

## Reviewer(s)' Comments to Author:

Reviewer: 1

### Comments to the Author

I enjoyed reviewing the article, and my judgment is that it fulfills most of the criteria for publication--i.e., relevance to the readers of the journal; writing that is clear and accessible; a good balance between theory and practice; some consideration of the application in other contexts; and up-to-date references to the literature. It convincingly establishes that there is a gap that this article aims to fill, and it does so quite successfully. There are only a few points that the author(s) might consider addressing.

1. Introduction: The statement that "since formulaic language is associated with identity" needs clarification. In addition, is there empirical evidence to substantiate the claim that "a lack of knowledge of multiword expressions can negatively affect motivation, perception of oneself in the L2, and, and ultimately, attainment of fluency"?
2. What is the reason why the authors did not acknowledge or discuss Lewis's lexical approach?
3. Why did the authors choose the British National Corpus and not other more readily available, freely accessible corpora as the main corpora for \*illustrative purposes\*? Suggestions that readers/practitioners use search engines like Google must be stated with the caution that the counts can be misleading/inaccurate for multi-counting occurrences with the same material.
4. Although the inherent context-specific, user-specific use of the FTF is a strength, as stated by the authors that "that is the point," one may also question whether the "guidance" that the FTF aims to provide "coursebook authors" or "test developers" or whether the need to address the lack of rationale for inclusion of multiword expressions by material developers would lead to any meaningful changes in practicality, because such a decision is necessarily instructor- (as the authors demonstrated how judgment of multiword items' frequency and transparency may or may not vary), learner-, and context-dependent.
5. For non-native teachers of English, who, as we all know, constitute a large portion of ELT practitioners, how does the use of Google, corpus-informed dictionaries, and online corpora, which may inform their judgment on the frequency continuum, help them with their FTF judgment on the transparency-opaque continuum?
6. Figure 4 needs to be clarified for readers who are not familiar with output generated from corpora.
7. It is a sound piece of advice to suggest inter-rater reliability as an additional "tool" to establish a group of multiword expressions, but in my view (as someone who also has over two decades of classroom experience), while the approach is sound in theory, its practicality in the real world (i.e., including rater training, rating of two instructors of all multiword expressions that they wish

to include and prioritize in class, the discussion of items that had low inter-rater scores, and the discussion of potential disagreement, etc.) must be considered.

8. The article provides readers an idea for a tool that is limited in that they may use it in deciding what multiword expressions to include and how they should be prioritizing those items in their teaching. I see it also as a learning tool that can be used to facilitate learners' development of self-regulated learning skills, in that pedagogical activities can be designed to, for example, help learners discover the use of corpora to determine the frequency of multiword expressions that they come across, and/or to involve learners in the use of the FTF to rate certain expressions or place those expressions on the frequent-infrequent, transparent-opaque grid to help them check their understanding and perceptions of those expressions, and so on. I think that considering learners' use of the tool would make the ideas presented in this article hold more promise and have a greater potential impact in bridging theory and practice.

Reviewer: 2

Comments to the Author

This is an interesting article and well written. The topic is relevant and generally well researched, and the suggested framework is well thought-out. The combination of frequency and opacity is explained and useful suggestions are made for how certain multiword expressions might fit into it.

However, I feel there are several key areas which require revision in order to make the article suitable for publication.

First, it is not sufficiently clear or explicit who and what the framework is for. Several times there are references to coursebook authors, teachers, and once to test developers, but it is not clear whether the framework is suggested as a way to look at the lexis which appears in texts and decide whether or not it is suitable for explicit focus, or as a way to approach the presentation of new lexis to students across a syllabus. If the latter, then there is little suggestion of where the author or teacher may begin in looking for these words and classifying them using the framework. This is especially important as the counterpoint raised (p.3 l.49) regarding number of multiword expressions is not followed up in any detail. I feel the recommendations for how and to what the framework may usefully be applied need to be much more explicit.

Second, given the interesting combination of frequency with opacity as determining factors, there is an imbalance between the attention given to means of determining frequency, and the attention given to opacity. It is a shame that raters' judgements of opacity are fictitious, as evidence of how real raters cope with determining opacity, even on a small scale, would be very interesting and would add weight to this part of the article. This would be particularly important when considering longer, and semi-fixed expressions, and how the notion of opacity might be determined when the expressions do not have a common feature like the examples with 'take' presented in the article.

Alongside these broader recommendations for revision, the following points may also be usefully taken into consideration:

Introduction: The initial paragraph is a little unspecific using the phrases "several researchers" and "some years". This could be more direct and defined and would also benefit from reference to some of the other key figures in this area e.g. Lewis (1993, 1996, 2000); Hoey (2005)

Introduction para.3: It would be useful to contrast this with what the coursebook publishers themselves say about their inclusion of lexis and multiword expressions.

P.3 I.23: This is overly hedged and doesn't add much to the discussion, even with the accompanying footnote.

P.3 I.57: This counterpoint is also worth brief discussion and referencing.

P.7 I.62: Is there an assumption here that at higher levels other expressions have already been learnt successfully?

P.8 I.17: Is 'good sense' here related to 'usefulness' earlier. If not, what does it involve?