

New-line and Run-on Guiding Devices in Print Monolingual Dictionaries for Learners of English

Bartosz Ptasznik, *University of Warmia and Mazury, Olsztyn, Poland*
(bartosz.ptaszniak@uwm.edu.pl)

and

Robert Lew, *Faculty of English, Department of Lexicography and Lexicology,
Adam Mickiewicz University, Poznań, Poland* (rlew@amu.edu.pl)

Abstract: The paper focuses on guiding devices in print monolingual dictionaries for learners of English. It aims to find answers to six research questions. The main aim is to investigate how the location of guiding devices within entries — starting from a new line versus run-on — affects consultation time and sense selection accuracy of dictionary users in entry navigation. In addition, the paper looks at the effect that part of speech (noun versus verb entries) has on consultation time and sense selection accuracy; further, the relationship between entry length and consultation time is investigated, as well as that between consultation time and sense selection accuracy.

Keywords: DICTIONARY, DICTIONARY CONSULTATION, DICTIONARY ACCESS, LEXICOGRAPHY, GUIDING DEVICES, SIGNPOSTS, SIGNPOSTING, MONOLINGUAL DICTIONARIES, LEARNERS' DICTIONARIES, ENGLISH, POLISH LEARNERS

Opsomming: Nuwereël- en deurloop-hulpmiddels vir gebruiksheiding in gedrukte eentalige woordeboeke vir Engelse leerders. In hierdie artikel word gefokus op gebruiksheidingsinstrumente in gedrukte eentalige woordeboeke vir Engelse leerders. Dit het die beantwoording van ses navorsingsvrae ten doel. Die hoofdoel daarvan is om te ondersoek hoe die plasing van gebruiksheidingsinstrumente in inskrywings — beginnende by 'n nuwe reël versus deurloop — die woordeboekgebruiker se konsultasietyd en akkuraatheid van betekenis-seleksie in inskrywingsnavigasie beïnvloed. Daarbenewens word daar ook in hierdie artikel gekyk na die effek wat woordsoort (selfstandige naamwoord- versus werkwoordinskrywings) op die konsultasietyd en akkuraatheid van betekenis-seleksie het. Vervolgens word die verband tussen die lengte van die inskrywing en die konsultasietyd ondersoek, asook die verband tussen konsultasietyd en akkuraatheid van betekenis-seleksie.

Sleutelwoorde: WOORDEBOEK, WOORDEBOEKKONSULTASIE, WOORDEBOEKTOEGANG, LEKSIKOGRAFIE, GEBRUIKHEIDINGSINSTRUMENTE, AANWYSERS, AANWYSING, EENTALIGE WOORDEBOEKE, AANLEERDERSWOORDEBOEKE, ENGELS, POOLSE LEERDERS

1. Background

1.1 Research on guiding devices in English monolingual learners' dictionaries

Empirical studies that deal with guiding devices in English monolingual learners' dictionaries are becoming increasingly common. One motivation for this is the pursuit of empirically motivated principles to inform the design of dictionary entries so that they satisfy the needs of dictionary users who regularly consult a dictionary, and who in the process of doing so repeatedly have to browse through lengthy entries that take up much of their time. Although the gradual transition to digital dictionaries makes this issue less relevant, in many parts of the world, print dictionaries are still far from disappearing. Thus far, two broad types of guiding devices have been used in English learners' dictionaries: (1) signposts, which briefly define their respective senses, with each signpost being placed next to its sense; and (2) menus, which form a list of sense cues appearing as a block at the top of the entry.

Signpost-related research goes back to Tono's (1992) study and his finding that less proficient English learners benefit from menus in entries to a greater extent than do more proficient learners. In another study (1997), Tono found that LDOCE3 signposts are superior to CIDE guide words, especially with regard to the accuracy of sense selection and time needed for consultation; however, not necessarily so in the case of longer entries. Further, LDOCE3 guiding devices appeared to convey semantically more meaningful information. One year later, Bogaards (1998) concluded that semantically-guided LDOCE3 and CIDE sense-navigation devices generally outperformed grammar-based COBUILD2 access devices, as well as the access devices adopted by OALD5 lexicographers. Interestingly, the subjects who were assisted by OALD5 guiding devices performed the worst, despite having been exposed to the shortest entries.

More recent research includes a series of studies carried out by Lew. Lew and Pajkowska (2007) stressed the need to conduct further research on meaning access devices in learners' dictionaries, and noted in their research that pre-intermediate English learners may benefit from signposts more when entries were of shorter length, whereas intermediate learners may benefit more from signposts in longer entries. Lew and Tokarek (2010) focused their attention on entry menus in web-based bilingual dictionaries and observed that by and large menus with highlighting of target senses were both more effective (accurate) and efficient (faster) than menus which appeared without highlighting. In the same year, Lew (2010) compared the signpost and menu systems (based on OALD7 entries) and found that signposts resulted in superior sense selection and translation accuracy compared with menus. One explanation offered for this apparent advantage of signposts was the location of the information within entries. Signposts appear next to their respective senses, which might make it

easier for dictionary users to consult the signpost along with the related definition and examples that appeared next to it. By contrast, in menu-based systems, dictionary users need to go over the list of sense cues located above the entry, and only then are they able to proceed to the actual visual scan of the senses. Thus, in the case of menus, the guiding information appears out of the immediate context of the respective senses: this could lead to dictionary users not being able to reach the type of information they are searching for in the senses, and might leave users confused between the menu and the actual senses.

Tono (2011) applied the eye-tracking technique to observe patterns of dictionary use and test the effectiveness of the signposting and menu systems. One finding was that, depending on how signposts were phrased, they could either be helpful or misleading to dictionary users. Furthermore, in view of earlier findings (Tono 1992), Tono expected that menus would be of greater benefit to less advanced English students. In the eye-tracking study, proficient students tended to ignore menus, but at the same time consulted signposts willingly. Less proficient students, however, did not use the signposts in entries, possibly because they had no clue as to what the purpose of the signposts was.

A study by Nesi and Tan (2011) confirmed Lew's (2010) finding of the superiority of signposts over menus. It also found that learners were able to select senses more accurately and quickly when the target senses were located at the beginning or towards the end of entries, while the middle of entries was the most problematic. Unexpectedly, sense selection success and consultation times were best for the target senses appearing at the end of entries. Nesi and Tan attributed these results to the so-called bathtub effect, which is a known effect of the early and final parts of words being easier to remember (Aitchison 1997); in the context of entry consultation, this would translate into facilitation of initial and final senses in entries. In terms of the advantage of final senses over initial senses, one explanation could be advanced learners starting their entry search with the final senses, rather than in a top-to-bottom fashion. The authors speculated that such a strategy may have arisen from the discovery through regular dictionary consultation that the senses placed at the beginning of entries tend to be familiar to advanced users, due to the fact that in many dictionaries they tend to be the frequent senses. Nesi and Tan (2011) also found that entry or definition length have no effect on accuracy of sense selection. Likewise, consultation time remained unaffected by entry length, while definition length was positively related to consultation time. An additional general finding was that adjective and verb entries were more problematic than noun entries for dictionary users. To some extent, this finding was confirmed by Ptasznik's research (2015), according to which consultation of verb entries took more time than consultation of noun entries.

Ptasznik and Lew (2014) compared entries equipped with a combination of signposts and menus against entries with signposts only. The study found that adding entry-initial menus to entries with signposts did not improve the

rate of correct sense selection, nor did it reduce the time needed for entry consultation. Hence, Ptasznik and Lew concluded that it is best when entries in paper dictionaries are equipped with signposts only.

The two most recent experimental studies of signposting are Dziemianko (2016) and Dziemianko (2017). The former study undertook an empirical comparison of the alternative methods of presenting signposts featured in three online dictionaries: LDOCE5 (white capitals on a blue background), OALD8 (crimson capitals above a crimson line), and OALD9 (black lower-case letters above a dark orange line). The LDOCE5 signpost highlighting strategy was found to be the most beneficial of all three strategies with regard to consultation time, while the OALD8 strategy was the least helpful. As far as accuracy of sense selection is concerned, all three methods achieved comparable scores, and so none of the signpost highlighting strategies came out as most beneficial. The study also tested meaning retention; for this outcome measure, white capitals on a blue background and lower-case letters above a dark orange line resulted in highest retention rates. This study also touched upon sense positioning in entries, with the finding consonant with Nesi and Tan (2011) that dictionary users consulted entry-final senses the fastest. However, signposts that appeared in crimson capitals above a crimson line and black lower-case letters above a dark orange line did not exhibit a significant effect of sense position on the time of consultation. The findings were not compatible with the interpretation that proficient learners would scan entries from the bottom up, as shorter times for entry-final senses were only recorded for senses equipped with signposts in the form of white capitals on a blue background. In addition, accuracy of sense selection was highest for senses located in the middle of entries, which contradicts Tono's (1992) finding of initial senses being the best in this respect. All three methods of signpost highlighting produced more or less similar sense selection success rates in all types of sense positions, which does not support Nesi and Tan's (2011) claim that initial and final senses lead to most successful results. Dziemianko (2016) concurs with Ptasznik (2015) in warning against simplistically assuming the advantage of homogenous form of signposts throughout the entry and stresses the importance of conducting further research on heterogeneous and homogenous sense-navigation devices, as well as the length of sense indicators.

Dziemianko's most recent study (Dziemianko 2017) did not confirm Nesi and Tan's (2011) finding regarding the bathtub effect. Dziemianko found no evidence of a bathtub effect on either sense selection, reception, or production. Furthermore, selection of senses, as well as decoding and encoding were not affected by the length of entries. On the other hand, the study found a tendency for participants to identify longer target senses with greater success.

1.2 Which dictionaries have new-line and run-on guiding devices?

By and large, guiding devices that start from a new line¹ have dominated in

print monolingual English learners' dictionaries (see **Table 1**). Presumably, the thinking was that starting senses on a new line would make it easier for English learners to distinguish one sense of an entry from another. The only English monolingual learners' dictionary that has departed from this general strategy in some of its editions was the *Cambridge Advanced Learner's Dictionary* [CALD]. In general, CALD3 entries have been equipped with run-on guidewords (guidewords are the specific brand of guiding devices that appear in CALD3), though in the most highly polysemous entries (i.e. those with the greatest number of senses) new-line guidewords are employed, which may suggest that CALD3 lexicographers agreed with the view of the competing dictionaries that adding a line break might be advantageous, but compromised in the interest of saving space. The next edition, however (CALD4), only included run-on guidewords, irrespective of entry length. **Table 1** gives the solutions adopted in the specific dictionary editions with respect to the presentation of senses from a new line or as run-on.

Dictionary edition	Year of publication	Guiding device position
<i>Longman Dictionary of Contemporary English (LDOCE3)</i>	1995	New-line
<i>Longman Dictionary of Contemporary English (LDOCE4)</i>	2003	New-line
<i>Longman Dictionary of Contemporary English (LDOCE5)</i>	2009	New-line
<i>Longman Dictionary of Contemporary English (LDOCE6)</i>	2014	New-line
<i>Cambridge International Dictionary of English (CIDE)</i>	1995	New-line
<i>Cambridge Advanced Learner's Dictionary (CALD1)</i>	2003	New-line
<i>Cambridge Advanced Learner's Dictionary (CALD2)</i>	2005	New-line
<i>Cambridge Advanced Learner's Dictionary (CALD3)</i>	2008	Run-on ²
<i>Cambridge Advanced Learner's Dictionary (CALD4)</i>	2013	Run-on
<i>Oxford Advanced Learner's Dictionary of Current English (OALD4)</i>	1989	New-line
<i>Oxford Advanced Learner's Dictionary of Current English (OALD5)</i>	1995	New-line
<i>Oxford Advanced Learner's Dictionary of Current English (OALD6)</i>	2000	New-line
<i>Oxford Advanced Learner's Dictionary of Current English (OALD7)</i>	2005	New-line
<i>Oxford Advanced Learner's Dictionary of Current English (OALD8)</i>	2010	New-line
<i>Oxford Advanced Learner's Dictionary of Current English (OALD9)</i>	2015	New-line
<i>Macmillan English Dictionary for Advanced Learners (MED1)</i>	2002	New-line
<i>Macmillan English Dictionary for Advanced Learners (MED2)</i>	2007	New-line

Table 1: New-line and run-on guiding devices in specific editions of print monolingual English learners' dictionaries

Although new-line guiding devices seem to be the dominant practice for print monolingual English learners' dictionaries (**Table 1**), we believe it is still impor-

tant to test if it is indeed the better way to structure senses when these start with signposts. Two outcomes are measured: consultation time and sense selection accuracy. The next section outlines the aims and methods, followed by results and discussion.

3. The study

3.1 Aims of the study

The aims of the present study were to investigate:

- (1) how CONSULTATION TIME is affected by GUIDING DEVICE POSITION (new-line guiding devices vs. run-on guiding devices) during entry navigation; we would expect that positioning the guiding devices from a new line should facilitate visual search and thus reduce the time needed to find the relevant sense
- (2) how CONSULTATION TIME is affected by PART OF SPEECH³ (nouns vs. verbs); we would expect verbs to be more challenging than nouns, the latter being *first-order words* (Piotrowski 1989: 102);
- (3) whether GUIDING DEVICE POSITION is a factor determining SENSE SELECTION ACCURACY in entry look-ups; one might expect that the new-line format would potentially allow the user to avoid more errors due to its less cluttered, unambiguous layout;
- (4) whether SENSE SELECTION ACCURACY varies by PART OF SPEECH; again, nouns might be expected to trigger fewer errors, due to their semantic primacy;
- (5) if CONSULTATION TIME depends on ENTRY LENGTH measured in words; we would expect longer entries to require longer consultation, thus a positive correlation between the variables would be expected;
- (6) whether SENSE SELECTION ACCURACY of lexical items is related to CONSULTATION TIME; we would expect a negative correlation, with longer times corresponding to lower accuracy; both being expected consequences of increased item difficulty.

3.2 Method, participants, and procedure

The study was carried out at the Faculty of Humanities at the University of Warmia and Mazury in Olsztyn. 100 students of English who were third and fourth year students participated in the study. Their English language proficiency level was assessed by their academic teachers as B2 to C1 by the Com-

mon European Framework of Reference for Languages standards; Polish was their native language.

Four variables were selected for the study, two predictor variables: GUIDING DEVICE POSITION and PART OF SPEECH; and two outcome variables: CONSULTATION TIME and SENSE SELECTION ACCURACY. GUIDING DEVICE POSITION had two levels: guiding devices that started on a new line (new-line guiding devices) and devices that continued within the entry without a line break (run-on guiding devices). The number of senses was not manipulated in this study, but was controlled and set at the reasonable number of seven senses. All entries selected for the study consisted of at least seven senses; if there were more, some irrelevant senses were removed, so that the final number of senses was always seven. Half of the entries were nouns and half were verbs. All subjects were exposed to the same target items (each task was the same for each subject, items had the same cue sentences, they were of the same part of speech and the lexicographic data were the same in all entries), but specific entries appeared with either new-line or run-on guiding devices. The order of presentation of items was randomized to one of two versions of the test, the versions then being assigned randomly (the effect of version was checked and it was not significant). CONSULTATION TIME and SENSE SELECTION ACCURACY were recorded for each subject and test item.

A pilot test was carried out on 10 subjects to see if the whole procedure worked well. The subjects had enough time for each task during their class hours and by and large the subjects achieved a SENSE SELECTION ACCURACY within the range of about 55–80%. None of the test items exhibited a floor or ceiling effect, and there were no other problems, and so the main study was conducted after the pilot testing phase.

Twenty test items were used in the main study: *bear, introduction, jump, hole, match, mark, slip, power, lift, weight, withdraw, wave, transfer, unit, beat, base, print, grip, cast, root*. The lexicographic data of the entries were taken from LDOCE6, while the cue sentences were taken from four dictionaries: OALD9, the free online version of LDOCE [LDOCEO], and — sporadically — from the *Macmillan English Dictionary Online* [MEDO] and *Cambridge Dictionary Online* [CDO]. Each single test item (from among the twenty test items mentioned above) that was chosen for the study needed to have at least seven senses with guiding devices and had to include at least one less common sense that the subjects would not be likely to be familiar with. These less common senses of words were the target senses in the cue sentences of the study, as the subjects had to be exposed to tasks that would depend on entry consultation for successful completion.

The test was administered in a classroom by the first author, in five groups of between fifteen and twenty students in each session, over the course of one week. Before the actual test, each group received step-by-step instructions on how to proceed. The instruction was delivered verbally from a printed script in Polish, the native language of all participants. The subjects were told

in their native language that they were expected to read the cue sentence. The cue sentences provided the subjects with some context in the target language and an underlined word (the target item), which appeared in a less-known sense (see Appendix A and B). Subsequently, they were instructed to identify the target item and its context, search for the meaning of the underlined target word (the target items in the cue sentences were provided with a dictionary entry underneath the cue sentences: see Appendix A and B), write down on the test sheet the number of the target sense in which the target item is used, record their own time and move on to the next task item. The participants were informed that each test had twenty test items on separate sheets and that there were no time restrictions for the completion of the test. Participants recorded their own time using the stopwatch function of their mobile phones (it would not be possible for the experimenter to record twenty participants at the same time, and the school only allowed testing in groups). The next step was a complete practice run on two items (these were not used in the data analysis), followed by the actual testing.

All participants were exposed to the same test items. New-line and run-on guiding devices were rotated across test versions. In any single test, half of the test items had new-line guiding devices, and the other half had run-on guiding devices. Also, in any single test, half of the test items were noun entries, and the other half were verb entries. Across the 100 participants, the same number of responses were recorded for new-line and run-on guiding devices, as well as noun and verb entries. CONSULTATION TIME and SENSE SELECTION ACCURACY were recorded for each subject and test item (that is 2,000 data points for each variable).

4. Results and discussion

4.1 Consultation time

Consultation time is markedly non-normal and skewed to the right (skew = 1.94, kurtosis = 5.1), which — as is usual for temporal variables such as human reaction time or time-on-task — normalizes well when the time values are expressed as natural log values. Our data is no exception: after logarithmizing the times, the distribution becomes symmetrical and nearly normal (skew = -0.16, kurtosis = 0.0). This is also evident in the normal Quantile–Quantile plots (see **Figure 1**). Therefore, we will be operating in logged values, but for the readers' convenience converting them back to raw time expressed in seconds. Likewise, average times will be calculated as mean logged values, which is equivalent to geometric means on the original scale.

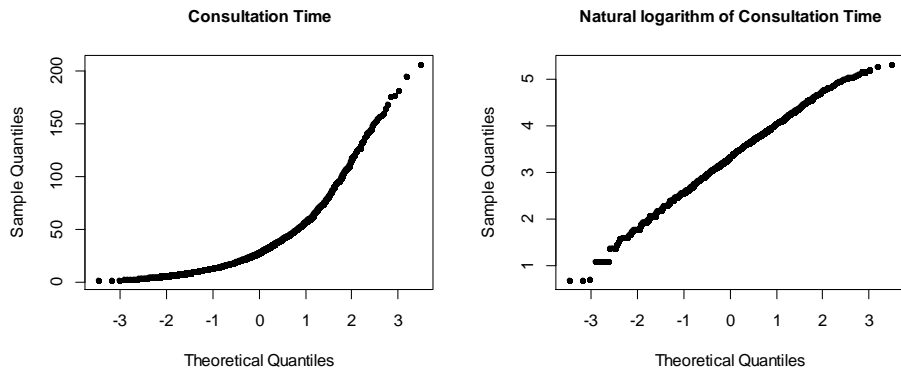


Figure 1: Normal Quantile–Quantile plots for CONSULTATION TIME: raw (panel on left) and logged values (panel on right)

The grand (geometric) mean for CONSULTATION TIME, averaged across all entry look-ups, was 27.7 seconds. When broken down by part of speech and guiding device position, the (geometric) means pattern as in **Table 2**.

	new-line	run-on	any position
Noun	23.1	25.5	24.3
Verb	32.8	30.6	31.7
any POS	27.5	27.9	27.7

Table 2: Average CONSULTATION TIME for combinations of GUIDING DEVICE POSITION and PART OF SPEECH (geometric mean given in seconds)

The marginal values in **Table 2** suggest a difference due to PART OF SPEECH, with verb entries being, on average, some thirty percent slower than noun entries. By contrast, the position of the guiding device does not seem to matter at all, since the CONSULTATION TIMES for the new-line and run-on conditions are virtually identical. This impression is confirmed in the so-called pirate plots (**Figure 2**) generated with the help of the `yarr` package (Phillips 2017). A pirate plot is an advanced plot that visually conveys information on (1) the central tendency (in our case, the geometric mean); (2) the detailed distribution of raw data through a jittered plot of all the data points (here, grey dots); and (3) a probability density estimate (the "beans"). In addition, the narrow boxes around the mean bar represent inference bands, in this case computed as 95% Bayesian Highest Density Intervals. The subjects in the study needed on average 32 seconds for the consultation of verb entries and about eight seconds less for noun entries (24 seconds).

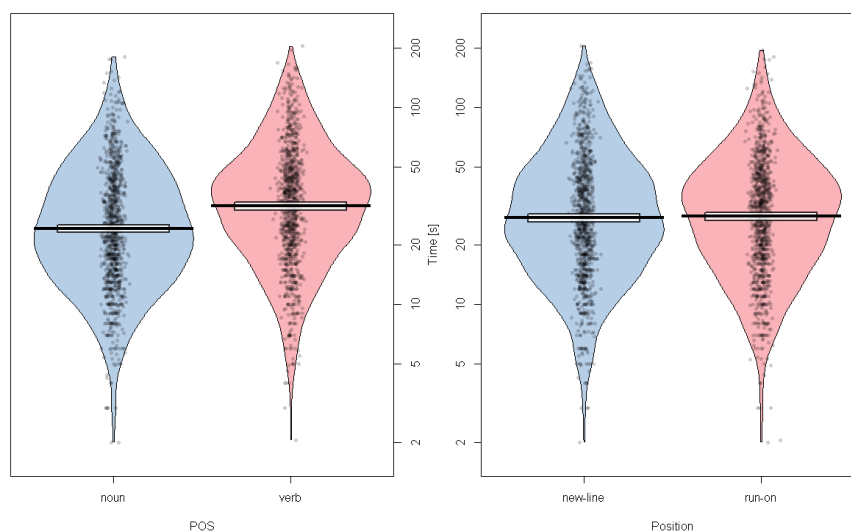


Figure 2: Pirate plots of CONSULTATION TIME by PART OF SPEECH (left) and GUIDING DEVICE POSITION (right). Inference bands around the means are 95% Bayesian Highest Density Intervals.

These results suggest that PART OF SPEECH affects CONSULTATION TIME in entry navigation, with verb entries requiring on average eight seconds more consultation time, which is about a third of the time more compared to noun entries. This may mean that verb entries tend to be more problematic for dictionary users than noun entries.

4.2 Best model for CONSULTATION TIME

A series of linear mixed models were fitted using `lme4::lmer` (Bates et al. 2015) and `afex::mixed` (Singmann et al. 2018) for logarithmized CONSULTATION TIME as the outcome variable, starting with complete models with interactions. By both BIC and AIC criteria, the best model included only PART OF SPEECH as a fixed effect (though not either LENGTH or GUIDING DEVICE POSITION), as well as random intercepts for SUBJECT and ITEM ($\log.\text{Time} \sim \text{POS} + (1|\text{Subject}) + (1|\text{Item})$). In this model, residuals had an approximately normal distribution (see **Figure 3**). Part of speech was marginally significant ($F_{(18,1)} = 4.26, p = 0.05$) using the Kenward-Roger approximation (Judd et al. 2012; the same method was used for computing p-levels in subsequent analyses). This confirms the significance of the effect of PART OF SPEECH on CONSULTATION TIME, with verbs taking on average more time to consult than nouns by about one third.

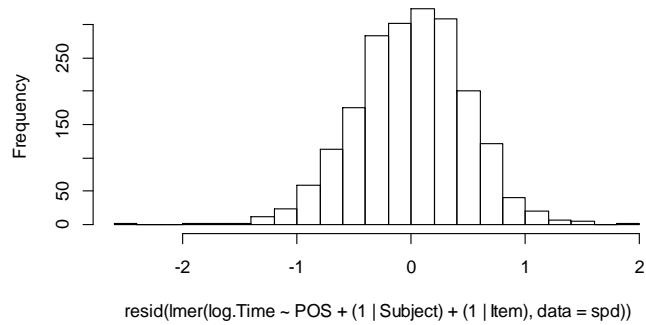


Figure 3: Histogram of residuals for the best model $\log.\text{Time} \sim \text{POS} + (1 | \text{Subject}) + (1 | \text{Item})$. The distribution of residuals is symmetrical and approximately normal.

4.3 Selection accuracy

In terms of SENSE SELECTION ACCURACY, the grand mean was 62 percent, and it was exactly the same in the case of run-on and new-line entries. Nouns exhibited a somewhat higher mean accuracy (66 percent) than verbs (58 percent). Broken down by PART OF SPEECH and GUIDING DEVICE POSITION, the means patterned as in **Table 3**.

	new-line	run-on	any position
Noun	67%	66%	66%
Verb	57%	59%	58%
any POS	62%	62%	62%

Table 3: Mean SENSE SELECTION ACCURACY for combinations of GUIDING DEVICE POSITION and PART OF SPEECH (in percent)

A series of binary logistic models with random factors was fitted with SENSE SELECTION ACCURACY as the outcome variable. The best model was the intercept-only model with random intercepts for SUBJECT and ITEM ($\text{Correct.sense} \sim 1 + (1 | \text{Subject}) + (1 | \text{Item})$). The best model that was not intercept-only was the model that included PART OF SPEECH as a predictor, with nearly the same *Akaike Information Criterion* value ($AIC = 2416.7$) as the intercept-only model ($AIC = 2416.4$). This may be taken to interpret that PART OF SPEECH may play some minor role in determining sense selection accuracy (with nouns yielding

more success on average than verbs). By contrast, GUIDING DEVICE POSITION appears not to matter for the accuracy of sense selection within entries.

4.4 Relationship between entry length and consultation time

In **Figure 4**, we plot the relationship between ENTRY LENGTH and CONSULTATION TIME for the individual test items (for CONSULTATION TIME, these are unlogged fitted values). The slope is not significantly different from zero ($p = 0.15$, n.s.), which suggests CONSULTATION TIME is not related to ENTRY LENGTH; this may be seen as somewhat surprising: a longer entry might be expected to require longer study time due to its sheer length. In addition, a longer entry could also mean that the lexical item itself might be a more challenging one, and that might again be expected to require longer consultation. No such effect, however, is evident in our data, and there is in fact a (non-significant) hint of a reverse trend.

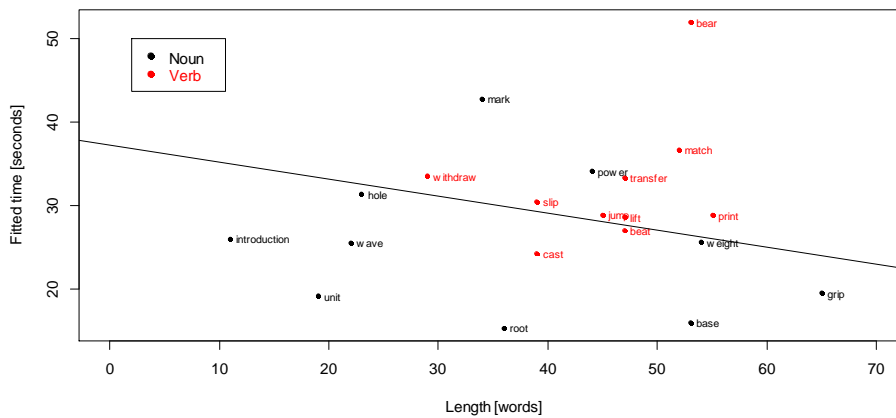


Figure 4: Plot of CONSULTATION TIME (fitted values) and ENTRY LENGTH for the individual items (nouns are shown in black, verbs in red font)

4.5 Relationship between selection accuracy and consultation time

The final question refers to the relationship between SELECTION ACCURACY and CONSULTATION TIME for our test items. We might expect that the relatively easy items would be dealt with relatively more quickly and with better success rates than the more challenging items. Therefore, we would expect a negative relationship between success rates for SELECTION ACCURACY and CONSULTATION TIME. In **Figure 5**, we plot the relationship between these two variables fitted in a linear mixed model with CONSULTATION TIME as outcome, SELECTION ACCURACY

as predictor and random intercepts for subject and item (this turned out to be the best model: POS was only marginally significant and was removed from the model, and the POS by SELECTION ACCURACY interaction was not significant). The slope of the best line of fit included in **Figure 5** is significantly different from zero ($p = 0.005$), and indicates a robust negative linear relationship between CONSULTATION TIME and SELECTION ACCURACY. Roughly speaking, an improvement of 0.1 in SELECTION ACCURACY corresponds to a decrease of 3.2 seconds in CONSULTATION TIME.

Looking at the plot, the noun item *base* has the best SELECTION ACCURACY, and is also the fastest; in contrast, the noun *mark* is the least successful and is the second slowest item, with the verb *bear* being the slowest of all.

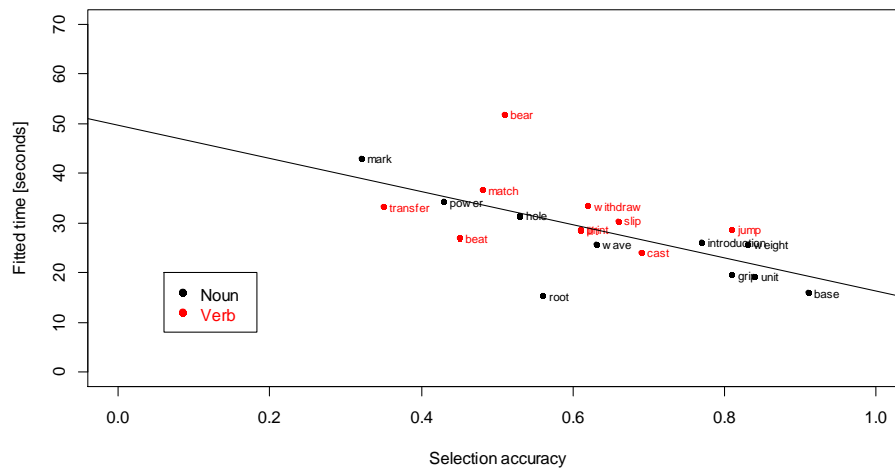


Figure 5: Plot of CONSULTATION TIME (unlogged fitted values) and SENSE SELECTION ACCURACY for the individual items (nouns are shown in black, verbs in red font)

5. Discussion and conclusion

The most important and practical research finding of this study is that no effect of GUIDING DEVICE POSITION on either CONSULTATION TIME or SENSE SELECTION ACCURACY was found. Although this is a null effect, the substantial sample size gives us some confidence that these hypothetical effects, even if present, are small in magnitude. In practical terms, this means that there is no evidence that adding new lines in an attempt to better signal new senses to the user adds further benefit, as the line breaks did not significantly improve either speed or success of consultation (they were actually the same). We might then tenta-

tively speculate that when entry senses are prefixed with signposts rendered in typography similar to that adopted in the present study (i.e. in framed capitals following sense numbers: see Appendix), the typography is salient enough to offer good sense discrimination in a visual search, and thus gets no further help from breaking the line for each sense. Since adding line breaks uses up extra space, then — based on the present findings — we would advise *against* their inclusion in dictionary entries, as long as sufficiently salient signposts are present, until real evidence is presented of any associated benefits. This finding is of direct relevance to lexicographic practice, since the variable manipulated is one that lexicographers can actually control in practical lexicographic work: they can decide to make entries with or without the extra line breaks. Our findings suggest that *without* might be good enough.

The remaining findings of the present study are interesting, though they cannot be directly translated into improved lexicographic practice. In terms of the part of speech of the entry, verb entries needed about thirty percent more time than noun entries, confirming findings from previous research (Nesi and Tan 2011; Ptasznik 2015) that verb entries tend to be more problematic than noun entries and that consulting verb entries takes more time. When it comes to the other outcome measure, accuracy of sense selection, nouns outperformed verbs by about fourteen percent. Although this difference did not turn out to be significant, with the intercept-only model offering the best explanation by the usual trade-off between model accuracy and parsimony, the direction of the difference again contributes to the overall tendency evident from other studies demonstrating verb entries to be more problematic, in general, than noun entries, and this can translate into differences in terms of both speed and success of dictionary consultation. This is no fault of lexicographers and nothing they can fix by improving dictionary entries, as the underlying cause lies in an inherent difference between nouns and verbs. Other things being equal, nouns are more fundamental to human experience (prototypically, they designate objects) than verbs (which prototypically represent actions). This insight was captured by Lyons (1977: 7.4, 8.1) and re-iterated in the lexicographic context by Piotrowski (1989: 102). The inherent relative difficulty of verbs presents a challenge to the definer: definitions of verbs tend to be longer than those of nouns, and, given the present and previous findings, presumably more challenging to process and comprehend. In this context, an interesting extension would be to test adjective entries in a similar manner and, even more interestingly, to see if the problem persists to the same extent if translation equivalents are used rather than same-language definitions.

As stated above, the more challenging lexical items typically require longer entries; however, our study did not find any indication of consultation time increasing with greater entry length. This finding invites further research into the phenomenon of information overload in the context of dictionary entries, as in our study there is no evidence that an entry of some 70 words in length is problematic in this regard. On the other hand, the two main outcome

variables in our study did correlate quite highly: consultation time was inversely related to the accuracy of sense selection, meaning that the longer the consultation, the lower the accuracy: a relationship that is to be expected.

This study has investigated entries in the print format. In digital dictionaries, sense navigation devices may work somewhat differently, and it may be of relevance to re-test some of the present research questions in digital dictionary environments.

Notes

1. New-line and run-on guiding devices are demonstrated in the appendices of the paper.
2. Only the most polysemous CALD3 entries have new-line guidewords.
3. Variables are written in small capital letters throughout this article (PART OF SPEECH, GUIDING DEVICE POSITION, CONSULTATION TIME, SENSE SELECTION ACCURACY).

References

Online dictionaries

- Cambridge Dictionary Online*. (<https://dictionary.cambridge.org/>). [CDO]
Longman Dictionary of Contemporary English Online. (<http://www.ldoceonline.com/>). [LDOCEO]
Macmillan English Dictionary Online. (<http://www.macmillandictionary.com>). [MEDO]

Print dictionaries

- Cowie, A.P. (Ed.)**. 1989. *Oxford Advanced Learner's Dictionary of Current English* (4th edition). Oxford: Oxford University Press. [OALD4]
Crowther, J. (Ed.). 1995. *Oxford Advanced Learner's Dictionary of Current English* (5th edition). Oxford: Oxford University Press. [OALD5]
Delacroix, L. (Ed.). 2014. *Longman Dictionary of Contemporary English* (6th edition). Harlow: Longman. [LDOCE6]
Deuter, M. (Ed.). 2015. *Oxford Advanced Learner's Dictionary of Current English* (9th edition). Oxford: Oxford University Press. [OALD9]
Gillard, P. (Ed.). 2003. *Cambridge Advanced Learner's Dictionary* (1st edition). Cambridge: Cambridge University Press. [CALD1]
Mayor, M. (Ed.). 2009. *Longman Dictionary of Contemporary English* (5th edition). Harlow: Longman. [LDOCE5]
McIntosh, C. (Ed.). 2013. *Cambridge Advanced Learner's Dictionary* (4th edition). Cambridge: Cambridge University Press. [CALD4]
Procter, P. (Ed.). 1995. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press. [CIDE]
Rundell, M. (Ed.). 2002. *Macmillan English Dictionary for Advanced Learners* (1st edition). Oxford: Macmillan Education. [MED1]

- Rundell, M. (Ed.)**. 2007. *Macmillan English Dictionary for Advanced Learners* (2nd edition). Oxford: Macmillan Education. [MED2]
- Summers, D. (Ed.)**. 1995. *Longman Dictionary of Contemporary English* (3rd edition). Harlow: Longman. [LDOCE3]
- Summers, D. (Ed.)**. 2003. *Longman Dictionary of Contemporary English* (4th edition). Harlow: Longman. [LDOCE4]
- Turnbull, J. (Ed.)**. 2010. *Oxford Advanced Learner's Dictionary of Current English* (8th edition). Oxford: Oxford University Press. [OALD8]
- Walter, E. (Ed.)**. 2005. *Cambridge Advanced Learner's Dictionary* (2nd edition). Cambridge: Cambridge University Press. [CALD2]
- Walter, E. (Ed.)**. 2008. *Cambridge Advanced Learner's Dictionary* (3rd edition). Cambridge: Cambridge University Press. [CALD3]
- Wehmeier, S. (Ed.)**. 2000. *Oxford Advanced Learner's Dictionary of Current English* (6th edition). Oxford: Oxford University Press. [OALD6]
- Wehmeier, S. (Ed.)**. 2005. *Oxford Advanced Learner's Dictionary of Current English* (7th edition). Oxford: Oxford University Press. [OALD7]

Other references

- Aitchison, Jean**. 1997. *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Blackwell.
- Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker**. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1-48.
- Bogaards, Paul**. 1998. Scanning Long Entries in Learner's Dictionaries. Fontenelle, Thierry, Philippe Hilgsmann, Archibal Michiels, André Moulin and Siegfried Theissen (Eds.). 1998. *EURALEX '98 Actes/Proceedings*: 555-563. Liège: English and Dutch Departments, University of Liège.
- Dziemianko, Anna**. 2016. An Insight into the Visual Presentation of Signposts in English Learners' Dictionaries Online. *International Journal of Lexicography* 29(4): 490-524.
- Dziemianko, Anna**. 2017. Dictionary Entries and Bathtubs: Does It Make Sense? *International Journal of Lexicography* 30(3): 263-284.
- Judd, Charles M., Jacob Westfall and David A. Kenny**. 2012. Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology* 103(1): 54-69.
- Lew, Robert**. 2010. Users Take Shortcuts: Navigating Dictionary Entries. Dykstra, Anne and Tanneke Schoonheim (Eds.). 2010. *Proceedings of the XIV Euralex International Congress*: 1121-1132. Ljouwert: Afûk.
- Lew, Robert and Julita Pajkowska**. 2007. The Effect of Signposts on Access Speed and Lookup Task Success in Long and Short Entries. *Horizontes de Lingüística Aplicada* 6(2): 235-252.
- Lew, Robert and Patryk Tokarek**. 2010. Entry Menus in Bilingual Electronic Dictionaries. Granger, Sylviane and Magali Paquot (Eds.). 2010. *eLexicography in the 21st Century: New Challenges, New Applications*: 193-202. Louvain-la-Neuve: Cahiers du CENTAL.
- Lyons, John**. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Nesi, Hilary and Kim Hua Tan**. 2011. The Effect of Menus and Signposting on the Speed and Accuracy of Sense Selection. *International Journal of Lexicography* 24(1): 79-96.
- Phillips, Nathaniel**. 2017. *Yarrrr!: The Pirate's Guide to R* (R Package Version 0.1.5).

- Piotrowski, Tadeusz.** 1989. Monolingual and Bilingual Dictionaries: Fundamental Differences. Tickoo, Makhan L. (Ed.). 1989. *Learners' Dictionaries: State of the Art*: 72-83. Singapore: SEAMEO Regional Language Centre.
- Ptasznik, Bartosz.** 2015. *Signposts and Menus in Monolingual Dictionaries for Learners of English*. Poznań: Adam Mickiewicz University Press.
- Ptasznik, Bartosz and Robert Lew.** 2014. Do Menus Provide Added Value to Signposts in Print Monolingual Dictionary Entries? An Application of Linear Mixed-Effects Modelling in Dictionary User Research. *International Journal of Lexicography* 27(3): 241-258.
- Singmann, Henrik, Ben Bolker, Jake Westfall and Frederik Aust.** 2018. *afex: Analysis of Factorial Experiments* (R Package Version 0.19-1).
- Tono, Yukio.** 1992. The Effect of Menus on EFL Learners' Look-up Processes. *Lexikos* 2: 230-253.
- Tono, Yukio.** 1997. Guide Word or Signpost? An Experimental Study on the Effect of Meaning Access Indexes in EFL Learners' Dictionaries. *English Studies* 28: 55-77.
- Tono, Yukio.** 2011. Application of Eye-Tracking in EFL Learners' Dictionary Look-up Process Research. *International Journal of Lexicography* 24(1): 124-153.

APPENDIX A: New-line guiding devices

Zapisać **NUMER** znaczenia podkreślonego wyrazu, który został użyty w podanym kontekście: _____

Zapisać **CZAS** wymagany do wykonania zadania: _____

Women are advised not to drink more than fourteen units of alcohol per week.

UNIT noun

1 **GROUP** a group of people working together as part of a larger group: *The man is in the hospital's **intensive care unit**.*

2 **MEASURING** an amount of something used as a standard of measurement: [**+of**] *The watt is a **unit** of electrical power.*

3 **PART** a thing, person, or group that is regarded as one single whole part of something larger: *a Russian **army unit** | [**+of**] *The family is the basic social unit of modern society.**

4 **PART OF A BOOK** one of the numbered parts into which a **TEXTBOOK** (= a book used in schools) is divided

5 **PRODUCT** *technical* a single complete product made by a company: *The factory's output is now up to 150,000 units each month.*

6 **PART OF A MACHINE** a piece of equipment which is part of a larger machine: **control / display / filter etc unit**

7 **SCHOOL/UNIVERSITY** *AmE* an amount of work that a student needs to do in order to complete a particular course

APPENDIX B: Run-on guiding devices

Zapisać **NUMER** znaczenia podkreślonego wyrazu, który został użyty w podanym kontekście: _____

Zapisać **CZAS** wymagany do wykonania zadania: _____

Women are advised not to drink more than fourteen units of alcohol per week.

UNIT noun

1 **GROUP** a group of people working together as part of a larger group: *The man is in the hospital's **intensive care unit**.* **2** **MEASURING** an amount of something used as a standard of measurement: **[+of]** *The watt is a unit of electrical power.* **3** **PART** a thing, person, or group that is regarded as one single whole part of something larger: *a Russian **army unit*** | **[+of]** *The family is the basic social unit of modern society.* **4** **PART OF A BOOK** one of the

numbered parts into which a **TEXTBOOK** (=a book used in schools) is divided **5** **PRODUCT** *technical* a single complete product made by a company: *The factory's output is now up to 150,000 units each month.* **6** **PART OF A MACHINE** a piece of equipment which is part of a larger machine: **control / display / filter etc unit** **7** **SCHOOL/UNIVERSITY** *AmE* an amount of work that a student needs to do in order to complete a particular course