

引用内容分析サービス sciteのご紹介

AIを使った文献評価 - 発見した文献の評価を引用文献から知る



2023年12月度：第407回ライフサイエンス分科会
2023年12月15日

株式会社サンメディア

scite サービス提供の経緯

scite_

- 2018年に共同出資者でCEOのJosh Nicholsonがニューヨークのブルックリンで創業したスタートアップ
- CEOのJosh Nicholsonは、ヴァージニア工科大学で細胞生物学の分野でPhDを取得し、癌における染色体分離に対する異数性の影響を研究
- sciteの創業前に、どのように科学者による出版と協力を改善するかを目的とした会社も設立
- sciteの目的は「研究者が関連論文の中身を読むための膨大な時間を節約し、また不適切な研究内容に依拠して自身の研究を進めてしまうことを防ぐこと」
- 既に様々な賞や、NSF（アメリカ国立科学財団）やNIH（アメリカ国立衛生研究所）の助成金を得るなど高い評価を獲得
- 2023年11月にResearch Solutionsのビジネスユニットの一つになる



scite CEO - Josh Nicholson



ISMTE People's
Choice Award



ALPSP Award
for Innovation in
Publishing



NSF Phase 1
SBIR Grant



NIH/NIDA SBIR
Fast-track grant



Vesalius Prize
Runner Up

<https://www.sunmedia.co.jp/solution-seminar-2023-site/>

“scite”とは？

sciteでできること

scite_

Sciteでは、**スマートサイテーション**による文脈を踏まえた引用文献の評価チェックができるだけでなく、そのスマートサイテーションを駆使して様々な機能を提供しています。研究者様のリサーチだけでなく、研究支援部門のご担当者様にもお役立ていただける機能をご用意しています。



引用文献の評価チェック

引用する論文の信頼性や妥当性を手早く簡単にチェックすることができます。



ジャーナルのリサーチ

評価スコアやscite独自のインデックスを使い、評価の高いジャーナルを調べることができます。論文の投稿先の選定等にお役立ていただけます。



研究機関のリサーチ

機関単位での評価をチェックすることができるため、自機関を含めた研究機関の業績評価を簡単に確認できます。



研究テーマのリサーチ

特定の研究テーマに関する重要な文献や主要な研究者を簡単に調べることができます。



研究助成機関のリサーチ

特定の研究者の業績の評価を手早くチェックすることができます。



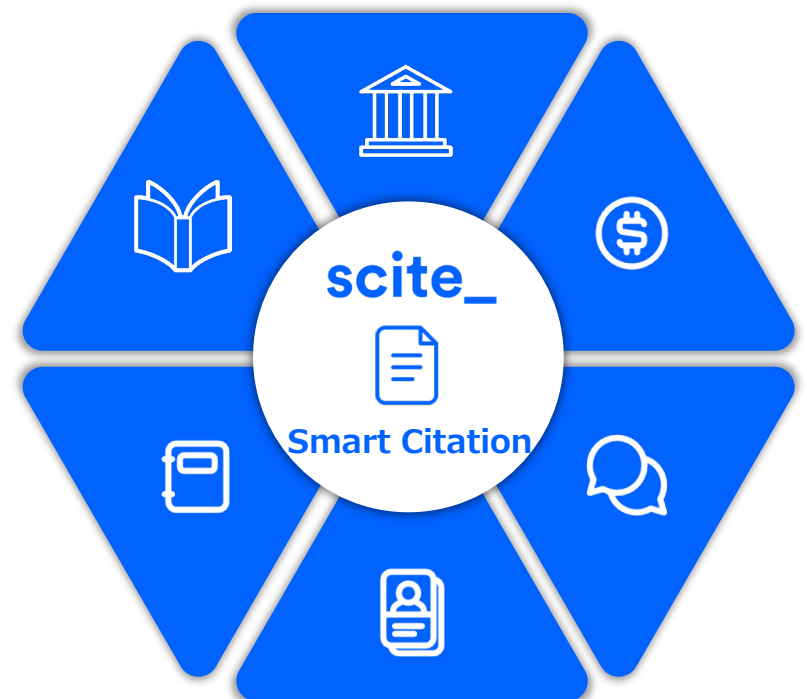
研究者のリサーチ

研究者の評価スコアを確認できます。共同研究者のリサーチを行うことができます。



Generative AI for Science

ChatGPTやClaudeをカスタマイズした機能を通じて、生成AIが回答の根拠となる学術情報を明示します。



- 見つけた多くの論文がどのように引用されているかを見ることで、その論文が自分にとって有用かどうかを簡単に知ることができる。
- 見つけた多くの論文の全てに目を通す必要がなく、時間の浪費や不要な文献の入手、ダウンロードを防げる。

その結果、

- 研究者は時間を有効に使って研究やマーケティングに専念できる
- 研究者が研究論文を検証し、答えを見つけ、クラス最高の同僚を見つけることもできる
- 意思決定者が自身の研究プログラムや競合他社を検証し、戦略的パートナーシップや主要なオピニオン リーダーを特定するのに役立てることも可能になる。

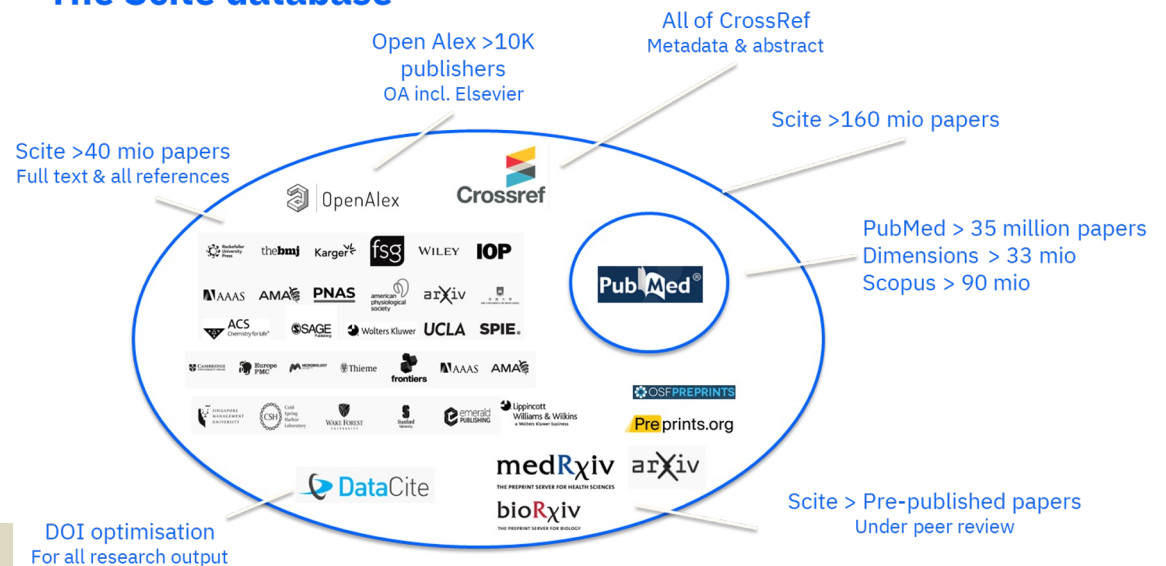
sciteのデータソース

scite_

scite には **1 億 8,700 万件**を超える**出版物**の記録があり、**3,450万件**を超える**学術論文の全文**が取り込まれています。
これらの論文から **20 億件**の**引用文**を抽出しています。（2023年12月現在）

- 査読付き出版物を発行する大手、中堅、学会出版社のリストを多数収録
- オープンアクセス+主要出版社とのインデックス契約によりオープン+クローズドの双方の論文の引用が利用可能
- 自然科学・社会科学・人文科学など全学問領域を幅広くカバー
- arXivなどプレプリントもカバー
- Crossrefのメタデータ、引用情報／被引用情報を搭載
- OurResearch (Unpaywall, etc.) との協力によりメタデータをエンリッチメント

The Scite database



最新の収載データ数を公開しています。

■ URL

<https://scite.ai/data-and-services>



sciteの対象分野

scite_

Subject	Citations	Citation Statements	Contrasting	Mentioning	Supporting
Biology	688.1m	512.3m	2.7m	476.1m	22.7m
Medicine	612.5m	443.4m	3.2m	404.6m	19.9m
Chemistry	526.7m	340.4m	1.5m	317.2m	16.0m
Physics	444.8m	309.1m	1.2m	290.0m	12.8m
Computer science	439.7m	295.7m	790.7k	280.1m	7.8m
Materials science	287.0m	172.8m	569.8k	162.7m	7.0m
Mathematics	274.8m	197.5m	577.7k	187.1m	5.9m
Engineering	257.7m	155.6m	458.5k	146.9m	4.7m
Psychology	214.2m	164.4m	1.1m	148.8m	8.0m
Economics	143.7m	101.2m	483.7k	93.6m	3.7m
Philosophy	127.5m	88.6m	336.8k	80.6m	2.9m
Sociology	122.7m	91.8m	546.9k	83.6m	3.7m
Political science	118.9m	83.0m	299.5k	75.8m	2.4m
Geology	109.2m	74.4m	308.9k	69.4m	3.4m
Geography	89.6m	63.5m	295.8k	58.5m	2.4m
Environmental science	89.1m	57.8m	297.8k	53.6m	2.5m
Business	77.8m	55.2m	203.2k	50.8m	1.7m
History	26.8m	16.1m	58.0k	14.9m	481.1k
Art	24.5m	15.6m	49.8k	13.4m	438.1k

Smart Citation – 引用文献の評価

sciteとは

scite_

sciteは「スマートサイテーション」という独自の引用内容分析情報を提供するサービスです。

The screenshot displays the scite_ website's search interface. At the top, there's a search bar with the placeholder text "Search by title, author, keywords or DOI". To the right of the search bar are navigation links: Assistant, Product, Discover, Pricing, Blog, and a user profile icon. Below the search bar, the "Search" section is active, with tabs for All, Citation Statements, Papers, and scite Assistant (marked as Beta). There are also links for Advanced search, a document icon, a share icon, a bell icon, and a download icon. Below the tabs, there are filters for Authors, Year, Sections, Types, Citations, Journal, Affiliation, Pub Type, Topics, Editorial Notice, MeSH, and More. The search results show 161,793,922 results in 3.82 seconds. A link "Help me understand these results" and a "Sort by: Relevance" dropdown are present. The first result is a German paper titled "Aktualisierte Vergleichstafeln für militärische Lastenklassen bei Straßenbrücken" by Patrick Forman, Christina Fust, and Peter Mark, published in 2012 in the journal "Beton- und Stahlbetonbau". The abstract is partially visible, showing text about load models and bridge structures. There are icons for copying, sharing, and downloading the abstract. At the bottom of the result card, there are buttons for "View full text", "Add to dashboard", and "Cite".

Smart Citation Demonstration

スマートサイテーションの例

scite_

書誌情報

抄録

②Citation Statement
(該当の論文の被引用論文における文中引用箇所およびその前後の文章を表示)

①どのように評価(引用)されているのかが分かるスコア

Brain and Language 2012 DOI: 10.1016/j.bandl.2010.08.003 [View full text] [Cite] [Set alert] [Twitter] [Facebook] [LinkedIn]

What artificial grammar learning reveals about the neurobiology of syntax

Karl Magnus Petersson^{a,*}, Vasiliki Folia^{a,*}, Peter Hagoort^a

Abstract: a b s t r a c t In this paper we examine the neurobiological correlates of syntax, the processing of structured sequences, by comparing fMRI results on artificial and natural language syntax. We discuss these and similar findings in the context of formal language and computability theory. We used a simple right-linear unification grammar in an implicit artificial grammar learning paradigm in 32 healthy Dutch university students (natural language fMRI data were already acquired for these participants). We predic... [Show more](#)

[Help me understand this report](#)

Search citation statements

Context, author(s), title e... Q

Order By: Relevance

Paper Sections

Select...

- ☐ Introduction 58
- ☐ Discussion 31
- ☐ Artificial Grammar Learning 17
- ☐ Neurobiological Hypotheses 7
- ☐ On the Proto-syntactic Learnin

Citation Types

- ☐ Supporting 19
- ☐ Mentioning 206
- ☐ Contrasting 2
- ☐ Unclassified 3

Cited by 159 publications (230 citation statements)

References 104 publications (185 reference statements)

"...The outcome of this comparison suggests that the left inferior frontal region (BA 44/45) is significantly related to implicit AGL and artificial syntax processing, independent of the fact that the simple right-linear unification grammar we investigated does not involve nested center-embedded non-adjacent dependencies or dependencies introduced by syntactic displacement (i.e., syntactic movement). These findings are similar to corresponding findings reported in Petersson et al. (2010) for the grammaticality instruction. Thus, in the context of artificial syntax processing, and more generally language processing, the left inferior frontal region is unlikely to be specific to the processing of or nested center-embedded structures or non-adjacent dependencies introduced by syntactic movement...."

Section: Discussion supporting confidence: 92%

"...This is the first fMRI study to investigate implicit AGL with preference classification in a proper learning design (i.e., including baseline measurements prior to grammar exposure). The main fMRI findings are consistent with previous grammaticality classification results (Petersson et al., 2004, 2010; Forkstam et al., 2006; Folia et al., 2011). Importantly, and central to this study, we show that the preference classification results are the outcome of the underlying implicit learning process...."

Section: Discussion supporting confidence: 91%

See 8 more Smart Citations

Implicit structured sequence learning: an fMRI study of the structural mere-exposure effect

Folia^a, Petersson^a 2014 *Front. Psychol.* Self Cite

21 6 37 0

View full text Add to dashboard Cite

"...These findings are consistent with previous reports on the activation of temporal cortices in various auditory SL paradigms (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2006; Plante et al., 2015). Human temporal regions are sensitive to processing of auditory sequential patterns in artificial languages, in addition to the classical frontal regions (Petersson et al., 2012; Wang et al., 2015), and are reliably activated during syntactic processing of natural languages (Fedorenko et al., 2020; Snijders et al., 2009)...."

Section: Discussion supporting confidence: 92%

該当の論文の
被引用文献一覧

スマートサイテーションの特徴①

scite_

①どのように評価（引用）されているのかが分かるスコア

スマートサイテーションでは、該当の論文の引用評価をSupporting, Mentioning, Contrastingという三つのカテゴリーに分類して表示しています。分類は、引用の文脈を三つのカテゴリーに分類するためにディープラーニングを使ってトレーニングされた人工知能によって行われています。

引用の回数による評価だけでなく、「どのように評価（引用）されているのか」が一目で分かるようになっています。

Citation Types ①

☐	Supporting	✓	19
☐	Mentioning	🕒	206
☐	Contrasting	?	2
☐	Unclassified	○	3

Supporting: 引用された研究の裏付けとなる証拠を提供する引用

Mentioning: 有効性の証拠を提供せずに言及された引用

Contrasting: 引用された研究の対照的な証拠を提供する引用

Unclassified: 判別不可の引用

スコアリングの仕組み

scite_

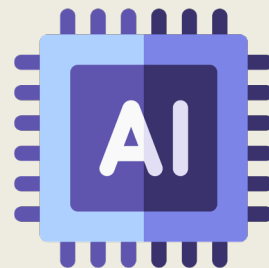
「スマートサイテーション」は、ある論文が「どのくらいの数の論文で引用されているのか」だけでなく、「**どのように引用・評価されているのか**」を独自のスコアで示します。また、その引用箇所の前後の文章を抽出し、「**どのような文脈で該当の論文が引用されているか**」を示します。

■ スコアリングの簡単な例

この実験結果は、●●による研究 [論文A] と共通の結果を示していた。

.....一方、
については先行研究 [論文B] とは相反する結果が示されていた。
.....

AIがフルテキストを解析



論文AのSupportingに +1

Citation Types ⓘ

☐	Supporting	✓	2
☐	Mentioning	⌚	15
☐	Contrasting	?	0

論文BのContrastingに +1

Citation Types ⓘ

☐	Supporting	✓	19
☐	Mentioning	⌚	206
☐	Contrasting	?	2

スコアリングの技術的な詳細は公開されています。

scite_

スコアリングの仕組みを含むsciteの技術的な詳細については、MIT PressのOA論文として公開されています。

The screenshot shows the MIT Press Direct website interface. At the top, there's a navigation bar with 'MIT Press Direct', a search bar, and links for 'Quantitative Science Stud', 'Advanced Search', 'Register', 'PACIFIC LUTHERAN UNIVERSITY', and 'Sign In'. Below this is a header for 'Quantitative Science Studies' with links for 'Issues', 'Online Early', 'About', and 'Submit'. The main content area displays the article 'scite: A smart citation index that displays the context of citations and classifies their intent using deep learning' by Josh M. Nicholson et al., published in November 05 2021. The article is from 'Quantitative Science Studies (2021) 2 (3): 882–898'. The page includes a 'Check for updates' button, 'Author and Article Information', and a list of authors. On the left, there's a sidebar with 'Volume 2, Issue 3', 'Summer 2021', and 'Article Contents' (Abstract, PEER REVIEW, 1. INTRODUCTION, 2. METHOD, 3. DISCUSSION, 4. CONCLUSIONS, ACKNOWLEDGMENTS). The abstract text is visible, starting with 'Citation indices are tools used by the academic community for research and research evaluation that aggregate scientific literature output and measure impact by collating citation counts.' On the right, there's an 'Email Alerts' section with options for 'Article Activity Alert', 'Closed Issue Alert', and 'Continuous Publishing Added Articles Alert'. Below this is a 'View Metrics' button with a '31' icon. At the bottom right, there's a 'scite_' widget showing a list of articles with a 'PDF' button and a 'Help' button. A large QR code is positioned to the right of the article page.

Josh M. Nicholson, Milo Mordaunt, Patrice Lopez, Ashish Uppala, Domenic Rosati, Neves P. Rodrigues, Peter Grabitz, Sean C. Rife; scite: A smart citation index that displays the context of citations and classifies their intent using deep learning. *Quantitative Science Studies*, MIT Press, 2021; 2 (3): 882–898. doi: https://doi.org/10.1162/qss_a_00146

スマートサイテーションの特徴②

scite_

②Citation Statement

該当の論文の被引用文献において、該当の論文が本文中でどのように引用されているのかを提示します。スコアを提示するだけでなく、その根拠となった実際の引用箇所やその前後の文章を示すため、利用者自身で該当の論文がどのような文脈で引用されているのかを効率的に判断することができます。

“...Given the possible slight difference between reference and the strain we used and possible mapping errors in repetitive regions, the genome recovery of the GMD samples is probably close to completion. The assembled results of E. coli data agreed with mapping results (Supplemental Table 3) using the SPAdes assembler and gave somewhat less coverage using our Velvet-based single-cell assembly pipeline (Bankevich et al 2012 ; data not shown). We found almost no contamination in E. coli data produced with either single cells or GMDs, and they do not differ significantly (Mann-Whitney test, P = 0.3939)....”

Section: Gmds Produce Better Sequencing Results Than Single Cells  **supporting** confidence: 66%

Nearly finished genomes produced using gel microdroplet culturing reveal substantial intraspecies genomic diversity within the human microbiome

Fitzsimons¹, Novotny², Lo³ et al. 2013 *Genome Res.*

 53 |  1 |  52 |  0

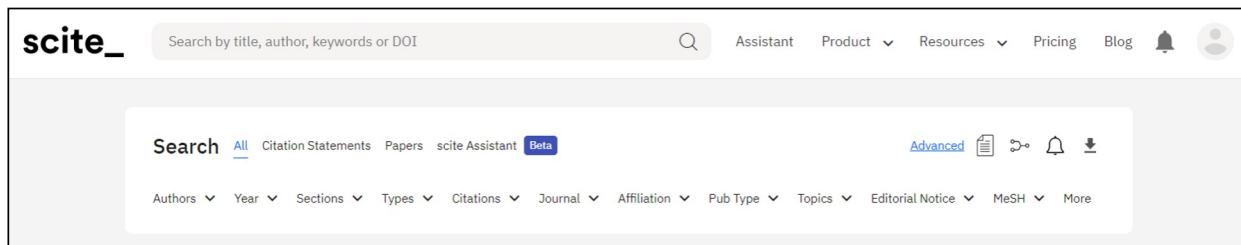
[View full text](#) [Add to dashboard](#) [Cite](#)

文中引用箇所とその前後の文章を抽出・表示

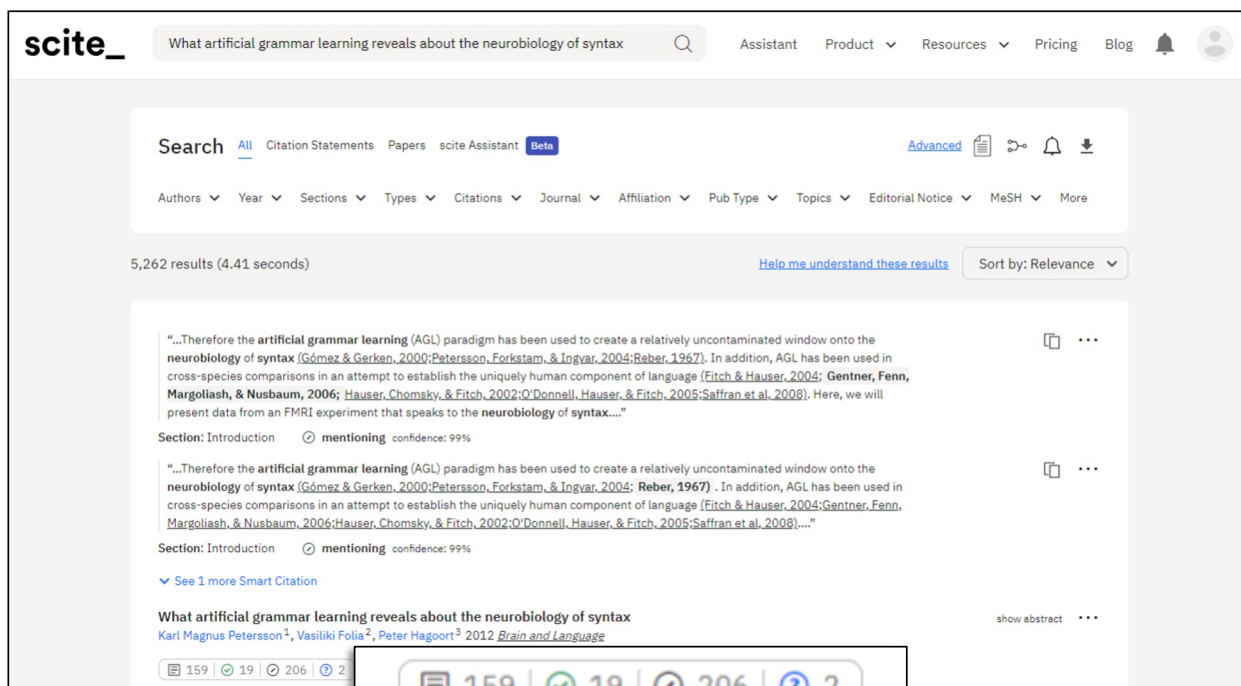


引用文献の評価チェック

scite_



検索画面で論文名や著者名、キーワードなどで検索します。



検索結果一覧で該当する論文をクリックします（この段階で、検索結果画面に論文の評価スコアが表示されています）。



引用文献の評価チェック

scite_

論文名をクリックすると、その論文のスマートサイテーションページが表示されます。

Brain and Language 2012 DOI: 10.1016/j.bandl.2010.08.003 [View full text](#) [Cite](#) [Set alert](#) [Twitter](#) [Facebook](#) [LinkedIn](#)

What artificial grammar learning reveals about the neurobiology of syntax

Karl Magnus Petersson¹, Vasiliki Folia², Peter Hagoort³

Abstract: a b s t r a c t In this paper we examine the neurobiological correlates of syntax, the processing of structured sequences, by comparing fMRI results on artificial and natural language syntax. We discuss these and similar findings in the context of formal language and computability theory. We used a simple right-linear unification grammar in an implicit artificial grammar learning paradigm in 32 healthy Dutch university students (natural language fMRI data were already acquired for these participants). We predic... [Show more](#)

[Help me understand this report](#)

Search citation statements

Context, author(s), title e...

Order By: Relevance

Paper Sections

Select...

- ☐ Introduction 58
- ☐ Discussion 31
- ☐ Artificial Grammar Learning 17
- ☐ Neurobiological Hypotheses 7 On the Proto-syntactic Learnin

Citation Types

- ☐ Supporting 19
- ☐ Mentioning 206
- ☐ Contrasting 2
- ☐ Unclassified 3

Cited by 159 publications (230 citation statements)

References 104 publications (185 reference statements)

"...The outcome of this comparison suggests that the left inferior frontal region (BA 44/45) is significantly related to implicit AGL and artificial syntax processing, independent of the fact that the simple right-linear unification grammar we investigated does not involve nested center-embedded non-adjacent dependencies or dependencies introduced by syntactic displacement (i.e., syntactic movement). These findings are similar to corresponding findings reported in Petersson et al. (2010) for the grammaticality instruction. Thus, in the context of artificial syntax processing, and more generally language processing, the left inferior frontal region is unlikely to be specific to the processing of or nested center-embedded structures or non-adjacent dependencies introduced by syntactic movement...."

Section: Discussion **supporting** confidence: 92%

"...This is the first fMRI study to investigate implicit AGL with preference classification in a proper learning design (i.e., including baseline measurements prior to grammar exposure). The main fMRI findings are consistent with previous grammaticality classification results (Petersson et al., 2004, 2010; Forkstam et al., 2006; Folia et al., 2011). Importantly, and central to this study, we show that the preference classification results are the outcome of the underlying implicit learning process...."

Section: Discussion **supporting** confidence: 91%

[See 8 more Smart Citations](#)

Implicit structured sequence learning: an fMRI study of the structural mere-exposure effect

Folia¹, Petersson² 2014 *Front. Psychol.* [Self Cite](#)

☐ 21 ☒ 6 ☐ 37 ☐ 0

[View full text](#) [Add to dashboard](#) [Cite](#)

"...These findings are consistent with previous reports on the activation of temporal cortices in various auditory SL paradigms (Cunillera et al., 2009; Karuza et al., 2013; McNealy et al., 2006; Plante et al., 2015). Human temporal regions are sensitive to processing of auditory sequential patterns in artificial languages, in addition to the classical frontal regions (Petersson et al., 2012; Wang et al., 2015), and are reliably activated during syntactic processing of natural languages (Fedorenko et al., 2020; Snijders et al., 2009)...."

Section: Discussion **supporting** confidence: 92%