

第二言語における

コロケーション形式への敏感さと母語の関係

名古屋大学大学院 草薙 邦広

1. 研究背景

1. 1 コロケーションの習得研究

近年、第二言語(L2)における習得・運用に関する研究では、複数の語からなる慣習的な言語表現(本稿では一般的総称として「コロケーション」と呼ぶ)も研究の対象となってきている。一般的に、コロケーションの使用は、流暢な言語使用に貢献し(e.g., Wray, 2002), 談話的・社会言語学的側面においても重要な機能を持つ(e.g., Skehan, 1998), と考えられており、コロケーションの習得はL2学習者の重要な目標のひとつであるといえよう。

心理言語学的な観点では、コロケーションは、コロケーションでない語の連鎖よりも速く読解され(e.g., Conklin & Schmitt, 2008), より正確に安定して処理される(e.g., Jiang & Nekrasova, 2007), という事が分かってきている。このような知見から、コロケーションは、複数の語を逐次的にではなく単一の単位として処理される(全体処理)と考えられている(Jiang & Nekrasova, 2007)。コロケーションの使用は、このような全体処理の恩恵によって話者及び聞き手の認知的負荷を軽減し、流暢で効率的な言語使用に貢献すると考えられている。

一方、L2学習者におけるコロケーションの産出傾向についての研究も盛んになされてきた(e.g., Hawarth, 1998)。これらの研究は、学習者による特定のコロケーションの過少使用・過剰使用や、誤りの傾向の分析を対象としてきた。また、厳密には誤りではないものの、明らかに個別言語の慣習的側面を逸脱した不自然なコロケーション(Malformed Collocation)を学習者が頻繁に産出することもよく知られている(Millar, 2010)。

Bilal Anwar and Khan(2012)は、パキスタン人の高熟達度英語学習者に対して、コロケーションに関する多肢選択課題とエッセイ課題を用いて、コロケーションにおける受容的能力、及び産出能力を調査した。結果、実験の被験者は、多肢選択課題においては優れた成績を示したものの、エッセイ課題では以下の例のような不自然なコロケーションを多数産出したと報告されている。

1. Huge help
2. Small accident
3. Hear the music

Bilal Anwar and Khan (2012, pp. 62-63)

1. 2 不自然なコロケーションに対する敏感さ

しかしながら、実際の言語使用において、不自然なコロケーションがどのような認知的影響を持つかということ、実証的に検証した研究は非常に限られている。コロケーションのオンライン処理という観点からは、対象言語の慣習的結びつきを逸脱する表現は、二語を一単位として扱う全体処理がなされることはなく、認知的負荷を高め、効率的な言語使用を阻害すると考えられる。

Millar (2010) は、日本語を母語とする英語学習者 (JLE) のエッセイを採録したコーパスから、母語話者が不自然と判断する代表的なコロケーションを抽出し、更に、それに対応する自然な英語表現のペアを作成した。ペアとなる自然な表現は BNC (British National Corpus) から頻度、相互情報量などの指標などと照らし合せて選出し、更に母語話者を対象とした予備実験を用いて、不自然なコロケーションがそれに対応する自然なコロケーションより低い親密度評定値を持つことを確認した。

その後、英語母語話者を被験者として、自然なコロケーション表現、不自然なコロケーション表現を含む英文を刺激材とした自己ペース読み課題を実施した。結果、英文の読解中において、不自然なコロケーションは統計的に有意な読解時間の遅れを生む事が明らかになった。この事から、Millar は、母語話者はコロケーションの慣習性に敏感であり、不自然なコロケーションは読み手の認知負荷を増加させ、流暢で効率的な言語使用を阻害すると結論付けている。

しかし、L2 学習者の場合には、Millar の実験結果とは異なり、読解時間の遅延をもたらさない可能性もある。Kusanagi, Leung, Bando, Fukuta, and Sugiura (in press) は、視線計測、移動窓方式の自己ペース読み課題、コロケーションの親密度評定の三つの手法を併用し、高熟達度 JLE 計 31 名に対して、Millar の追行実験を行った。視線計測の結果では、用いられた全ての視線計測測定値 (初回注視時間、注視継続時間、総注視時間) において、自然・不自然なコロケーション間で差が統計的に有意ではなかった。また、自己ペース読み課題における対象領域の平均反応時間の比較でも結果は同様であった。更に、文脈の無いフレーズ単位提示での親密度評定においても、自然・不自然なコロケーション間において平均値の差は有意ではなかった。この事から、Kusanagi, et al.

は、母語話者とは異なり、(実験に参加した)学習者はコロケーションの不自然さに対して鈍感(insensitive)である、と結論づけている。

1. 3 コロケーションの習得における母語の影響

コロケーション習得における母語の影響については、Yamashita and Jiang(2010)の研究がある。第一に、彼女らは母語と L2 の心的辞書において、それぞれ異なる形式が意味と結びつきを持ち得ると捉えている。そして、L2 コロケーション知識の発達とは、L2 における形式と意味の直接的な結びつきを得ることである、と述べている(p. 652)。

Yamashita and Jiang (2010)は、母語話者、外国語環境(EFL)で学習する JLE、第二言語環境(ESL)で学習する JLE に対して、母語との逐語的翻訳関係において対応があるコロケーション(Congruent Collocation)と対応がないコロケーション(Incongruent Collocation)を刺激としたフレーズ性判断課題を用いた。正答率と反応時間の分析の結果は、ESL 学習者も EFL 学習者も両方、母語の対応関係の影響を受けていることが明らかになった。しかし、EFL 学習者は正答率においても、反応時間においても有意な正答率の低下、反応時間の遅れを生んだのに対し、インプットを大量に得る機会を持つ ESL 学習者は反応時間においては影響がなかった。この事は、言語間で対応関係の無いコロケーションは、母語の干渉から習得が困難であるが、一度習得された場合、母語の影響が無くなることを示している(p. 661)。

一方、コロケーションの産出における母語の影響について、Wolter (2006)は、不自然なコロケーションは母語との表現に対応関係が無い場合に、学習者が母語のコロケーション知識に頼る結果である、と示唆している(p. 742)。しかしながら、低頻度語や上位概念語での置き換えをするという学習者の一般的傾向も考えられる。その為、必ずしも不自然なコロケーションの全てが母語の対応関係による影響であるということはないが、言語間におけるコロケーション表現の対応関係は、コロケーションの産出においても重要な要因であろう。

したがって、Kusanagi, et al. (in press)の結果では、学習者はコロケーションの不自然さに鈍感であると示されているが、その鈍感さは、母語の影響によるものなのか調査する必要があるといえるだろう。特に、Kusanagi, et al. は、不自然なコロケーションを産出した学習者と同じ母語を持つ被験者による実験しか行っていない為、異なる母語を持つ学習者と、直接的にその敏感さを比較する交差型実験が必要であろう。

2. 研究課題

上記の先行研究を踏まえ、以下のように研究課題を設定する。

研究課題: 不自然なコロケーションに対する敏感さには母語の影響があるか。

本研究では、JLE が産出する不自然なコロケーションを、中国語を母語とする英語学習者 (CLE) と JLE 自身がその表現の自然さを評価するという実験を行う。日本語と中国語において、類型論的に個々のコロケーション表現の対応性を予測することは容易ではないが、各個別言語間における意味と形式の対応関係の違いが、各項目に対する反応の差を生むことも考えられる。

3. 実験

3. 1 被験者

本研究の被験者は 32 名であり、実験のデザイン上、二群に分けられている。被験者の基本的な情報は表 1 の通りである。

被験者群 1 は、日本語を母語として使用する 16 名の英語学習者である。全員が日本の大学に所属する大学院生であり、学術目的の為に英語を使用、学習している。被験者が有する TOEIC スコアから、全員の英語熟達度は相対的であるものの、高熟達度であると判断できる。

被験者群 2 は、中国語 (普通話) を母語として使用する英語学習者 16 名で構成されている。被験者群 1 と同様に、全員が日本の大学に所属する大学院生であり、学術目的で英語を使用、学習している。この群における全ての被験者は、高等教育において第二外国語として日本語を学習している。更に、実験実施の時点において、日本語が英語と比べて優勢であると報告した者が、全体に対して 62.50% 含まれている。被験者の英語熟達度については、CET (Chinese English Test) における取得級で報告している。被験者群 1 とは直接的な比較が出来ないものの、中級以上の英語熟達度を有していると考えられる。

また、全ての被験者に対して、5 件法を用いて四技能における自己評定を聴取した。2 つの群の、年齢、英語学習歴、及び四技能それぞれの平均値の差において、効果量 (d)¹ を産出したところ、年齢、及び話す能力で効果量中、その他の変数では効果量小を示した。この事から、被験者の背景における二群の差は全体的には比較的小さいと考えられる。²

表 1. 被験者の背景情報

	JLE (<i>n</i> = 16) <i>M</i> [<i>SD</i>]	CLE (<i>n</i> = 16) <i>M</i> [<i>SD</i>]	Cohen's <i>d</i>
年齢 (歳)	25.17 [2.86]	27.09 [2.48]	0.71
英語学習歴 (年)	11.50 [2.24]	12.18 [3.99]	0.21
聞く能力	3.08 [1.00]	2.82 [1.08]	0.25
話す能力	2.83 [0.94]	2.36 [0.92]	0.52
読む能力	3.25 [0.75]	3.09 [1.04]	0.18
書く能力	2.83 [1.03]	2.73 [0.90]	0.10
熟達度	TOEIC : 802.03 [76.53]	CET : 四級 37.50% 六級 25.00% 八級 37.50%	n.a.

3. 2 課題

実験は二つの紙面上での課題を以て実施された。

一つ目の課題は、「コロケーション選択課題」である。この課題では、被験者は 24 文の項目を読み、空欄の部分について、より自然で、文脈によりよく適合する語を、2つの選択肢の中から選択する。二つの選択肢は、日本語を母語とする学習者が産出した不自然なコロケーションと、それに対応する母語話者らしい自然なコロケーションの組である(以下に例を挙げる)。下例では、made が自然な表現であり、made を選択した場合に 1 点を得点したものとした。

The cafe was so busy that the waitress () mistakes and gave customers the wrong orders. [made / did]

二つ目の課題は、「親密度評定課題」である。この課題では、被験者は文脈の無い句提示のコロケーションを読み、「その表現がどれだけ自然に感じるか」を 5 段階(リッカート方式)で評定する。本研究では、評定値を間隔尺度とみなす。評定する項目の数は一被験者につき、24 フレーズである。

親密度評定課題における不自然・自然な項目の組み合わせについて、被験者間でカウンターバランスを行った。更に、項目の提示順序は 2 要因 2 個ずつ(自然 2, 不自然 2)を 1 ブロック(ブロック内の順序と組み合わせは固定)として、3 ブロックの全組み合わせにて、計 6 フォーム作成し、各被験者に対して割り

当てた。また、課題の練習効果の観点から、全ての被験者に対して、親密度評定課題を先に実施した。上記の2課題の他に、課題前に被験者の背景情報に関する情報を、質問紙を用いて記入させた。実験にかかった時間は最大で一人当たり25分程度であった。

3.3 刺激材

実験に用いた刺激は、Millar(2010)、及びKusanagi, et al. (in press)で使用された、日本語を母語とする英語学習者のコーパスから抽出された不自然な表現とそれに対応する英語母語話者らしい自然な表現の組である。Kusanagi, et al. は、Millar が母語話者に対する実験の為に開発した項目を、L2 学習者用に使用語彙の難易度、文長の統制の為に、24 組(48 文)を選定し、高難易度語彙の置き換え等を行い、実験を実施した。本研究は、Kusanagi, et al. で使用された刺激を採用した。刺激の一部(全体の約5分の1にあたる5組)を本稿の付録に記載している。

3.4 分析

文提示コロケーション選択課題の分析は、等分散を仮定しない Welch の方法による t 検定を用いて、被験者群の成績を比較した。親密度評定課題の分析は、親密度評定値を従属変数とした、二元配置分散分析(混合計画)を用いて行った。実験計画は、第一要因を被験者群として二水準(JLE・CLE)、刺激の種類を第二要因を二水準(自然・不自然)とした。その後、各群の評定値に対して相関分析を行った。

4. 結果

最初に、コロケーション選択課題の記述統計を表2に示す(図1にも示している)。全体的な傾向としては、両群の正答率はともにチャンスレベルよりも高いが、両群とも、3割前後不自然なコロケーションを選択してしまっている事が分かる。

表2. コロケーション選択課題の記述統計

	<i>M</i>	<i>SD</i>	Maximum	Minimum
JLE	0.77	0.05	0.88	0.71
CLE	0.68	0.08	0.79	0.54

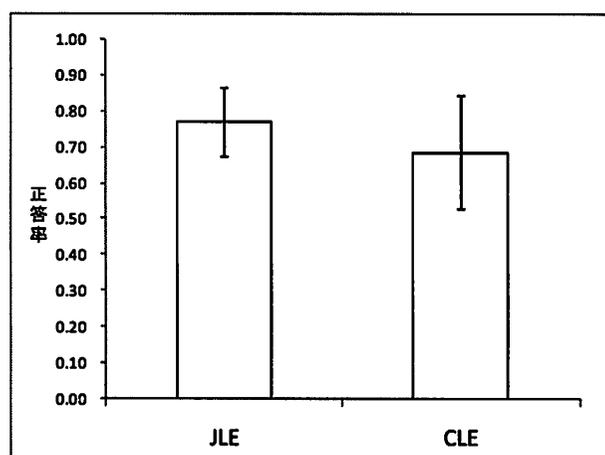


図1. コロケーション選択課題の結果³

被験者群間における t 検定の結果では、成績の差が統計的に有意であった ($t(25) = 3.67, p < .01, d = 1.35$)。この結果は、JLE は CLE よりも自然なコロケーションを選ぶ傾向が強いことを示している。

次に、親密度評定課題の記述統計を表3に示す(図2にも示している)。

表3. 親密度評定課題の記述統計

	不自然なコロケーション				自然なコロケーション			
	M	SD	Maximum	Minimum	M	SD	Maximum	Minimum
JLE	2.74	0.41	3.17	1.97	3.22	0.29	3.92	2.92
CLE	2.94	0.56	3.75	1.89	3.29	0.59	4.00	1.58

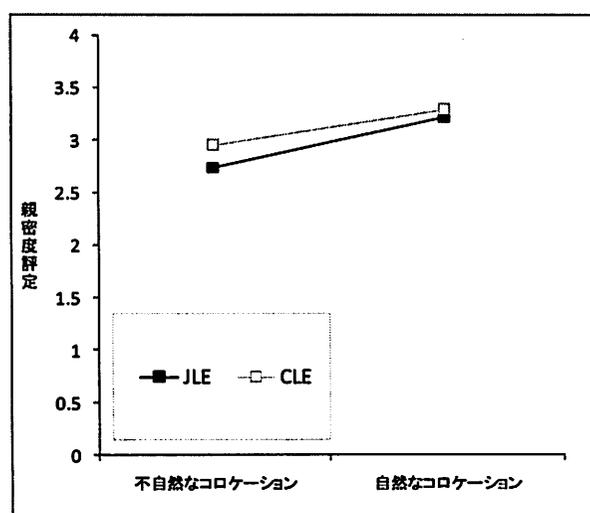


図2. 親密度評定課題の結果

二元配置の分散分析では、被験者群の主効果が統計的に有意でなく ($F(1, 30) = 0.80, p = .38, \eta_p^2 = .03$)、刺激の種類の主効果が統計的に有意であった ($F(1, 30) = 36.94, p > .01, \eta_p^2 = .55$)。また、2要因の交互作用も統計的に有意ではなかった ($F(1, 30) = 1.07, p = .31, \eta_p^2 = .03$)。この結果は、母語に関わらず、両群で不自然なコロケーションの親密度を低く評定していることを示している。

次に、各項目における JLE 及び CLE 間の平均評定値の相関分析を、自然なコロケーション、不自然なコロケーションの2種類に分割して行った(散布図はそれぞれ図3及び図4に示している)。

まず、不自然なコロケーションにおける相関分析では、 $r = .15$ (Pearsonの積率相関係数)であり、無相関検定の結果は、5%水準で有意ではなかった。また、回帰直線の傾きも小さく、両群の評定値に相関関係は無い、と考えることができよう。

一方、自然なコロケーションにおける相関分析では、 $r = .66$ であり、無相関検定の結果は1%水準で有意であった。また、回帰直線の傾きは0.73であった。このことから、二群の評定値に中程度の相関関係がある、と判断できる。

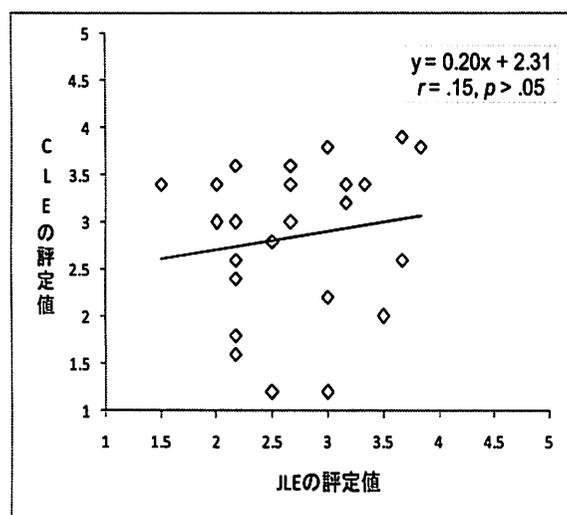


図3. 不自然なコロケーションにおける二群の評定値の散布図 ($k = 24$)

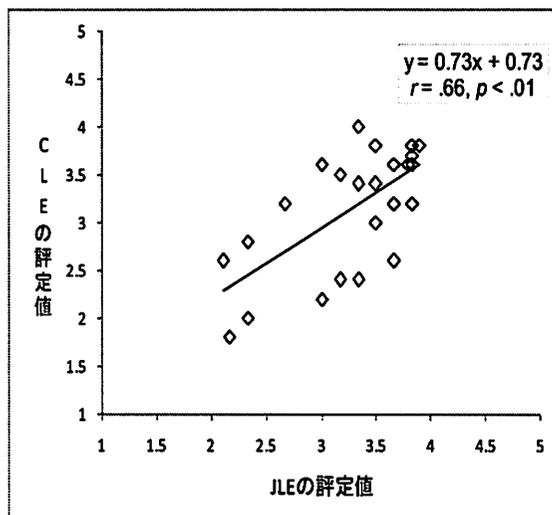


図4. 自然なコロケーションにおける二群の評定値の散布図 ($k = 24$)

5. 議論

最初に、コロケーション選択課題の結果について議論する。結果の記述統計から、両群の成績は、明らかに偶然生じる確率(50%)を超えるものではあったものの、JLE及びCLEは、不自然なコロケーションを自然でより適切な表現だと誤判断してしまう傾向があるといえよう。また、被験者群間における成績

の差の検定は有意であり、寧ろ、母語による負の影響が存在すると考えられる JLE の成績の方が高かったことが示された。このことから、不自然なコロケーションに対する鈍感さに影響するのは、母語のみでないと考えられる。

しかしながら、本実験の場合、同一指標での直接的な比較が出来なかったが、JLEの方が、熟達度が比較的高かったという可能性も考えられる。この課題は文脈付きの条件で行われていた為、熟達度が高い JLEの方が、文脈から類推する能力や、明示的な表現の知識をより豊富に持っている可能性もある。更に、自然な表現は、語単位では頻度や親密度が必ずしも高くはなく、学習者が産出した語の方が語単位での理解しやすさから、選択されやすいという可能性もあり得る。

一方、親密度評定課題の結果は、JLE・CLE間の評定値の差が有意ではなく、両群とも不自然なコロケーションの方に低い親密度評定をつけている。しかし、不自然なコロケーションは対象言語の発話等において頻出することはなく、個別言語の慣習に逸脱しているにも関わらず、両群ともその平均評定値は5件法における中間に近い。一方、Millar(2010)の結果では、母語話者は均質的に二極的な評定をつけている。本課題の結果では、不自然なコロケーション形式に対する敏感さは両群とも同程度であり、学習者は母語話者のような敏感さを示さないといえよう。

ここまでの結果の解釈では、コロケーション形式への敏感さに対する母語の影響は小さいと考察できるが、以降は各項目ごとにおける反応の違いについて議論する。まず、親密度評定値の相関分析において、不自然なコロケーションの評定値に相関関係が無かった。この結果は、中国語、日本語、英語、それぞれにおける慣習的表現の対応関係に起因すると考えられる。例として、日本語と英語において対応関係の無い表現であっても、中国語と日本語の対応表現があり、JLEとCLEの評定が同様に高くなるといった可能性がある。また、英語と中国語に対応関係があった場合、CLEはJLEが産出した不自然なコロケーションに相対的に敏感であり、低い評定を下すことも考えられる。このように、不自然なコロケーションに対する敏感さは、「一元的に、同じ母語を持つ学習者が産出した表現の不自然さに鈍感である」のではなく、個々の表現における、母語との対応関係や熟達度によって大きく左右される可能性があると考えられる。母語とは独立した、学習者のコロケーション形式への敏感さに影響する要因の解明も望まれるところである。

一方、自然なコロケーションにおける評定値の相関分析では、JLEとCLEは同様の評定をつける傾向があることが示された。自然・不自然なコロケーシ

ョンにおけるそれぞれの結果は、対象言語として自然な表現を自然である（accept）と判断する知識と、不自然な表現を不自然である（reject）と判断する知識に質的な違いがあることを示唆している。今後、このような差についても理論的な観点から、詳細に検討すべきであろう。

6. 結論と今後の展望

本研究ではコロケーション形式への敏感さにおける母語の影響について調査した。本研究の結果は、コロケーション形式への敏感さに対する母語の影響は、個別の項目の性質に依存する、とまとめられよう。しかしながら、本研究は、各三言語の言語表現を個別に直接分類、検証したものではなかった。このような対照言語学的視点を取り入れることが今後の課題の筆頭といえよう。また、本研究には、以下のような問題点がある。

1. 標本のサイズが小さいこと
2. 熟達度が同一指標で比較されていないこと
3. 被験者の第三言語の影響を排除できていないこと

このような問題点を踏まえ、今後の調査には、(1)熟達度を統制した母語交差型実験、(2)個別言語の各項目の対応関係をパラレルコーパスなどで統制した実験、または、(3)全体的処理を視野に入れた課題の反応時間や課題中の眼球運動などに着目した心理言語学的実験、などが望まれるであろう。

注

1. 全ての変数において平均値の差の検定を行う事は、検定の多重性による第一種の誤謬の可能の高さを考慮すると、望ましくない。本研究では検定統計量に依存しない Cohen's *d* を用いて両群の性質を検討した。
2. 年齢、及び話す能力の自己評定が、コロケーション形式への敏感さに及ぼす影響が大きいとは考えにくい為、二群は、母語、及び熟達度以外では比較的均質であったと考えられよう。
3. エラーバーは標本値の 95% 予測区間 (Prediction Interval) を示す。

参考文献

Bilal Anwer, R. M., & Khan, L. A. (2012). Collocations and second language use: Errors made by advanced learners in Pakistan. *Scottish Journal of*

Arts, Social Sciences and Scientific Studies, 1, 47–73.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29, 1–18.

Hawarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 24–44.

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal*, 91, 433–445.

Kusanagi, K., Leung, C. Y., Bando, T., Fukuta, J., & Sugiura, M. (in press). L2 learners' online insensitivity to malformed collocations: A study using eye-tracking technique and self-paced reading tasks. *Kyklos: International Communication*, 10.

Millar, N. (2010). The processing of malformed formulaic language. *Applied Linguistics*, 32, 128–148.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, 27, 741–747.

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647–668.

付録 (刺激の例)

	不自然なコロケーション	自然なコロケーション
1	<i>birth country</i>	<i>native country</i>
2	<i>country language</i>	<i>national language</i>
3	<i>did mistakes</i>	<i>made mistakes</i>
4	<i>expensive wages</i>	<i>high wages</i>
5	<i>feel cultures</i>	<i>experience cultures</i>