Mining Large Corpora for Social Information: The Case of Elderly
Author(s): Gerlinde Mautner
Source: *Language in Society*, Vol. 36, No. 1 (Feb., 2007), pp. 51-72
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/4169561
Accessed: 19-04-2017 08:41 UTC

Language in Society **36**, 51–72. Printed in the United States of America
DOI: 10.1017/S0047404507070030

Mining large corpora for social information: The case of *elderly*

GERLINDE MAUTNER

Institute for English Business Communication
Vienna University of Economics and Business Administration
(Wirtschaftsuniversität Wien)
Nordbergstrasse 15, 1090 Vienna, Austria
gerlinde.mautner@wu-wien.ac.at

ABSTRACT

Using a large, computerized corpus, this study aims to provide lexicogrammatical evidence of stereotypical constructions of age and aging. It focuses on *elderly*, a word that is pivotal to the domain in question and whose associative meaning is contested. The collocational profile drawn up on the basis of corpus evidence shows that *elderly* is primarily associated with discourses of care, disability, and vulnerability, emerging less as a marker of chronological age than of perceived social consequences. In addition to making a contribution to discourse-oriented aging research, the article also demonstrates the use of corpus linguistic methods within a sociolinguistic framework. (Ageism, ageist language, corpus linguistics.)*

INTRODUCTION: AIM OF THE INQUIRY AND KEY RESEARCH QUESTIONS

In communication about age and aging, *elderly* is a key expression. Clues from various genres suggest that it is also contested. Dictionaries, for example, disagree about its evaluative meaning. Thus, the *Longman Dictionary of Contemporary English* tells its readers that *elderly* "is used as a polite way of saying that someone is old," whereas the *Macmillan English Dictionary for Advanced Learners* advises users that "many people now think that this word is offensive." Trawling the Internet, one also finds its contested nature emerging in self-reported identity, as in the comment made by a 75-year-old, "I am not elderly. I am old and proud of it" (Hay 1998) – an assertion that clearly would not make sense if, as some dictionaries (e.g., the *Concise Oxford*) claim, *elderly* were simply a synonym of *old*. Finally, *elderly* has been picked up as "ageist" by institutional arbiters of politically correct language. The *Publication Manual of the American Psychological Association*, for example, states, "Elderly is not acceptable as a noun and is considered pejorative by some as an adjective" (American Psychological Association 2001:69).

With this initial evidence in mind, and against the backdrop of current research into aging (reviewed in the next section), evidence from a set of large reference corpora was sought to investigate the following questions:

- What is it about *elderly* that makes some users, individual as well as institutional, reject it as "ageist"?
- What attributes and activities are older people commonly associated with, and what dominant roles and identities emerge as a result?

The aim of this article is twofold. On the substantive level of sociolinguistic inquiry, it aims to contribute to the growing body of research on discourse and aging, which will be reviewed, albeit briefly, in the next section. This article's contribution consists in drawing up a collocational profile of an expression that plays a key role in discursive representations of older people. On the reflective level of methodology, on the other hand, the article aims to demonstrate that large reference corpora are rich repositories of social information and thus offer considerable potential for research in sociolinguistics and discourse analysis. In making this case, I shall be drawing on existing, commercially available data and the software that comes with them. I shall focus on how these can be used by scholars who may have neither the resources to assemble large teams of researchers nor the time and computing know-how to develop and use new tools.

DISCOURSE AND AGING

Aging is a field of research spanning many disciplines, reflecting the fact that the process of aging has biological, psychological, and social implications. While it is true that sociolinguistics has been paying comparatively less attention to age than to other factors – enough so for Coupland (2001:185) to refer to age as "sociolinguistics' under-developed social dimension" – there is now a sizable body of research dealing with communicative aspects of aging. The key themes involved include intergenerational communication (Williams & Nussbaum 2001, Williams & Harwood 2004, Barker, Giles & Harwood 2004), and the communicative construction of relationships (Mares & Fitzpatrick 2004), health communication (Thompson, Robinson & Beisecker 2004), as well as age communication in the media (Nussbaum, Pecchioni, Robinson & Thompson 2000:85–112, Robinson, Skill & Turner 2004) and in marketing (DeRenzo & Malley 1992, Balazs 2004, Stern & Mastro 2004). An angle that is of particular interest here is ageism, defined as "a complex ideological formation – a structured, historically formed set of myths or discourses which endorse the subordinate or marginal positions and qualities of the old" (Coupland 2001:197). Butler 1969 is generally credited with coining the term *ageism* (Bytheway & Johnson 1990:27), and Nuessel 1984 with the first attempt at cataloguing its lexicon. Nuessel's collection comprises not only overtly denigrating terms such as *bat* and *codger* but also seemingly neutral expressions such as *generation gap* (focusing, as it does, "on the polarization of young and old people in our youth-oriented society";

Nuessel 1984:21). The range of spuriously neutral lexis, according to Nuessel's 1984 paper, also includes *elderly*, which he calls "stigmatic because this word has been employed by media reporters who have traditionally portrayed this social subdivision in a negative and derisive fashion" (Nuessel 1984:18). However, in a paper published two years earlier (Nuessel 1982), he had still commended *elderly* for its neutral and non-stereotyping qualities. Williams & Giles 1998, on the other hand, point out that older people "are addressed in a depersonalized (often patronizing) manner as members of a negatively valenced social category," and they add, "i.e. as 'elderly'." Both their use of distancing scare quotes around *elderly* and Nuessel's uncertainty point in the same direction as do the primary data quoted in the Introduction: that the evaluative meaning of *elderly* is complex and problematic. This is the issue at the core of this article.

A common thread running through current research on aging is the idea that age and aging are socially and culturally constructed (Hareven 1995:121) – witness, for example, the telling title of Hepworth's 2003 paper, "Aging bodies: Aged by culture." In the case of linguistic work in this field, the contention is, specifically, that age is DISCURSIVELY constructed. The carving up of the lifespan into distinct stages, the way in which these stages are "boundaried" (Williams & Giles 1998:155), and the attribution of value to each stage: All this happens through language and is sustained, negotiated, or challenged in discourse. People are put (and put themselves) into lifespan categories by labels, being referred to, for example, as *young*, *middle-aged, older, old*, *elderly*, or, to mention coinages from gerontology, as *young-old*, *old-old*, or *oldest-old* (Garfein & Herzog 1995). The labeling expressions, in turn, shape both people's identities and interpersonal relationships. In the process, both individuality and non-age-related group identities are backgrounded or forfeited altogether.

If we accept that labeling plays a crucial role in categorization, boundary drawing, and stereotyping, then the study of age labels emerges as a worthwhile contribution to aging research. If we wish to substantiate intuitions and ad hoc evidence about any specific label, we need to turn to representative textual data. Combining both conditions leads us to the corpus-driven investigation reported on in this article.

CORPUS LINGUISTICS

What does it mean to make a research design "corpus-driven"? The key characteristics of corpus linguistics highlighted in recent programmatic statements include (i) its empirical approach and focus on authentic data, (ii) its contextual and functional approach to meaning, and (iii) its use of information technology (IT) to store and process language data (Tognini-Bonelli 2001:2; Teubert 2005). Although these principles can be applied fruitfully to corpora of all sizes, corpus linguistics methodology comes into its own mainly with very large volumes of data.

Corpus linguistics is probably most widely known for its contributions to lexicography and descriptive grammar, informing such groundbreaking projects as the *Collins COBUILD Advanced Learners English Dictionary* (first published in 1987, now in its 4th edition) and the *Longman Grammar of Spoken and Written English*, respectively. However, awareness has been increasing, over the past ten years or so, that corpus linguistic techniques can also be harnessed profitably for uncovering relationships between language and the social. Examples include Stubbs & Gerbig's (1993) inquiry into verbs used in geography textbooks, Krishnamurthy's (1996) comparative analysis of three ideologically loaded words (*racial*, *ethnic*, and *tribal*), Alexander's (1999) corpus-based analysis of ecological issues in business texts, Hardt-Mautner's (1995), Teubert's (2000), and Mautner's (2000) investigations of British "Euro" discourse, Fairclough's (2000) work on the language of Tony Blair's "New Labour" party, Piper's (2000a, 2000b) corpus-based investigation of the discourse of lifelong learning, Orpin's (2005) analysis of words related to corruption (such as *sleaze* and *bribery*), and Baker & McEnery's (2005) study of texts on refugees. Reflections on the dialogue between discourse studies and corpus linguistics can be found in Mautner 2000, 2001, Hunston (2002:109–23), O'Halloran & Coffin 2004, and Koller & Mautner 2004. Mautner 2005, also pursuing the keyword approach, builds what she calls a "discursive profile" of an expression that has a pivotal function in current higher education discourse. What unites all of these studies is the firm belief that connotations are not merely subjective associations that lexical items trigger in individual speakers, but that, like denotation, "connotations are also widely shared within a speech community" (Stubbs 2001:35). Corpora running to tens or even hundreds of millions of words allow researchers to capture such shared connotations in ways that no other method of data gathering can. Specifically, for those working within a discourse analytic paradigm, the value of accessing large computer-held corpora lies in boosting the empirical credence of analyses. This serves to counteract the charge that individual texts have been cherry-picked to suit the researcher's own political agenda (Koller & Mautner 2004:225; Partington 2004b:13; Orpin 2005:38).

Although fairly recent, the projects mentioned above are related, in spirit if not necessarily through explicit attribution, to the early British school of linguistics associated with J. R. Firth, to whom many corpus linguists acknowledge considerable scholarly debt (Sinclair 2003, Stubbs 1996). In his 1935 paper, Firth mapped out an agenda for what he termed "sociological linguistics" (Firth 1935 [1957]:27), calling specifically for "research into the detailed contextual distribution of sociologically important words, what one might call *focal or pivotal* words" (Firth 1935 [1957]:10; original italics). The essentially social (rather than cognitive) orientation of corpus linguistics also surfaces in contemporary explorations of the discipline's theory and method. "Corpus linguistics," Teubert (2005:8) points out, "looks at language not from a psychological, but from a social perspective." In a similar vein, and also following the Firth-

ian tradition, Stubbs 2001 takes up Raymond Williams's (1976) idea of "cultural keywords" and explores case studies that include *heritage*, *care*, and *community*. Keywords, Stubbs argues, are "nodes around which ideological battles are fought" (2001:188). Within the discourse of aging, *elderly* certainly fits the bill.

CORPUS AND METHOD

Corpus size and composition

The primary source of data for this study was the 57-million-word online spin-off from the more than 500-million-word *Bank of English*. *Wordbanks Online*, as it is called (http://www.cobuild.collins.co.uk/Pages/wordbanks.aspx), consists of twelve subcorpora dating from the early to mid 1990s: the British daily newspapers *The Times* (5.76 million words), *The Sun* (5.82 million) and *Today* (5.2 million), British books (5.35 million), magazines (4.9 million), and ephemera such as flyers and mail-order catalogues (3.21 million); unscripted spoken British English (9.27 million) and BBC radio (2.6 million); U.S. books (5.62 million), U.S. ephemera (1.22 million), and National Public Radio broadcasts (3.13 million); and Australian newspapers (5.34 million).

The Wordbanks corpora are part-of-speech tagged, making it possible to search, for example, for cases where the node word *elderly* is coordinated with other adjectives, or for occurrences of *elderly people* followed by a verb form.

Combining quantitative and qualitative approaches

Wordbanks Online comes with its own software, which generates keyword-in-context (KWIC) concordances, calculates absolute and joint frequencies, and computes two measures of relative statistical significance – t-score and Mutual Information (MI) score – for each collocate (based on Church & Hanks 1990). T-score measures certainty of collocation, whereas MI score measures strength of collocation (Hunston 2002:73; McEnery & Wilson 2001:86). T-score directs our attention to high-frequency collocates such as grammatical words (and is thus likely to be more useful to the grammarian or lexicographer than to the sociolinguist or discourse analyst), whereas MI score highlights lexical items that are relatively infrequent by themselves but have a higher-than-random probability of co-occurring with the node word (Clear 1993:281). The two scores are useful, above all, in ranking collocations (Manning & Schütze 1999:166).

The automated statistics generally produce results interesting in their own right, and certainly did in the case of *elderly*, but their more substantive value lies in identifying key leads worth pursuing qualitatively, according to the tried and tested principle of corpus linguistics, "Decide on the 'strongest' pattern and start there" (Sinclair 2003:xvi). Quantitative indicators highlight particularly promising entry points into the data. Moreover, the ensuing qualitative investigation also benefits considerably from the software, with KWIC concordances

making it possible to scrutinize volumes of textual data so large that they would otherwise preclude qualitative analysis.

In the present study of *elderly*, the analysis of concordance lines focused on adjectival and verbal collocates, as well as on the concepts of "semantic preference" and "semantic prosody" (Partington 2004a). Following Stubbs, "semantic preference" is understood as the co-occurrence with "a class of words which share some semantic feature" (Stubbs 2001:88). This concept was invoked, for example, to capture the association of *elderly* with lexis referring to disability, frailty, and care (which may or may not be realized through the actual items *disabled*, *frail*, and *care*, showing that the statistics related to specific collocates, though valuable, are no substitute for qualitative analysis). "Semantic prosody," on the other hand, refers to "the consistent aura of meaning with which a form is imbued by its collocates" (Louw 1993:175; see also Sinclair 1987, 1998; Stubbs 2001; Partington 2004a). This is related to what Hunston (2004:157) refers to as "evaluative meaning," defined as "an indication that something is good or bad," and what Channell (2000:41) calls "evaluative polarity." Through semantic preferences and prosodies, respectively, it is possible to identify both the social domains that lexical items are associated with (Orpin 2005:49) and the societal value judgments they carry. With both factored in, the collocational profile of a keyword is semantically enriched and acquires the potential to contribute to sociolinguistic investigations.

FINDINGS

Quantitative evidence

The investigation began by examining the three collocation lists for *elderly*, ordered by joint frequency, t-score and Mutual Information Score. Among the 50 items with the highest joint frequency (i.e., co-occurring within a span of four words on either side of the node), there were ten lexical words all of which also had high t-scores, of 5 and above (see Table 1).[1]

That *elderly*, an adjective with an almost exclusively [+human] orientation, should co-occur frequently with generic [+human] nouns such as *people* and *couple* is hardly surprising. Among these nouns, though, there is the oddity of *children* featuring in the top ten. Given the semantic incompatibility of *children* and *elderly*, the high joint frequency obviously cannot be due to pre-modification. To clarify matters, *children* was earmarked for closer scrutiny in the second, concordance-based part of the analysis. The remaining three items in the top ten lexical collocates are *disabled*, *care*, and *sick*, which are consistent not only in terms of social domain (and hence semantic preference), but also have a distinctly negative evaluative load,[2] which, to borrow Louw's (1993) phrase, "imbues" *elderly* with negative prosody.

This is further corroborated by the collocation list ordered by MI score, which measures the difference between expected and observed frequencies and hence

TABLE 1. *Lexical collocates of* elderly *with the highest joint frequencies among the top 50 collocates.*

Collocate	Joint frequency	t-score
people	255	15.095201
woman	75	8.373628
disabled	55	7.385917
man	51	6.310005
care	51	6.882761
couple	41	6.148697
children	41	5.661666
lady	37	5.939457
women	35	5.259221
sick	32	5.580488

the strength of the collocational bond. This time, as is usually the case with MI rank orders, the top 50 are all lexical items.[3] To clear the list of chance co-occurrences likely to result in wild-goose chases,[4] it is necessary to disregard items with very low joint frequencies.[5] With items that have joint frequencies of 5 or less deleted, the collocates for *elderly*, ordered by decreasing MI value, are as shown in Table 2.

In addition to the [+ human] nouns already identified by the joint frequency list, there are several new ones. Crucially, they include not only more of the fairly general expressions such as *relatives* and *population*, but also domain-specific nouns that once again reflect semantic preferences clustering around discourses of illness and care, namely *patients* and *homes* (and, as the concordances show, all 14 occurrences of *homes* in the vicinity of *elderly* refer to the residents of care homes). The list of statistically significant adjectival and adverbial collocates has also become longer, and these new additions, too, come from the same (prosodically negative) domains: *infirm*, *frail*, *handicapped*, *mentally*, and *blind*.[6] However, the MI ranking also points to a different kind of vulnerability, with *victims* showing up as one of the stronger collocates (with an MI score of 4.58, a t-score of 3.18 and joint frequency of 11).

At this point, it is appropriate to refine the search and ask whether this collocational behavior is typical of *elderly* in all syntactic roles. Intuitively, one would expect *elderly* to have a comparatively more negative semantic prosody when it functions as a noun phrase head (*the elderly*) than when it serves in an attributive or predicative syntactic role. As far as adjectival collocates are concerned, this is confirmed by the data (see Table 3). When *elderly* is followed by a noun, the list of strongest collocates includes only one negatively loaded adjective,

TABLE 2. *Lexical collocates of* elderly *with the highest Mutual Information scores among the top 50 collocates.*

Collocate	Joint frequency	MI score
infirm	10	10.497061
disabled	55	7.936899
frail	9	7.838440
handicapped	8	6.954731
gentleman	13	6.731002
relatives	20	6.463236
sick	32	6.211547
Britons	6	6.110565
mentally	7	5.813919
widow	7	5.676567
neighbour	9	5.571578
lady	37	5.408101
elderly	10	5.309478
population	18	5.021355
woman	75	4.917658
patients	16	4.908641
couples	7	4.812007
care	51	4.787522
passengers	8	4.761318
blind	7	4.672328
couple	41	4.653912
residents	9	4.610377
homes	14	4.598588
victims	11	4.584622

TABLE 3. *Adjectival collocates with an MI score of at least 3 and a joint frequency of at least 5.*

elderly+NOUN			*the elderly*		
	joint frequency	MI score		joint frequency	MI score
sick	9	6.196509	*infirm*	8	12.224016
			handicapped	6	8.588568
			disabled	17	8.291749
			sick	6	5.845184
			poor	5	4.395648
			young	7	3.226197

whereas *the elderly* has five such adjectives among its statistically significant collocates[7] (apart from *young*, which, as the concordances show, is present for the same reason as *children*, indicating the shared fate of two vulnerable social groups).

The picture that emerges, then, is of *elderly* being strongly associated with lexical items related to disability, care, victimhood and vulnerable social groups. The association is particularly noticeable with the nominalized adjective *the elderly*. With this quantitative evidence as a guide, we can now move on to examine concordances, which show the node in its immediate textual environment.[8]

Qualitative evidence: Refining the collocational profile

Adjectival and nominal collocates: Constructing group identities. The statistics reported in the preceding subsection tell us that adjectives with a negative semantic load from particular semantic domains have a higher than random probability of occurring in the vicinity of *elderly*.[9] The concordances, on the other hand, show us what the nature of this strong association is. Also, they allow us to spot collocates that contribute to the semantic preference and prosody of *elderly*, because they come from the same or related domains and share the same evaluative load as the adjectives identified by the quantitative discovery procedures, but may not be frequent enough individually to pass the threshold of statistical significance. In what follows, I shall be quoting from different concordances for *elderly*, including a random selection of 500 examples (out of the 1,506 total occurrences in the 57-million-word corpus). In addition to the following quotations from this selection, three sample concordances are reproduced in the Appendix.

Focusing first of all on adjectival collocates, we can see from the concordance[10] that *infirm*, *disabled*, *handicapped* and *sick* may occur near *elderly* because the texts talk about the elderly segment of a particular group. For example:

(1)

a class she had for **elderly** disabled people
to provide sitting services for the **elderly** infirm
making arrangements for the care of **elderly** visible minority people

Alternatively, as the following selection of concordance lines illustrates, the co-occurrence may be due to the elderly being mentioned in the same breath as other disadvantaged groups:

(2)

projects to help the sick, **elderly** and handicapped
inevitably the **elderly** and infirm are proving
unattractive customers
customers with special needs such as the **elderly**, disabled and low income
households
social safety nets for the **elderly** and the underprivileged
health care to the **elderly** and the impoverished

When we extend our view from adjectives that are statistically significant to the full set, even more such "lumping together" becomes apparent, as does the negative polarity of many of the labeling expressions involved (including *arthritic, disadvantaged, geriatric*, and *unemployed*). A few new domains emerge, too: lack of mobility, loneliness, and poverty (cf. collocates *housebound*, *lonely*, *poor*). Even among the few positively loaded examples (e.g., *handsome*, *sprightly*, *dear*, *independent*) there are some that turn out to be negative when viewed as part of coherent discourse. Consider, for example, the case of *Myron Taylor, whom Evelyn Waugh described as elderly, handsome, obtuse*, with the positive evaluation conveyed by the second adjective being deflated by the negative one that follows. Similarly, the full concordance citation for *independent* shows that, while its evaluative meaning is certainly positive, the context is once again that of care provision (*It's thanks to Community Services Agency and special people like you that help make our elderly feel independent*[11]), not to mention the slightly patronizing tone of *our elderly* and the fact that the elderly are only made to *feel* (rather than actually become) independent. Other co-occurrences of *elderly* with apparently positive adjectives must also be taken with a grain of salt. If a couple is referred to as *elderly but sprightly*, this may be telling us something positive about said couple, but it is at the same time telling us a lot about the negative expectations that are built into *elderly*. The word *but* says quite clearly that sprightliness is not considered a default quality of the elderly. There is a similar ambivalence about *a dear elderly lady*. While *dear* may in itself be positive, and in some adjective+noun combinations certainly is (as in *a dear friend*), it assumes a rather patronizing quality when teamed up with *elderly* or *old* (cf. *a dear old man*), conjuring up images of gentle, unassuming sweetness and innocence. These are positive stereotypes about older people, but stereotypes nonetheless (Hummert, Garstka, Ryan & Bonnesen 2004:93) and, as such, socially constraining (Williams & Giles 1998:156–57). Try using *elderly lady*, with or without *dear*, for, say, a retired U.S. secretary of state, or indeed *elderly man* for someone like the former chairman of the Federal Reserve Bank, and you will find it does not work, irrespective of both Ms. Albright's and Mr. Greenspan's advanced age.

That *elderly* is predominantly a social rather than a chronological label is also evident from its nominal collocates. In the quantitative part of the analysis (see the preceding section), *residents*, *victims*, and *children* turned out to be among the nouns co-occurring with *elderly* at statistically significant levels. The concordances for *elderly residents*, then, point not only to the issue of institutionalized care (cf. *the charity-funded home for up to 56 elderly residents*), but also explore the theme of vulnerability to crime:

(3)

Teens stabbed in wild street brawl. **Elderly** residents hid behind locked doors
three attacks on **elderly** residents in the city
charged with the murder of two **elderly** residents

This is confirmed by the concordances that feature *elderly* referring to *victims*: of a total of eleven instances, only two refer to disease (*elderly heart attack victims*) or disaster (*Some of the victims, mainly elderly, were trampled and others fell to their deaths*). The remaining nine talk about crime (*some of our elderly have become victims of crime*).

In the case of *children*, the corresponding concordances reveal how co-occurrence and, in particular, coordination constructs shared group identities. Some expressions including *children* have the ring of a stock phrase to them – *women, children and the elderly*, for example, occurs six times in identical form – and of the 41 joint occurrences of *children* and *elderly*, 38 are cases where *elderly* and *children* are part of a coordinated noun phrase, linked by *and*, *or*, or a comma. The concordances for the collocate *children* also draw our attention to the fact that this bracketing together of vulnerable groups relies heavily on the nominalized and collectivized expression *the elderly*. In fact, 34 out of the 41 cases in which *children* collocates significantly with *elderly* feature the adjective in nominal use, confirming evidence that was revealed by the automated procedures reported on in the previous section (cf. Table 3).

Finally, the concordance for *elderly* followed by a noun also includes several instances of *customers* and *investors*, suggesting that there at last we may find some evidence of economic empowerment and independence. Any such hopes are quickly shattered, though. As the complete set of citations for these collocates shows, "elderly" customers are, once again, closely linked with disabled customers and are shown to be lacking basic skills (hiding their money under mattresses and in socks). They are poor (buying cigarettes by the piece), victims of accidents and beneficiaries of charitable work. Elderly investors do not fare much better. They suffer financial hardship, have to be protected against repossession, and are duped into making dubious investments:

(4)

```
                    Our free guide for DISABLED or elderly customers
Help the Aged on projects that further benefit our elderly customers
   Accidents. Most victims are children, staff and elderly customers
          he's selling single cigarettes to some elderly customers for
                                        35 cents.
 in a sock, under your mattress, as some of our elderly customers do.
           the financial hardship of the mostly elderly investors who
                        purchased junk bonds
                              We want to see elderly investors restored
                 to the position they were in
                              a deal to help elderly investors who were
                   mis-sold home income plans
              regulator said the agreement that elderly investors will not
                              be repossessed
```

Having established what qualities and social groups the elderly are generally associated with, let us now move on to activities.

Verbal collocates: What the elderly do. What, then, are elderly people seen to be doing in the corpus? Are they represented as shaping their own lives? What

we are looking for primarily, following a systemic-functional approach (Halliday 2004), are cases where *elderly* is part of a noun group occupying the role of "Actor" in a "material process." This is not the only way in which agency can be expressed (van Leeuwen 1996:33), but it is certainly an important one, and one that can be accessed reasonably well through concordance output. Material clauses "express the notion that some entity 'does' something" (Halliday 2004:181), and the Actor "is the one that does the deed" (Halliday 2004:179). There is no query in *Wordbanks Online* that will yield up this information ready-made, but it is possible to extract a suitably focused selection of examples from the corpus by searching for *elderly* followed by any type of verb, with up to two items in between.[12] Having established that stereotyping is particularly noticeable when *elderly* is used nominally and collectively (see the preceding subsection), it made sense to run a separate search for this pattern, in addition to that capturing attributive adjectival use.

It turns out that, in a random 100-line sample of *elderly* as an attributive adjective, just over a quarter of instances (26) include a material process with an elderly Actor. Among those, the activities described vary in terms of how much conscious shaping of events they actually involve and in how far they reflect stereotypical role models. Citations such as the following are at the most active end of the spectrum:

(5)

a group of **elderly** people who took regular exercise
the **elderly** shareholder took directors to task

Others are closely related to the semantic preferences identified in previous sections:

(6)

an **elderly** resident would shuffle out
elderly people who have retired

When *the elderly* is used collectively,[13] the proportion of cases where they feature as the Actor in material processes drops to 9% of the randomly sampled citations. Encouragingly, these cases include examples such as the following:

(7)

The elderly eat better than the young
the elderly, who come for education courses

On the other hand, there are citations that clearly tie in with the negative semantic prosody that is projected onto *the elderly* by its adjectival collocates (see the subsection on quantitative evidence):

(8)

where **the elderly** meet to sing and play dominoes
the elderly who tend to approach the district nurses usually

The remaining hope for the analyst looking for positive constructions of aging was that elderly people might be cast in the role of agents in passive construc-

tions, as the complement of *by*, a pattern not included in the query described above. More disappointment was in store. Out of a total of 609 occurrences of *elderly* preceded by a preposition, *by* accounts for a mere 3.9%. *For* and *of* make up around 30% each, and a variety of other prepositions the remaining 30%.

Perhaps unexpectedly, even these instances of the most common prepositions, such as *of*, can contribute to the discursive profile of *elderly*, and the social reality reflected and constructed by it. Among the noun phrases including *of*, one notices clusters of *care of* and *carers of*, as well as of *number of* together with *rising* and *growing*. This mirrors society's preoccupations in dealing with "the elderly," reflecting the fact "that the aging population is 'problematized' (that is, brought into prominence as a dilemma)" (Bailey & Corner 2003:177, referring to Katz 1996). The evidence is very similar for phrases including *for*, with the high incidence of *care for*, *centre for*, *home for*, and *clubs for* indicating what the priorities are. Examples conveying a different image – such as *the power of the elderly lobby* – do exist, but they remain isolated and are drowned out by representations that are negative and/or focus on institutionalized care schemes for people who have lost (or are being denied) their independence.

CONCLUSION

Elderly and the discourse of aging: Summary and contextualization of results

Let me return to the research questions posed in the third section and summarize how the corpus evidence suggests they should be answered.

First, the ageist quality of *elderly* (and thus its rejection by individual users and lobbying groups) can be traced back to its collocational profile. Dominant semantic preferences include disability, illness, care, and vulnerability to crime, with the associated semantic prosodies often negative. We look in vain for the lexis of independence, initiative, and empowerment. Other blind spots are education, jobs, experience and the potential arising from these.

Second, at various points in the investigation – in connection with adjectival collocation, for example, and in the study of agency – evidence emerged that the negative polarity of *elderly* is particularly strong when the adjective is used as a nominalized, collective expression (*the elderly*).

Third, the elderly are portrayed as a predominantly passive group, on the receiving end of care as well as of crime. Rarely are they seen to be actively shaping their own lives. In combination with the dominant semantic preferences mentioned above, a picture emerges of *the elderly* being conceptualized as dependent and vulnerable.

The associative meaning that is bestowed on *elderly* through its semantic preferences and prosody takes us from the mere facts of chronological age to the social sphere, in which age is used as the basis for a whole chain of stereotypical assumptions about individuals' attributes, needs, and life chances. Who is described as *elderly* depends less on the date given on their birth certificate than on

whether they are perceived to fit the stereotype. We can conclude from corpus evidence that someone who is old and fit, old and wealthy, or old and powerful, or a combination of all these, is much less likely to be referred to as *elderly* than are those who are old and disabled, old and poor, or old and dependent (or, as is quite likely, a combination of all these). Education and money offer no protection against becoming old, but they do against being made to become elderly.

Returning briefly to the quotations mentioned in the introduction, one has to reject the notion, upheld by nearly all current dictionaries, that *elderly* can be accurately glossed simply as "old or aging" (*Concise Oxford English Dictionary* 2004). Its associative meaning clearly ranges much more widely. The evaluative semantic load it carries is predominantly negative, which explains why even people who ARE old or aging do not wish to be referred to as *elderly*, and why organizations advocating non-ageist language have largely abandoned *elderly* in favor of *older*. Thus, while *elderly* may have been a euphemism for *old* once, there are definitely no grounds now for describing it as "a polite way of saying old" (*Longman Dictionary of Contemporary English* 2003). Instead, dictionaries ought to follow Macmillan's lead in including a warning that the word is now considered offensive by many and should be avoided (*Macmillan English Dictionary for Advanced Learners* 2002). The findings support the policy of those publications and organizations already avoiding and advising against the use of *elderly*, and should encourage those that do not yet have such a policy to follow suit. Nonetheless, a caveat is in order. Important though the promotion of non-ageist language is, the social issue at the heart of this article is the way in which older people are negatively stereotyped, and for THAT to change it takes more than a few guidelines on non-offensive lexical choices. Certainly, given the dialectic between language and social life, ageist language not only reflects ageism in society but also contributes to promoting it. Yet we need to remind ourselves that ageism cannot be eradicated by non-ageist language alone. As other "isms" have shown, adding another word to the blacklist of politically incorrect usage, important though that is, does not automatically change mindsets and policies, or improve people's lives.

Finally, two qualifying remarks. First, many older people evidently do require and should receive care, are disabled or poor, or become victims of crime. This cannot be denied, nor should public and voluntary sector programs intended to alleviate the ensuing disadvantage and suffering be belittled. The problem is, rather, that other types of activity – those that focus, for example, on empowerment, older people's own initiative, and their (re)integration into the job market – are comparatively underrepresented.

On the other hand, and this is the second qualification, some skepticism is warranted about alternative discourses of "successful aging," centering on what Katz & Marshall (2003:4) refer to as "a culture of obsessive self-improvement." Ideologies built on the pathologization of normal aging – and the concomitant creation of pressures on individuals to "do something about it" – should not be

allowed to become hegemonic either. The normative goal should be policies and discourses that recognize diversity among the older generation, assuming neither that all older people are frail and in need of care, nor that they are all fit and independent.

Using corpus methodology: A critique

In keeping with this article's two-pronged aim, thematic as well as methodological, some concluding reflections on methodology are also in order. I have made a case for tracing the discursive construction of a particular group of people by accessing a large, computer-held corpus, and combining such quantitative and qualitative approaches as the software allowed.

Like all other research designs, using large corpora and concordance programs opens up specific potential but also has limitations. To some extent, the latter are the flip side of the former. On the plus side, there is representativeness. Millions of words from diverse genres allow one to be reasonably confident about generalizability, especially when the item under investigation, as is the case with *elderly*, has a high total frequency in the corpus concerned. What large-scale data are not well suited for, on the other hand, is making direct, text-by-text links between the linguistic evidence and the contextual framework it is embedded in. In *Wordbanks Online*, the corpus used here, there is some source information for each occurrence, but by sociolinguistic and discourse analytic standards this is rather sparse. For the newspaper data, for example, we are given the name of the paper and the date of publication, but there is no information on which journalistic subgenre (e.g., an editorial or a news feature) each extract comes from. Also, as *Wordbanks* and similar corpora work with text-format only, the accompanying visuals are lost. In the "Spoken" section of the corpus, speakers' gender is identified, but that is the only demographic or situational information that is given. We are not dealing with complete interactions, but with a data aggregate that features only extracts from situated talk. The maximum stretch of text that can be viewed is 512 characters. In the vast majority of cases this provides enough co-text for the analyst to work out the meaning of the search word in each instance, to identify the semantic "aura" around the word (Louw 1993), and to pinpoint the many instances of contestation that do in fact become manifest within a relatively confined collocational space. Still, the cap on accessible co-text means that we cannot trace how meaning unfolds over longer stretches of discourse, or how meaning is negotiated between interactants. To unravel discursive patterning, one has to go elsewhere for data and tackle them with different tools. However, like other approaches, corpus linguistic methods should not be blamed for not being able to do what they were never meant to do in the first place.

While I believe that findings such as those reported in this paper are in themselves valuable, the type of approach illustrated here will not come into its own unless it is integrated into a research design that combines the collocational analy-

sis of a data aggregate with the in-depth analysis of a small sample of longer stretches of coherent discourse. In a more comprehensive project, the idea would be to oscillate between observations derived from longer, complete texts on the one hand and the search for patterns across large reference corpora on the other (Mautner 2000:104–6), with corpus data playing a crucial role in providing a safeguard against "overinterpretation" and "underinterpretation" (O'Halloran & Coffin 2004). Evidence gleaned from reference corpora becomes the foil against which usage in individual texts can be judged. As Sinclair (2004:29) puts it, "When a reliable description of the regularities has been assembled, then individual texts can be read against it, and at that time the individual instance will make a balanced impact by comparison with the norms."

Essentially, the analysis can move in two directions, from quantitative to qualitative, or the other way round. In practice, one is likely to want to go back and forth several times during an investigation. One can start from situated text and talk, and then put questions about norms to the reference corpus, or begin the iterative process at the other end, interrogating the corpus first and then using the results of automated discovery procedures as clues to guide the qualitative analysis.

In either case, it is important to note that keyword-in-context concordances provide a strong link between quantification and interpretation, allowing the researcher to see beyond phenomena identified as significant by computational statistics, and yet do so on the basis of a very large number of citations. Effectively, what the method outlined here does is mediate between a macro and a micro approach to language, allowing access to data aggregates as well as close attention to detail. Crucially, the corpora and the accompanying software enable individual researchers to combine these approaches singlehandedly, without recourse to additional manpower, which may be hard to obtain in research environments increasingly squeezed for funding. There is no miracle cure for the perennial tension between macro and micro levels of analysis, and between quantitative and qualitative procedures, but enriching projects in sociolinguistics and discourse analysis with corpus methodology can go some way toward turning this tension into fruitful dialogue.

APPENDIX: SAMPLE CONCORDANCES FOR *ELDERLY*

1. Random selection of 50 concordance citations of *elderly* followed by an adjective, with up to one word in between.

people I work with most of them are	**elderly**	⟨F01⟩ Right. What are their
from prescribing such drugs to the	**elderly**.	⟨p⟩ Older people have enough
Laura, about west of Cairns, has 16	**elderly**	and long-time Cape living in
⟨p⟩ In The Sitters, Miller's	**elderly**	and unnamed artist-narrator meets
⟨p⟩ Offers support and advice to the	**elderly**	and housebound in the Govanhill
This is particularly useful for the	**elderly**	and disabled. ⟨p⟩ Low-energy light
⟨p⟩ The Mission Dine Club for the	**Elderly**	and Disabled invites you to join
needs of the sick, children, the	**elderly**	and pregnant women. ⟨p⟩ While
hospital bills for about 34 million	**elderly**	and disabled Social Security
is long-term health care for the	**elderly**	and disabled. President Clinton
such as Counsel and Care for the	**Elderly**	and Elderly Accommodation Counsel
designed to help the low-income	**elderly**	and disabled maintain their own

areas and cannot be reached by the **elderly** and disabled. ⟨p⟩ Grocer and
and the MURDER of infants and the **elderly** are acceptable behavior 44 ⟨p⟩
was top of the list. ⟨p⟩ Even the **elderly** are likely to be hit as ministers
hospitality of its people. The **elderly** Armenian proprietor of Baron's
are cutting back on key services for **elderly**, blind and disabled gas customers.
of care centers for thousands of **elderly** blind and disabled residents.
particularly women, children and the **elderly** - feel safer and more secure in
people like you that help make our **elderly** feel independent. "⟨p⟩ Daughter of
rise 'manager from hell' by the **elderly** Gold Coast unit owners he stole
a writer. ⟨p⟩ Then the death of an **elderly** local businessman and a discovery
all the cliches about vicars and **elderly** maiden aunts throwing caution to
themselves. Many of these owners are **elderly**, many housebound, others disabled.
the myth of sexual apathy among the **elderly**: most older adults continue to be
disease hits again NEW YORK: An **elderly** New York woman has died of the so-
ending of tax relief for the **elderly** on private health insurance.
within a community such as women, **elderly** or young people. ⟨p⟩ Successful
toughest roots, yet light enough for **elderly** or arthritic hands to use. The
such as travel passes for the **elderly**, or disabled parking badges, will
your needs it can also lead to the **elderly** or young mums having ⟨24⟩ no
cause any problems, for example with **elderly** or sick residents. Even at this
for vulnerable groups such as the **elderly** or disabled, who cannot easily use
of confidence in our efforts to keep **elderly** people warm and comfortable is
Co-ordinator, told Asian Times that **elderly** people, lonely and isolated in
of torture. It also notes that four **elderly** political prisoners were executed
the vandalism was ordered by that **elderly** public-maintained ornament, Prince
at home full-time for a disabled or **elderly** relative. Many have given up work
had grown up with-pleading with an **elderly**, robed, bearded, white male figure
safety breaches were putting the **elderly**, sick and those with impaired
discourage the monsters who prey on **elderly** single women who live alone. It
saw things for the handicapped, the **elderly**, special rooms for kids They had
The vast majority of suicides of **elderly**, terminally ill, or even disabled
benefiting groups such as the **elderly**, the unemployed, children and
it was giving priority to the **elderly**, the disabled and families with no
could be spent on looking after the **elderly**, the sick and the infirm. It could
properly to mind: he had been too **elderly**, too amiable, too vague to be
sith brightlly coloured rucksacks, **elderly** Tory gentry and taxi drivers. They
making arrangements for the care of **elderly** visible minority people by
at all ⟨M01⟩ No. ⟨F01⟩ she was an **elderly** white-haired woman from the north

2. Random selection of 50 concordances of *elderly* followed by a noun (singular or plural) and a verb (any form), with up to two words in between

so it was about 2.30am before the **elderly** 386dx was turned off and we both
to keep the oldsters mobile. The **elderly** athlete says all those old star
to wrong side. ⟨p⟩ October 26. ⟨p⟩ **Elderly** Brisbane couple killed when
as they thought. "Last week, two **elderly** brothers were beaten with hammers
Fifties Focus, representing 60,000 **elderly** citizens, said a jump in jobless
yesterday. ⟨p⟩ ⟨b⟩ SMITH B ⟨/b⟩ **Elderly** couple hit by raid AN elderly
of the nearby town of Pakrac, an **elderly** couple is trying to clean up the
of-nowhere location. ⟨p⟩ Like the **elderly** couple who went to Poland and
Most victims are children, staff and **elderly** customers, hit in the legs or back
You owe Rhodry and Yraen coin. "The **elderly** dwarf howled and turned his face
dressed as lamb, in other words **elderly** ewes worn out by breeding passed
aided by a valuable bequest from an **elderly** friend, sets about enjoying her
The author recently chatted with two **elderly** gentlemen who appeared the
the financial hardship of the mostly **elderly** investors who purchased junk bonds
refuse invitations unless a dear **elderly** lady called Muriel Gunn will have
and now he can see them again. "An **elderly** lady examined by Harry's wife
in shiny black jacket, the other an **elderly** lady wearing a green raincoat.
memory. ⟨p⟩ ⟨b⟩ SMITH B ⟨/b⟩ **Elderly** man murdered AN elderly man was
is about the national average. ⟨p⟩ **ELDERLY** MAN playing dice Two ones and a
fall;Jim Armfield ⟨/h⟩ ⟨p⟩ AN **ELDERLY** man died after falling off a
⟨/h1⟩ ⟨dt⟩ 25 April 1998 ⟨/dt⟩ AN **ELDERLY** man is fighting for his life after
government health-care plan for the **elderly**, Medicare, and cut education and
weather sometimes contrasts with **elderly** men standing for long periods with
that slowly crushed her spine, our **elderly** mother needed to admit him to the
with the house - did we mind? An **elderly** neighbor had died, leaving the cat
categories, about fifty are sick or **elderly**, others include those caught on
that sons and daughters would rather **elderly** parents spent money on themselves
beneficial and cost-effective for **elderly** patients. The scientists reported
sulpha-based antiobiotics to **elderly** people were working. ⟨p⟩ Although
⟨p⟩ What worries me is the number of **elderly** people who have turned up at my

might be dangerous for children and **elderly** people. ⟨p⟩ Use lower wattage
as a neighbourhood resource for **elderly** people, whilst seeking to ensure
er you were saying about you know er **elderly** people kissing and cuddling what
spend a bit more time talking to the **elderly** people. I think ⟨ZF1⟩ that ⟨ZF0⟩
supporters including women and **elderly** people, pelting them with stones
Journal ⟨f⟩ showed that a group of **elderly** people who took regular exercise
found frozen on her doorstep. ⟨p⟩ **Elderly** people were warned to stay indoors
of theft but of violence against **elderly** people. ⟨p⟩ That made Watson's
⟨p⟩ It's somewhere for a single **elderly** person to call home. ⟨p⟩ PLASTIC
some blocks that had a high **elderly** population that still didn't have
⟨p⟩ BRANCH CLOSURE HITS BONUS ⟨p⟩ AN **elderly** reader who prefers not to be named
If one waited for long enough, an **elderly** resident would shuffle out and
in Brisbane yesterday, with one **elderly** shareholder determined to have his
discharged and then the team the **elderly** team go in they do an assessment.
they can be.' We have guns," an **elderly** villager announced, proudly
squaddie who savagely attacked an **elderly** widow escaped jail yesterday when
79, to death By SEAN PARNELL AN **elderly** woman was mauled to death by a
not to stop being afraid to ask." An **elderly** woman says, 'Age puzzles me. I
Artillery shell doorstop ⟨/h⟩ ⟨p⟩ AN **elderly** woman who used an old artillery
in 36 Mardi Gra attacks. And an **elderly** woman died months after she was

3. Random selection of 50 concordances of *the elderly* used collectively

other groups like children for **the elderly** ⟨F02⟩ ⟨tc text=coughs⟩ ⟨F01⟩ erm
have erm the task of caring for **the elderly** ⟨ZF1⟩ and ⟨ZF0⟩ and nobody quite
from prescribing such drugs to **the elderly**. ⟨p⟩ Older people have enough
but can also affect children and **the elderly**. ⟨p⟩ The disease commonly attacks
healthy to long-term care for **the elderly**. ⟨p⟩ What the hospitals want fits
depression particularly in **the elderly**. ⟨p⟩ The psychological benefits
through Leasehold Schemes for **the elderly**, a system under which homes have
running housing charities for **the elderly**, Abbeyfield, cares for the older
heart strong. It also protects **the elderly** against the brittle bone disease
from the newborn baby to **the elderly** and frail. Whether you are
spend an awful lot of money with **the elderly** and never get near any other
at Asian people's houses and **the elderly** and they p they're ⟨ZGY⟩ if they
dry season. Then they trained **the elderly** and the handicapped to make
the roof of a high-rise home for **the elderly** and handicapped. People closed
State officials, advocates for **the elderly** and industry representatives seem
of social safety nets for **the elderly** and the underprivileged. A
restrictions on providing for **the elderly** and disabled. Mike Dawson,
deep concern including abuse of **the elderly** and hospital waiting lists for
said services were failing **the elderly** and mentally ill and more
It has been well documented that **the elderly** are particularly susceptible tothe
care. But rising numbers of **the elderly** are bound to claim a bigger share
it's little wonder that **the elderly** are seen as sexless. (After all,
of dependent-care centers for **the elderly**, blind and disabled. It's
with post-menopausal women and **the elderly** but is now recognized to occur in
Robinson. ⟨p⟩ Moving houses for **the elderly** can be very traumatic," Ms Wallace
East London: 'The interests of **the elderly** come first ⟨p⟩ Nathan admitted at
⟨F01⟩ Er they've a nice way with **the elderly** er not afraid to stop and chat and
can do erm like the care of **the elderly** erm and the care of children erm
leading to considerable anxiety. **The elderly**, for their part, need established
to the old. ⟨p⟩ Concern for **the elderly** has resulted in an increasing
man's home in accommodation for **the elderly** in Dublin's Arbour Hill area, then
federal spending programs on **the elderly**, Medicare, and the poor on
Kingston like this one, where **the elderly** meet to sing and play dominoes.
wireless intercom is ideal for **the elderly** or infirm, those who live alone
your needs it can also lead to **the elderly** or young mums having ⟨24⟩ no
for vulnerable groups such as **the elderly** or disabled, who cannot easily use
does. ⟨p⟩ ⟨o⟩ Psychiatry of **the elderly** (psychogeriatrics) deals with
stagnant housing market has hit **the elderly**, says Sara mcconnell ⟨p⟩ Growing
Mm ⟨F01⟩ I enjoy working with **the elderly** so whilst MX is at school I now go
something about nutrition and **the elderly**. So they looked at erm shopping
er physician with an interest in **the elderly** that initially set up the six beds
for the harmless flirtations of **the elderly** that I suspect he went in for when
Council, he Youth Services, **the elderly**, the Council for Voluntary Service
nurse Linda Bishop has cared for **the elderly**, the housebound and the terminally
in particular, the care of **the elderly**. Those involved in the
how an insulation project with **the elderly** was managed - project reports. ⟨p⟩
the cost of long-term care for **the elderly** weighs heavily on every one of us.
⟨p⟩ The women, children, and **the elderly** were forcibly bused to the town of
number of people and the old and **the elderly** when I see a man in a wheelchair
P ⟨F0X⟩ Mhm. ⟨M04⟩ and there is **the elderly** who tend to approach the district

NOTES

* I am grateful to several colleagues for their comments on a draft version, including most notably Susan Hunston, Virpi Ylänne-McEwen, Kieran O'Halloran, Angie Williams, and two anonymous reviewers. Veronika Koller contributed to the initial corpus searches and also provided useful feedback on the first draft. Eva Kerbler's support during the revision phase was invaluable and much appreciated. I am indebted to Wirtschaftsuniversität Wien and the Austrian National Bank (OeNB) for financial support while spending a sabbatical at the Centre for Language and Communication Research of Cardiff University. Material from the *Bank of English*® reproduced with the kind permission of HarperCollins Publishers Ltd.

[1] In studies using the *Bank of English*, the cutoff point above which t-scores are assumed to signal statistical significance is reported to be 2 (Hunston 2002:72).

[2] For the record, I ought to stress that the association of "elderliness" with disability is not only ageist but also "able-ist." To young people with a disability, who may share some but by no means all needs with the elderly, it must surely be objectionable to have their lifespan discursively condensed by routine coupling with the elderly (not to mention the factual inadequacy of many joint-provision care schemes).

[3] This is so because grammatical words combine so freely with many different items that even when there is a statistically significant likelihood of their co-occurrence with the node, it does not follow that the association is also strong (i.e., unexpected) in a statistically significant way (cf. Hunston 2002:74).

[4] For example, the collocate with the tenth highest MI score is *deng*, owing to *elderly* being used as an epithet for Chinese leader Deng Xiao Ping. Its joint frequency, however, is only 3.

[5] What "very low" means in this context is essentially a matter of judgment, but suggestions in the literature range from 3 to 5 as suitable cutoff points (Stubbs 1995:40).

[6] These collocates all have MI scores of 4.5 or more and are thus well above the threshold of 3, which is generally taken to indicate significance (Hunston 2002:71).

[7] The statistics for *elderly*+NOUN were calculated on the basis of a selection of 450 occurrences (out of a total of 950), from which all stray collective uses ("false positives" resulting from errors in the tagging software, such as *Did the elderly cause this deficit?*) had been manually deleted. Similarly, the concordances for *the elderly* (again a selection of 450, out of a total of 475) had to be manually scanned and non-collective uses (e.g. *the elderly owner had died*) eliminated before compiling the collocation list.

[8] *Wordbanks* offers between 80 and 512 characters of context.

[9] These findings were also confirmed by comparative data from the British National Corpus (BNC), using "Word Sketch," a tool that provides "one-page automatic, corpus-based summaries of a word's grammatical and collocational behaviour" (Kilgarriff, Rychly, Smrz & Tugwell 2004:1). In the BNC, too, the ten adjectives linked most frequently with *elderly* through *and/or* were *disabled, frail, infirm, vulnerable, sick, ill, handicapped, housebound, confused*, and *dependent*.

[10] The search command used was one that filtered out all occurrences of *elderly* followed by an adjective and with up to one word in between (to allow for cases of coordination with conjunctions). A 50-line selection (out of the total 150) is reproduced as Figure 1 in the Appendix.

[11] This comes from the spoken section of the *Wordbanks* corpus, hence the idiosyncratic syntax.

[12] This allows for simple and compound tenses, as well as for the presence of "mid-position" adverbs. See Figure 2 in the Appendix for a sample of 50 concordance lines with this pattern.

[13] For a sample of 50 concordance lines exhibiting this pattern see Figure 3 in the Appendix.

REFERENCES

Alexander, Richard J. (1999). Ecological commitment in business: A computer-corpus-based critical discourse analysis. In Jef Verschueren (ed.), *Language and ideology: Selected papers from the 6th International Pragmatics Conference* 1:14–24. Antwerp: International Pragmatics Association.

American Psychological Association (2001). *Publication manual of the American Psychological Association*. 5th ed. Washington, DC: APA.

Bailey, Chris, & Corner, Jessica (2003). Care and the older person with cancer. *European Journal of Cancer Care* 12(2):176–182.

Baker, Paul, & McEnery, Tony (2005). A corpus-based approach to discourses of refugees in UN and newspaper texts. *Journal of Language and Politics* 4(2):197–226.
Balazs, Anne L. (2004). Marketing to older adults. In Nussbaum & Coupland (eds.), 329–51.
Barker, Valerie; Giles, Howard; & Harwood, Jake (2004). Inter- and intragroup perspectives on intergenerational communication. In Nussbaum & Coupland (eds.), 139–65.
Butler, Robert N. (1969). Age-ism: Another form of bigotry. *Gerontologist* 9:243–46.
Bytheway, Bill, & Johnson, Julia (1990). On defining ageism. *Critical Social Policy* 10(2):27–39.
Channell, Joanna (2000). Corpus-based analysis of evaluative lexis. In Susan Hunston & Geoff Thompson (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 38–55. Oxford: Oxford University Press.
Church, Kenneth W., & Hanks, Patrick (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics* 16(1):22–29.
Clear, Jeremy (1993). From Firth principles: Computational tools for the study of collocation. In Mona Baker et al. (eds.), *Text and technology: In honour of John Sinclair*, 271–92. Amsterdam & Philadelphia: John Benjamins.
Collins COBUILD Advanced Learners English Dictionary (2004). 4th ed. London: Collins COBUILD.
Concise Oxford English Dictionary (2004). 11th ed. Oxford: Oxford University Press.
Coupland, Nikolas (2001). Age in social and sociolinguistic theory. In Nikolas Coupland et al. (eds.), *Sociolinguistics and social theory*, 185–211. Harlow, UK: Longman.
DeRenzo, Evan G., & Malley, Jim (1992). Increasing use of ageist language in skin-care product advertising: 1969 through 1988. *Journal of Women and Aging* 4(3):105–26.
Fairclough, Norman (2000). *New Labour, new language?* London: Routledge.
Firth, John R. (1935 [1957]). The technique of semantics. In his *Papers in Linguistics 1934–1951*, 7–33. London: Oxford University Press.
_____ (1957 [1968]). A synopsis of linguistic theory, 1930–1955. In Frank R. Palmer (ed.), *Selected papers of J. R. Firth 1952–1959*, 168–205. London: Longman.
Garfein, A. J., & Herzog, A. Regula (1995). Robust aging among the young-old, old-old, and oldest-old. *Journals of Gerontology ser. B: Psychological Sciences and Social Sciences* 50(2): S77–S87.
Halliday, M. A. K. (2004). *An introduction to functional grammar*. 3rd ed. London: Arnold.
Hardt-Mautner, Gerlinde (1995). Only connect: Critical discourse analysis and corpus linguistics. http://www.comp.lancs.ac.uk/computing/research/ucrel/papers/techpaper/vol6.pdf.
Hareven, Tamara K. (1995). Changing images of aging and the social construction of the life course. In Mike Featherstone & Andrew Wernick (eds.), *Images of aging: Cultural representations of later life*, 119–34. London & New York: Routledge.
Hay, Henrietta (1998). Here come the woopies! http://www.henriettahay.com/modern/jul1098.htm. [Accessed 30 Nov. 2005].
Hepworth, Mike (2003). Aging bodies: Aged by culture. In Justine Coupland & Richard Gwyn (eds.), *Discourse, the body, and identity*, 89–106. Basingstoke: Palgrave Macmillan.
Hummert, Mary L.; Garstka, Teri A.; Ryan, Ellen B.; & Bonnesen, Jaye L. (2004). The role of age stereotypes in interpersonal communication. In Nussbaum & Coupland (eds.), 91–114.
Hunston, Susan (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
_____ (2004). Counting the uncountable: Problems of identifying evaluation in a text and in a corpus. In Alan Partington et al. (eds.), *Corpora and discourse*, 157–88. Bern: Peter Lang.
Katz, Stephen (1996). *Disciplining old age: The formation of gerontological knowledge*. Charlottesville: University Press of Virginia.
_____, & Marshall, Barbara (2003). New sex for old: Lifestyle, consumerism, and the ethics of aging well. *Journal of Aging Studies* 17:3–16.
Kilgarriff, Adam; Rychly, Pavel; Smrz, Pavel; & Tugwell, David (2004). The sketch engine. http://www.sketchengine.co.uk/ [Accessed 2 Oct. 2004].
Koller, Veronika, & Mautner, Gerlinde (2004). Computer applications in Critical Discourse Analysis. In Ann Hewings et al. (eds.), *Applying English grammar*, 216–28. London: Arnold.
Krishnamurthy, Ramesh (1996). Ethnic, racial and tribal: The language of racism? In Carmen Rosa Caldas-Coulthard & Malcolm Coulthard (eds.), *Texts and practices: Readings in Critical Discourse Analysis*, 129–49. London & New York: Routledge.
Longman Dictionary of Contemporary English (2003). 4th ed. London: Longman.
Longman Grammar of Spoken and Written English (1999). London: Longman.

Louw, Bill (1993). Irony in the text or insincerity in the writer? – The diagnostic potential of semantic prosodies. In Mona Baker et al. (eds.), *Text and technology: In honour of John Sinclair*, 157–76. Amsterdam & Philadelphia: John Benjamins.

Macmillan English Dictionary for Advanced Learners (2002). Oxford: Macmillan.

McEnery, Tony, & Wilson, Andrew (2001). *Corpus linguistics*. 2nd ed. Edinburgh: Edinburgh University Press.

Manning, Christopher D., & Schütze, Hinrich (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Mares, Marie-Louise, & Fitzpatrick, Mary A. (2004). Communication in close relationships of older people. In Nussbaum & Coupland (eds.), 231–49.

Mautner, Gerlinde (2000). *Der britische Europa-Diskurs: Methodenreflexion und Fallstudien zur Berichterstattung in der Tagespresse*. [British discourse on Europe: Reflections on methodology and case studies on reporting in daily papers.] Wien: Passagen.

_____ (2001). Menügesteuert. Konkordanzprogramme im Dienste qualitativer Diskursanalyse. [Menu-driven. Using concordance programs in qualitative discourse analysis.] In Helmut Gruber & Florian Menz (eds.), *Interdisziplinarität in der Angewandten Sprachwissenschaft. Methodenmenü oder -salat?*, 161–88. Frankfurt/M.: Peter Lang.

_____ (2005). The entrepreneurial university. A discursive profile of a higher education buzzword. *Critical Discourse Studies* 2(2):95–120.

Nuessel, Frank (1982). The language of ageism. *Gerontologist* 22:273–76.

_____ (1984). Ageist language. *Maledicta* 8:17–28.

Nussbaum, Jon F., & Coupland, Justine (eds.) (2004). *Handbook of communication and aging research*. 2nd ed. Mahwah, NJ & London: Lawrence Erlbaum.

_____; Pecchioni, Loretta L.; Robinson, James D.; & Thompson, Teresa L. (2000). *Communication and aging*. 2nd ed. Mahwah, NJ & London: Lawrence Erlbaum.

O'Halloran, Kieran, & Coffin, Caroline (2004). Checking overinterpretation and underinterpretation: Help from corpora in critical linguistics. In Ann Hewings et al. (eds.), *Applying English grammar*, 275–97. London: Arnold.

Orpin, Debbie (2005). Corpus linguistics and critical discourse analysis. *International Journal of Corpus Linguistics* 10(1):37–61.

Partington, Alan (2004a). Utterly content in each other's company: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1):131–56.

_____ (2004b). Corpora and discourse, a most congruous beast. In Alan Partington et al. (eds.), *Corpora and discourse*, 11–20. Bern: Peter Lang.

Piper, Alison (2000a). Lifelong learning, human capital, and the soundbite. *Text* 20(1):109–46.

_____ (2000b). Some have credit cards and others have giro cheques: 'Individuals' and 'people' as lifelong learners in late modernity. *Discourse & Society* 11(4):515–42.

Robinson, James D.; Skill, Tom; & Turner, Jeanine W. (2004). Media usage patterns and portrayals of seniors. In Nussbaum & Coupland (eds.), 423–46.

Sinclair, John (1987). *Looking up*. London: Collins.

_____ (1998). The lexical item. In Edda Weigand (ed.), *Contrastive lexical semantics*, 1–24. Amsterdam & Philadelphia: John Benjamins.

_____ (2003). *Reading concordances: An introduction*. London: Pearson Longman.

_____ (2004). *Trust the text: Language, corpus and discourse*. London & New York: Routledge.

Stern, Susannah R., & Mastro, Dana E. (2004). Gender portrayals across the life span: A content analytic look at broadcast commercials. *Mass Communication & Society* 7(2):215–36.

Stubbs, Michael (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2(1):23–55.

_____ (1996). *Text and corpus analysis*. Oxford & Cambridge, MA: Blackwell.

_____ (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford & Malden, MA: Blackwell.

_____, & Gerbig, Andrea (1993). Human and inhuman geography: On the computer-assisted analysis of long texts. In Michael Hoey (ed.), *Data, description, discourse*, 64–85. London: HarperCollins.

Teubert, Wolfgang (2000). A province of a federal superstate, ruled by an unelected bureaucracy: Keywords of the Eurosceptic discourse in Britain. In Andreas Musolff et al. (eds.), *Attitudes towards Europe: Language in the unification Process*, 45–86. Aldershot: Ashgate.

_____ (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics* 10(1):1–13.
Thompson, Teresa L.; Robinson, James D.; & Beisecker, Analee E. (2004). The older patient-physician interaction. In Nussbaum & Coupland (eds.), 451–77.
Tognini-Bonelli, Elena (2001). *Corpus linguistics at work*. Amsterdam & Philadelphia: John Benjamins.
van Leeuwen, Theo (1996). The representation of social actors. In Carmen Rosa Caldas-Coulthard & Malcolm Coulthard (eds.), *Texts and practices: Readings in critical discourse analysis*, 32–70. London & New York: Routledge.
Williams, Angie, & Giles, Howard (1998). Communication of ageism. In Michael L. Hecht (ed.), *Communicating prejudice*, 136–60. Thousand Oaks, CA: Sage.
_____, & Harwood, Jake (2004). Intergenerational communication: Intergroup, accommodation, and family perspectives. In Nussbaum & Coupland (eds.), 115–37.
_____, & Nussbaum, Jon F. (2001). *Intergenerational communication across the lifespan*. Mahwah, NJ: Lawrence Erlbaum.
Williams, Raymond (1976). *Keywords: A vocabulary of culture and society*. New York: Oxford University Press.

(Received 21 February 2005; revision received 30 November 2005; accepted 12 December 2005; final revision received 3 April 2006)