NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

CHAPTER 3. PRELIMINARY RESEARCH

3.1 Introduction

The previous chapter discussed the distinctive features of a corporate Web site. As a Web site is quite different from more traditional media, current methods for developing a typology, based on traditional communication theories, may not work well in predicting dynamic interaction between firms and consumers. It is clear that, despite continued research efforts on the Web, to date systematic empirical research on a Web site typology has been deficient. To address this issue, preliminary empirical research was conducted. Accordingly, the purpose of this chapter is to describe a scheme for classifying corporate Web sites into meaningful groups in terms of content and structure.

This chapter describes an empirical investigation on a Web site typology based on classification schemes this study explored in the previous chapter. This chapter commences by discussing the methods used to classify Web sites including a general description of content analysis, the sample, pretesting of the coding procedure, and statistical analysis. Then, the results of the content analysis and the non-parametric Mann-Whitney *U* test are discussed.

3.2 Methods

3.2.1 Content Analysis

To develop the typology, the Web sites of a sample of Australian and Korean companies were content analysed. A corporate Web site is an outcome of a firm's effort to communicate with customers, so a study of the Web site can provide useful insights into how marketing communication activities are used to provide information for customers.

Content analysis is a purely descriptive technique that enables the systematic evaluation of communication contents (Aaker, Kumar, and Day, 2001; Berelson, 1952; Cartwright, 1953; Holsti, 1969; Palmer and Griffith, 1998). It is one of the most widely used methods for evaluating communication media including advertising (e.g., Yale and Gilly, 1988). As Rourke and Anderson (2002) argue, content analysis is:

> "systematic in the sense that a theoretical, a priori set of categories is constructed into which communication content is classified. It is objective in the sense that classification is rule-based and the reliability of classification is tested by having multiple coders classify the same content. Its quantitative character is evident in the process of converting communication content into discrete units and calculating the frequency of occurrence of each unit " (p.7).

Understanding what kinds of messages are presented and the hypermedia used on corporate Web sites is a necessary step toward the evaluation of consumer attitudes toward these marketing activities. Therefore, content analysis can provide a solid foundation for theoretical development in marketing communication.

As Web sites range in size from one page to thousands of pages (Nowak, Shamp, Hollander, and Cameron, 1999; Shneiderman, 1997), this study has focused on the home page (i.e., the first page of a Web site that users encounter) because this is the "gateway" to a Web site and as such plays a critical role for users much like an index or a table of contents for a book (Dalal *et al*., 2000; Esrock and Leichty, 2000; Shiva, 1997). Focusing on the home page allows the standardised comparison of Web sites of different companies.

In addition, the inclusion of both Australian and Korean Web site home pages of companies, which were enlisted in the international financial resources directory (e.g., Wright Investors' Service), not only reflects the international nature of the Web, but also allowed the study to encompass a wide range of industries and company sizes. Furthermore, the two countries play important roles in the Asia Pacific region and they represent different perspectives on culture and industry structure, so the inclusion of both countries enhanced the quality of the total sample frame helping to eliminate any unintentional biases in the selection process.

Each Web site was evaluated by two judges who were experienced Web users and bilingual Korean-English speakers. One judge was trained by the other to increase inter-

coder reliability (e.g., Kolbe and Burnett, 1991). The two judges coded the sampled sites independently. Inter-coder reliability was determined through pre-testing.

On the basis of a literature review of Web-mediated marketing studies (e.g., Cockburn and Wilson, 1996; Dholakia and Rego, 1998; Huizingh, 2000; McNaughton, 2001; Palmer and Griffith, 1998; Perry and Bodkin, 2000; Simeon, 1999), a coding frame of categories covering both Web site content (i.e., advertising, public relations, sales promotion, online selling, and customer service) and Web site structure (i.e., text, images, video, audio, and links) was generated as shown in Table 3.1 and 3.2. All categories were mutually exclusive, with precise and unambiguous definitions (e.g., Holsti, 1969). As discussed in the previous chapter, advertising was defined as messages related to a product or service to persuade potential customers to choose a specific product or service. Public relations was defined as activities that use corporate-related information to enhance corporate image or goodwill with broad publics. Sales promotion was defined as an attempt to induce a customer to perform a specific action within a short period of time. Online selling was defined as an online transaction such as direct selling and buying products or services via a Web site. Customer service was defined as the activities of answering customers, responding to customers, and solving customer problems.

As shown in Table 3.1, the Web site content typology included five general categories (i.e., advertising, public relations, sales promotion, online selling, and customer service) with 27 items. The Web site structure typology encompassed five general categories (i.e., text, images, video, audio, and links) with 13 items (Table 3.2). The coding of Web site content was carried out between September and October 2001 (384 sites were

accessible), whereas that of Web site structure was carried out between November and December 2001 (383 sites).

Table 3.1 Coding Categories and Items for Web Site Content

Category	Item
Advertising	Product/service description
	Product catalogue
	Product/service news
	Product package
Public Relations	Company news
	Company history
	Mission statement
	Company organisation
	Financial information
	Affiliated company
	Human resources
	Recruiting
	CEOs (Chief Executive Officers)
	Social activity
	Company profile
Sales Promotion	Contest
	Financial incentives
	Sampling
	Product trial
Online Selling	Transaction process
	Online payment
Customer Service	Branch locations
	After sales service
	FAQs (Frequently Asked Questions)
	Technical information
	Customer feedback
	Contact

Table 3.2 Coding Categories and Items for Web Site Structure

Category	Item
Text	Headline
	Description
Image	Logo
	Photo
Video	Moving text
	Moving image
	Video
	Video icon
Audio	Music
	Sound
	Audio icon
Links	Internal links
	External links

3.2.2 The Sample

(1) Sample Frame

Wright Investors' Service (http://www.corporateinformation.com) was used to locate and examine corporate home pages. This directory has listings of over 22,000 major companies around the world, including 644 Korean and 446 Australian companies at the time of the preliminary study. As the directory excludes foreign companies in each country, local companies (i.e., locally operated and owned) could be compared. Furthermore, most of the companies listed in this directory had Web addresses.

(2) Sampling Procedure and Sample Size

The preliminary study utilised cluster analysis, which involves a large sample with more than 100 units (Malhotra, 1996). To ensure a wide variation in company sizes and types of industries, approximately 200 Web sites were selected from each country with equal numbers of Australian and Korean companies (e.g., Berelson, 1952; Carney, 1972). 408 companies were selected by systematic sampling (every second company was selected from the Australian list and every third from the Korean list). For the companies that had no Web site addresses (URLs) in the *Wright Investors' Service* directory, local search engines were used to try to find their corporate Web sites. The Australian search engines used were *Webwombat* (http://www.webwombat.com.au) and *OzSearch* (http://www.ozsearch.com.au), whereas the Korean search engines were *Simmani* (http://www.simmani.com) and *Yahoo! Korea* (http://kr.yahoo.com). As can be seen in Appendix 3.1, a total of 386 Web site homepages were identified, including 190 Australian (43% of the Australian companies) and 196 Korean Web sites (30% of the Korean companies). All Web sites were treated as one sample in cluster analysis.

(3) Sample Profile

Table 3.3 and 3.4 shows the distribution of company size and industry. First of all, company size was measured by annual sales, which varied from under US $1,000 to over $1 billion. Around half of the Australian firms had annual sales of over $50 million, compared with 83.6 % of the Korean companies (Table 3.3). In terms of industry classification, the *International Standard Industrial Classification (ISIC)* was adopted. More of the firms in the Korean sample were in the manufacturing sectors than

in the Australian sample (68% vs. 14%), and the distribution of firms in each sample broadly reflected the industrial structure in the two countries. That is, the Australian economy was dominated by the service sector while the Korean economy was much more focused on the secondary sector especially manufacturing (Korea Federation of Industries, 2001; The World Bank Group, 2001a and b). Overall, the manufacturing sector was the largest in the sample (41.5%), followed by mining (12.4%) and finance (11.4%). Agriculture (0.8%), hotel (0.5%), education (0.5%), and health (0.2%) showed very low rates. Concerning the two countries, Korean firms were dominantly linked to manufacturing (68.7%). On the contrary, Australian companies were dispersed across various sectors such as mining (24.9%), real estate (18.0%), manufacturing (13.8%), and finance (13.8%). In terms of company size and industry, the sample appeared to be representative of the population of interest.

Table 3. 3 Distribution of Company Size by Country

Annual Sales	Australia[a]		Korea		Total	
	Count	%	Count	%	Count	%
$1,000	19	10.1	0	.0	19	4.9
$500,000	3	1.6	0	.0	3	.8
$1,000,000	4	2.1	1	.5	5	1.3
$2,500,000	9	4.8	0	.0	9	2.3
$5,000,000	8	4.3	0	.0	8	2.1
$10,000,000	18	9.6	3	1.5	21	5.5
$20,000,000	23	12.2	28	14.3	51	13.3
$50,000,000	18	9.6	32	16.3	50	13.0
$100,000,000	53	28.2	81	41.3	134	34.9
$500,000,000	9	4.8	22	11.2	31	8.1
$1,000,000,000	24	12.8	29	14.8	53	13.8
Total	188	100.0[b]	196	100.0[b]	384	100.0

[a]. The annual sales for 2 companies were unknown.
[b]. Rounding error

Table 3.4 Distribution of Industry by Country

ISIC	Australia		Korea		Total	
	Count	%	Count	%	Count	%
Agriculture	2	1.1	1	.5	3	.8
Mining	47	24.7	1	.5	48	12.4
Manufacturing	26	13.7	134	68.4	160	41.5
Electricity	3	1.6	5	2.6	8	2.1
Construction	14	7.4	19	9.7	33	8.5
Wholesale	11	5.8	8	4.1	19	4.9
Hotel	1	.5	1	.5	2	.5
Transport	17	8.9	6	3.1	23	6.0
Finance	26	13.7	18	9.2	44	11.4
Real Estate	34	17.9	1	.5	35	9.1
Education	1	.5	1	.5	2	.5
Health	1	.5	0	.0	1	.2
Community	7	3.7	1	.5	8	2.1
Total	190	100.0	196	100.1	386	100.0*

*. Rounding error

3.2.3 Pretest of the Coding Procedure

After developing a draft coding book to provide a general guide for the coding procedure (Thompsen, 1994), a series of pretests were conducted. The importance of pre-testing has been emphasised by many researchers (e.g., Gorden, 1992; Miles and Huberman, 1994; Perreault and Leigh, 1989; Weber, 1985; Wimmer and Dominick, 1997). Pre-testing allows the researchers to identify and correct unclear definitions and coding rules, thereby enhancing the level of inter-coder agreement.

Inter-coder reliability is defined as the degree to which different coders, coding independently, reach the same coding decisions (Rourke *et al*., 2001). For estimating inter-coder reliability, Thompsen (1994) suggests that at least 10 per cent of the total sample should be analysed by multiple coders, while Wimmer and Dominick (1997)

insist that between 10 % and 25 % should be coded. For this study 40 Web sites were selected for Web site content and 75 sites for Web site structure, accounting for approximately 10 % and 20% of the total sample.

One of the most widely-used measures for assessing inter-coder reliability has been the percentage of agreement between coders. While it is very easy to calculate, this method does not take into account any chance agreement among coders (Harris, Pryor, and Adams, 1997; Hughes and Garrett, 1990). To overcome this shortcoming, researchers have developed new measures, among them Cohen's kappa, which is a measure of agreement among coders after chance agreement has been eliminated (Becker, 1999; Hughes and Garrett, 1990; Rourke *et al*., 2001). The formula for Cohen's kappa is

$$\text{Kappa} = \frac{F_o - F_c}{N - F_c}$$

where Fo is the observed number of agreement, F_c is the expected number of agreement by chance alone, and N is the total number of units coded by each coder. Cohen's kappa values range from 0 to 1, where 1 means perfect agreement and 0 no agreement between coders other than chance agreement. In general, values higher than 0.6 or 0.7 are regarded as acceptable (e.g., Harris *et al*., 1997; Krause, Muller-Benedict, and Wisemann, 2000; Landis and Koch, 1977). Cohen's kappa values were obtained through a Crosstab Analysis (see Table A3.1 to A3.6).

For Web content, coding items were divided into two groups: dichotomous items (presence and absence, 14 items) and three-point scale items (description, category only, and absence, 16 items). The kappa values obtained (0.802 and 0.775, respectively)

demonstrated substantial levels of inter-coder reliability (e.g., Stemler, 2001). On the basis of the pretest, coding rules were slightly modified and one item was eliminated. The finalised coding book for Web site content is shown in Table A2.1.

The kappa value of Web site structure items (all of which were dichotomies) was 0.856, which showed an almost perfect agreement (e.g., Stemler, 2001). The draft coding book was therefore not modified following this pretest. The final coding book for Web site structure is shown in Table A2.2. On the basis of the pretesting, and subsequent revision to the coding books, it can be concluded that the inter-coder reliability of both Web site content and structure were acceptable.

3.2.4 Statistical Analysis

To develop the functional and structural typologies of corporate Web sites, the content analysis results were subjected to a cluster analysis in order to identify meaningful homogeneous groupings of Web sites. The differences between the groups so obtained were tested for statistical significance using the non-parametric Mann-Whitney *U* test. For this test, the significance was set at the 0.05 level for a two-tailed test of significance. All analyses were performed using SPSS for 11.0 Windows.

3.3 Results

3.3.1 Web Site Content

(1) Functions of a Corporate Web Site

In general, corporate Web sites were found to deliver five functions of marketing communication: advertising, public relations, sales promotion, online selling, and customer service. The analysis of home page functions found that while advertising, public relations, and customer service were frequently used, online selling, and sales promotion functions were much less frequently used. As shown in Table 3.5, public relations (97.4% of the home pages sampled used more than 1 item) was the most frequently used function, followed by customer service (94.5%) and advertising (83.6%). In contrast, online selling (21.1%) and sale promotion (13.3%) were still used infrequently. In the category of public relations, there were more individual items than for any other functions; that is, around 60% of corporate Web sites used more than four items.

Table 3.5 Frequency of Web Site Content Items by Function

Function	Number of Items	Frequency		
		Count	Percent	1+
Advertising	0	63	16.4	
	1	223	58.1	83.6%
	2	69	18.0	
	3	27	7.0	
	4	2	.5	
Public Relations	0	10	2.6	
	1	27	7.0	
	2	56	14.6	97.4%
	3	69	18.0	
	4	87	22.7	
	5	62	16.1	
	6	32	8.3	
	7	24	6.3	
	8	13	3.4	
	9	3	.8	
	10	1	.3	
Sales Promotion	0	333	86.7	
	1	39	10.2	
	2	6	1.6	13.3%
	3	6	1.6	
Online Selling	0	303	78.9	
	1	62	16.1	21.1%
	2	19	4.9	
Customer Service	0	21	5.5	
	1	136	35.4	94.5%
	2	127	33.1	
	3	59	15.4	
	4	21	5.5	
	5	15	3.9	
	6	5	1.3	

Table 3.6 shows the use of specific items under each function. Items of customer service, public relations and advertising were the most frequently used, i.e., contact information (84.9%), company profile (81.5%), and company news (71.1%) were the most frequently used items. These were followed by financial information (57.8%), product catalogue (47.9%), customer feedback (42.4%), and product description (39.3%). From these results it can be concluded that a corporate Web site is more of an

integrated marketing communication tool rather than a sole advertising tool or a basic information source.

Table 3.6 Frequency of Web Site Content Items

Function	Item	Frequency (Count/%)	
		Yes	Description[a]
Advertising	1. Product/service description	151(39.3)	20 (5.2)
	2. Product catalogue	184 (47.9)	28 (7.3)
	3. Product/service news	62 (16.1)	34 (8.9)
	4. Product package	53 (13.8)	NA
Public Relations	5. Company news	273 (71.1)	156 (40.6)
	6. Company history	73 (19.0)	3 (0.8)
	7. Mission statement	98 (25.5)	67 (17.4)
	8. Company organisation	112 (29.2)	6 (1.6)
	9. Financial information	222 (57.8)	26 (6.8)
	10. Affiliated company	120 (31.3)	6 (1.6)
	11. Human resources	44 (11.5)	3 (0.8)
	12. Recruiting	121 (31.5)	5 (1.3)
	13. CEOs	79 (20.6)	11 (2.9)
	14. Social activity	50 (13.0)	1 (0.3)
	15. Company profile	313 (81.5)	94 (24.5)
Sales Promotion	16. Contest	30 (7.8)	NA
	17. Financial incentives	18 (4.7)	NA
	18. Sampling	16 (4.2)	NA
	19. Product demonstration	5 (1.3)	NA
Online Selling	20. Transaction process	72 (18.8)	NA
	21. Online payment	28 (7.3)	NA
Customer Service	22. Branch locations	61 (15.9)	NA
	23. After sales service	78 (20.3)	NA
	24. FAQs	47 (12.2)	NA
	25. Technical information	75 (19.5)	NA
	26. Customer feedback	163 (42.4)	NA
	27. Contact	326 (84.9)	NA

[a.] More detailed information
NA = Not applicable. Only presence and absence were counted.

(2) Classification of the Web Site Content

The grouping of the corporate Web sites was performed by cluster analysis, based on 27 Web site content variables. For the cluster analysis all items were coded using a dichotomy; i.e., 1 for presence and 0 for absence of a feature. A number of hierarchical and non-hierarchical clustering methods were explored (e.g., Ward method, average linkage, single linkage, centroid method, and k-means clustering). As this study did not set a predetermined number of clusters, an hierarchical method was preferred (Malhotra *et al*., 1996; Punj and Stewart, 1983). It was found that the hierarchical Ward's procedure produced the more stable and useable cluster solution (see also Klastorin, 1983; Punj and Stewart, 1983), so it was decided to use this method based on the squared Euclidean distance between the variables. A two-cluster solution was selected as most appropriate because, as shown in Figure 3.1, a dendrogram demonstrated a big jump in the distance between clusters (Hair *et al*., 1995; Malhotra *et al*., 1996). The larger cluster included mainly the public relations criterion such as company profile (PRO), contact (CTC), company news (NWS), and financial information (FIN), on the contrary, the smaller group included integrated information such as advertising (i.e., DES, CAT, and PAC), public relations (i.e., HIS, AFF, REC, etc.), sales promotion (i.e., CON, INC, etc.), online selling (i.e., TRA and PAY), and customer service (i.e., FEE, ASS, etc.).

Figure 3.1 Dendrogram Using Ward Method

```
                        Rescaled Distance Cluster Combine

  C A S E      0         5        10        15        20        25
Label     Num  +---------+---------+---------+---------+---------+

INC        17   -+
TRI        19   -+
SAM        18   -+
PAY        21   -+-+
CON        16   -+ I
HUM        11   ---+-+
SOC        14   ---+ +-+
PAC         4   ---+ I I
FAQ        24   ---+-+ +---------+
NEW         3   ---+   I         I
BRA        22   ---+-+ I         I
TEC        25   ---+ +-+         I
TRA        20   ---+-+           +-------------------------------+
AS         23   ---+             I                               I
DES         1   -------+---+     I                               I
FEE        26   -------+   I     I                               I
HIS         6   ---+-+     I     I                               I
CEO        13   ---+ I     +-----+                               I
ORG         8   -----+---+ I                                     I
MIS         7   -----+   +-+                                     I
AFF        10   -------+-+ I                                     I
REC        12   -------+   I                                     I
CAT         2   -----------+                                     I
PRO        15   -----+-+                                         I
CTC        27   -----+ +---------------------------------------+
NWS         5   -------+
FIN         9   -------+
```

The Mann-Whitney *U* test was used to test significance between each pair. As shown in Table 3.7, “type 1” represented 38.5% and “type 2” 61.5 % of the total sample. Type 1 used all the functions and most possible items from each function except product description and company profile. On the other hand, type 2 contained product description and company profile exclusively with couple of common items. Accordingly, type 1 sites likely to have more central features than type 2 sites. For four

items (i.e., company news, financial information, human resources, and contact information) there was no statistically significant difference in usage between the two clusters.

Based on the results of the Mann-Whitney *U* test, type 1 was named the "integrated communication" type, which used more items of all the functions (i.e., advertising, public relations, sales promotion, online selling, and customer service) excluding product/service description and company profile. Type 2 was titled the "basic information" type, which used more items in advertising (i.e., product/service description) and public relations (i.e., company news, financial information, human resources, and company profile from public relations). Company news, financial information, and contact were commonly included in both types. Specific categories and items of each type are as follows:

- **Type 1 = "Integrated communication" type (38.5%)**: This type included advertising (i.e., product catalogue, product/service news, and product package), public relations (i.e., company news, company history, mission statement, company organisation, financial information, affiliated company, human resources, recruiting, CEOs, and social activity), sales promotion (i.e., contest, financial incentives, sampling, and product demonstration), online selling (i.e., transaction process and online payment), and customer service (i.e., branch locations, after sales service, FAQs, technical information, customer feedback, and contact).

- **Type 2 = "Basic information" type (61.5%)**: This type encompassed advertising (i.e., product/service description), public relations (i.e., company news, financial information, human resources, and company profile), and customer service (i.e., contact).

Table 3.7 Result of the Mann-Whitney *U* Test for Web Site Content

Function and Item		Type (yes %)		Mann-Whitney *U* test[a]
		1 (n=148)	2 (n=236)	
Advertising	1. Product/service description	31.3	44.5	**
	2. Product catalogue	69.6	34.3	***
	3. Product/service news	37.8	2.5	***
	4. Product package	20.9	9.3	**
Public Relations	5. Company news	75.7	68.2	NS
	6. Company history	37.8	7.2	***
	7. Mission statement	31.1	22.0	*
	8. Company organisation	37.8	23.7	**
	9. Financial information	59.5	56.8	NS
	10. Affiliated company	41.9	24.6	***
	11. Human resources	9.5	12.7	NS
	12. Recruiting	43.9	23.7	***
	13. CEOs	39.9	8.5	***
	14. Social activity	18.9	9.3	**
	15. Company profile	70.3	88.6	***
Sales Promotion	16. Contest	17.6	1.7	***
	17. Financial incentives	9.5	1.7	***
	18. Sampling	7.4	2.1	*
	19. Product demonstration	3.4	0.0	**
Online Selling	20. Transaction process	31.8	10.6	***
	21. Online payment	14.2	3.0	***
Customer Service	22. Branch locations	25.0	10.2	***
	23. After sales service	28.4	15.3	**
	24. FAQs	23.6	5.1	***
	25. Technical information	28.4	14.0	**
	26. Customer feedback	74.3	22.5	***
	27. Contact	86.5	83.9	NS

[a]. * $p < .05$ ** $p < .01$ *** $p < .001$
NS = No significant difference

3.3.2 Web Site Structure

(1) Structure of a Corporate Web Site

Table 3.8 presents the overall use of the Web design items in corporate home pages. As discussed in the previous chapter, text was defined as written language including alphabets and numbers, expressed in a static visual form. In the study, text included headlines (i.e., short words, slogans, and catch phrases) and descriptive information about a product or a company. An image was defined as non-lingual, static, and visual element of a Web site, including a company logo, photograph, picture, and drawing. Video was defined as a dynamic visual element such as a film and an animated image including animated text. When there was an icon for video elements (e.g., flash, media player, and etc.), it was counted as the presence of video. Audio was defined as any kind of sound including music and other special effects. Audio icons such as speaker, real audio, and so on were recognised as the presence of audio in a Web site. A link was defined as a connection between two or more sets of information in a Web site, and included both internal and external links.

In general, corporate home pages are found to deliver all forms of hypermedia items: text, static images, video, audio, and links. The most commonly used elements were images (around 90 %), and text (around 85%). Animated images and the links were moderately applied (up to 58%). On the other hand, audio was very low in its use.

Table 3.8 Frequency of Web Site Structure Items

Category	Item	Yes (%)
Text	1. Headline	47.8
	2. Description	37.9
Image	3. Logo	92.2
	4. Photo	90.3
Video	5. Moving text	58.0
	6. Moving image	54.8
	7. Video	.5
	8. Video icon	12.8
Audio	9. Music	2.3
	10. Sound	3.7
	11. Audio icon	13.1
Links	12. Internal links	57.2
	13. External links	47.0

(2) Classification of the Web Site Structure

The grouping of the Web site structure type was examined by cluster analysis, based on 13 Web variables. Like the classification of the Web site content, a number of hierarchical and non-hierarchical clustering methods were explored. As there was no predetermined number of clusters, a hierarchical clustering method was selected. It was found that the Average Linkage (within group) method produced the closest and more stable solution. Accordingly, it was decided to use the Average Linkage method based on the squared Euclidean distance between the variables. A two-cluster solution was selected as most appropriate because, as shown in Figure 3.2, a dendrogram showed a big jump in the distance between clusters. The larger cluster included short text (HEA), moving images (i.e., MTE and MIM), static images (i.e., LOG and PHO), and internal links (ILI). In contrast, the smaller group contained descriptive text (i.e., DES), audio (i.e., SIC, MUS, and SOU), video (i.e., VIC and VID), and external links (i.e., ELI).

Figure 3.2 Dendrogram Using Average Linkage (Within Group)

```
                        Rescaled Distance Cluster Combine

  C A S E      0         5        10        15        20        25
Label     Num  +---------+---------+---------+---------+---------+

VIC         8   -+-------+
SIC        11   -+       +-----------+
VID         7   -+-+     I           I
MUS         9   -+ +-----+           +-------+
SOU        10   ---+                 I       +-------------------+
DES         2   ---------------------+       I                   I
ELI        13   ------------------------------+                   I
MTE         5   ---------------------+---------+                 I
MIM         6   ---------------------+         +-------+         I
HEA         1   --------------------------------+       +---------+
LOG         3   -------------+-------------------+     I
PHO         4   -------------+                   +-----+
ILI        12   ----------------------------------+
```

The Mann-Whitney *U* test was used to test significance between each pair. As shown in Table 3.9, “type 1” represented 55.1% and “type 2” 44.9 % of the total sample. While type 1 used all the hypermedia items except descriptive text, type 2 utilised mainly descriptive text. Therefore, type 1 sites likely to have more central features than type 2 sites. For three items (i.e., logos, video, and external links) there was no statistically significant difference in usage between the two clusters. Logos and external links were commonly frequently used, whereas video was commonly little used.

Table 3.9 Result of the Mann-Whitney *U* Test for Web Site Structure

Variable		Type (yes/%)		Mann-Whitney *U* test
		Type1 (n=211)	Type2 (n=172)	
Text	1. Headline	86.7	.0	***
	2. Description	4.7	78.5	***
Image	3. Logo	94.3	89.5	NS
	4. Photo	95.3	84.3	***
Video	5. Moving text	76.8	34.9	***
	6. Moving image	73.9	31.4	***
	7. Video	0.9	.0	NS
	8. Video icon	17.1	7.6	**
Audio	9. Music	4.3	.0	**
	10. Sound	5.7	1.2	*
	11. Audio icon	18.0	7.0	**
Links	12. Internal links	67.3	44.8	***
	13. External links	45.5	48.8	NS

* $p < .05$ ** $p < .01$ *** $p < .001$ NS = No significant difference

Reflecting the results of the Mann-Whitney *U* test, type 1 was named the "hypermedia" type, which includes almost all the hypermedia elements (i.e., text, static images, animated images, audio, and internal links) except some items description (from text) and video (from video). Type 2 was called the "static image" type, which mainly utilises the descriptive text and static images. Logos, photographs and external links were commonly included in both types. Specific categories and items are as follows:

- **Type 1 = "Hypermedia" type (55.1%)**: This type included text (i.e., headline), images (i.e., logos and photos), video (i.e., moving text, moving images, video, and video icons), audio (i.e., music, sound, and audio icons), and links (i.e., internal links and external links).
- **Type 2 = "Static image" type (44.9%)**: This type utilised text (i.e., description), images (i.e., logos and photos), and links (i.e., external links).

3.4 Summary

This chapter described an empirical study to develop a Web site typology in terms of content and structure. To develop the typology, 386 Web site home pages of Australian and Korean companies were selected in the analysis using systematic sampling, which ensured a wide variation in company sizes and types of industries. On the basis of a literature review of Web-mediated marketing studies, a coding frame was prepared. The Web site content typology included five general categories (i.e., advertising, public relations, sales promotion, online selling, and customer service) with 27 items. The Web site structure typology encompassed five general categories (i.e., text, static images, video, audio, and hyperlinks) with 13 items. After developing a draft coding book, a series of pretests were carried out, which allowed slight modification of coding rules.

The content analysis identified two meaningful homogeneous groups (i.e., the "integrated communication" and " basic information" type) of Web site content, which was also confirmed by the Mann-Whitney *U* test. The integrated communication type, which used all the functions and most of possible items from each function except product description and company profile, represented 38.5% of the total sample. The basic information type, which utilised more items in advertising (i.e., product/service description) and public relations (i.e., company news, financial information, human resources, and company profile from public relations), represented 61.5% of the total sample. Company news, financial information, and contact were commonly used in both types.

Web site structure was classified into two types: “hypermedia” and “static image”. The hypermedia type, which included almost all the hypermedia elements (i.e., text, static images, animated images, audio, and internal links) except some items (i.e., descriptive text and video), represented 55.1%. Whereas, the static image type, which mainly used descriptive text and static images, represented 44.9%. Company logos, photos and external links were commonly included in both types.

Based on this typology, four-types of Web site (2 content x 2 structure types) will be created and used to investigate the effects of the Web site typology on communication through social interaction in the second part of the study.