

あなたの梅子度はどのくらい？ 親密度効果と大学生の属性推定への応用

How much Umeko-effects do you get?

Familiarity and its application for personal property estimation of university students

小野 創[†], 柳野 祥子[‡], 小川 萌子[†], 大塚 亜未[†]
Hajime Ono, Ami Otsuka, Shoko Yanagino, Moeko Ogawa

[†] 津田塾大学, [‡] 株式会社ジャストシステム
Tsuda University, JustSystems Corporation
hajime@tsuda.ac.jp

Abstract

It has been well known that the response times toward a word modulate due to its lexical frequency and word familiarity. The word familiarity available in the literature is collected often through a kind of questionnaire from a large group of the participants. On the other hand, it is conceivable that the familiarity of a given word varies depending on the interests or experience from a particular individual. We measured the lexical decision latency for words that are frequently used by students from Tsuda University. Such a reaction time latency was compared with that for words used widely and commonly, and the latency difference is termed as “Tsuda-familiarity effect”. We also report that the Tsuda-familiarity effect was observed in students from Tsuda University, but not in participants who are not from Tsuda University. We conclude that the familiarity effects observed in this study can be used for estimating a phase of personal property of university students.

Keywords — Lexical decision, familiarity, personal property estimation, reaction times

1. はじめに

単語に対する語彙判断に要する時間は様々な要因によって変化することは広く観察されており (for a general overview, Traxler 2012, etc.), 語彙頻度や親密度は代表的な要因として知られているものの一部である (Wolk, et al. 2004; Connine, et al. 1990; その他多数). 語彙頻度や親密度の高い単語は、それらが低い単語に比べて、一般的に反応速度が速くなる。その際拠り所とする親密度は、多くの被験者を用いたアンケートなどを通して得られた評定値であることが多い (Nusbaum & Dedina 1985, Whittlesea, et al. 1990, 天野・近藤・笠原 2008)。

このような親密度評定値は、母語話者集団全体としての総体的な親密度を反映したものということもできるかもしれない。当然ながら、個人の体験は集団の平均とは完全に一致するわけではなく、その個人の生活

パターン、趣味、嗜好、経験などの様々な要因によって、どのような単語に対して親しみを持つか（接触の機会が多いか）は変化する。個人の体験などによって親密度が異なるというこのような観察を個人の属性推定に生かす取り組みは、情報工学分野でも活発に研究が進んでいる (Abel, et al. 2011, Conover, et al. 2011; その他多数)。マーケティングを通して、消費者一人一人に対してのカスタマイズや、語学学習の際の学習者一人一人の学習スタイルに合わせたカリキュラムの提案などへの応用も望まれている。

本研究では、親密度ベースの個人の属性推定が、語彙判断に要する反応時間という指標からどの程度可能であるのかを検討することを目的とする。特定の大学に属する学生が日常頻繁に接する単語を選定し、それらを親密度の操作として用いた語彙判断実験を実施した。そして、そのような親密度指標が属性推定に一定程度有効であることを明らかにした。

2. 実験

語彙判断課題を実施した。被験者は津田塾大学の学生 35 名（すべて女性）と他大学の学生など津田塾大学に属していない日本人母語話者 34 名（うち女性 19 名）であった（被験者の平均年齢は 22 歳）。ターゲット刺激語として、津田塾大学内で学生が日常の大学生活で接する機会の多い単語で、一般的にも日本語語彙として認知されている語を 60 語選定した（津田語：例「女性」「ジェンダー」「多文化」）。加えて、広く一般で使われている単語で、かつ上記の津田語と語彙頻度が大きく異ならない単語も 60 語選定した（非津田語：例「人参」「防寒」）。津田語と非津田語の間は、平均文字数や平均モーラ数が異ならないよう、統制された。また、津田語と非津田語の間で、一般的な親密度に違いがないように、天野他 (2008)をもとにそれぞれの語に対す

る親密度を調査し、違いがないように統制した(表1)。つまり、津田語と非津田語の間には、一般的な意味での親密度の差はないと考えられる。また付属的な分析の目的のために、それらのターゲット刺激は、語彙頻度によって高頻度語と低頻度語に分けられた(表2)。語彙判断課題を実施するにあたり、非単語フィラーを120 試行分用意した(例「むんご」「落符」)。

試行は、凝視点(1000ms)と空白(500ms)の後、文字列が表示された。被験者には、呈示された文字列が日本語の単語であるかどうかをできるだけ速く正確に、キーボードのキーを押すことで回答するように教示した。刺激呈示の制御には PsychoPy (Pierce, 2007) が用いられた。

表 1. 刺激語の平均親密度(7 点満点)と標準偏差

	平均親密度	標準偏差
津田語	4.83	1.62
非津田語	5.00	1.35

表 2. 各条件の刺激語における語彙頻度(対数頻度) カッコ内は標準偏差

	高頻度	低頻度	計
津田語	2.977 (0.589)	1.272 (0.566)	2.122 (1.028)
非津田語	2.517 (0.396)	1.564 (0.544)	2.040 (0.673)

3. 結果と考察

反応時間が他の被験者に比べて著しく遅かった被験者 1 名(非津田生)のデータならびに正答率の悪かった 5 つの刺激語(非津田語)のデータを除外した。残りのデータに基づいて、被験者属性と津田親密度別の平均正答率を算出した。それらの要因に基づいて分散分析を実施したところ、交互作用が観察され($F(1,66)=53.57, p<.001$)、津田生は津田語に対する正答率が非津田語に対する正答率よりも高く、非津田生では逆のパターンがあることがわかった。

表 3. 被験者属性ごとの正答率(カッコ内は標準偏差)

	津田生	非津田生
津田語	98.2 (3.1)	87.2 (7.4)
非津田語	93.3 (4.7)	92.5 (4.5)

残りのデータから誤答の試行データを除外しターゲット刺激語に対する平均反応時間を算出した(表4)。その上で、被験者の属性(津田生、非津田生)という被験者間要因、語彙頻度(高頻度語、低頻度語)と津田親密度(津田語、非津田語)の 2 つの被験者内要因に基づいて分散分析を実施した。

表 4. 被験者属性・語彙頻度・津田親密度ごとの平均反応時間(カッコ内は標準偏差)

	津田生	非津田生
津田語		
高頻度語	670 (100)	690 (118)
低頻度語	800 (139)	940 (244)
非津田語		
高頻度語	744 (122)	763 (138)
低頻度語	824 (154)	837 (174)

語彙頻度による主効果が観察された($F(1,66)=175.26, p<.001$)。また、被験者の属性と津田親密度の交互作用が観察され($F(1,66)=16.46, p<.001$)、単純主効果の検定から津田親密度効果が津田生には観察された($p<.03$)一方、非津田生にはそのような効果が観察されなかったこと(n.s.)が明らかになった。具体的には、津田生は非津田語に比べて津田語に対する反応時間が有意に速かったが、非津田生にはそのような差がなかった。

本研究は、母語話者全体として観察される親密度ではなく、特定の大学(津田塾大学)に属する集団が日常接する単語を親密度の高い語として選定し語彙判断課題を実施した。その親密度効果は津田塾大学の学生には観察されたが、そこに属さない被験者群では観察されなかった。この結果は、特定の単語群に対する反応時間を用いることで被験者の属性の推定に用いることが可能であることを示唆する。この津田語に対する親密度効果(津田語と非津田語の反応時間の差)の大きさを、大学創設者の津田梅子にちなんで「梅子度」として指標化し、今後簡便に測定を実施するために、iOS と Android で動作可能なアプリケーションを開発した。現在アプリケーションについては、公開をするために処理を行っている(追加の情報は第 1 著者のウェブサイトに掲載予定)。

このような仕組みを通して、学生一人ひとりの大学への定着度の推定や学生生活に対する不適応への早期介入のヒントを捉えることが可能かもしれない。この

ような推定について、十分な効果を得るためのサンプル数の同定、またどのような刺激語が安定的に推定に寄与するのかなどについて今後検討することで、このような類いの親密度効果の背景となる要因について明らかにする端緒となることが期待される。

参考文献

- [1] F. Abel, Q. Gao, G. J. Houben and K. Tao, (2011) “Semantic Enrichment of Twitter Posts for User Profile Construction on the Social Web”, *The Semantic Web, Research and Applications*, Springer, Berlin Heidelberg, pp. 375–389.
- [2] C. M. Connine, J. Mullennix, E. Shernoff, and J. Yelen, (1990) “Word familiarity and frequency in visual and auditory word recognition”, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 1084–1088.
- [3] M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini and F. Menczer, (2011) “Predicting the Political Alignment of Twitter Users”, *Proc. of the International Conference on Social Computing*, 192–199.
- [4] H. Nusbaum and M. Dedina, (1985) “The effects of word frequency and subjective familiarity on visual lexical decisions”, (Research on Speech Perception, Progress Report No. 11). Bloomington, IN: Indiana University.
- [5] J. W. Peirce, (2007) “PsychoPy - Psychophysics software in Python”, *Journal of Neuroscience Methods*, 162(1-2), 8–13.
- [6] M. J. Traxler, (2012) *Introduction to psycholinguistics*. Wiley-Blackwell, Chichester, UK.
- [7] B. W. A. Whittlesea, L. L. Jacoby, and K. Girard, (1990) “Illusions of immediate memory: evidence of an attributional basis for feelings of familiarity and perceptual quality”, *Journal of Memory and Language*, 29, 716–732.
- [8] D. A. Wolk, D. L. Schacter, A. R. Berman, P. J. Holcomb, K. R. Daffner, and A. E. Budson, (2004) “An electrophysiological investigation of the relationship between conceptual fluency and familiarity”, *Neuroscience Letters*, 369, 150–155.
- [9] 天野成昭・近藤公久・笠原要, (2008) “NTT データベースシリーズ: 日本語の語彙特性第4期(第9巻)”, 東京, 三省堂.