Single-clause when-definitions: Take three

Robert Lew & Anna Dziemianko

Keywords: *definition, folk defining, syntactic information, learner's dictionary, definition format.*

Abstract

In our EURALEX 2006 contribution (Dziemianko and Lew 2006), we focused on the practice of defining certain abstract nouns by means of a *when*-clause, which seems to have gained much popularity in recent years in some major monolingual English learners' dictionaries. We tested the hypothesis that a definition of this format would fare worse than the classic analytical definition in terms of conveying information on the syntactic class of the lemma. Experiments with Polish high-intermediate and advanced learners of English provided strong empirical support for this hypothesis. However, the testing instruments employed in the 2006 study used a relatively restricted microstructure, with just headwords and definitions. In the present follow-up study, we attempt to verify the results using a more complete microstructure to assess the strength of the effect of single-clause *when*-definitions on syntactic class identification in the presence of other potential indicators of syntactic class. Below we summarize the findings of the whole series of studies of this contentious defining format.

1. Background

Single-clause *when*-definitions have recently enjoyed a surge of popularity in English monolingual dictionaries for language learners, even though their use in English-language lexicography goes back to the 16th and 17th centuries (Osselton 2007; Stein 2011: 72). The attraction of this dictionary format is seen in the avoidance of the somewhat pedantic *genus-differentia* structure while being usefully shorter than the (double-clause) full-sentence definition format (Hanks 1987). Advocates of this format argue that it is used outside lexicography: in conversation and folk defining. Indeed, there is some evidence of the occasional use of single-clause *when*-definitions in spontaneous defining (Fabiszewski-Jaworski 2011). In Fabiszewski-Jaworski's study of spontaneous defining, the item most commonly defined through this format (13.5% of all definitions) was the noun *envy*. Fittingly, it is also abstract nouns that recent editions of dictionaries have treated with the single-clause *when*-definition most readily.

The single-clause *when*-definition — in the form most common in modern English learners' dictionaries — can be characterized as a stand-alone relative clause introduced with the relative word *when*, as in the following definition of *renown* taken from LDOCE online:

> **re·nown** [uncountable] *formal*
> when you are famous and a lot of people admire you for a special skill, achievement, or quality

Although the *when*-definition appears to be rather favourably regarded by Japanese learners of English (Ichikawa et al. 2005; Kanazashi 2008; Kanazashi et al. 2009), there are concerns that its characteristic structure may mislead the user with regard to the syntactic category being defined. This is because the clausal structure of this format does not signal in any overt way that the word being defined is a noun. In contrast, consider the traditional analytical definition of *renown* taken from MED online, similar on the lexical level, and yet structured differently:

renown – definition
NOUN [UNCOUNTABLE] FORMAL
the state of being famous and admired for a special skill or achievement

This analytical-type definition has the form of a post-modified noun phrase, with its ultimate head being the noun *state*. Presumably, the salient status of *(the) state* signals to the user that the word defined is likewise a noun ('renown is a type of state'), while no such overt signal seems to be present in the single-clause *when*-definition. One way in which dictionary users confronted with a single-clause definition might recognize that the definition defines a noun would be through their familiarity with the convention of using this definition type to explain nouns. The question is, however, to what extent this actually *is* a convention: can we be sure, for example, that such definitions are never used to define adjectives or verbs? There is no evidence to tell us this, but even if it were true for native speakers of English, it is still another thing to expect foreign learners of English to recognize the convention. This line of thinking inspired the studies described in Dziemianko and Lew (2006) and Lew and Dziemianko (2006a, b), which we will briefly summarize before proceeding to report on the latest study.

2. Previous studies of *when*-definition

2.1. *Study 1*

The first study of the effectiveness of single-clause *when*-definitions for conveying part-of-speech information was Lew and Dziemianko (2006a). In this study, 129 Polish students majoring in English, mostly in their first year of college, were presented with definitions in two formats, single-clause *when* and analytical, and asked to supply Polish translations of the headwords, which had been replaced with invented words.

Compared with the analytical definition, the *when*-definition did very poorly (see Fig. 1, Study 1): only a third of the syntactic classes were recognized correctly, which is probably close to baseline chance level, given that items of three syntactic classes were used in the study: noun, verb, and adjective. The analytical definition did significantly better, with two-thirds of the syntactic classes being recognized correctly. However, it needs to be noted that since the goal was to gauge the actual potential of the *when*-definition to inform (or perhaps misinform) learners as to the noun status of the word being defined, the first study did not include any explicit syntactic class information — just the definitions.

This led to the question of what would happen if such definitions, clearly deficient in their potential to convey syntactic class information, were combined with explicit part-of-speech labels, and if the task were made more syntax-focused.

2.2 Study 2

This question was addressed in the first follow up study (Dziemianko and Lew 2006; Lew and Dziemianko 2006b). Here, the authors used a more complete microstructure, this time taking care to include the syntactic class labels in the form most popular in MLD's (*noun*, *verb*, *adj*). The test sheets also included other labels (mainly syntactic codes and some usage labels) and example sentences.

The task was also different: three alternative Polish translations of each headword were provided, and the students were asked to select one. The three Polish translations were all

based on the same morphological stem and differed mainly in terms of their syntactic class. The level of English of the 238 subjects used in the study was lower, most being intermediate and receiving instruction at one of several secondary schools.

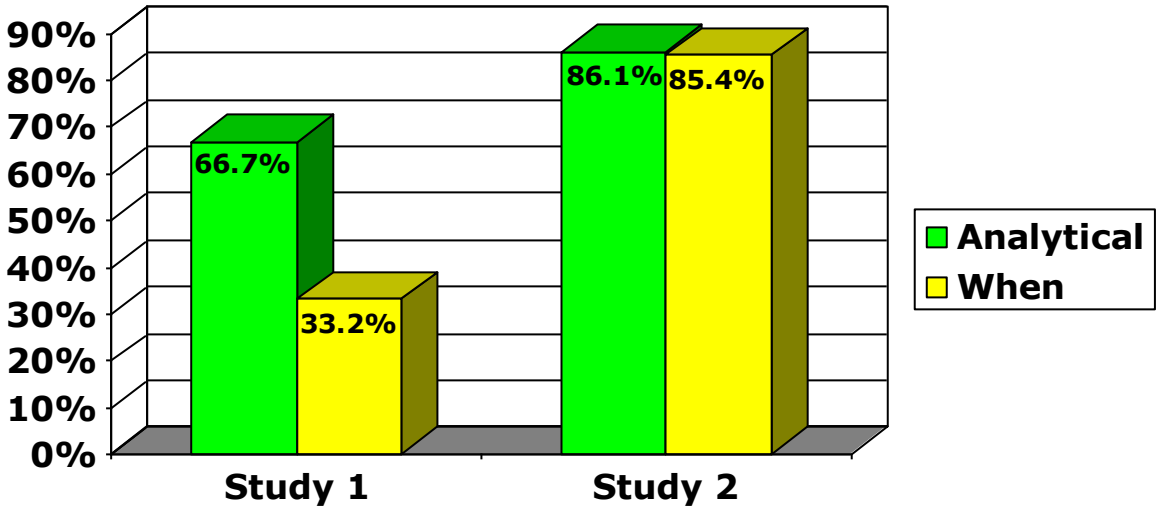

Figure 1. POS identification success rates in Study 1 and Study 2.

Results of Study 2 differed rather dramatically from Study 1 (see **Figure 1**). Here, students using both definition formats were equally successful in identifying the syntactic class of the headword, disguised as in Study 1. (In addition, the success rates were markedly higher than in Study 1, but this latter effect can easily be explained by the less open-ended format of the task.)

The difference between Study 1 and Study 2 suggests that the inclusion of part-of-speech labels, perhaps combined with the more syntax-focused task, made a real difference to the success of entries with *when*-definitions. A confirmation of this comes from the analysis of the parts of the entries underlined by the participants as helpful during the task. Most participants underlined the part-of-speech label, and only a few the definition.

Such a pattern of experimental entry consultation suggests that most participants, upon seeing the equivalents provided in the test, recognized that they needed to focus on syntactic class, and then homed in on the one microstructural element that held exactly this information. Thus, participants tended to approach the task as a type of *meta*lexicographic exercise requiring a selective reading of a dictionary entry. While the high success rate paints an optimistic picture of the reference skills of these secondary school students, as they were able to isolate and focus on the microstructural element most relevant for the task, this study does not tell us much about the effect of the format of the definition itself, as most participants probably did not really process the definition attentively, if at all. In their conclusion, Lew and Dziemianko (2006b: 289) suggest the need for a second follow-up study:

> We would still like to know, though, if such compensation [of the poor POS-conveying potential of the when-definition by other elements of the entry] would remain to be effective under less syntax-focused task conditions, and when the salience of the syntactic class label were reduced by separating it from the lemma sign with the phonetic transcription in its customary location. Another follow-up study is needed to fully clarify this issue.

This follow-up study was undertaken and is reported below.

3. Study 3

3.1. *Aim, design, materials and participants*

The aim of the study is to verify the findings of the previous two studies summarized above under more naturalistic conditions. This means a typical meaning-focus task (translation into L1) and a complete microstructure of the test entries, with phonetic transcription now occupying the usual place between the headword and the POS label, as opposed to the artificial salience of POS labels in Study 2. Other than the introduction of the transcription, the same test materials were used as in Study 2.

Thus, each participant received a sheet with 20 items, including 10 target noun entries and 10 filler verb and adjective items, supplemented in order to lower the salience of the target items. Half of the target noun definitions were in the analytical format, and the other half were single-clause *when*-definitions. There were two forms of the test, with half of the target entries assigned an analytical definition, and the other half defined through the single-clause *when*-definition. In the two forms, the assignment of definition type to target item was reversed. In this way, each participant was exposed to both definition types in equal measure, and every target item was accompanied by each definition type about half of the time, so as to minimize confounding effects of item and subject.

There were 134 participants taking part in the study, all native speakers of Polish, with their English proficiency in the upper-intermediate to advanced range. They were all asked to provide a Polish equivalent based on experimental dictionary entries provided. As in Study 1 and Study 2, actual headwords had been replaced with invented non-words, morphologically neutral so as not to suggest any particular part of speech.

Participants were asked to supply a single-word Polish equivalent in the space provided, following the measure used in Study 1 (participants in Study 1 were also asked to compose a sentence, but the two measures performed similarly and we recommended equivalent provision as a preferable measure).

Each Polish equivalent supplied was assigned to a syntactic category. Whenever a Polish noun was given as an equivalent, this was counted as a correctly recognized item. If an equivalent of a different syntactic class was supplied, this was counted as incorrectly recognized. For each subject, POS recognition rates were computed separately for each defining format. To better control for the ceiling effect (high overall recognition rates) and any subject-related variation, per-subject differences in POS recognition rates were calculated between the two defining formats by subtracting the individual recognition rate for *when*-definition items from the rate for analytical items, with each subject acting as his or her own control.

3.2. *Results and discussion*

Mean POS identification success rates for analytical and *when*-definition types with their standard deviations are given in Table 1. Analytical definitions outperform the *when* type, but only by a small margin: 90 percent as against 87 percent, respectively.

Table 1. POS identification success rates for analytical and *when*-definitions.

Definition type	Mean	SD
Analytical	90%	15%
When	87%	19%

This difference was found to be marginally significant (one-way GLM ANOVA, $F_{(1,132)}=3.95$, $p=0.049$). However, the difference of three percentage points in this context is of little *practical* significance and effect size is very small (partial $\eta^2 = 0.03$).

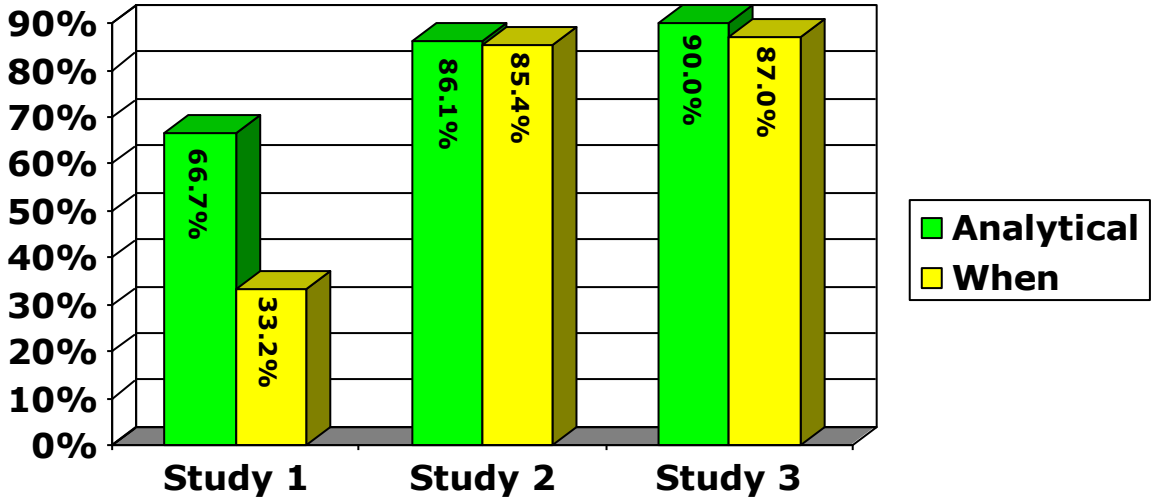

Figure 2. POS identification success rates in Studies 1, 2, and 3.

In Study 3, no systematic information was collected on the sections of the microstructure consulted by the participants, as we did not want to compromise the relatively naturalistic character of the task by asking subjects to underline parts of entries. However, participants' response patterns suggest that example sentences may have been an important source of information.

We also scrutinized closely the structure and wording of individual definitions for clues as to when definition format might make a difference, and one such possible pattern emerges from individual results. The presence of an indefinite pronoun *someone* or *something* right after the word *when* (e.g.: *when someone is easily upset or offended by things that people say*) appears to be particularly misleading: as many as five out of six entries with *when* followed immediately by *someone* or *something* had lower POS recognition rates than their analytical counterparts. In contrast, single-clause definitions with the word *when* followed by a regular personal pronoun or a nominal phrase performed no worse than the corresponding analytical definitions. This finding, however, must be seen as tentative, as it is based on a post-hoc analysis of a rather limited number of items.

4. Conclusions

The findings from the three studies invite a number of general conclusions. First, in a naturalistic dictionary entry and a meaning-based task there exists a small effect of definition type on POS recognition of abstract nouns, with the single-clause *when*-definition being at a

slight disadvantage; however, this is hardly of any practical significance for skilled dictionary users, as long as the POS labels are salient enough within the entry.

Second, even though the single-clause *when*-definition by itself gives Polish learners of English very few (if any) clues as to the syntactic class of the headword, evidence points to an important compensatory role of a complete microstructure in POS identification, and in particular the presence of syntactic class labels and example sentences, which get noticed and are by and large appropriately used by learners.

Finally, to offer some recommendations for dictionary-making, single-clause *when*-definitions may be acceptable in English monolingual learners' dictionaries only when embedded in a rich, complete microstructure; otherwise, they can be severely misleading with respect to what they tell users about the syntactic category of the headword. However, since the compensatory effect described above has only been attested with relatively proficient users, who would have no serious problems with the analytical definition type anyway, perhaps the single-clause *when*-definition is best avoided. Particularly problematic are definitions with *when* followed by an indefinite pronoun such as *someone* or *something*.

References

Dziemianko, A. and R. Lew 2006. 'When You Are Explaining the Meaning of a Word: The Effect of Abstract Noun Definition Format on Syntactic Class Identification.' In E. Corino, C. Marello and C. Onesti (eds.), *Atti del XII Congresso Internazionale di Lessicografia : Torino, 6-9 settembre 2006*. Alessandria: Edizioni dell'Orso, 857–863.

Fabiszewski-Jaworski, M. 2011. 'Spontaneous Defining by Native Speakers of English.' In K. Akasu and S. Uchida (eds.), *Asialex2011 Proceedings Lexicography: Theoretical and Practical Perspectives*. Kyoto: Asian Association for Lexicography, 102–109.

Hanks, P. 1987. 'Definitions and Explanations.' In J. Sinclair (ed.), *Looking Up: An Account of the Cobuild Project in Lexical Computing*. London - Glasgow: Collins, 116–136.

Ichikawa, Y., A. Shimizu, R. Takahashi, T. Kanazashi and Y. Ishii 2005. 'An Analysis of *Longman Dictionary of Contemporary English*, Fourth Edition.' *Lexicon* 35: 1–126.

Kanazashi, T. 2008. 'An Analysis of Two Business English Dictionaries for Learners.' *Lexicon* 38: 29–45.

Kanazashi, T., T. Otani, A. Nonmiya and M. Ryu 2009. 'An Analysis of the *Longman Advanced American Dictionary*, New Edition: A Pedagogical Viewpoint.' *Lexicon* 39: 18–99.

Lew, R. and A. Dziemianko 2006a. 'A New Type of Folk-Inspired Definition in English Monolingual Learners' Dictionaries and Its Usefulness for Conveying Syntactic Information.' *International Journal of Lexicography* 19.3: 225–242.

Lew, R. and A. Dziemianko 2006b. 'Non-Standard Dictionary Definitions: What They Cannot Tell Native Speakers of Polish.' *Cadernos de Traduçao* 18: 275–294.

Osselton, N. E. 2007. 'Innovation and Continuity in English Learners' Dictionaries: The Single-Clause When-Definition.' *International Journal of Lexicography* 20.4: 393–399.

Stein, G. 2011. 'The Linking of Lemma to Gloss in Elyot's *Dictionary* (1538).' In O. Timofeeva and T. Säily (eds.), *Words in Dictionaries and History. Essays in Honour of R.W. Mcconchie*. Amsterdam: John Benjamins Publishing Company, 55–79.