

生成式大模型与人类词汇联想比较研究

——基于词频效应与联想不对称性

关思扬

苏州大学外国语学院, 江苏 苏州

收稿日期: 2026年5月6日; 录用日期: 2026年5月26日; 发布日期: 2026年6月4日

摘要

本研究从词汇联想范式出发, 比较生成式大模型与人类在语义表征结构上的差异, 重点考察联想多样性与联想不对称性。结果表明, 人类词汇联想在多样性与不对称性上均高于生成式大模型, 呈现出更为复杂的语义组织结构; 两类数据均表现出词频相关规律, 但生成式大模型对词频的依赖更强。此外, 联想多样性与不对称性之间的关系在人类数据中更为紧密。总体而言, 生成式大模型虽能在一定程度上复现人类词汇联想的统计特征, 但其语义表征的结构复杂性仍存在差距。

关键词

生成式大模型, 词汇联想, 词频效应, 联想不对称性

A Comparative Study of Lexical Associations between GenAI and Humans

—Based on Word Frequency Effects and Associative Asymmetry

Siyang Guan

School of Foreign Languages, Soochow University, Suzhou Jiangsu

Received: May 6, 2026; accepted: May 26, 2026; published: June 4, 2026

Abstract

This study compares lexical association patterns between GenAI (Generative Artificial Intelligence) and humans to examine whether GenAI exhibits human-like semantic organization. Focusing on association diversity and asymmetry, the results show that human data demonstrate higher diversity and stronger directional asymmetry than GenAI-generated associations. While both systems display

systematic relationships between word frequency, GenAI shows a stronger dependence on frequency. In addition, the relationship between diversity and asymmetry is more tightly coupled in human data. Overall, the findings suggest that GenAI can approximate certain statistical regularities of human lexical associations, but their semantic representations remain less structurally complex than those of the human mental lexicon.

Keywords

Generative AI, Lexical Association, Word Frequency Effect, Association Asymmetry

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

词汇联想任务长期被视为刻画人类语义结构的重要工具，最常见的词汇联想任务会要求被试对于给定的刺激词(stimulus)给定一个至几个最先想到的反应词(response)，通过对这些反应词进行分析，就可以间接探索词汇在心理词库中的组织方式。而以 SWOW (Small World of Words [1])为代表的大规模词汇联想常模库建设，则使这一研究范式可以从小样本的实验进步到大样本的统计分析。

近年来，生成式人工智能在大规模文本语料的训练以及概率生成机制的驱动下，已经体现出了与人类似的(human-like)语言输出表层形式。然而，我们无法直接研究这种与人类似的特征是否仅体现在表层的语言输出，还是已经形成了与人类类似的心理词库的组织结构。人类心理词库的组织具有多样特征，我们可以选择其中的特定角度进行切入，而已经经过词汇联想研究验证的两个关键结构特征，即词汇联想多样性和词汇联想不对称性，可以为这一方向的研究提供一定的经验证据。

基于以上目的，本研究采取 SWO [1]人类词汇联想常模库和构建方法相同的大模型词汇联想常模库 LWW (LLM World of Words) [2]为基础，系统比较两者在词汇联想多样性以及两项不对称性上的异同，以尝试探讨生成式大模型是否已经在表层以下呈现出了与人类似的语义组织形式。

2. 文献综述

生成式大模型的语言能力来源于其对大规模文本语料的基于分布的学习。现有主流大模型大多基于 Transformer 架构[3],通过对高质量人类语言文本进行拆分,进行词元(token)与词元之间的条件概率计算,实现了与人类似文本的成功输出。这一过程不依赖显性的语义规则,而是纯粹地基于对词元间共现关系的编码和调整,使语义相似性、句法依赖乃至更复杂的御用信息在高维度参数空间中以不可见的方式进行了表征[4]。进一步研究显示,现有生成式大语言模型的表征结构在一定程度上能够模拟人类语言加工中的语义相似度判断、类比推理等行为[5]。然而我们需要指出,这种一次性的训练方式不同于人类对语义表征的渐进的学习过程,因此生成式大模型的表征能力是否真的趋近于人类的语义表征,仍然是当下研究中的核心议题。

在对于人类语义表征的研究中,词汇联想范式提供了大量的经验性证据。研究结果表明,语义表征并不是简单的相似性整合,而是也呈现出复杂的、动态的网络结构。研究者使用词汇联想的方法,已经发现词频效应、联想多样性、联想不对称性是人类词汇联想中的稳定发生、具有解释力的现象。联想多样性指的是一个刺激词能够引发的反应词的词目数和离散程度,一般而言高频词往往据欧更多的联想多

样性,即,高频词的反应分布更加分散,有更多的语义临域[1][6]。网络研究的结果也证明了,在人类的语义网络中,高频词倾向于吸引更多的新词进行连接,这一过程也称为“优先连接”[7];同时,低频词的语义范围通常更具体,与整个语义网络的连接较为稀疏,所以它们的反应词就会更少。另一方面,联想不对称性反应了语义关系的方向性特征,即从刺激词到反应词的联想强度通常与反应词到刺激词的联想强度不同。例如,刺激词“医生”到反应词“医院”的联想概率往往比“医院”到“医生”的要高,这一现象也具有高度的稳定性。联想的不对称被认为是人类在构建语义网络的过程中扩散激活……机制的体现,即不同的词汇表征在网络中的位置也影响了信息的传播路径和强度。总的来看,联想多样性与联想不对称性共同反应了人类心理词库、语义网络组织的基本数量、方向特征。

出于上述人类/生成式大模型的语义网络构建方式差异,通过词汇联想的方法将两者进行对比就具有了明确理论意义。已有研究表明,大模型在语义相似性判断、词义消歧等任务中能够达到甚至超过人类水平[8],能够自行构建数据中没有显性标注的成分间关系[9]。但对其内部的语义组织形式研究则较少,词汇联想任务则在此有明显优势,其输出不受显式任务约束,更能反映语义的内在结构[1]。如果生成式大模型能够在联想多样性与不对称性上复现人类数据中的结构模式,则在这两个维度上为其与人类类似的表征结构提供了一定证据;反之,如果模型在这些维度上表现出系统性偏差,则可以认为其语义表征仍在一定程度上停留在统计层面,在此两个维度上不支持其“具有与人类类似的语义表征”的推断。因此,从词汇联想角度切入检验人类与生成式大模型在联想多样性和不对称性上的表现,能够为我们的核心问题提供具有解释力的证据。

基于以上的问题和方法,本研究选取人类与生成式大模型的词汇联想数据,从词汇联想多样性和不对称性的角度,对比研究生成式大模型在这两个维度上是否已经有了与人类类似的表征结构。具体研究问题如下:

- 1) 人类与生成式大模型在词汇联想多样性上是否存在差异?如果有,这些差异是什么?
- 2) 人类与生成式大模型在词汇联想不对称性上是否存在差异?如果有,这些差异是什么?

人类与生成式大模型的词汇联想多样性、不对称性结果是否具有交互?如果有,交互作用的差异是什么?

3. 研究设计

3.1. 实验材料

本研究用于对比人类与生成式大模型词汇联想模式的数据由同一方式构建。人类词汇联想常模数据来自建立的 SWOW-EN 常模库。作为一系列项目中的一个,SWOW-EN 是 SWOW (Small World of Words) 项目的英语部分。De Deyne 等人[1]使用滚雪球的方法,通过社交媒体、电子邮件等方式共招募了 88,722 名受试(81%母语者),通过 12,217 个刺激词共得到 3,686,400 个反应词。与 USF [10]、EAT [11]等词汇联想常模库相比,SWOW-EN 在词汇判断、词汇命名和语义判断等任务中均表现均相对优异,且构建过程中的数据收集流程、受试的异质性更佳。Abramski 等人[2]使用同样的范式,构建了 LWOW 常模库(LLM World of Words),该常模库使用同样的刺激词表,要求三种不同的生成式大模型在每次均独立的对话中进行反应。这三种生成式大模型(LLaMA3、Haiku 和 Mistral)中,LLaMA3 是唯一开源,且独特反应数接近 SWOW-EN 中的人类数据(人类 116,640 个,LLaMA 3,105,367 个,作为对比,Haiku 和 Mistral 仅有 15,275、41,369 个),考虑到本研究研究的对象就包括联想多样性,不应在材料选取阶段出现明显差异,因此研究选取的是 LLaMA3 数据所谓生成式大模型数据(后文均称生成式大模型数据)。本研究通过应用 SWOW-EN 和 LWOW 数据,在控制构建方法的同时提供了最大的数据量。

考虑到 SWOW 和 LWOW 两个常模库的构建方式虽然相同,但主体不同,SWOW 中的人类数据是随机分发,很多刺激词的反应词少于 100 个;而生成式大模型构建的 LWOW 则控制每个刺激词均有 100 个反应词,因此,为了进一步对齐两者规模,研究选择筛除 SWOW 中不足 50 个反应的刺激词,将两者均在 50 个反应词处截取。为了进一步帮助进行数据分析,选择使用 van Heuven 等人[12]的母语者词频标注数据进行二次筛选,筛除其中所有没有出现过的刺激词,以及有反应词没有标记的刺激词,最终得到了人类/生成式大模型两组完全相同的、所有刺激词/反应词均有词频标记的数据组(共 829 个相同的刺激词,每个刺激词有 50 个反应词),用于进一步的数据分析。

3.2. 数据分析

为了回答研究问题,研究使用上述筛选的刺激-反应词对作为数据集,采用两项核心指标对生成式大模型与人类词汇联想进行比较。词汇联想多样性以香农熵进行刻画,针对联想不对称性,则按照……的方法构建联想不对称性指数。

词汇联想任务的联想多样性用于刻画刺激词的反应词分布离散程度。本研究选择香农熵……对其进行可检验的测量。本研究将其定义为 $H(s) = -\sum p(r_i|s) \log p(r_i|s)$, 其中 s 为刺激词,其对应的反应词为 r_n , 每一个反应词的概率为 $p\left(\frac{r_i}{s}\right)$ 。它直观反应了反应词的多样性,如果一个刺激词的反应词高度集中于少数几个词(如某一反应类型占据绝大多数反应),则熵值较低,而当反应类型较多时,熵值较高。在本研究的实际计算中,首先统计两个常模库中每一个刺激词对应所有反应词的出现概率,然后将其归一化为概率分布,随后对两个常模库中的每个刺激词分别计算香农熵,得到可以用进行差异检验的词汇多样性数据组。统计时,使用 van Heuven 等人[12]标记的英语母语者词频数据进行对应,然后将对数词频数据与两组中的词汇多样性数据进行皮尔逊相关分析,相关系数 r 大于 0 时,词频与联想多样性呈正相关,相关系数 r 小于 0 时,说明词频与联想多样性呈负相关,而当 r 趋近于 0 时,则说明词频与联想多样性之间不存在显著的相关关系。最后,本研究使用线性回归模型,检验人类和大语言模型在词汇联想多样性上的差异。

词汇联想的不对称性指标用于刻画刺激词-反应词对之间的联想关系的方向性差异。本研究中的词汇联想不对称性指数定义为 $A(s,r) = \log\left(\frac{P(rs)}{P(sr)}\right)$, 其中 A 表示一个刺激词-反应词对的不对称性,当 $A(s,r)$ 大于 0 时,说明刺激词到反应词的联想强度强于反向路径,而当其小于 0 时,说明刺激词到反应词的联想强度小于反向路径,以上两者都说明刺激词与反应词之间存在联想不对称。仅有当 $A(s,r)$ 约等于 0 时,代表着刺激词-反应词之间联想对称。由于研究选择的是两个不同的词汇联想常模库,两者不可能完全对齐,因此研究在两个常模库中分别选取了正反刺激词-反应词对,用于计算不对称性指数。在数据分析中,取所有 A 的绝对值,并进行单一样本 t 检验以验证其是否整体偏离 0,如有偏离则说明其存在词汇联想的不对称性。然后,将两组中的 A 绝对值($|A|$),进行独立样本 t 检验,即能够比较两者词汇联想不对称性(如有)的强度差异。

最后,为了进一步分析人类与生成式大模型的词汇联想多样性与联想不对称性的潜在交互效应,本研究计算组内的联想多样性与联想不对称性相关系数,然后计算人类与大模型的词汇联想多样性-不对称性关系是否存在统计意义上的差异。首先,选择具有词汇不对称系数的词(即在一个常模库中同时作为过刺激词和反应词的词),提取它们的词汇的联想多样性系数 $H(s)$,然后计算这些词不对称性指数 A 绝对值的平均值,作为刺激词层面的平均不对称性强度 $\bar{A}(c)$,从而实现了其与词汇联想多样性的意义对应。然后,首先分别计算人类/生成式大模型组中 $H(s)$ 和 $\bar{A}(c)$ 的皮尔逊相关系数。当相关系数 r 大于 0 时,说明联想越多的词,其联想关系越具有不对称性;当 r 小于 0 时,说明联想的多样性与方向性之间存在

负相关，而当 r 约等于 0 则说明两者基本独立，研究进而可以了解人类与生成式大模型在词汇联想多样性、不对称性上的潜在交互。

4. 研究结果

4.1. 人类与生成式大模型联想多样性对比

本研究首先对人类词汇联想数据与生成式大模型词汇联想数据的联想多样性进行了描述统计分析，结果如表 1 所示。人类数据的平均香农熵为 2.42 ($SD = 0.63$)，而生成式大模型的平均香农熵为 2.27 ($SD = 0.46$)。整体来看，人类联想数据的熵值水平高于生成式大模型，表明其在相同刺激条件下具有更为分散的反应词分布结构。两组数据的样本量均为 829 个刺激词，保证了后续统计比较的可比性。

进一步对联想多样性与词频之间的关系进行相关分析见表 1，结果显示两组数据均呈现显著负相关关系。如表 1 所示，人类数据中联想多样性与词频的相关系数为 $r = -0.17$ ($p < 0.001$)，生成式大模型数据中的相关系数为 $r = -0.28$ ($p < 0.001$)。结果表明，在两种数据来源中，较高词频均对应较低的联想多样性水平，且该关系在生成式大模型中更为明显。

在控制词频因素的基础上，进一步构建线性回归模型，以联想多样性为因变量，组别与词频为自变量进行联合预测分析。结果显示，模型整体显著 ($F = 56.05, p < 0.001$)， $R^2 = 0.06$ 。具体而言，组别变量对联想多样性具有显著预测作用 ($\beta = -0.27, t = -9.07, p < 0.001$)，词频变量同样达到显著水平 ($\beta = -0.23, t = -8.89, p < 0.001$)，表明在控制词频影响后，两类数据在联想多样性上仍存在稳定差异。

Table 1. Correlation between entropy and word frequency in human and GenAI's Word association

表 1. 人类与生成式大模型词汇联想香农熵与词频相关分析

组别	M	SD	r	p
人类	2.42	0.63	-0.17	0.000***
生成式大模型	2.27	0.46	-0.28	0.000***

注：***: $p < 0.000$ ，下同。

4.2. 人类与生成式大模型联想不对称性对比

本研究首先对词汇联想不对称性指数进行有效样本统计，结果显示人类数据中可计算的词汇联想对数量为 99,673 对，而生成式大模型数据中为 15,323 对，表明在人类数据中可观测的双向联想关系显著多于生成式大模型数据，这一差异反映了两类数据在联想网络连通结构上的规模差异。

进一步对不对称性强度进行描述统计分析，结果如表 2 所示，人类数据的平均绝对不对称性指数为 1.357 ($SD = 0.97$)，生成式大模型的平均绝对不对称性指数为 0.87 ($SD = 0.77$)，表明人类词汇联想关系在刺激词与反应词之间的方向性差异整体更为显著，而生成式大模型的联想关系相对较为对称。单一样本 t 检验结果显示，两组数据的 A 绝对值均显著偏离 0 (人类数据: $t = 443.72, p < 0.001$; 生成式大模型: $t = 138.96, p < 0.001$)，说明在两类语料中均普遍存在词汇联想的不对称性结构。

Table 2. Difference between human and GenAI's word association asymmetry

表 2. 人类与生成式大模型词汇联想不对称性差异检验

组别	N	M ($ A $)	SD ($ A $)	t	p
人类	99,673	1.36	0.97	443.72	0.000***
生成式大模型	15,323	0.86	0.77	138.96	0.000***

在组间比较方面,独立样本 t 检验结果表明,两组在不对称性强度上存在显著差异($t = 70.99, p < 0.001$),即人类词汇联想的不对称性显著高于生成式大模型。整体结果如表所示。

4.3. 人类与生成式大模型联想多样性、不对称性交互

本研究首先对词汇联想多样性与不对称性之间的关系进行组内相关分析,结果显示人类数据中两者呈显著负相关($N = 7076, r = -0.36, p < 0.001$),生成式大模型数据中同样呈显著负相关($N = 2161, r = -0.28, p < 0.001$),说明在两类数据中,联想多样性较高的刺激词通常对应较低的联想不对称性强度。

进一步对两组相关系数进行差异检验, Fisher Z 转换结果表明两组之间存在显著差异($z = -3.52, p < 0.001$),说明人类与生成式大模型在联想多样性与不对称性之间的耦合关系强度上存在统计学意义上的差别。

对相关结构进行置信区间估计的结果如表 3 所示,人类数据的相关系数为 $r = -0.36$ (95% CI [-0.38, -0.34]),生成式大模型数据的相关系数为 $r = -0.28$ (95% CI [-0.32, -0.24]),表明两组数据的负相关关系均较为稳定,但人类数据中的耦合强度整体更高。

Table 3. Correlational analysis between entropy and asymmetry in human and GenAI word association

表 3. 人类与生成式大模型词汇联想香农熵与不对称性的交互相关分析

组别	N	r [95% CI]
人类	7076	-0.36 [-0.38, -0.34]
生成式大模型	2161	-0.28 [-0.32, -0.24]

在结构强度补充分析中,人类数据的平均联想多样性为 $\text{Mean } H = 2.64$,平均不对称性强度为 $\text{Mean } |A| = 1.09$;生成式大模型数据的对应值分别为 $\text{Mean } H = 2.42$ 与 $\text{Mean } |A| = 0.75$,结果显示人类数据在联想多样性与不对称性两个维度上的整体水平平均高于生成式大模型,且两者在结构层面呈现一致性差异特征。

5. 讨论

研究发现,人类在词汇联想多样性上显著高于生成式大模型,且词频对生成式大模型的生成具有一定约束,表明其语义网络结构受到一定局限,且有频率相关的倾向。人类数据的平均香农熵高于生成式大模型,证明人类在给定同一刺激词时能够产生更离散的反应。这一结果与人类语义网络被广泛证明的无标度与小世界特性[13]相吻合,这种动态结构允许人类在提取词汇时超越简单的概率,而是形成了复杂多样的连接。相比之下,以 LLaMA3 为代表的生成式人工智能在联想生成时存在明显向高概率典型反应收敛的趋势。进一步检验表明,词频对联想多样性具有显著的负向预测作用,且对生成式大模型的影响更为显著。在人类心理词库中,高频词由于优先连接的倾向[7],不仅拥有核心的强语义连接,同时还在长期使用中建立了很多弱连接。然而,基于同样的倾向,生成式大模型虽成功建立了很多强连接,但在弱连接上表现较差。这种词频敏感性侧面说明,尽管大模型在分布式学习上取得了巨大成功,但其表征还没有复刻人类的心理词库的全部复杂性。

联想不对称性方面,人类数据不仅具备更密集的双向联想连通性,且表现出远高于生成式大模型的不对称性。两组数据虽然都表现出一定的不对称性,但生成式大模型中的联想不对称性远弱于人类。长久以来,联想不对称性被视为人类扩散激活机制的核心证据之一[14]。人类语义网络中,信息从局部向整体的激活往往比反向路径更为迅速和强烈,这种方向性受概念层级、习得顺序等多种因素制约。然而,

生成式大模型语义表征的底层逻辑是向量空间模型[15]，从本质上讲，高维空间中词元的距离是对称的，尽管 Transformer 架构引入了注意力机制[3]，能够在一定程度上捕捉前后文的单向性，但当脱离句子环境仅进行词元提取时，生成式大模型不得不更多依赖词元的空间距离。联想不对称性上的差异表明，无指导的大规模预训练尚不足以建立一个脱离了句子的与人类相似的心理词库。

词汇联想多样性与不对称性的交互结果进一步表明，尽管人类与生成式大模型均表现出多样性越高、不对称性越低的规律，但人类数据中这一耦合关联更为紧密，反映出更复杂的结构特征。人类数据中这种更为紧密的耦合关系，证明了人类的语义组织是一个高度自组织的复杂动态系统[16]。而对大语言模型来说，这一种脱离句子仍存在的、复杂动态的系统没有得到实质上的复制。生成式大模型凭借大量语料在表层上呈现了符合统计的联想模式，但由于机制与人类的语义网络建构截然不同，在本研究的维度上暂时没有成功复制人类的语义网络结构。

6. 结论

本研究通过对比人类与一种开源大模型(LLaMA3)的英语词汇联想网络，发现，至少对于该种模型来说，其虽能模拟词频效应和非对称特征，但仍与人类联想结果呈现显著差异，呈现出依赖统计分布的倾向，证明其深层语义表征尚未具有与人类相似的组织。随着大模型的快速迭代，以及对于人类语言间差异，不同的组合可能会呈现出不同的结果。

未来研究可进一步考察不同参数规模模型在语义网络构建上的演进，深入探讨非具身模型走向与人类类似认知能力的可能路径。本研究为大模型语义表征的内部研究提供了经验证据，也为评价人工智能的认知真实性提供了新视角。

参考文献

- [1] De Deyne, S., Navarro, D.J., Perfors, A., Brysbaert, M. and Storms, G. (2019) The “Small World of Words” English Word Association Norms for over 12,000 Cue Words. *Behavior Research Methods*, **51**, 987-1006. <https://doi.org/10.3758/s13428-018-1115-7>
- [2] Abramski, K., Improta, R., Rossetti, G. and Stella, M. (2025) The “LLM World of Words” English Free Association Norms Generated by Large Language Models. *Scientific Data*, **12**, Article No. 803. <https://doi.org/10.1038/s41597-025-05156-9>
- [3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., et al. (2017) Attention Is All You Need. *Advances in Neural Information Processing Systems*, **30**, 1-11.
- [4] Mikolov, T., Chen, K., Corrado, G.S. and Dean, J. (2013) Efficient Estimation of Word Representations in Vector Space. 2013 *International Conference on Learning Representations*, Scottsdale, 2-4 May 2013.
- [5] Ethayarajh, K. (2019) How Contextual Are Contextualized Word Representations? Comparing the Geometry of BERT, Elmo, and GPT-2 Embeddings. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, November 2019, 55-65. <https://doi.org/10.18653/v1/d19-1006>
- [6] Brysbaert, M. and New, B. (2009) Moving beyond Kučera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behavior Research Methods*, **41**, 977-990. <https://doi.org/10.3758/brm.41.4.977>
- [7] Hills, T.T., Maouene, M., Maouene, J., Sheya, A. and Smith, L. (2009) Longitudinal Analysis of Early Semantic Networks: Preferential Attachment or Preferential Acquisition? *Psychological Science*, **20**, 729-739. <https://doi.org/10.1111/j.1467-9280.2009.02365.x>
- [8] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, 2-7 June 2019, 4171-4186.
- [9] Xiao, B., Duan, X., Haslett, D.A. and Cai, Z. (2025) Human-Likeness of LLMs in the Mental Lexicon. *Proceedings of the 29th Conference on Computational Natural Language Learning*, Vienna, July 2025, 586-601. <https://doi.org/10.18653/v1/2025.conll-1.38>

-
- [10] Nelson, D.L., McEvoy, C.L. and Schreiber, T.A. (2004) The University of South Florida Free Association, Rhyme, and Word Fragment Norms. *Behavior Research Methods, Instruments, & Computers*, **36**, 402-407. <https://doi.org/10.3758/bf03195588>
- [11] Kiss, G.R., Armstrong, C., Milroy, R., Piper, J., Aitken, A., Bailey, R. and Hamilton-Smith, N. (1973) *The Computer and Literary Studies*. University Press.
- [12] van Heuven, W.J.B., Mandera, P., Keuleers, E. and Brysbaert, M. (2014) SUBTLEX-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, **67**, 1176-1190. <https://doi.org/10.1080/17470218.2013.850521>
- [13] Barabási, A.L. and Albert, R. (1999) Emergence of Scaling in Random Networks. *Science*, **286**, 509-512. <https://doi.org/10.1126/science.286.5439.509>
- [14] Collins, A.M. and Loftus, E.F. (1975) A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, **82**, 407-428. <https://doi.org/10.1037/0033-295x.82.6.407>
- [15] Turney, P.D. and Pantel, P. (2010) From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research*, **37**, 141-188. <https://doi.org/10.1613/jair.2934>
- [16] De Deyne, S. and Storms, G. (2008) Word Associations: Network and Semantic Properties. *Behavior Research Methods*, **40**, 213-231. <https://doi.org/10.3758/brm.40.1.213>