Researching users and uses of dictionaries

Nesi, H.

Author post-print (accepted) deposited in CURVE October 2015

Original citation & hyperlink:

Publisher statement: This copy is not the final print version and is not to be cited.

Copyright © and Moral Rights are retained by the author(s) and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

CURVE is the Institutional Repository for Coventry University
http://curve.coventry.ac.uk/open
3.3. Researching users and uses of dictionaries

Dictionary use was not a popular research topic until fairly recently. Welker (2010) summarises 320 empirical dictionary use studies, but only six of these were conducted before 1980. In the 1980s there was an upsurge of interest, however, and an increasing number of studies have taken place in each decade since then, in an ever-wider range of dictionary-using contexts. Much of the latest research focuses on electronic dictionaries produced locally or regionally for specific user groups, as evidenced for example in the proceedings of the two eLex conferences (Granger & Paquot 2010, Kosem & Kosem 2011) and recent AFRILEX, ASIALEX and EURALEX conferences.

The aim of all studies of dictionary use is to discover ways to increase the success of dictionary consultation. This involves the identification of users’ needs and skills deficits, and the making of appropriate matches between types of dictionary, types of dictionary user, and types of dictionary use. The following research questions are pertinent:

- Who are the users of dictionaries?
- What kinds of activity prompt dictionary use?
- What kinds of dictionary do users prefer to use?
- What kinds of information do dictionary users look for?
- What consultation strategies do dictionary users employ, and how successful are these strategies?

This chapter looks at some of the various ways in which these questions have been addressed.
Welker (2010) identifies six main methods of investigating dictionary use:

1. Questionnaire surveys
2. Interviews
3. Observation
4. Protocols
5. Tests and experiments
6. Log files

All these methods are intended to shed light on the way dictionary users consult dictionaries for their own purposes, under non-experimental conditions. Gathering such information is not easy, however, because the data-gathering instruments often rely heavily on users’ ability to explain their consultation behaviour, and because they also tend to steer users towards uncharacteristic patterns of use.

The questionnaire survey is the most common approach, and can be an effective means of gathering data from large numbers of respondents regarding their long-term dictionary-using habits and attitudes (Lew 2002). Questionnaires have come in for criticism, however, for example by Hatherall (1984), Wiegand (1998), Nesi (2000a) and Tarp (2009), because they sometimes place unreasonable demands on users’ powers of recall, and because users and questionnaire designers may not share the same concepts and terminology.

Interviews and observations are used less frequently in dictionary research, and generally with only a few participants because of the cost in terms of time and expertise. Neubach and Cohen (1988) interviewed only six dictionary users, for example, and East (2008) observed
groups of six and five users. Interviews and observations can, however, be more successful than questionnaire surveys as a means of probing dictionary-using behaviour. Interview participants can ask each other for clarification if unexpected aspects of dictionary use come to light, and observations can reveal behaviour without the need for users to describe it at all. The data are therefore less likely to be coloured by misunderstandings and misconceptions, although they do not always reveal natural look-up behaviour because the interviewer or observer may unintentionally influence the outcome, especially if participants believe that researchers approve of certain strategies, and disapprove of others.

Recently, new laboratory-based methods have enabled researchers to observe in detail the way users interact with dictionary information on the computer screen. Heid and Zimmerman (2012), for example, adopted usability testing methods from the field of information science to record keystroke patterns for search routes through different types of dictionary interface. Tono (2011), Kaneta (2011) and Simonsen (2011) employed eye-tracking techniques previously used in the fields of cognitive science, psycholinguistics and human-computer interaction to discover what areas of the dictionary entry users view, in what order, and for how long. Inevitably research of this kind is conducted in an artificial setting with relatively small numbers of participants, but the techniques reveal aspects of dictionary use that probably cannot be discovered by a human interviewer or observer.

Completely natural look-up behaviour is difficult to record because it is a private activity that occurs spontaneously rather than to order. A researcher might spend a very long time observing a potential user as they read or write for their own purposes, before catching the moment when dictionary consultation occurs. Tests or tasks which prompt dictionary use are useful as a means of generating a lot of structured data in a short amount of time, particularly
for comparative purposes, for example to identify the conditions under which users look up the most words, take the least time, achieve the highest comprehension scores or retain the most vocabulary. It is not always possible to extrapolate information about natural dictionary consultation from these findings, however – for example the dictionaries available at the site of the experiment may not be the same as the ones users normally consult, and the task may not bear much resemblance to the users’ normal reading, writing or translating activities.

Protocols, or self-reports, can shed light on users’ understanding and decision making, either during spontaneous dictionary use, or whilst completing a task set by the researcher. Oral protocols are recordings of participants’ thoughts, spoken aloud throughout the consultation process. User behaviour is thus open to examination without the distortion of faulty recall or re-interpretation, but usually relates to only a small number of participants because of the special skills needed to think aloud, and the amount of time required to gather and analyse spoken data. Nesi and Boonmoh (2009), for example, chose 17 participants from a cohort of 580 students to train in think-aloud techniques. They collected data from just eight of these, the ones who proved most capable of verbalising their dictionary use.

Written protocols can be either freely-written or structured using a format prepared by the researcher, perhaps with multiple choice options. Typically they record a reason for each dictionary search, the information searched for, the dictionary used, and an evaluation of the success of the consultation. The method is suitable for use with multiple participants: Müllich (1990) collected 108 written protocols from language learners, for example, and Harvey and Yuill (1997) collected 211. Retrospective protocols are problematic because users quickly forget the details of the consultation, but the process of completing a protocol while using a dictionary is quite disruptive. Atkins and Varantola (1997) asked participants to work in
pairs to reduce disruption, one member using a dictionary, and the other recording the process. With all forms of protocol it is likely that some behaviours will go unrecorded or misrecorded, however, because consultation processes cannot always easily be described.

Log files observe users’ interactions with their computers in an unobtrusive way. They can be used to record experimental data (e.g. Lew & Doroszewska 2009), but are also a good way of capturing information about the searches users make online, when they are engaged in their normal activities, perhaps over an extended period of time. Like other forms of observation, however, log files alone cannot provide much insight into the context or purpose of dictionary consultation, unless the dictionary is linked to an online text or task. Moreover although a log file may be able to indicate whether consultation led to the information the user was searching for, when analysed in isolation it cannot record whether the user considered the consultation successful.

Dictionary users, uses, and contexts of use can all vary enormously, making it unsafe to generalise from the findings of individual studies. In some other fields of research large-scale controlled trials can test how effectively a given treatment works, but the effectiveness of a dictionary cannot usually be investigated by this means because it is difficult to enlist the aid of a representative sample of all potential users (Welker 2010: 13). Studies therefore tend to focus on the behaviour of smaller and more specific groups, representing dictionary users of one particular type, in one particular context. To facilitate the comparison of findings from different studies researchers sometimes try to adopt similar methods; Welker (2010:13) cites a number of studies utilising similar questionnaire formats, for example, and Dziemianko (in press) traces a sequence of dictionary use replication studies. Researchers also seem
increasingly likely to adopt a mixed method approach. This helps to compensate for the inevitable limitations of each individual method, and increases the reliability of the findings.

Multiple studies of different groups, using similar or complementary methods, may gradually enable us to build up a complete picture of “how dictionaries are used…. who the users are, where, when and why they use dictionaries, and with which result” (Tarp 2009: 279).

Who are the users of dictionaries?
Varantola (2002:33) divides dictionary users into three broad categories: language learners, non-professional users, and professional users, these last being those who ‘normally use a dictionary to perform a task that they get paid for’. Other user variables that are likely to affect behaviour are:

- age
- mother tongue
- second or foreign language
- language proficiency level
- educational level
- level of skill in dictionary use
- role (as a teacher, learner, translator, traveller, player of word games etc.)
- location (geographically, and within the home, place of work or educational institution).

The geographical location of online dictionary users is ascertainable from log files. Otherwise questionnaires are typically considered the best way to obtain factual information relating to these variables. Questionnaire items to establish user profiles were used in studies by Atkins
and Varantola (1998), for example, and Hartmann (1999), and are often included in larger surveys of user wants and needs involving multiple data collection methods. Thus Law and Li (2011) triangulated findings from a questionnaire survey with findings from interviews, and Ptaszynski and Sobkowiak (2011) combined a questionnaire survey with protocols produced during translation and writing tasks, and post-task interviews.

Research participants are usually selected on the grounds that they are available, willing to take part, and reasonably representative of the types of user the researchers are most concerned with. This means that they are often university students, as researchers are usually based in universities. It also means that people in locations where little research takes place tend to be under-represented in studies of dictionary use. Lew (2011) points out that there is a deficit of information about many contexts of use, for example by tourists, or families doing crossword puzzles at home.

What kinds of activity prompt dictionary use?

Most dictionary consultations are undertaken when the user is engaged in another activity, in order to ‘solve a context-dependent problem’ (Varantola 2002). Dictionary use is typically classified as ‘receptive’ (i.e. to help with text decoding tasks) or ‘productive’ (i.e. to help with text encoding tasks), although dictionaries can also be treated as resources for learning new vocabulary or finding out about a language. Table One summarises the broad range of activities generally associated with dictionary use.

<table>
<thead>
<tr>
<th>Written medium</th>
<th>Receptive</th>
<th>Productive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>Writing</td>
<td></td>
</tr>
<tr>
<td>Translating from L2 to L1</td>
<td>Translating from L1 to L2</td>
<td></td>
</tr>
<tr>
<td>Spoken medium</td>
<td>Listening</td>
<td>Speaking</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Interpreting from L2 to L1</td>
<td>Interpreting from L1 to L2</td>
</tr>
<tr>
<td></td>
<td>Gathering language information</td>
<td></td>
</tr>
</tbody>
</table>

**Table One: Activities associated with dictionary use.**

Traditionally the large monolingual dictionaries have focussed on the receptive needs of native speakers, while learners’ dictionaries, bilingualised dictionaries and L1-L2 bilingual dictionaries also support language production by providing translations and/or more grammar, phraseology, usage and pronunciation information. Most, but not all, surveys have found that dictionaries are more often used receptively, whilst reading (Marello 1987, Hartmann 1999, Stark 1999) or translating from L2 to L1 (Tomaszczyk 1979). Battenburg (1991) found that lower level students used their dictionaries more whilst reading, and advanced level students used their dictionaries more whilst writing. Tomaszczyk’s survey respondents reported using dictionaries for speaking and listening activities, but he concluded that they might have been referring to the preparation of oral reports. Dictionary use prior to the advent of mobile e-dictionaries was generally associated with activities in the written medium.

‘Reading’, ‘writing’, ‘speaking’ and ‘listening’ are very broad activity types. Some questionnaires make finer distinctions; Ripfel (1990) lists reading newspapers or magazines, listening to the radio or watching television, explaining word meanings to children, doing homework, and writing letters, for example, and Hartmann (1999) includes playing word games, writing assignments, and reading for study and pleasure. Presumably users’ needs change according to the type of receptive or productive activity they are engaged in: translators need to understand every word in the text, while learners reading for pleasure may only read for gist, ignoring many of the words they do not know. Activity type can also affect
users’ choice of dictionary format. Nesi (2010) records a complex picture of e-dictionary preferences, with students using computer-based dictionaries when reading and writing at the computer, and portable electronic dictionaries when reading and writing with paper-based materials. Her participants also preferred mobile e-dictionaries for speaking and listening activities, because of their accessibility and audio pronunciation features.

The secondary ‘knowledge-oriented’ use of dictionaries (Bergenholtz & Tarp 2003) has most often been studied in connection with vocabulary learning and retention. This is usually measured by testing participants after they have completed a task under various conditions, for example with or without a bilingual and/or monolingual dictionary, in print and/or in electronic form (e.g. Luppescu and Day 1993, Wingate 2002, Dziemianko 2010). Some items in questionnaire surveys explore the secondary use of dictionaries in more natural surroundings (e.g. Marello 1987, Chi 1998, Hass 2005). Ronald (2002) examined vocabulary acquisition as a result of using a dictionary while reading. Nesi (2010) investigated the way users created and annotated their own wordlists, using the latest e-dictionary resources. It seems that the e-dictionary format encourages browsing for general interest, especially when words within one entry are hyperlinked to other entries (Nesi 2000b).

**What kinds of dictionary do users prefer to use?**

The typology of reference works is complex, but the basic choices facing users are between monolingual and bilingual, hard-copy and electronic. Studies suggest that although users can distinguish these broad categories, many fail to make finer distinctions in terms of types of reference work, the different user groups they are intended for, and the relative merit of comparable titles. Participants in surveys and interviews are often unable to give precise information about the publishers and titles of their dictionaries (Nesi & Haill 2002, Law & Li
2011), and generally the investigation of dictionary preferences is hampered by users’ ignorance of the details concerning the dictionaries they own and use, and of the different types of dictionaries that exist. In many educational contexts dictionary skills are not systematically taught (Atkins and Varantola 1997, Bae 2011). When they are taught they rarely include the skills of selection and criticism.

One simple but seemingly underused way of establishing users’ preferences is to ask them to evaluate various kinds of dictionary material. MacFarquhar and Richards (1983) used this method to compare users’ impressions of different defining styles, and Kanazashi (2011) reports studies comparing users’ responses to formatting and entry features. User evaluations do not prove that one lexicographical approach is superior to another, but they are a useful supplement to the comments of dictionary reviewers, who often do not belong to the user group the publisher is targeting.

Publishers can obtain information about the popularity of their dictionaries through log files and sales figures, but such commercially sensitive information is rarely made available to external researchers, and most published log-file studies of natural dictionary use, such as those of Bergenholz and Johnsen (2007) and Hult (2007), relate to the use of individual experimental reference works. User preferences are more often investigated by means of questionnaire surveys: results generally indicate that language learners prefer bilingual or bilingualised dictionaries (see, for example, Tomaszczyk 1979, Baxter 1980, Atkins and Varantola 1997, Lew 2004), although monolinguals tend to be used progressively more at more advanced levels of study. Monolingual dictionaries are also often regarded as superior in quality, especially by teachers (Boonmoh & Nesi 2008). Chan (2011) found that users mistakenly believed bilingualised dictionaries to contain less usage information. Some
surveys have investigated dictionary purchasing choices: the decision sometimes rests with teachers (Béjoint 1981, Hartmann 1999, Boonmoh & Nesi 2008) and may not reflect users’ real preferences.

The market for print dictionaries is declining rapidly, as increasing numbers of users access information via the internet, usually for free. E-dictionary packages and portals often contain a wide range of dictionaries of varying quality, some of which contain unattested headwords and idioms, presumably included to increase the extent of coverage and impress unsophisticated users (Nesi 2012). However despite the growing number of studies of e-dictionary ownership and user preferences, little research has been undertaken to evaluate the content of popular e-dictionary sites, and little information is available to help users choose and use e-dictionaries appropriately.

What kinds of information do dictionary users look for, and what consultation strategies do they employ?

Unsurprisingly, surveys of native speaker users (e.g. Quirk 1975, Jackson 1988, Hartmann 1999, Chatzidimou 2007) and language learners (e.g. Tomaszczyk 1979, Béjoint 1981, Battenburg 1991, Bishop 1998) indicate greatest interest in information that can be applied immediately to a receptive or productive task, such as meaning and spelling, rather than knowledge-oriented information such as the etymology of the look-up word.

Because dictionary use generally occurs whilst users are busy doing something else, they generally want to find information quickly, with as little disruption as possible to the task they are undertaking. Several studies have investigated users’ misinterpretations of dictionary information, through failing to read the entire entry (Miller & Gildea 1985, Nesi & Meara
1994), failing to understand grammatical information (Chan 2012), or consulting the wrong entry or subentry (Nesi & Haill 2002). The first definition in polysemous entries is the one that immediately catches the user’s eye, and it also usually represents the most familiar meaning, so alternative definitions lower down the entry are often ignored (Tono 1984, Bogaards 1998). Some studies have explored the role of ‘signposts’ within long entries as a means of helping users find the right, contextually appropriate, definition (Tono 1992, 1997, Bogaards 1998, Lew and Pajkowska 2007, Lew 2010, Nesi & Tan 2011).

One of the reasons why e-dictionaries are becoming so popular is that they provide faster access than print dictionaries. Initially researchers worried that the speed of consultation might affect the quality of the experience (Taylor and Chan 1994, Zhang 2004, Stirling 2003), but experiments with comparable print and e-dictionaries have either recorded no significant difference in task performance (Nesi 2000b, Koyama and Takeuchi 2003, 2007) or significantly better performance by e-dictionary users (Shizuka 2003, Dziemianko 2010).

Many popular online bilingual dictionaries translate in a fairly primitive way, however, without information or labels to indicate register differences or restrictions on use (Nesi 2012). Thus they might encourage a tendency, noted by Hatherall (1984), to look for one-word equivalents of search terms, translating word-for word rather than considering the context.

Researching e-dictionary use is particularly problematic because the content of e-dictionary packages and portals is changeable and poorly described. Commercial products can also be prohibitively expensive; unlike print publishers, e-dictionary developers are not in the habit of offering review copies or discounts on class sets. Shizuka (2003), Koyama and Takeuchi (2007) and Diehr (undated) acquired pocket electronic dictionaries from Casio for use in
experiments in Japan and Germany, but other researchers such as Chen (2010) complain that they lacked the resources they needed to fully investigate e-dictionary use.

**Conclusion**

Research into dictionary use is ultimately intended to help users consult dictionaries more successfully. Progress in this respect has been patchy, however. In some cases research has informed teacher training, and the teaching of dictionary skills (see, for example, Bae 2011), but it does not seem to have greatly affected the choices made by commercial dictionary publishers. ‘Few modifications to the learners’ dictionary design are supported by published results of experimental research on how learners really use dictionaries’, as Lew and Dziemianko (2006:277) point out.

The decline of the print dictionary has led to a reduction in the size of lexicographical teams in mainstream publishing houses, greater reliance on automatic dictionary compilation procedures, and the rise of online dictionary sites created and managed by computer experts rather than lexicographers. This might suggest that there will be fewer opportunities for user research to influence design in years to come. Fortunately, however, the technology is also enabling many university-based research groups to experiment with new presentation techniques and dictionary content. The proceedings of the two recent eLex conferences (Granger & Paquot 2010, Kosem & Kosem 2011) are full of descriptions of small and specialised e-dictionaries designed to meet the needs of particular groups of users. Development teams for dictionaries of this sort are familiar with the research methods described in this chapter, and have the means to conduct their own user research, building on prior research findings. Their dictionaries will not be ‘block-busters’ like the famous print
dictionaries of the past, but they do offer the hope of further, fruitful, user-research-informed design.

References


Bergenholtz, H. & Johnsen, M, (2007) Log files can and should be prepared for a functionalistic approach. *Lexikos* 17 1-20


Chatzidimou, K. (2007) Dictionary use in Greek educaton: an attempt to track the field through three empirical studies. *Horizontes de Lingüística Aplicada* 6 (2) 91-104


