

A lexical analysis of 1995, 2000 and 2005 ascilite conference papers

Craig Zimitat

Griffith Institute for Higher Education

Griffith University

Papers from the 1991, 2000 and 2005 ascilite conferences were analysed to identify key themes and concepts that have emerged from the thinking and research of Australian academics working with learning technologies. In 1995, themes were related to student learning and software / multimedia development. In 2000 there was a focus on student learning and specific products. In 2005, the online learning environment, learning activities and outcomes were key themes. Key themes in the whole collection of papers were the examination of most facets of “going online” and learning activities/assessment.

Keywords: lexical analysis, students, learning, learning technology, research themes

Introduction

The Australasian Society for Computers in Learning in Tertiary Education is getting close to the milestone of a quarter of a century of existence. During that time computers have shrunk from room size to pocket size, software has grown in complexity and educational paradigms have multiplied. The Society has held regular conferences since the early 1980s, and regularly published proceedings of those conferences. Since the mid 1990s, these conference proceedings have been made available on the internet as a searchable archive. This collection of documents is a useful resource for an examination of key themes and trends in the use of learning technologies in Australia over the last decade.

Content analysis can be considered as the “study of recorded human communications, such as books, web sites, paintings and laws” (Babbie, 2004). Descriptive content analysis involves the examination of large volumes of information, often using some theoretical framework as a scaffold for the identification of themes and concepts emerging from the data. The quality of the resultant analysis can be limited by inappropriate frameworks and lack of objectivity. Grounded theory approaches (Glaser & Strauss, 1967) involve being open to discovery of propositions; they do not test hypotheses. Through processes of note-taking, coding and memos, concepts and themes are sorted into categories and the ‘theory’ is emergent rather than a particular theory being imposed on the data.

Leximancer is a software tool used to support lexical-text analysis, consistent with grounded theory methodology. The body of text is examined and a ranked list of terms is generated by an analysis of frequency of use and related occurrence. These terms then feed into a thesaurus builder, which creates a set of classifiers by iteratively extending terms through identifying more distant co-occurrence. This results in the formation of *concepts* that are related to chunks of text, usually 2–3 sentences in length. Leximancer calculates the relative co-occurrence of concepts to generate a matrix which in turn is used to generate a visual display that illustrates the connectedness of concepts. Each concept is linked to the original reference text. This allows the user to revisit the analysis and impose specific limitations on the analysis, or seed the analysis with key terms or concepts. The advantage of using Leximancer, over hand coded descriptive or grounded theory analysis, is that large amounts of text can be subject to analysis in a routine manner, using consistent methodology that generates a repeatable outcome. Multi-dimensional visualisation techniques facilitate understanding of the relationships between concepts, as well as the strength of those relationships.

The purpose of this exploratory research was to (i) demonstrate the potential use of Leximancer to undertake lexical analysis of a limited amount of ascilite archival material, and (ii) through that analysis, identify key and emergent themes in a body of largely Australian research on learning technologies.

Method

Over 200 documents, refereed papers from ascilite conferences, were analysed to identify key concepts and themes that have emerged in the practice of academics using learning technologies over the last decade. Conference papers (full papers and short/concise papers) were downloaded from the ascilite website (www.ascilite.org.au) or Conference CD-ROM in the supplied file format. The abstract, figures, tables, list of references, acknowledgements, biographical notes and copyright statements were deleted from each paper and the documents were converted to text files for analysis. The list of authors and their affiliations were extracted from each file and saved by year. Similarly, lists of references from each paper were extracted from each file and saved in files by year of publication. References were checked for conversion failures and edited as necessary into a useable format. These lists were analysed to identify the most frequently cited journals and authors for each year.

Lexical analysis was undertaken using Leximancer v. 2.1 (www.Leximancer.com). Some details on the approach to using Leximancer can be found in Watson, Smith and Watter (2005). The default settings were used for analysis, with the following changes: (i) analysis blocks were set at two sentences; (ii) bigram sensitivity was set at 3 to treat hyphenated words as a single entity; (iii) language testing was turned off to exclude tables and lists from analysis; (iv) boilerplate was inactivated to exclude commonly used or 'templated' blocks of text; and (v) the learning threshold was set at default (10). These settings were preserved in the initial `lex.config.pm` file and were used for all subsequent analyses. The concept dictionary arising from the analyses was edited using the following processes: words and their plural forms were combined, words with their related tenses were combined, identical words in capitalised and non-capitalised forms were combined, and words with English and American spellings were combined. Abbreviations were combined with their long hand term, and special terms such as `www`, `edu` and `au` were deleted from the list of concepts. Other terms such as students and learners were not combined because of their potential to relate to other terms in different ways.

Results and discussion

The spatial map and list of concepts for the analyses of the 1995 ($n=65$), 2000 ($n=61$) and 2005 ($n=85$) papers are showing in Figure 1. The spatial map illustrates three important characteristics of the text. First, the frequency of concept in the document collection is related to the boldness of the text – the bolder (or brighter) the concept, the more often it appears in the text. The brightness of the links between concepts reflects the co-occurrence of those concepts. Finally, the proximity between concepts in the map reflects their closeness in terms of appearing in related conceptual contexts within in the original text. The list of concepts is ranked, showing the top 10 or so concepts based upon frequency.

The primary ranked concepts in 1995 were students and learning. As illustrated in the spatial map, there were three clustered concepts in the 1995 papers – student learning, multimedia development and computer-based information and materials. This was a period of time bridging the CAL movement and Internet. The relative frequency of the use of terms online and internet was less than 5%, growing to 9% in 2000. The most frequently used terms in 2000 were students and learning, followed by internet, online and teaching. Three main themes in the 2000 papers included students learning online; the development of teaching and learning materials, and learning outcomes e.g. knowledge and development of skills.

In 2005, the most frequently used terms were learning, online and technology, and teaching and development. Themes emergent from the lexical analysis were: the online learning environment; social dimensions of online experience; and learning outcomes (skills, knowledge, assessment and feedback). Within the online learning environment theme were issues of access, support, academic staff development and the student experience.

Overall, across the 204 full papers spanning the decade 1995–2005, the most frequently used terms were: development, online, teaching and technology. The three broad themes in the papers include: going online (the work involved in online teaching and teaching with technology); products (materials, information and resources), engagement and learning outcomes (including learning activities, design for learning and assessment).

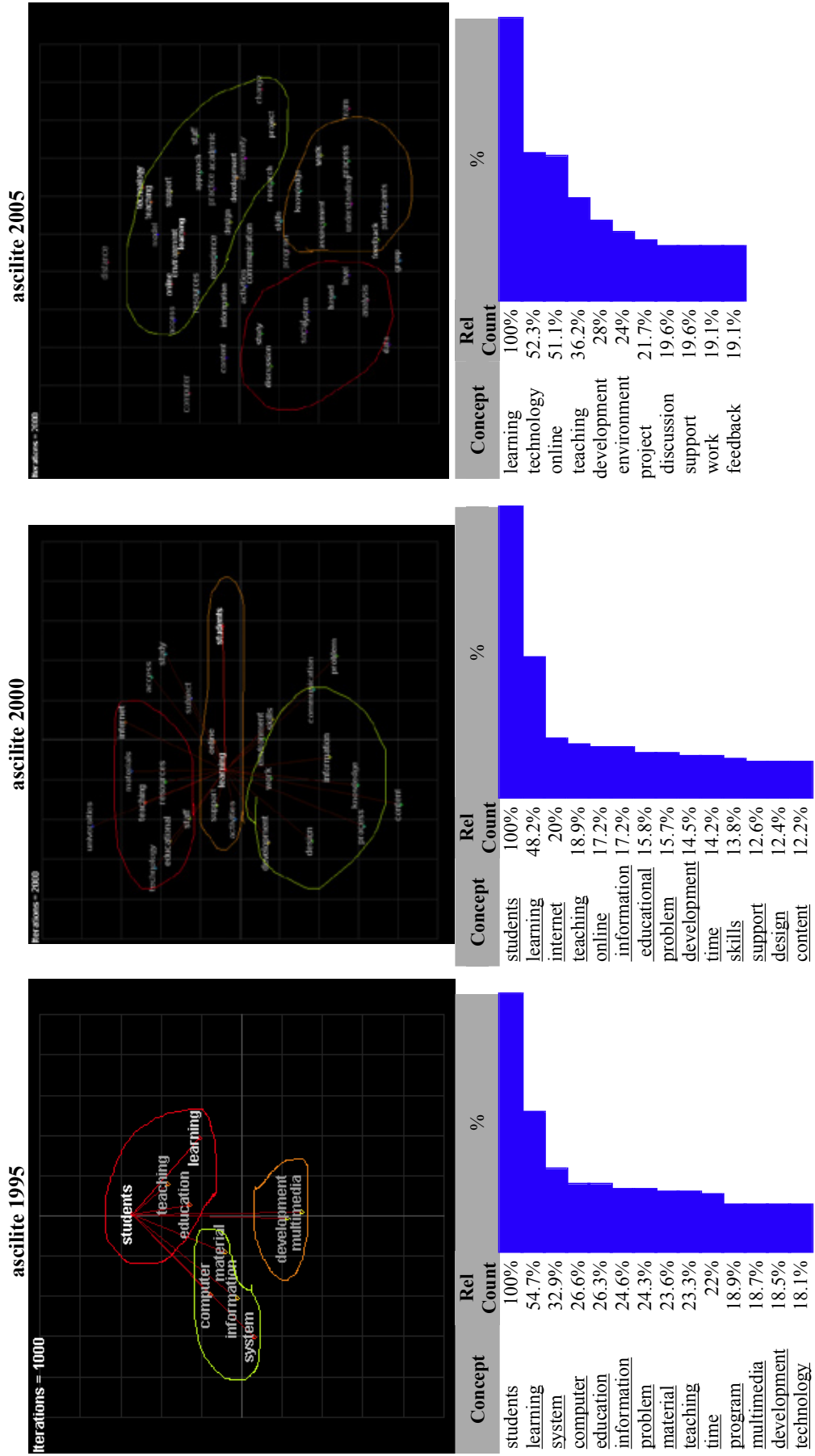


Figure 1: Lexical analysis of papers from each of the 1995, 2000 and 2005 ascilite conferences

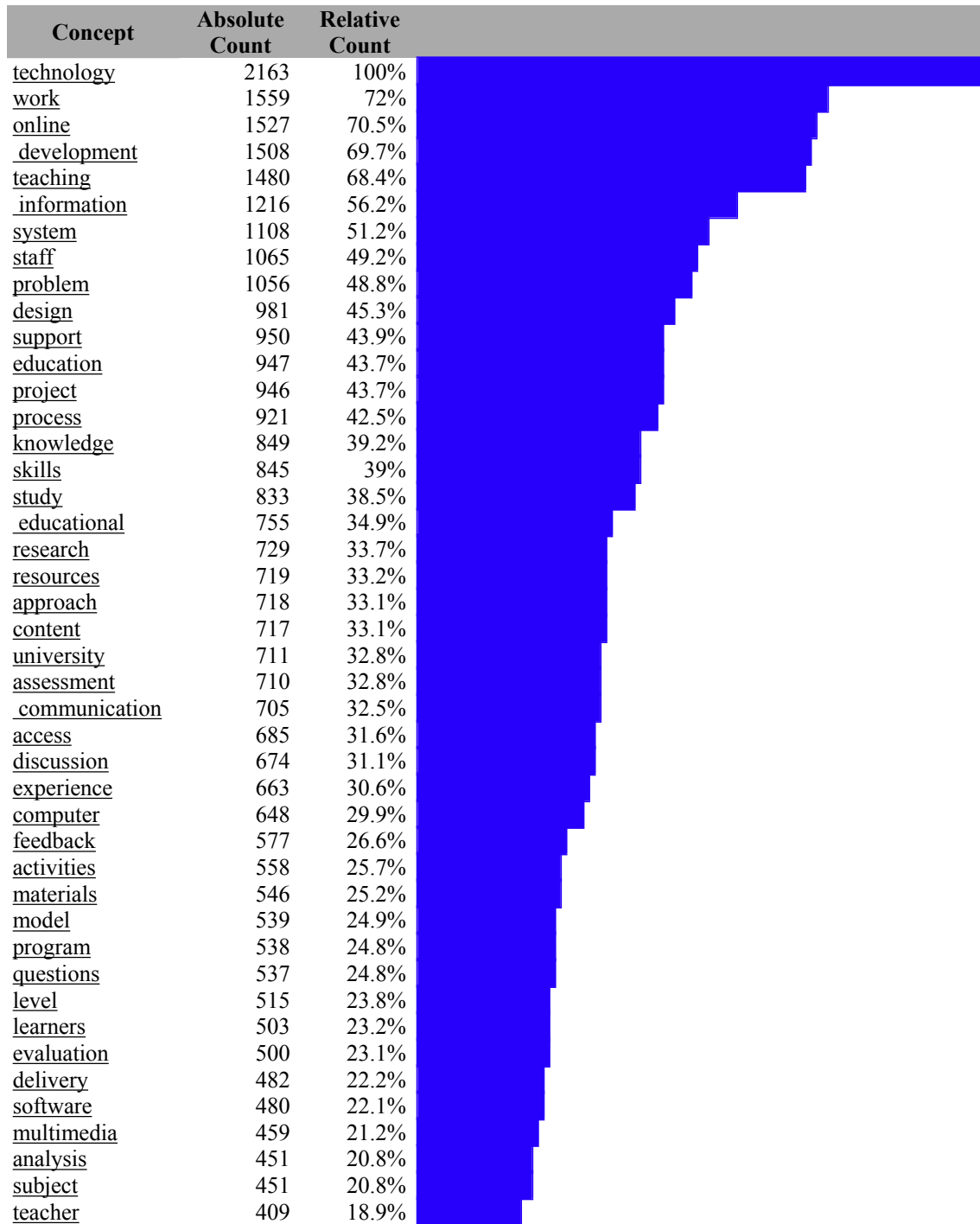


Figure 2: The ranked concept list for all papers 1995, 2000 and 2005

Further work

This exercise has demonstrated the capability of the Leximancer software to analyse archival conference papers. A fuller and larger analysis of the whole of the ascilite archives, including all conference papers, concise papers and historical papers not on the internet, and other higher education teaching-learning and information and communications technology (ICT) conferences would be a useful exercise.

In tandem with this, an examination of the key reference sources used for these papers would yield further information on important conceptual and philosophical influences on writing, key monographs and key journals informing thinking, activity and research.

