modifies Morgenthaler's list in some respects. Some words have been removed, because they seem to have been listed mistakenly as *hapaxes* (e. g. ἐᾶν, ἐκχύνεσθαι, νουμηνία, ὅτε, τάχος, ὑπέρ). Other words, including semantic *hapaxes*, have been added.

equation $\hat{y} = mx + b$, whereby (m) is the slope and (b) the intercept, it is possible to estimate or predict the dependent variable ($\hat{y}$) from the independent variable (x). Figure 1 expresses this relationship:
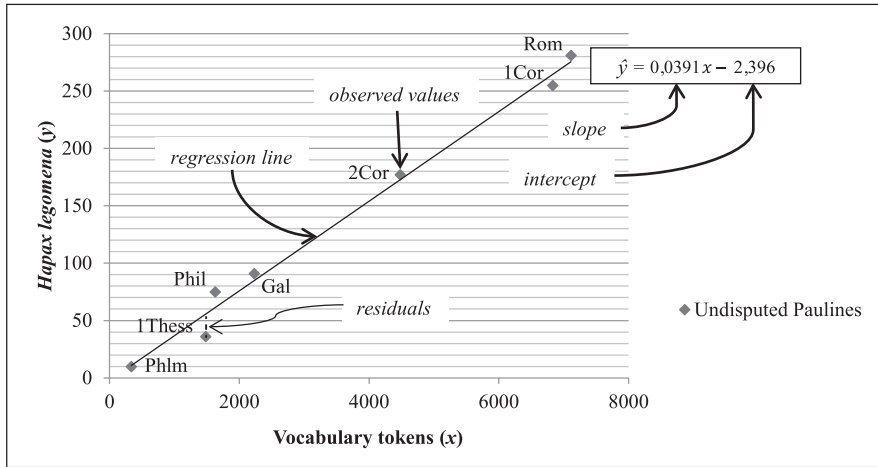

Fig. 1: Basic elements of a scatter diagram

The graph in figure 1 is called a scatter diagram. Since the undisputed Paulines are used as a touchstone for the disputed Paulines, at this moment only data of the undisputed Paulines are plotted. The regression line expresses the linear equation $\hat{y} = 0{,}0391x + 2{,}396$, and represents all predicted values for y. The differences between the observed values and the predicted values are called "residuals". The regression line is calculated in such a way that the sum of all residuals is 0. Residuals basically are "left-overs", i. e. the unexplained parts of the regression analysis. Table 2 shows which number of *hapaxes* one may expect ($\hat{y}$) for each undisputed Pauline letter:

Tab. 2: Expected *hapax legomena* in the undisputed Paulines

	Vocabulary tokens (x)	*Hapax legomena* (y)	Expected *hapax legomena* ($\hat{y}$)
Rom	7111	281	276
1Cor	6830	255	265
2Cor	4477	177	173
Gal	2230	91	85
Phil	1629	75	61
1Thess	1481	36	55
Phlm	335	10	11

The second step in this linear regression analysis involves the determination of a so-called "prediction interval" in order to see whether or not the observed *y*-values of the disputed Paulines fit the predicted values as determined by the observed *y*-values of the undisputed Paulines. A prediction interval can be plotted by means of the following formula:

$$\hat{y} + t_{\alpha/2}SE\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$

In table 3, all symbols with their corresponding formulas are explained:

Tab. 3: Prediction interval and additional statistics for a linear regression analysis

	Symbol	Equation	Explanation
Prediction interval	$\hat{y}$	$= mx + b$	estimated y-value
	$t_{\alpha/2}$	$(\alpha = 0,05; df = 5)$	T-statistic (from student's t-Distribution)
	SE	$= \sqrt{\dfrac{SS_{res}}{df}}$	standard error
	n	$= \sum (x, y)$	number (of observed values)
	x	$= (x_0), (x_1), ...$	observed x-value
	$\bar{x}$	$= \dfrac{\sum x}{n}$	average of all x-values
	SS_{xx}	$= \sum (x - \bar{x})^2$	sum of all squared deviations (x-values)
Additional statistics	m	$= \dfrac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$	slope
	b	$= \bar{y} - m\bar{x}$	intercept
	df	$= n - 2$	degrees of freedom
	SS_{res}	$= \sum (y - \hat{y})^2$	sum of the squared residuals (unexplained variation)
	SS_{regr}	$= \sum (\hat{y} - \bar{y})^2$	sum of squared regression (explained variation)
	F	$= \dfrac{SS_{regr}}{SS_{res}}$	F-statistic (from Snedecor's F-Distribution)
	R^2	$= \dfrac{1 - SS_{res}}{SS_{yy}}$	R-squared value
	SS_{yy}	$= \Sigma (y - \bar{y})^2$	sum of all squared deviations (y-values)

Prediction intervals provide the minimal (min.) and maximal (max.) value for the expected number of *hapaxes*. Once the prediction interval formula is applied, the following data are found for the undisputed Paulines:

Tab. 4: Prediction interval for *hapax legomena* in the undisputed Paulines

	Vocabulary tokens (*x*)	*Hapax legomena* (*y*)	Expected *hapax legomena* (*ŷ*)	Prediction interval	
				Min.	Max.
Rom	7111	281	276	238	313
1Cor	6830	255	265	227	302
2Cor	4477	177	173	139	207
Gal	2230	91	85	51	119
Phil	1629	75	61	27	96
1Thess	1481	36	55	21	90
Phlm	335	10	11	−26	47

Prediction intervals indicate with 95 % probability that it is not by accident that particular *y*-values lie within or outside the range. These intervals are expressed by polynomial lines in the scatter diagram:

The diagram shows the equations:

$$\hat{y} + t_{\alpha/2} SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$

$$\hat{y} = 0{,}0391x - 2{,}396$$

$$\hat{y} - t_{\alpha/2} SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$
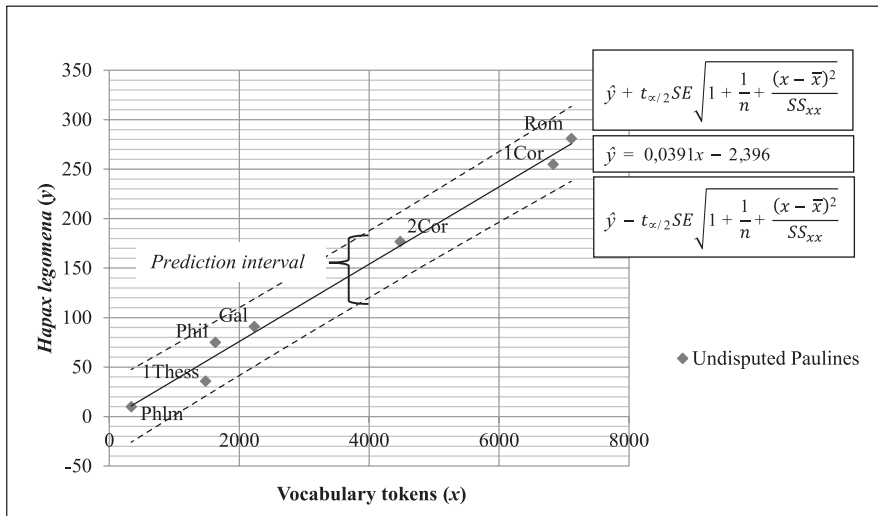
Fig. 2: Prediction interval in a scatter diagram

The third and final step in this regression analysis is to see whether the observed number of *hapaxes* for the disputed Paulines (see table 1) fit the minimal and maximal values as expressed by the prediction interval. Figure 3 shows the final result:
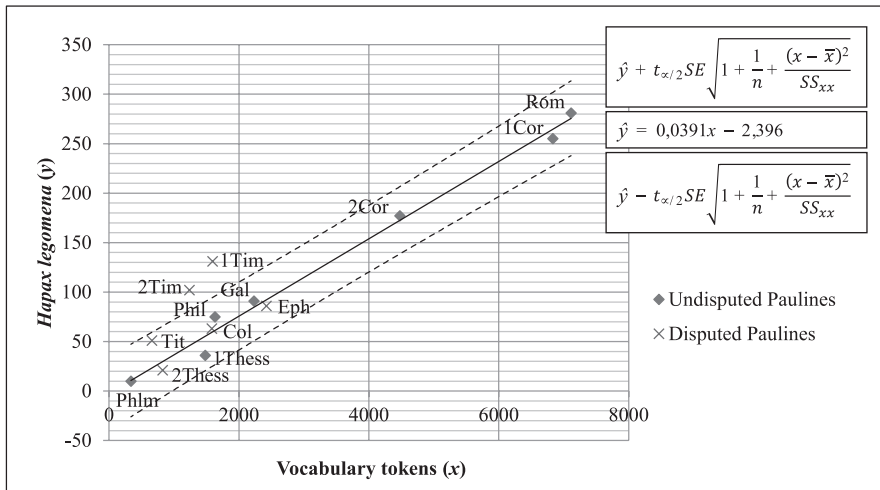

Fig. 3: *Hapax legomena* in the *Corpus Paulinum*

Figure 3 indicates that Colossians, Ephesians, and 2 Thessalonians neatly fit the undisputed Pauline pattern (as expressed by the regression line) when it comes to the use of *hapaxes*. Titus among the Pastorals also fits this pattern, be it close to the maximal prediction value. At this stage, it appears that only the numbers of *hapaxes* in the Timothy correspondence are significantly higher than in the other Paulines. This is a unique result among stylostatistical studies, challenging comments such as "[t]he vocabulary of Ephesians is [...] somewhat unusual,"[28] or "[t]he vocabulary of 2 Timothy is close to that of the other Pauline letters, whereas the diction in Titus and 1 Timothy diverges more markedly."[29] It also falsifies Kenneth Neumann's claim that *hapaxes* "are among the weakest and most ineffective stylistic criteria"[30] between the undisputed and disputed Paulines.

---

28 Raymond F. Collins, Letters That Paul Did Not Write: The Epistle to the Hebrews and the Pauline Pseudepigrapha (GNS 28), Wilmington (DE) 1988, repr. Eugene (OR) 2005, 142.
29 Luke Timothy Johnson, Letters to Paul's Delegates: 1 Timothy, 2 Timothy, Titus (NTC), Valley Forge (PA) 1996, 11.
30 Neumann, Authenticity (see n. 22), 214.

Yet linear regression analysis is not without limitations. Figure 3 shows nothing but the numerical differences between the uses of *hapaxes* in the *Corpus Paulinum*. As Douglas Montgomery, Elizabeth Peck, and Geoffrey Vining note:[31]

> [I]t is important to remember that regression analysis is part of a broader data-analytic approach to problem solving. That is, the regression equation itself may not be the primary objective of the study. It is usually more important to gain insight and understanding concerning the system generating the data.

Whereas for some scholars the results of figure 3 will affirm the pseudonymous status of the Timothy correspondence, others might want to ask what explanation(s) can be given for the differences between the linguistic data. This requires the consideration of some factors affecting the use of *hapaxes* in the *Corpus Paulinum*.

3 Critical Factors

3.1 Proper Nouns

One factor to consider is the use of proper nouns (including names and places). They are usually considered a distinct category among *hapax legomena*. Morgenthaler includes proper nouns in his New Testament vocabulary list, but also lists them in a separate table.[32] Even Harrison, who advocated the (semi-) pseudonymity of the Pastorals mainly on the basis of their disproportionate high number of *hapaxes*, excluded them from his analysis.[33] Without giving any explanation, he probably did so because there are no or very few semantic alternatives to proper nouns. Overall, there are 113 of such *hapaxes* in the *Corpus Paulinum*: 46 in Romans,[34] 6 in 1 Corinthians,[35] 5 in 2 Corinthians,[36] 12 in Gala-

---

31 Montgomery, Peck, and Vining, Regression Analysis (see n. 25), 5.
32 Morgenthaler, Statistik (see n. 27), 174.
33 Harrison, Problem (see n. 7), 20.
34 Ἀμπλιᾶτος, Ἀνδρόνικος, Ἀπελλῆς, Ἀριστόβουλος, Ἀσύγκριτος, Βάαλ, Γόμορρα, Ἐπαίνετος, Ἑρμᾶς, Ἑρμῆς, Ἠλίας, Ἡρῳδίων, Ἠσαΐας, Ἠσαῦ, Ἰακώβ, Ἰάσων, Ἰεσσαί, Ἰλλυρικόν, Ἰουλία, Ἰουνιᾶς, Κεγχρεαί, Κούαρτος, Λούκιος, Μαρία, Νάρκισσος, Νηρεύς, Ὀλυμπᾶς, Οὐρβανός, Πατροβᾶς, Περσίς, Ῥεβέκκα, Ῥοῦφος, Σάρρα, Σιών, Σόδομα, Σπανία, Στάχυς, Σωσίπατρος, Τέρτιος, Τρύφαινα, Τρυφῶσα, Φαραώ, Φιλόλογος, Φλέγων, Φοίβη, and Ὡσηέ.
35 Ἀχαϊκός, Κρίσπος, Στεφανᾶς, Σωσθένης, Φορτουνᾶτος, and Χλόη.
36 Ἀρέτας, Βελιάρ, Δαμασκηνός, Κορίνθιος, and Μακεδών.

tians,[37] 7 in Philippians,[38] 7 in Colossians,[39] 1 in 1 Thessalonians,[40] 2 in 1 Timothy,[41] 19 in 2 Timothy,[42] 6 in Titus,[43] and 2 in Philemon.[44] If these numbers are subtracted from the total number of Pauline *hapaxes* (see table 1), the following data apply:

Tab. 5: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns)

		Vocabulary tokens (x)	Hapax legomena (y)	Expected hapax legomena ($\hat{y}$)	Prediction interval	
					Min.	Max.
Undisputed Paulines	Rom	7111	235	251	209	293
	1Cor	6830	249	241	200	283
	2Cor	4477	172	158	120	196
	Gal	2230	79	78	40	116
	Phil	1629	68	57	18	95
	1Thess	1481	35	51	12	90
	Phlm	335	8	11	−30	52
Disputed Paulines	Eph	2422	86		t	2,570581836
	1Tim	1591	129		$\alpha/2$	0,025
	Col	1582	56		SE	13,65845554
	2Tim	1238	83		n	7
	2Thess	823	21		$\bar{x}$	3441,857143
	Tit	659	45		SS_{xx}	44266212,86

---

37 Ἀγάρ, Ἀραβία, Γαλάται, Ἱεροσόλυμα, Ἰουδαΐζειν, Ἰουδαϊκῶς, Ἰουδαϊσμός, Ἰωάννης, Κιλικία, Πέτρος, Σινᾶ, and Συρία.
38 Ἐπαφρόδιτος, Εὐοδία, Καῖσαρ, Κλήμης, Συντύχη, Φαρισαῖος, and Φιλιππήσιοι.
39 Ἱεράπολις, Ἰοῦστος, Κολοσσαί, Λαοδίκεια, Λαοδικεύς, Νύμφα, and Σκύθης.
40 Ἀθῆναι.
41 Πιλᾶτος and Πόντιος.
42 Δαλματία, Ἑρμογένης, Εὔβουλος, Εὐνίκη, Ἰαμβρῆς, Ἰάννης, Ἰκόνιον, Κάρπος, Κλαυδία, Κρήσκης, Λίνος, Λύστρα, Λωΐς, Μίλητος, Ὀνησίφορος, Πούδης, Τρόφιμος, Φίλητος, and Φύγελος.
43 Ἀρτεμᾶς, Ζηνᾶς, Ἰουδαϊκός, Κρής, Κρήτη, and Νικόπολις.
44 Ἀπφία and Φιλήμων.

The revised data of table 5 can be plotted in a new scatter diagram:
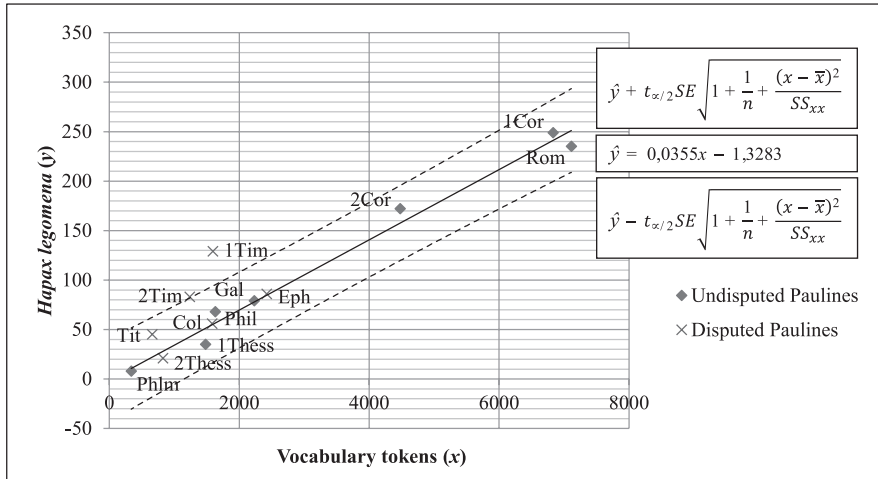

The chart plots *Hapax legomena (y)* against *Vocabulary tokens (x)*, with data points labeled for 1Cor, Rom, 2Cor, 1Tim, 2Tim, Gal, Eph, Col, Phil, Tit, 1Thess, 2Thess, Phlm. It includes the equations:

$$\hat{y} + t_{\alpha/2} SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$

$$\hat{y} = 0{,}0355x - 1{,}3283$$

$$\hat{y} - t_{\alpha/2} SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$

Legend: ◆ Undisputed Paulines; × Disputed Paulines

Fig. 4: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns)

Figure 4 shows that when proper noun *hapaxes* are not taken into consideration, no significant results are found except for the fact that 2 Timothy now almost equals the maximal predicted value; 1 Timothy still has a significantly high number of *hapaxes* in its text.

3.2 Explicit Quotations

Another critical factor to consider is the use of *hapaxes* taken from traditional materials. Many scholars have tried to account for the Pastorals' overuse of *hapaxes* by arguing that Paul borrowed them from the LXX.[45] Ceslas Spicq, for example, observed that about 200 *hapaxes* in the Pastorals also occur in the

---

45 So, for example, R. St. John Parry, The Pastoral Epistles, Cambridge 1920, cxvii; F. R. Montgomery Hitchcock, Tests for the Pastorals, JTS 30 (1929) 278–279; Francis John Badcock, The Pauline Epistles and the Epistle to the Hebrews in Their Historical Setting, London/New York 1937, 116–117; Pietro de Ambroggi, Le Epistole Pastorali di S. Paolo a Timoteo e a Tito (SB), Turin/Rome 1953, 35; Bonaventura Mariani, Introductio in libros sacros Novi Testamenti, Rome 1962, 369; Donald Guthrie, The Pastoral Epistles (TNTC), Leicester/Grand Rapids (MI) ²1990, 228–229; George W. Knight III, The Pastoral Epistles, Grand Rapids (MI)/Carlisle 1992, 42; William D. Mounce, The Pastoral Epistles (WBC 46), Nashville (TN) 2000, cxii.

LXX.[46] But this argument supposes, as Harrison noted, "that Paul must have been familiar with every Greek word in the LXX and Apocrypha."[47] For the sake of fairness, the argument should also take into account the number of LXX *hapaxes* found in the other Paulines.

Still, it is a matter of fact that some Pauline *hapaxes* clearly are derived from pre-Christian texts, whether biblical (e. g. ὕψος in Eph 4,8 = LXX Ps 67,19) or non-biblical (e. g. Κρής and θηρίον in Tit 1,12b = Callimachus, *HymnJov* 8; Epimenides, *De Oraculis*). In order to accommodate Harrison's hesitation to presuppose Paul's familiarity with such texts, it is better to take into account explicit quotations and not implicit ones (e. g. allusions or "echoes"). Overall, 92 Pauline *hapaxes* are found in explicit quotations, most of which are introduced by formulas such as καθὼς γέγραπται, γέγραπται γάρ, λέγει ἡ γραφή, etc.:[48] 52 in Romans,[49] 11 in 1 Corinthians,[50] 10 in 2 Corinthians,[51] 11 in Galatians,[52] 5 in Ephesians,[53] 1 in 1 Timothy,[54] and 2 in Titus.[55] If these numbers in addition to proper noun *hapaxes* (see table 5) are also subtracted from the total number of Pauline *hapaxes* (see table 1), the following data apply:[56]

46 Ceslas Spicq, Saint Paul. Les Épitres Pastorales (2 vols., Ebib), Paris ²1969, 1:187.
47 Harrison, Problem (see n. 7), 66.
48 Based on Dwight Moody Smith, The Pauline Literature, in: D. A. Carson/H. G. M. Williamson (eds.), It is Written: Scripture Citing Scripture, Cambridge 1988, 267–272, and Hans Hübner, *Vetus Testamentum in Novo*. Vol. 2: *Corpus Paulinum*, Göttingen 1997.
49 Αἰνεῖν, ἄμμος, ἄνθραξ, ἀπείθειν, ἀπέναντι, ἀριθμός, ἀσπίς, ἀσύνετος, ἀφαιρεῖν, ἀχρειοῦσθαι, Βάαλ, γέμειν, Γόμορρα, διαγγέλλειν, δολιοῦν, ἐκζητεῖν, ἐκκλίνειν, ἐκπεταννύναι, ἐμφανής, ἐπικαλύπτειν, ἐπιπίπτειν, ἥκειν, Ἠσαῦ, Ἰακώβ, Ἰεσσαί, ἰός, κατάνυξις, κατασκάπτειν, μοιχεύειν, νικᾶν, νῶτος, οἰκουμένη, οἰκτείρειν, ὁμοιοῦν, ὀνειδίζειν, πέρας, πρόβατον, σαβαώθ, Σάρρα, Σιών, σκοτίζειν, Σόδομα, συγκάμπτειν, σύμβουλος, συντελεῖν, συντέμνειν, σύντριμμα, σφαγή, ταλαιπωρία, τάφος, ὑπολείπεσθαι, and φονεύειν.
50 Ἀροτριᾶν, αὔριον, ἐξαίρειν, ἦθος, κέντρον, κημοῦν, ὁμιλία, παίζειν, πόρνη, προσκυνεῖν, and συνετός.
51 Βοηθεῖν, δότης, εἰσδέχεσθαι, ἐλαττονεῖν, ἐπακούειν, θυγάτηρ, ἱλαρός, παντοκράτωρ, πένης, and σκορπίζειν.
52 Βοᾶν, ἐκβάλλειν, ἐμμένειν, ἐνευλογεῖσθαι, ἐπικατάρατος, κρεμάννύναι, παιδίσκη, ῥησσεῖν, στεῖρος, τίκτειν, and ὠδίνειν.
53 Εὖ, μακροχρόνιος, ὀργίζεσθαι, προσκολλᾶσθαι, and ὕψος.
54 Φιμοῦν.
55 Θηρίον and Κρής.
56 *Hapaxes* relating to both factors (e. g. Γόμορρα in Romans 9,29 is a proper noun but is also used in a quotation) have been counted as 1.

Tab. 6: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns and *hapaxes* in explicit quotations)

		Vocabulary tokens (x)	Hapax lego- mena (y)	Expected hapax legomena ($\hat{y}$)	Prediction interval	
					Min.	Max.
Undisputed Paulines	Rom	7111	189	223	153	293
	1Cor	6830	238	214	145	284
	2Cor	4477	162	142	79	205
	Gal	2230	69	72	9	136
	Phil	1629	68	54	−11	118
	1Thess	1481	35	49	−15	114
	Phlm	335	8	14	−54	82
Disputed Paulines	Eph	2422	81		t	2,571
	1Tim	1591	128		$\alpha/2$	0,025
	Col	1582	56		SE	22,71
	2Tim	1238	83		n	7
	2Thess	823	21		$\bar{x}$	3441,86
	Tit	659	44		SS_{xx}	44266212,86

The revised data of table 6 can be plotted in a new scatter diagram, again:

$$\hat{y} + t_{\alpha/2}SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$

$$\hat{y} = 0,0309x + 3,5554$$

$$\hat{y} - t_{\alpha/2}SE \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_{xx}}}$$
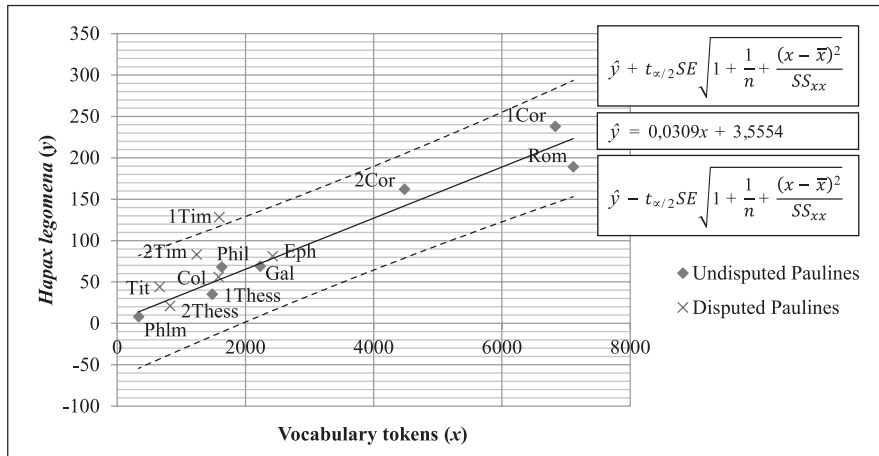
◆ Undisputed Paulines
✕ Disputed Paulines

Fig. 5: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns and *hapaxes* in explicit quotations)

This time figure 5 reveals that 2 Timothy fits well into the prediction interval; 1 Timothy still has a significant high number of *hapaxes* in its text, but it now also comes close to the prediction interval.

3.3 Prepositional Compounds

A final factor affecting the use of *hapaxes* in the *Corpus Paulinum* relates to compound words. In contemporary corpus linguistics this phenomenon is known as "productivity", referring to "the statistical readiness with which an affix enters into new combinations"[57] (neologisms). The concept of productivity is associated with Harald Baayen, who assumed a close relationship between *hapax legomena* and productivity. Baayen's theory essentially suggests that "the number of *hapaxes* of a given morphological category correlates with the number of neologisms of that category, so that the number of *hapaxes* can be seen as an indicator of productivity."[58] Baayen (with others) developed several formulas to measure productivity, demonstrating that some affixes generate more neologisms than others.[59]

Whereas Baayen's measures of productivity are of less importance here, his insight that specific Indo-European affixes (whether prefixes, infixes, or suffixes) enable language users to create new words is of particular interest in connection with the observation of modern classicists that in Antiquity "[h]apaxes are usually compound words."[60] In the *Corpus Paulinum*, it is typical to find *hapaxes* that are constructed by means of a preposition (ἀνά-, ἀντί-, ἀπό-, διά-, εἰς-, ἐκ-, ἐν-, ἐπί-, κατά-, μετά-, παρά-, περί-, πρό-, πρός-, σύν-, ὑπέρ-, and/or ὑπό-). Baayen's theory allows for the possibility that any author was able to create new words with prepositions. If so, one should be hesitant to include in a comparative study of Pauline vocabulary prepositional compounds of which upon removal of the preposition a lexeme is found that is used elsewhere in the *Corpus Paulinum*.[61] Should, for example, σύνδουλος (κτλ) in Colossians (1,7 and 4,7) have discriminating force if δοῦλος (κτλ) is used throughout the Pauline letter corpus (e. g. Rom 1,1; 1Cor 7,21–23; 2Cor 4,5; Gal 1,10; Eph 6,5–8; Phil 1,1; Col 3,11; 1Tim 6,1; 2Tim 2,24; Tit 2,9; Phlm 1,16)? Overall, there are 165 of such words in the *Corpus*

---

57 Alexander Haselow, Typological Changes in the Lexicon: Analytic Tendencies in English Noun Formation (TiEL 72), Berlin/New York 2011, 87.

58 Ingo Plag, Christiane Dalton-Puffer, and R. Harald Baayen, Morphological Productivity across Speech and Writing, JELL 3 (1999) 215.

59 Cf. R. Harald Baayen and Rochelle Lieber, Productivity and English Derivation: A Corpus Based Study, Lin 29 (1991) 801–843.

60 Agnieszka Kotlińska-Toma, Hellenistic Tragedy: Texts, Translations and a Critical Survey (BCSM), London/New York 2015, 35.

61 At least in two letters; otherwise this lexeme would qualify as a Pauline *hapax*.

Paulinum: 36 in Romans,[62] 22 in 1 Corinthians,[63] 25 in 2 Corinthians,[64] 18 in Galatians,[65] 10 in Ephesians,[66] 5 in Philippians,[67] 7 in Colossians,[68] 4 in 1 Thessalonians,[69] 5 in 2 Thessalonians,[70] 22 in 1 Timothy,[71] 5 in 2 Timothy,[72] 3 in Titus,[73] and 3 in Philemon.[74] If these numbers in addition to proper noun *hapaxes* (see table 5) and *hapaxes* in explicit quotations (see table 6) are also subtracted from the total number of Pauline *hapaxes* (see table 1), the following data apply:[75]

---

62 (ἀν)άγειν, (ἀνα)ζῆν, (ἀντι)στρατεύεσθαι, (ἀπο)τολμᾶν, (δια)πορεύεσθαι, (ἐγ)καλεῖν, (ἐκ)ζητεῖν, (ἐν)τυγχάνειν, (ἐπ)αναμιμνήσκειν, (ἐπ)αναπαύεσθαι, (ἐπι)πίπτειν, (ἐπι)τυγχάνειν, (ἐπι)φέρειν, (ἐπ)ονομάζεσθαι, (καθ)ορᾶν, (κατ)άγειν, (κατα)καυχᾶσθαι, (κατά)κριμα, (κατα)νοεῖν, (μετ)αλλάσσειν, (παρα)κεῖσθαι, (προ)γίνεσθαι, (προ)γινώσκειν, (προ)δίδόναι, (προ)ηγεῖσθαι, (πρό)θυμος, (συγ)κάμπτειν, (συμ)μαρτυρεῖν, (συμ)παρακαλεῖσθαι, (σύμ)φάναι, (συν)αγωνίζεσθαι, (συν)αναπαύεσθαι, (συν)δοξάζειν, (συν)τελεῖν, (συ)στενάζειν, and (ὑπερ)φρονεῖν.

63 (ἀνα)κρίνειν, (ἀπ)άγειν, (ἀπ)ελεύθερος, (ἀπο)φέρειν, (εἰσ)ακούειν, (ἐκ)δέχεσθαι, (ἐκ)νήφειν, (ἐκ)πειράζειν, (ἔν)νομος, (ἐπ)άνω, (ἐπί)κεῖσθαι, (κατά)κεῖσθαι, (κατα)χρᾶσθαι, (παρ)άγειν, (πάρ)οδος, (περι)άγειν, (περι)τιθέναι, (συγ)γνώμη, (συ)ζητητής, (συν)άγειν, (συν)έρχεσθαι, and (συν)εἰδέναι.

64 (ἀπό)κριμα, (ἐγ)γράφειν, (ἐγ)κρίνειν, (εἰσ)δέχεσθαι, (ἐμ)περιπατεῖν, (ἐξ)απορεῖσθαι, (ἐξ)ιστάναι, (ἐπ)ακούειν, (ἐπ)ενδύεσθαι, (καθ)αίρεσις, (κατά)κρισις, (μετα)νοεῖν, (παρα)φρονεῖν, (παρ)εκτός, (παρ)έρχεσθαι, (προ)αιρεῖσθαι, (προ)αμαρτάνειν, (προ)ενάρχεσθαι, (προ)έρχεσθαι, (προ)καταρτίζειν, (πρό)κεῖσθαι, (προσ)αναπληροῦν, (πρόσ)καιρος, (συμ)πέμπειν, and (συν)αποστέλλειν.

65 (ἀνα)τίθεσθαι, (ἀν)έρχεσθαι, (δια)μένειν, (ἐμ)μένειν, (ἐν)έχειν, (ἐξ)αποστέλλειν, (ἐπι)διατάσσεσθαι, (ἐπί)τροπος, (κατα)γινώσκειν, (κατα)σκοπεῖν, (μετα)τίθεσθαι, (παρα)τηρεῖν, (προ)ευαγγελίζεσθαι, (προ)ἰδεῖν, (προ)καλεῖσθαι, (προσ)τιθέναι, (συμ)παραλαμβάνειν, and (συ)στοιχεῖν.

66 (ἐκ)πορεύεσθαι, (ἐξ)ισχύειν, (ἐπ)έρχεσθαι, (πάρ)οικος, (προ)ελπίζειν, (προσ)κολλᾶσθαι, (συγ)καθίζειν, (συν)οικοδομεῖν, (ὑπερ)άνω, and (ὑπο)δεῖσθαι.

67 (ἐξ)ανάστασις, (παρα)μένειν, (παρα)πλήσιον, (συλ)λαμβάνειν, and (συμ)μιμητής.

68 (ἀντ)αναπληροῦν, (ἀπο)κρίνεσθαι, (παρα)λογίζεσθαι, (προ)ακούειν, (συλ)αγωγεῖν, (σύν)δουλος, and (ὑπ)εναντίος.

69 (ἀνα)μένειν, (εἴσ)οδος, (ἐκ)διώκειν, and (προ)πάσχειν.

70 (ἐγ)καυχᾶσθαι, (ἐν)δοξάζεσθαι, (κατ)αξιοῦν, (περι)εργάζεσθαι, and (ὑπερ)αυξάνειν.

71 (ἀντι)λαμβάνεσθαι, (ἀπό)δεκτος, (ἀπο)θησαυρίζειν, (εἰσ)φέρειν, (ἐκ)ζήτησις, (ἐκ)φέρειν, (ἐμ)πίπτειν, (ἐπ)αρκεῖν, (ἐπι)λαμβάνεσθαι, (ἐπ)ίστασθαι, (ἐπι)τίθεναι, (κατα)λέγεσθαι, (παρα)δέχεσθαι, (περί)εργος, (περι)έρχεσθαι, (περι)ποιεῖσθαι, (προ)άγειν, (πρό)δηλος, (πρό)κριμα, (προσ)έρχεσθαι, (προσ)μένειν, and (ὑπερ)πλεονάζειν.

72 (ἀνα)νήφειν, (ἔκ)δηλος, (ἐπι)τιμᾶν, (κατα)φθείρειν, and (μετα)λαμβάνειν.

73 (ἐγ)κρατής, (ἐπι)φαίνειν, and (περι)φρονεῖν.

74 (ἀνα)πέμπειν, (ἐπι)τάσσειν, and (προσ)οφείλειν.

75 *Hapaxes* relating to more than one factor (e. g. εἰσδέχεσθαι in 2 Corinthians 6,17 is used in a quotation but is also a prepositional compound) have been counted as 1.

Tab. 7: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns, *hapaxes* in explicit quotations, and prepositional compounds)

		Vocabulary tokens (x)	*Hapax lego-mena* (y)	Expected *hapax legomena* ($\hat{y}$)	Prediction interval	
					Min.	Max.
Undisputed Paulines	Rom	7111	157	193	118	268
	1Cor	6830	214	185	112	259
	2Cor	4477	138	122	55	189
	Gal	2230	51	62	−6	129
	Phil	1629	63	45	−23	114
	1Thess	1481	31	41	−28	110
	Phlm	335	5	11	−62	83
Disputed Paulines	Eph	2422	72		t	2,570581836
	1Tim	1591	105		$\alpha/2$	0,025
	Col	1582	49		SE	24,19299487
	2Tim	1238	78		n	7
	2Thess	823	16		$\bar{x}$	3441,857143
	Tit	659	41		SS_{xx}	44266212,86

Once more, a new scatter diagram can be plotted on the basis of the revised data in table 7:
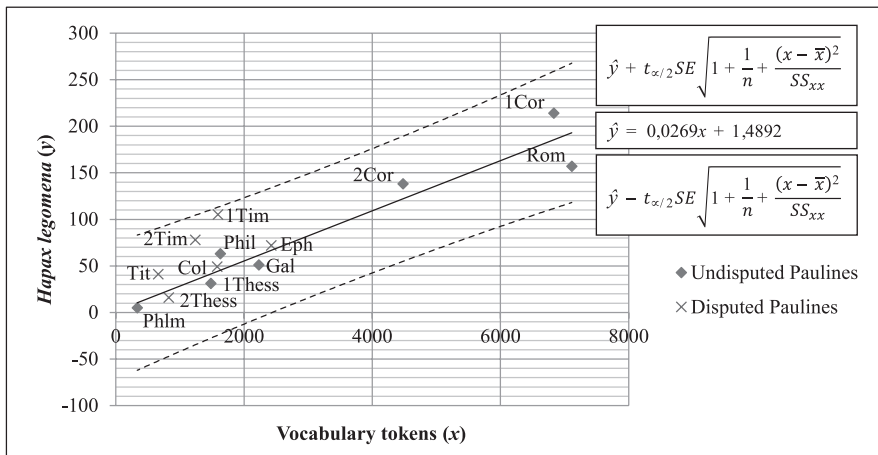

Fig. 6: *Hapax legomena* in the *Corpus Paulinum* (excluding proper nouns, *hapaxes* in explicit quotations, and prepositional compounds)

This time, figure 6 indicates that all disputed Paulines fit the predication interval, even though 1 Timothy still has a considerable high number of *hapaxes* in its text.

Conclusion

All of the above considerations and results may serve as a call to scholars to be more reserved in using *hapax legomena* as a criterion in Pauline authorship trials. Once properly defined and analyzed by means of (simple) linear regression analysis, *hapaxes* are not overused in Ephesians, Colossians, 2 Thessalonians, and Titus. A significant number of *hapaxes* is found in 1 and 2 Timothy, but there are some factors which may account for this result. Three factors have been discussed, showing that *hapaxes* used in explicit quotations as well as proper nouns and prepositional compounds should perhaps not have discriminating force. If these factors are not taken into consideration, it appears that none of the disputed Paulines use significantly more *hapaxes* than the undisputed Paulines. Accordingly, it is suggested to debate questions of authorship in terms of the Paulines' literary, historical, and/or theological differences in the future.