

Contrastive analysis of keywords in discourses: intégration and integration in French and German discourses about migration

Article

Accepted Version

Schroeter, M. ORCID: <https://orcid.org/0000-0001-9636-245X> and Veniard, M. (2016) Contrastive analysis of keywords in discourses: intégration and integration in French and German discourses about migration. *International Journal of Language and Culture*, 3 (1). pp. 1-33. ISSN 2214-3157 doi: <https://doi.org/10.1075/ijolc.3.1.01sch> Available at <https://centaur.reading.ac.uk/47347/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1075/ijolc.3.1.01sch>

Publisher: John Benjamins

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Contrastive Analysis of Keywords in Discourses

Intégration and Integration in French and German discourses about migration

Abstract

This article suggests a theoretical and methodological framework for a systematic contrastive discourse analysis across languages and discourse communities through keywords, constituting a lexical approach to discourse analysis which is considered to be particularly fruitful for comparative analysis. We use a corpus assisted methodology, presuming meaning to be constituted, revealed and constrained by collocation environment. We compare the use of the keyword *intégration* and *Integration* in French and German public discourses about migration on the basis of newspaper corpora built from two French and German newspapers from 1998 to 2011. We look at the frequency of these keywords over the given time span, group collocates into thematic categories and discuss indicators of discursive salience by comparing the development of collocation profiles over time in both corpora as well as the occurrence of neologisms and compounds based on *intégration/Integration*.

1 Introduction

Comparing discourses across languages and discourse communities is a complicated undertaking. First, discursive phenomena extend over a variety of different genres, semiotic codes and pragmatic strategies realized through a number of linguistic features. Particularly for comparisons, the linguistic phenomena under question need to be specified and their formal, functional and semantic equivalence needs to be critically evaluated. Secondly, discourses are determined by broader social and historical contexts which need to be taken into account.

In order to tackle these two predicaments for a meaningful and systematic discourse comparison, we consider the study of discourse keywords (DKW) a good solution. DKWs are salient lexical items that occur frequently in certain discourse contexts, for example *global warming* in the context of the discourse about climate and climate change or *austerity* in the context of the financial crises since 2008. Our view on DKWs is informed by two broader

research strands. Firstly, it is informed by research in the area of cultural keywords (Williams 1976; Wierzbicka 1997; 2006; 2010) and conceptual history (Brunner et al. 1972ff.). This tradition is characterized by a methodologically qualitative-hermeneutic approach that investigates keywords as formal linguistic representatives of key cultural concepts – for example, *freedom* (Wierzbicka 1997), *class* (Williams 1976) or *justice* (Bennett et al. 2005). Such conceptualizations, as they become tangible through lexis, are investigated in some historical depth. Most interesting for us in this respect is Wierzbicka's (1997) comparative work on keywords across languages. She details the semantic differences of what may look like lexical equivalents across languages through close semantic analysis of examples of usage. However, these more historically orientated accounts tend to focus on key texts from intellectual or literary discourse and the rationales for selecting the textual data which provides evidence of usage as well as the methodology for analysis are not always clear or transparent.

Secondly, our approach to studying key words is informed by corpus assisted studies of lexis and discourse (see Baker 2006; Mautner 2009; Partington et al. 2013). Here we find explication of methodology and empirical validation, but more often than not there is a lack of historical depth and of contrastive analysis. While Corpus Assisted Discourse Analysis (CADS) is mostly based on an analysis of lexis in the relevant corpora, the notion of keyword in corpus based studies is slightly different from ours. Keywords are determined with the aid of statistical calculation in corpus linguistic approaches, both by frequency and comparison; they are words that occur significantly more (positive keywords) respectively less (negative keywords) often in one text corpus than in another reference or comparison corpus. We are leaning more towards the notion of 'cultural keywords' and, as Wierzbicka remarks:

There is no finite set of such words in a language, and there is no ‘objective discovery procedure’ for identifying them. To show that a particular word is of special importance in a given culture (or discourse, MV/MS), one has to make a case for it. (Wierzbicka 1997: 16)

However, we realize the potential of corpus based approaches to establish patterns of usage across a large amount of textual data for the sake of empirical validity.

More generally, we agree with Teubert (2010) that meaning is not a stack of cognitive concepts neatly stored in speakers’ minds to be accessed when triggered by language use. Teubert posits that our conceptualizations can be constituted and also changed only in and through language use. Meaning needs to be understood as a *fait social*, and it is constituted in and determined by language use within speech communities. As any attempt to disambiguate polysemy easily shows, context is crucial for making sense. Hence, in order to study the meaning of words, empirical studies of their occurrence in real texts is crucial. This can best be achieved with a corpus linguistic approach that considers patterns of usage, of lexical co-occurrence and recurring structural embedding of lexical items in large text corpora.

In a plea towards deeper analysis of the syntagmatic dimension of language and especially phraseology, Sinclair advocates a lexical grammar where meaning can be associated with single words, “but corpus study suggests that the influence of the verbal environment (the cotext) on the occurrence of a word is so strong that many lexical items typically consist of more than one word, and often several.”(Sinclair 2007: 182). Thus, the realization of meaning is shared through units of varying sizes, which are, for Sinclair, collocations, colligations (co-occurrence within a grammatical structure), semantic preference and discourse (or semantic) prosody (Sinclair 2004). We draw on the stance that “meaning and co-text are inter-related in such a way that involves at least partial co-selection” of elements that co-occur on the pragmatic axis, so much so that “the knock-on effect of a paradigmatic choice will be felt on the syntagmatic axis” (Sinclair 2004: 170). If this is true

in a lexical-semantic perspective, it is even more true in a discursive perspective (Baker 2006; Née & Veniard 2012). Methodologically, we will establish a comprehensive *lexical profile* (Stubbs 2001: 84) or a *collocation profile* (Wierzbicka 2010: 395) which will summarize the characteristic uses of the keywords and facilitate comparisons.

2 A corpus-assisted, lexical approach to contrastive discourse analysis

Corpus-based approaches on the one hand usually impose limitations regarding the historical development of the keywords due to the problem of availability of digitized comparable historical sources. On the other hand, they offer the opportunity to work with large text corpora and to empirically validate current usage – and difference between usages – of DKWs. We understand DKWs as semantic nodes in thematic discourses and we are dealing primarily with the lexico-semantic side of discourse; compare:

At the current state of play, and considering the limitations of those tools that are sufficiently widely available, there is a very strong bias in favour of the individual lexical item and clusters thereof. Put simply, ‘the word’ is the peg that everything else is hung on. (Mautner 2009: 124)

Mautner writes from a corpus linguistic perspective on discourse analysis and if she is right, then there is a need for a pragmatic understanding of lexical semantics and in particular for conceptualization of salient lexical items in discourses. It has indeed been acknowledged from a more lexicological than discourse analytical view that “the meaning of a word is (some kind of summation of) the conceptual content made accessible by the use of that word (as opposed to any other) in particular contexts” (Cruse 2000: 30), i.e. “there is no such thing as ‘the meaning of a word’ in isolation from particular contexts: decontextualization of meaning is variable, and in principle, always incomplete” (ibid.: 51). Similarly, Teubert (2007: 70) maintains that “[w]hat [...] lexical words [...] mean, is what we learn about them in the discourse” and that “(a)ll that has been said about a discourse object contributes

towards its meaning” (Teubert 2007: 68). For integration this means that whatever has been said about it in one language will add to and be reflected in the meaning of the word.. We understand DKWs as lexical items (see Schröter & Storjohann 2015)

- a) that occur frequently especially in periods of salience of the discourse it belongs to (e.g., *austerity* in the discourse about the financial crisis since 2008)
- b) that function as semantic nodes in discourses which upon deeper analysis of their contexts of usage unravel a part of the history and ideology of the underlying discourse
- c) that are usually part of an ensemble of other lexical items that feature prominently in the same discourse; typically there are a number of DKWs (e.g. in discourses about migration: *multicultural society, integration, communities*) and they might be associated with certain points of view (e.g. *fortress Europe* as a criticism of European migration policy)
- d) that sometimes signify controversially debated issues; controversies can include the creation of concurring DKWs; for example, the DKW *illegal immigrants* emphasizes the illegality of the action of a group of people, whereas *illegalized immigrants* was positioned against the former DKW to shift the emphasis from the people to the system and the conditions under which people are ‘outlawed’.

While we do not suppose that discourse analysis and discursive and semiotic complexity can or should be boiled down to the level of lexis, we see some advantages in focusing on DKWs especially for contrastive discourse analysis. Points a)-d) above suggest that analyses of DKWs allow insights into the discourses within which they occur (Mahlberg 2007; Née & Veniard 2012). Wierzbicka (1997) uses the metaphor of a ball of wool:

Using ‘key words’ as an approach to the study of culture (or discourse, MS/MV) may be criticized as an ‘atomistic’ pursuit, inferior to ‘holistic’ approaches targeting more general cultural patterns (...). A key word (...) is like one loose end which we have managed to find in a tangled ball of

wool: by pulling it, we may be able to unravel a whole tangled ‘ball’ of attitudes, values, and expectations, embodied not only in words, but also in common collocations, in set phrases, in grammatical constructions, in proverbs, and so on. (16f.)

This quote suggests studying the lexical and structural environment of words, for which corpus tools have proven so useful. For the purpose of contrastive analyses, a corpus-assisted approach offers the advantage of employing the same analytical steps. It still entails problems of collating comparable corpora and arguing semantic and formal equivalence of words, but consistency in the methodological procedure is a great advantage for collaborative comparative research.

Given that our approach to *Integration/intégration* as DKWs is more aligned with the notion of cultural key words, it is necessary to include contextual knowledge about the discourse in order to ‘make the case’ (Wierzbicka 1997: 16) for studying it, to situate its occurrence in a specific discourses at specific times and to explain and interpret the results from corpus analyses. It is the social reality rather than anything within the language system that explains why e.g. *terrorism*, *nuclear* and *global* are used with increasing and decreasing frequencies and in shifting contexts over time.

Taking the historical, social and political context into account is necessary especially for contrastive discourse analyses across languages and discourse communities, because the notion of lexical equivalence requires some caution. Where similar events or debates (such as migration, debt crisis, climate change) occur, a contrastive approach can involve looking at individual key words, their role and their semantics in the relevant discourse. The advantages of studying DKWs in a discourse comparative project are their salience and frequency of occurrence in a variety of texts and genres in the public realm, their phenomenologically distinct form – as opposed to e.g. the notion of a persuasive strategy – and their ubiquity; every thematic discourse features keywords as semantic nodes. This means that keywords,

even though not always equivalent ones, can be found across different discourses and different languages.

To conclude, we align ourselves with the notion of ‘cultural keywords’ in that we are looking primarily at the historical context to identify DKWs in migration discourses. We use corpus tools to investigate the salience and semantics of integration in German and French migration discourses.

Discourses about immigration offer a good opportunity to investigate the notion of keyword and to test our methodology since they have occurred in many countries inside and outside of Europe for a number of decades, in some countries leading to rather intense debates at certain times. Thus, migration discourses are likely to bring out DKWs. Moreover, migration debates have a national as well as transnational (e.g., European Union) level (Wodak and Boukala 2015) and, whereas migration discourses have been treated from a national perspective (see, among many others, Baker *et al* 2008, 2013, Hart 2010 for Great Britain), comparative perspectives are not often pursued in discourse analysis (but cf. Benson 2013, Vollmer 2014). With our approach of a comparative discourse keyword analysis, we aim to highlight similarities and differences in the French and German discourses about immigration. Differences or convergence of national migration discourses between European countries are also of interest with a view on overarching European policies and debates about migration as illustrated by the current political unease around the European freedom of movement objective in Great Britain.

3 Historical and political background to the use of Integration and intégration

Considering both socio-historical and discursive contexts regarding the use of *integration*, we can say that *integration* is widely used in discourses about migration in Germany and France, but their developments as DKWs in German and French migration discourses are somewhat

different. Early in the new millennium in Germany, *Integration* gradually replaced the concept and the positive associations of *multicultural society* that was favored by left-liberal groups in the second half of the nineties. In France, *intégration* has during the 1990s come to be largely used instead of formerly *assimilation*. *Intégration* was seen as a more liberal alternative to the more rigid concept of *assimilation* which is still favored and used by right wing politicians.

3.1 Integration in Germany

In Germany, mass immigration had become a fact of life decades ago particularly in West Germany. Successive governments hesitated to acknowledge the presence of immigrant communities as a permanent change to the German society and failed to react politically and constructively to a diversifying society. When the recruitment of guest workers was terminated in the aftermath of the 1973 oil crisis, politicians emphasized that Germany needed to consider the limitation of the number of immigrants that the country could bear (Herbert 2003).

However, especially former Turkish guest workers had lived in Germany for so long that they acquired residence rights and were also entitled to be joined by their families, so that many decided to settle in Germany. It was not until the end of the 1990s and through a change in government from 16 years of conservative Kohl governments to a coalition between the social democratic and green parties that a distinctive shift took place in the discourse by finally acknowledging Germany's multicultural society as social reality and by turning to the question of how to deal with it (Green 2004).

Whereas left-liberal groups and milieus promoted the idea of multiculturalism, the conservatives succeeded in making it appear as a naïve *laissez faire* approach to immigration (Schröter 2013), with an effective pejorization especially of the short form *Multikulti*. At the

same time, parts of the conservatives attempted to promote the concept of *Leitkultur* (guiding culture) (Eitz 2010). The idea behind this was to declare the German culture as a ‘guiding culture’ for everyone living in Germany. However, this was criticized for an inappropriate conceptualization of ‘German culture’ and as an attempt to promote the assimilation of immigrants by asking them to orientate towards German guiding culture and to give up their cultural heritage. (Göktürk/Gramling/Kaes 2007, 243-330). However, assimilation as such was never officially promoted through the use of the key word *Assimilation*. Instead, conservatives as well as more left-liberal parties largely embraced *Integration* as a middle ground between multiculturalism (which is seen to encourage segregation) and assimilation (which was regarded too rigid and inhumane). There is a broad consensus between the political parties in Germany around the need for integration with an ongoing debate about how to achieve it. In 2004 and 2007 legislation was passed with the aim of integration, most notably with the introduction of integration classes for new immigrants, citizenship tests and language requirements for resident immigrants applying for a German passport.

3.2 *Intégration in France*

In France, the main word used to refer to the process of foreigners settling down from the 19th century until the 1970s was *assimilation* (Aprile and Dufoix 2009: 31-34). It meant that immigrants should blend into their new country and were expected to adopt the French language, customs and way of life. This is related to the fact that the French nation sees itself as an egalitarian and culturally homogeneous community. However, nowadays only the far-right Front National exclusively uses *assimilation*. Other political parties and the media widely use *intégration*, even if some variation remains as well as contestation on how integration should be brought about. The development of immigration policies and the use of pertaining keywords are closely related to the rise of far-right ideas in the public sphere.

Other political parties reacted either by taking these ideas up or by hushing them. This political background tended to be a strong driver of news coverage (Benson 2013: 101).

In the 1970s, immigration in the media was treated from a social perspective (Bonnafous 1991). The main newspapers looked at the living and working conditions of immigrants who came as cheap labor to work in car factories. Left-wing media were more concerned with the question of immigration than conservative and far-right media (Bonnafous 1991: 44). The deteriorating economic situation after the 1973 oil crisis gave way to far-right positions that reached out to an audience considerably larger than its supporters. Subsequently, newspapers focused more on issues of coexistence and (supposed) resulting difficulties, gradually constructing immigrants as a ‘problem’ and questioning their ability and willingness to blend in (Benson 2013:115). At the same time, the label *assimilation* became contested by left wing groups who advocated cultural diversity and anti-racist positions. These claims framed part of French journalistic construction of the immigration debate at the time (Benson 2013: 100).

In the 1980s, the electoral breakthrough of the far-right party in the 1983 local elections created a new configuration of positions in the political field which were soon amplified in the intellectual and journalistic field (Benson 2013:112). Left wing politicians and intellectuals turned away from anti-racist and diversity stances to promote a revival of republicanism “via the notion of integration” (Benson 2013:113). Immigration became a political issue as the debate was fuelled by the “headscarves” affair (a debate over women wearing the veil in public places) and several riots in underprivileged suburbs. *Intégration* occurred in speeches by the French Presidents from 1983 (Barats 1994) and was institutionalized in 1989 with the creation of the *Haut Conseil à l’intégration* (Higher Council for Integration). This was the beginning of the politicizing of immigration.

From the 1990s until now, immigration has been institutionalized when the left-wing government created a Ministry of Social Affairs and Integration in 1991 as part of an effort to prevent the rise of far-right ideas. These, however, have spread out widely, which became apparent in the 2002 presidential election when immigration and supposedly resulting delinquency have been at the heart of the electoral debate (Née 2012). The far-right candidate Jean Marie Le Pen stood in the second round of presidential elections alongside with the conservative Jacques Chirac, instead of the expected socialist candidate. Chirac won the election by a landslide, supported by right-wing as well as left-wing voters, willing to withstand the far-right candidate. This event contributed to a series of political measures, such as the creation of an integration contract in 2002, to be signed by every immigrant coming to settle in France, or such as the creation of the *Office français de l'immigration et de l'intégration* (French Agency for Immigration and Integration). In 2007, a Ministry for Immigration, Integration, National Identity and Sustainable Development was created by the conservative president Nicolas Sarkozy. However, it was soon considered too discriminatory to devote a ministry to immigrants only and in 2010, the management of integration and immigration matters was attributed to the Ministère de l'Intérieur (Home office). This change of attribution attests the historical shift from a social to a law-enforcement perspective. In parallel, between 2003 and 2011, no less than seven laws related to immigration and integration were passed (11/2003, 12/2003, 07/2004, 07/2006, 11/2006, 11/2007 and 06/2011).

4 Corpus and methodology

Within the historical and political context characterized above, we are interested in the semantics of integration associated with its use in the two languages and speech communities.

We are interested in whether a close analysis of their usage in the public discourse about migration in both languages reveals differences. Such differences can concern

- a) frequency – are there differences in the relative frequency with which the word is used?
- b) distribution – are there differences in the relative frequency with which the word is used in different time spans or by different newspapers?
- c) collocation – are there differences in the contexts in which the word is used as far as these can be indicated by their lexical environment through the study of collocations?
- d) patterns – are there differences in the structural embedding of the word?

While newspapers should not be seen to entirely represent ‘the public discourse about migration’, they are still a good place to start such an investigation. Firstly, they offer a snapshot of a reasonably widespread hegemonic discourse like hardly any other coherent source. Even if hard copy newspapers have seen a notable decline, usually very similar content is available online. This still attracts and addresses a wide - if more fragmented -, readership, not least through possibilities of disseminating links to news stories via social media networks (Bednarek & Caple 2012: 30ff.).

We chose two French and two German broadsheet newspapers, one each with a left-liberal political orientation (*Libération* and *tageszeitung*) and one each with a conservative orientation (*Die Welt* and *Le Figaro*). This choice was partly determined by the availability and accessibility of digitized newspaper material. In order to obtain some empirical evidence regarding the development of usage and distribution, our corpus covers the years between 1998 and 2011. The data was collected in 2012. At that point, the year 1998 was the first year for which material from all of the above named newspapers was available in databases and online archives.

It would not have been feasible or helpful to download all articles containing *intégration* and *Integration* respectively because of polysemy in both languages. To

disambiguate, and also to allow further comparative studies of other migration key words, we used search words that we considered to be general and indicative of the topic of migration. For French, we downloaded all texts containing the words *immigration*, *immigré(s)* and *immigrant(s)*. In German, there is more lexical variety, so *Einwanderer*, *Zuwanderer*, *Migranten* and *Einwanderung*, *Zuwanderung* and *Migration* were used as search words to download the corpus.

- Table 1 -

These search words will not be investigated as keywords, but they serve to collect a corpus on a specific theme. The aim was to collect a thematic corpus of immigration discourse which would yield keywords within this discourse and serve as the textual basis for investigating the use of these discourse keywords, such as *intégration/Integration*, but also for example *multiculturel/multikulturell*, *diversité/Diversität*, *communautarisme/Parallelgesellschaften*.

The verb *integrieren/intégrer* also occurs in the corpus. Space prohibits detailing its use here, but it is clear that there is a strong prevalence of the key word in its nominal form in both languages: as a noun, *Integration* occurs 5.175 times in the German corpus, whereas the verb occurs 1.260 times. In the French corpus, *intégration* occurs 4.714 times, whereas the verb occurs 1.934 times.

Both corpora were annotated for date and source. The German corpus was only available via the Corpus Workbench database to use with the Corpus Query Processor (CQP) (Hardie 2012), while there were not enough resources available to add the French corpus to the same database for use with the same tools. So for the French corpus, we used AntConc (Laurence 2011). However, because the corpora were both annotated with the same metadata

and because the two different tools provide the same statistical measure for calculating collocations, we could perform the same analytical steps.

5 Analysis

5.1 Frequency and distribution of Integration

The emerging consensus around Germany as a country of immigration and of integration as the favored approach to a diverse society is reflected in the rise in frequency since the year 2000, the year in which the politicizing of integration began.

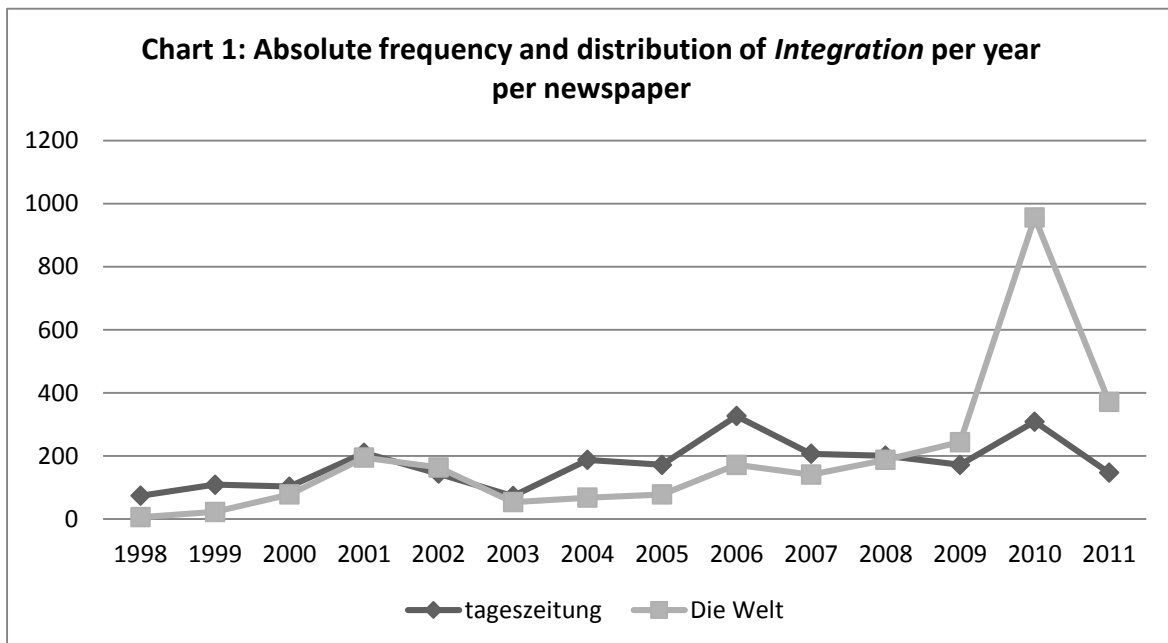
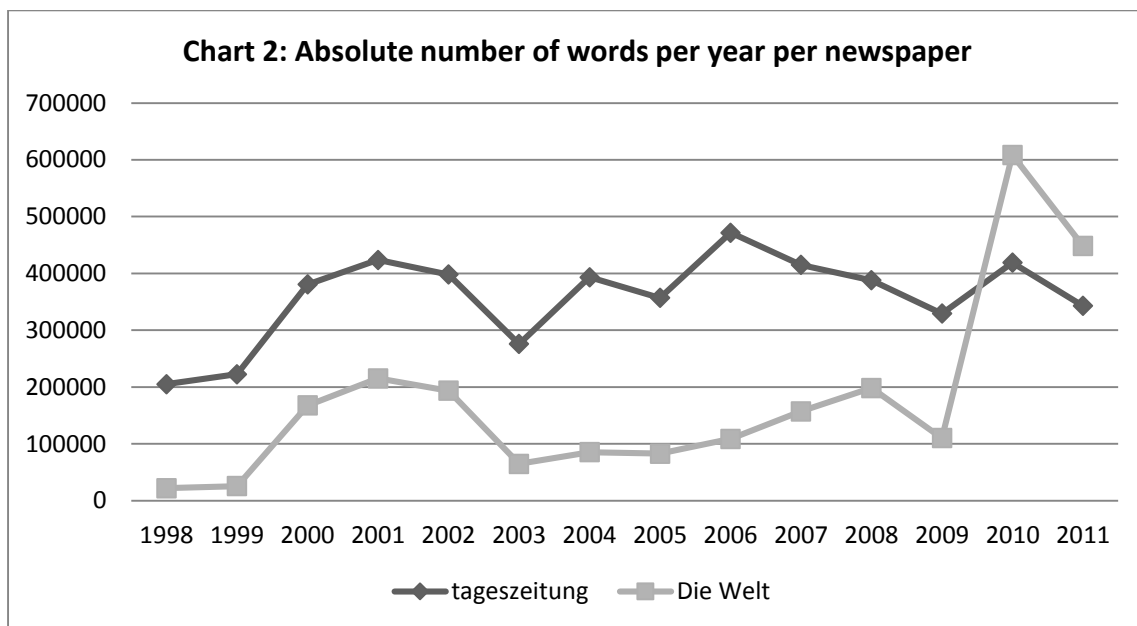
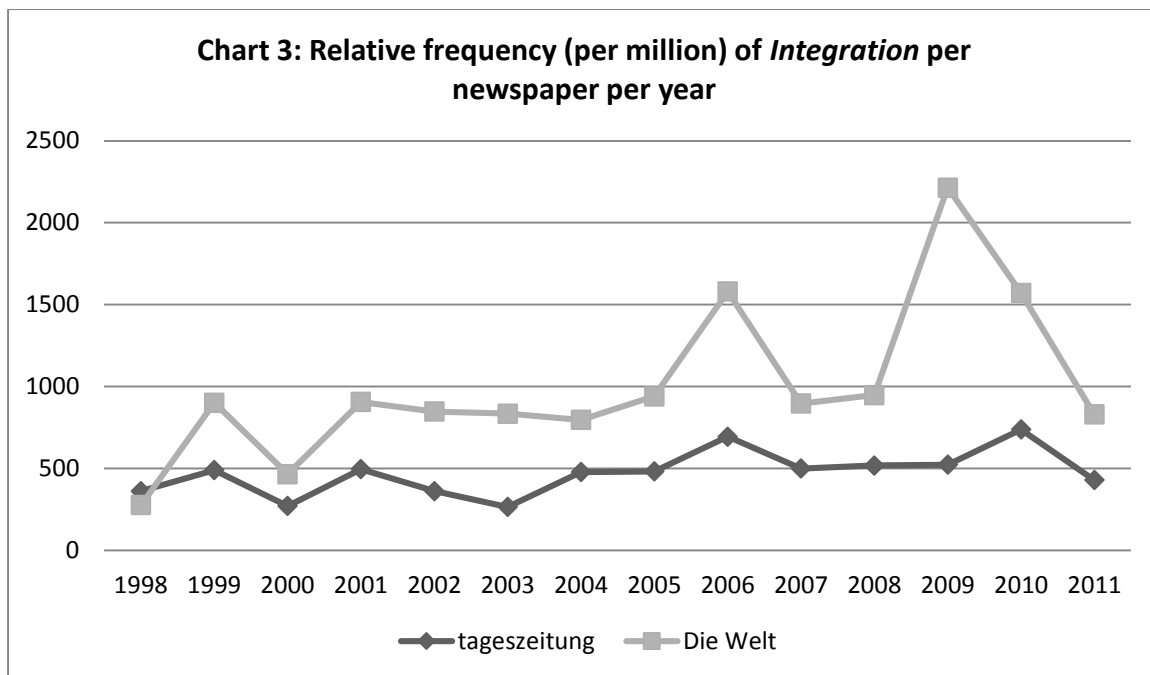


Chart 1 indicates that up to the year 2009, *Integration* occurs more often in the left-liberal *tageszeitung* than in the conservative *Die Welt*. However, there is a remarkable peak in the use of *Integration* in *Die Welt* in 2010. This picture is supported by a view on the absolute number of words per year per newspaper in the corpus (chart 2). While the frequency curve is shaped in a similar way in both newspapers, indicating general tides of topical currency, the number of words in the *Die Welt* subcorpus is consistently lower in frequency than in

tageszeitung. Here, too, *Die Welt* shows a peak in 2010, suggesting that during this year, it was particularly concerned with immigration.



Taking both newspapers together, the frequency per million words (PMW) of *Integration* in the German corpus across all years is 688.98. Looking at the frequency PMW in every year, the years 2001, 2006 and 2010 stand out. 633.55 PMW constitute a first peak in 2001 when the figures between 1998 and 2005 are otherwise below 600 PMW. The second peak rises more steeply in 2006 to 859.66 PMW, and the third to 1231.18 PMW in 2010.



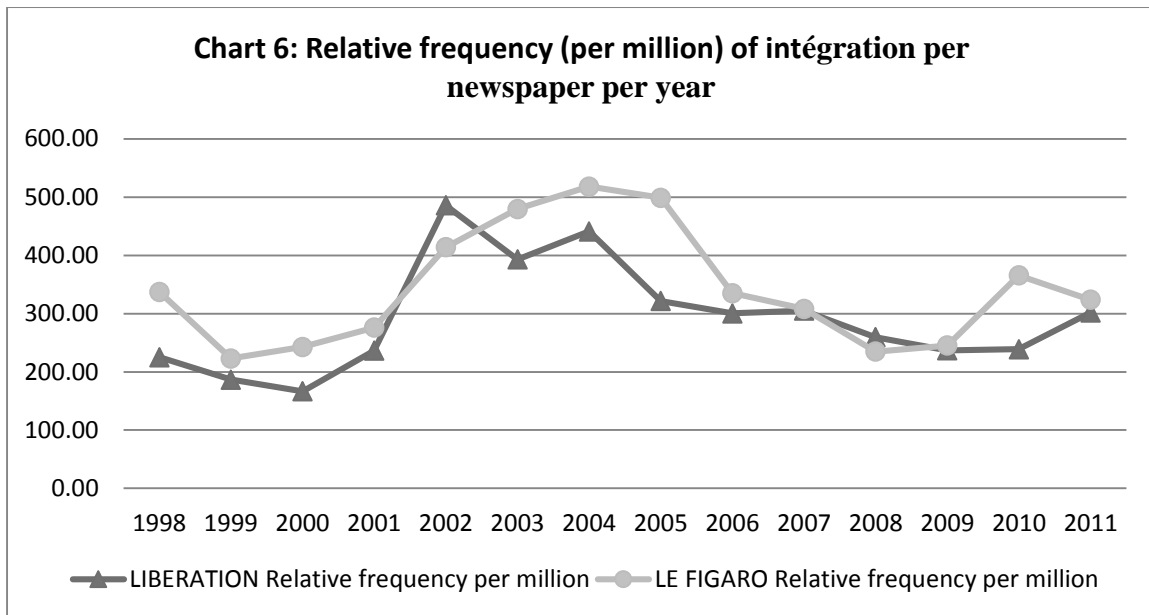
The dominant peak apparent in charts 1-3 for 2010 in the conservative *Die Welt* was caused by the publication of a book by former the former finance minister of the Berlin city government and member of the executive board of the German Bundesbank, Thilo Sarrazin. In this book he claimed that Muslim immigrants failed to integrate. The book became a bestseller instantly, but also sparked an intense and also critical debate about its central claims. Selected contributions to the debate from German public figures were also published as a book still in the same year (Deutschlandstiftung *Integration* 2010). The book was published in July and from January until July 2010, there is an average of 40 occurrences of *Integration*, rising to 114 in August, 247 in September and 190 in October. Even the overall frequency of words in *Die Welt* is divided in this year: Overall it amounts to 608.6 PMW in 2010. From January until and including June there are 159.000 words from *Die Welt*, rising to 449.6 between July until and December.. This is a remarkable demonstration of how frequency analyses can reveal how a certain discursive event – Sarrazin’s book – emerges as a discourse determining force.

The first peak in 1999 reflects the discussion around the citizenship reform which gave descendants of immigrants that fulfil certain residency requirements the automatic entitlement to German citizenship. The second peak in 2001 needs to be seen in the light of the discussion of reports from two government commissions which recommend a clearer regulation of immigration and which were in favor of the integration of long-term immigrants and their descendants as well as new immigrants, which was also the starting point for integration legislation.

In the year of the third peak in 2006, the occurrence of *Integration* in both papers clusters between February and April and in July. One notable event initiating newspaper coverage in the early Spring was the publication of a letter written by teachers from a school which predominantly taught students descendant from (mostly Muslim) immigrants about the situation at their school which they described as highly problematic. The coverage in July is determined by the first ‘integration summit’, initiated and chaired by Angela Merkel which took place in this month.

5.2 Frequency and distribution of *intégration*

Even if the word *intégration* has been used in French political vocabulary since the early 1990s, its relative frequency remains under approximately 200 PMW until 2002. This was the time when immigration started to be considered in the public debate as something having to be dealt with through concrete political measures. Frequency rises between 2002 and 2007, with peaks in 2002, 2004 and 2007 that were followed by smaller peaks in 2010 (conservative *Le Figaro*) and 2011 (left liberal *Libération*).



The relative frequencies of *intégration* presented on Chart 6 are parallel across the two newspapers, but that must not hide the fact that absolute numbers vary quite significantly.

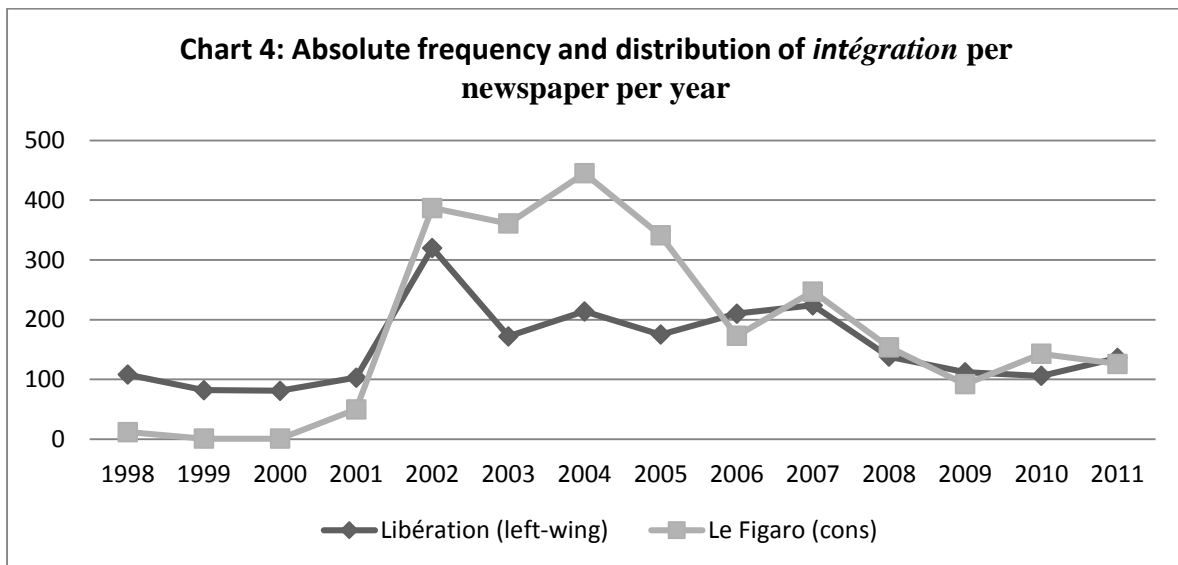
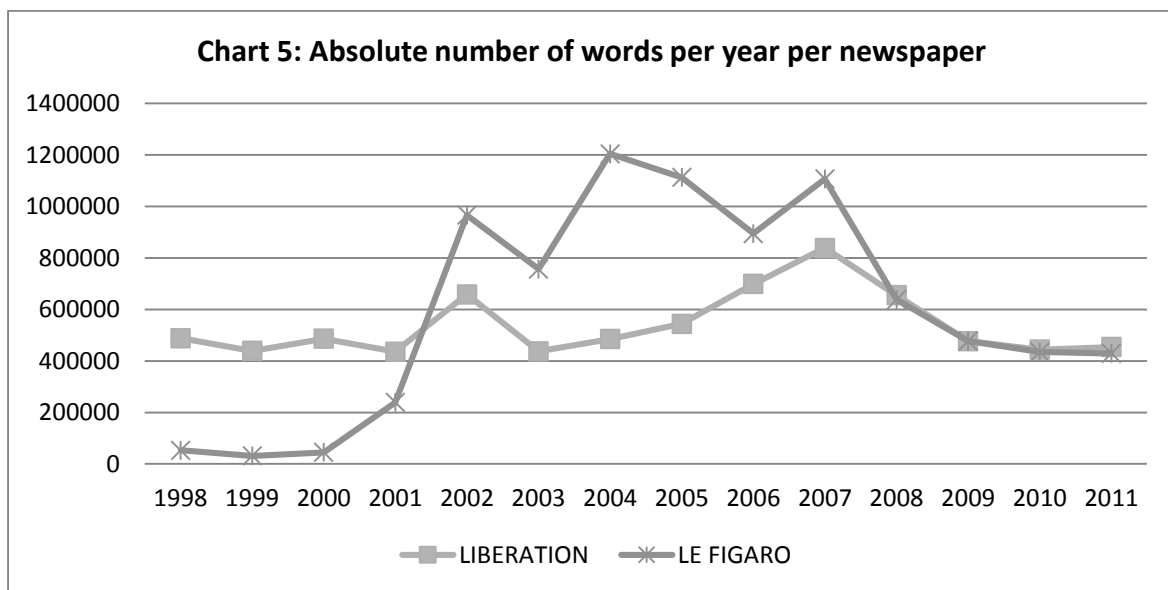


Chart 4 reveals that *Le Figaro* uses *intégration* very scarcely before 2001, but then the frequency rises higher and for longer period of time than in *Libération*. Absolute numbers are meaningful: *Le Figaro* uses the word 18 times in 1998 (110 for *Libération*) and frequency reaches 624 in the highest peak of 2004. It is considerably higher than the peak in *Libération*:

320 occurrences in 2002. In fact, it is not so much a low frequency of *intégration* itself than a low number of words in the *Figaro* sub-corpus prior to 2001. Chart 5 presents the absolute number of words devoted to immigration by each newspaper every year and shows that the number of words in *Le Figaro* has gone up by twenty-two times between 1998 and 2004 (from 53 000 to 1.2 million), whereas it remained stable in *Libération* (about 490 000 in 1998 and 2004), except for the 2002 peak.



Frequency of *intégration* clearly soars in 2002, as the 103 occurrences in *Libération* and 66 in *Le Figaro* from 2001 rise to reach 320 and 400 respectively in 2002. Reading through the concordance table and article extracts reveals that this rise is related to a series of political and media events. Two periods of high frequency have to be distinguished. The first, covering April, May and June (84, 109 and 76 occurrences) results from electoral debates revolving around insecurity, delinquency and as implicit causes, immigration and alleged unsuccessful integration. These debates were very vivid when the far-right candidate passed the first round of election on April 22nd. Other events are mentioned in the articles containing *intégration*, such as a football match where young people with immigration background hissed the national anthem (May 13th), the June 2002 Sevilla European Summit which

focused on asylum and immigration (occurrences of *Séville* suddenly rise up to 151 in June, starting from nil just two months before), or the discussion in Germany of a law related to immigration.

The peak in October and November (139 and 108 occurrences) is clearly related to the creation of a mandatory integration contract to be signed by every immigrant coming to settle down in France (*contrat d'accueil et d'intégration*). The n-gramm *contrat d'* ('contract of') occurs a hundred times in 2004, among them 66 in October and 19 in November, and all of these occurrences refer to the integration contract.

The frequency of *intégration* reaches another peak in 2004 with 659 occurrences (versus 505 in 2003). One third of these occurrences are clustered in January and February (187 occurrences). As in 2002, several political events explain this rise. A debate about affirmative action (*discrimination positive*: 48 occurrences in January, then declining) was initiated by the Home Office Minister and ended by the Higher Council for Integration who released a report ruling it out. At the time, there were also demonstrations against *la loi sur la laïcité*, a law discussed at Parliament prohibiting obvious symbols of religious allegiance in public, especially the veil (hijab). This is also reflected in the peak in frequency of *laïcité* in January of that year.

The peak in 2007 is related to the presidential election debates, notably the question of *l'identité nationale* ('national identity') which was brought about by one of the candidates, the then Home Secretary Nicolas Sarkozy, to gain on far-right voters. During this year, both uses of *intégration* and *identité nationale* rise to their highest point in March (respectively 59 and 96 occurrences). This coincides with a media event, Sarkozy's televised declaration to create a Ministry for immigration and integration if elected.

The peak in 2010 is mainly visible in *Le Figaro* (141 occurrences in *Le Figaro* and 106 in *Libération*) and takes place in September and October (47 and 48 occurrences),

coinciding with several events. In France, there were the discussion of a bill about immigration in Parliament and the release of a report by the Higher Council for Integration. Events in Germany are also mentioned, namely the acknowledgment by Angela Merkel that the multicultural model was a failure (the word *Allemagne*, ‘Germany’, peaks at 54 and 65 in September and October, otherwise, it is below 25 occurrences) and Thilo Sarrazin’s book mentioned earlier (*Sarrazin* reaches 31 of the 52 occurrences in September).

One last and smaller peak can be noted in 2011, lasting between March and July (more than 20 occurrences every month). The peak in April (49 occurrences) is once again due to the release of a report by the Higher Council for Integration (frequency of *Haut*, ‘Higher’ 14 in April, otherwise below 8). Other events are mentioned, such as the discussion in Sénat of another bill relating to immigration and the release of a book by Malika Sorel-Sutter (8 occurrences of *Sutter* in April and May, otherwise 0). She was at the time a member of the Higher Council for Integration and she claimed that immigrants should make more efforts to integrate. The frequency of *intégration* was still 25 in May due to a media event, when Home Secretary Claude Guéant declared on television that integration was a failure in France. He specifically mentioned supposedly bad school results of children with migrant background. This statement was much commented upon and *échec* (‘failure’), which is a collocate of *intégration* in 2011, peaks in May and June (14 occurrences each month, below 10 otherwise).

5.3 Collocations of Integration

After frequency and distribution have revealed trends of usage of *Integration/intégration* over the years, collocations will offer more detailed insights into the semantics of these DKWs. Drawing on Firth’s and Sinclair’s conception of meaning as a result of co-occurring words,

discourse analysis posits that "in order to understand the meanings of words, we have to compare them in relation to other words" (Baker 2006: 95).

We used T-score to calculate collocations for this research, available both on AntConc and The Corpus Workbench. We preferred T-score to MI (Mutual Information) because MI tends to give high scores to relatively low frequency words (Baker 2006: 102). For instance, in the French corpus, the strongest collocate of *intégration* according to MI is *démasquée* ('uncovered'), which has a frequency of two. T-score provides a better balance between frequency and salience by placing more emphasis on frequent words, which also means that grammatical words appear with a strong collocation value. This was not considered as a drawback since it does not exclude content words from the list. The content word with the strongest value of collocation with *intégration* in the French corpus is *immigration*, which occurs as many as 693 times near the key word (T-score of 25.9). With an interest in the semantics rather than the grammar of integration, we followed an established procedure in corpus assisted discourse analysis to focus on the lexical words in the list of collocations. We set the collocation scope at lexical items that occur up to five positions left or right of our search word, with a required minimum of five occurrences in the corpus.

Overall, there are 7.491 different words that collocate with *Integration* in the German corpus over both newspapers and over the whole time span and 6106 collocations of *intégration* in the French corpus. We consider those collocates with the highest statistical collocation value the most salient indicators of repeated contexts in which the word is used. It would be a very unwieldy procedure to analyze each of these collocates in detail, so we restricted our exploration below to the 100 collocates with the strongest collocation value according to T-score, the measure used. T-score values range overall from 18.3 to 3.4 for the 100 strongest lexical collocations in the German corpus and from 25.9 to 7.0 for the 100 strongest lexical collocations in the French corpus. Raw frequencies of content words among

the strongest 100 collocates range from 285 to 13 in the German corpus, and in the French corpus from 693 to 52. We will analyze the collocations by indicating their rounded T-score values, since the procedure for retrieving them is based on this statistical measure.

In a first step, we assigned the content words within the first hundred collocations to emergent thematic categories. This is only a rough way of seizing and comparing data since collocates might play different roles in slightly different patterns of co-occurrence. However, we can then check the concordance lines and positions in which the collocations occur so as to support our assumptions about the dominant role played by the collocation. For some – but for reasons of space limitations not for all – collocates, we demonstrate through examples patterns of co-occurrence of the collocations with the search words.

In the German corpus, the 100 strongest collocates contain 16 function words (auxiliary verbs, prepositions, articles) and 84 content words (nouns, verbs, adverbs, adjectives). Because the search was conducted for word forms and not for lemmas in order to remain sensitive for contexts of usage – variation of which might be indicated by grammatical form especially in German with its case system – some content words appear with different endings due to grammatical case or in singular and plural; for example (all plural): *Zuwanderer* (T-score 6.5), *Zuwanderern* (dative case; T-score 10) and *Zugewanderten* (participle, plural; T-score 4). Moreover, *Zuwanderer* is synonymous with *Einwanderer* (T-score 5.5) which also appears separately in the dative case, but not with the participle form. Both words mean ‘immigrants’ and are synonymous with *Migranten* (masculine form; T-score 13.9) and *Migrantinnen* (feminine form; T-score 4) which also appear among the 100 strongest collocates. *Ausländer* (foreigner; T-score 5.9), a semantically related word, also appears in two forms. Taking a closer look at *Zuwanderer* vs. *Zuwanderern* reveals that both mostly occur on the second position to the right of *Integration* as *Integration der Zuwanderer* and *Integration von Zuwanderern*. Both translates ‘integration

of (the) immigrants', but the second uses a preposition which requires a dative ending instead of an article which itself indicates the dative case.

In 63% of their altogether 842 co-occurrences with *Integration*, the labels for immigrants mentioned above appear two positions to the right of the search word in these two patterns. This supports the initial assumption that this group of words indicates the subjects and target groups of Integration as one emergent thematic category. A table that lists all content word collocations of *Integration* with their translation and T-score values and which indicates the thematic categories to which they were allocated can be found in the appendix 1. The collocate *lebenden* (present participle of to live; T-score 7.3) needs to be considered in the same thematic category because to 87% it occurs in the range of between 3 and 5 to the right of *Integration* in the patterns *Integration der hier lebenden/der bereits hier lebenden/der in Deutschland lebenden/der bereits in Deutschland lebenden [Ausländer, Migranten etc.]* (integration of [foreigners, migrants etc.] (already) living here/in Germany).

Collocates such as *Gesellschaft* (society; T-score 5.8), *Arbeitsmarkt* (labour market; T-score 5.6), *gesellschaftliche/soziale* (social; T-scores 4.8/4.8), *Einbürgerung* (naturalisation; T-score 3.9) and *kulturelle* (cultural; T-score 3.4) point towards target domains of integration, i.e., in what respects immigrants could be, or become integrated.

Another thematic category in the context of integration, as indicated by the collocations, is the state of integration – whether or not or to what extent integration is seen to have been achieved. Related collocates are: *gelungene/gelungener/gelungenen* (T-scores 8.3/4.3/3.5), *erfolgreiche* (successful; T-score 5.8), *bessere/besseren* (better, T-scores 5/4.8) and *mangelnde* (lacking T-score 3.7), which occur in 89% of all cases one position left of *Integration*, in addition to related verbs like *funktionieren* (to work/function T-score 3.4), *scheitern* (to fail), *gelingen* (succeed; T-score 5.6), *erschweren* (to inhibit; T-score 4) and *verbessern* (to improve; T-score 3.9) as well as the nouns *Probleme* (problems, T-score 7.3),

Stand (state (of integration); T-score 4) and *Beispiele* (examples; T-score 3.9), mostly in the context of ‘examples of successful integration’. *Probleme* occurs in 48 of 78 cases 2 or 3 to the left of *Integration* in the pattern *Probleme der* or *Probleme mit/bei der Integration* (problems of/with integration); similarly, *Fragen* (T-score 3.7) occurs in 17 of 36 cases two left of *Integration* in the pattern *Fragen der Integration* (questions of integration).

Collocates indicating metacommunication, such as *Thema/Themen* (topic/topics; T-scores 11.4 and 5, respectively), *Debatte* (debate; T-score 8.1), and *Diskussion* (discussion; T-score 4.5) as well as the above mentioned *Fragen* and *Sachen* (T-score 6.8) point towards an awareness of integration as a subject of public debate. *Sachen* co-occurs to 80% in one position to the left of *Integration* within the pattern *in Sachen Integraion* (with regard to/concerning integration).

The collocates *Anstrengungen* (T-score 3.8), *Bemühungen* (T-score 3.5) (both translate: efforts), *Bereitschaft* (readiness; 4.4), *Aufgabe* (task; T-score 4), *Voraussetzung* (precondition; T-score 5), *fördern* (promote/facilitate; T-score 6.7), *Engagement* (T-score 4.8), *Beitrag* (contribution; T-score 4.6), *Schlüssel* (key; T-score 7.1) point towards the conceptualization of *Integration* as a state that can be achieved by making a conscious effort to integrate – rather than something that happens as a natural process.

In this context, a related category emerges which reflects concerns about requirements for integration to succeed. *Schlüssel* occurs in 68% of cases two left of *Integration* in the pattern *Schlüsssel zur Integration* (key to integration); related to this pattern is the use of the collocate *Sprache* (language; T-score 3.6) in the vicinity of *Integration* (five to the left) in 38% occurrences of *Sprache*; *Sprache ist der Schlüssel zur Integration* (language is the key to integration); *Bildung* (education) occurs in the same position alternatively in 17% of the occurrences of *Bildung*, i.e., education is the key to integration. The collocates *Bildung* (education; T-score 5.5), *Einbahnstraße* (one-way-road; T-score 4) *braucht* (needs, verb; T-

score 3.7) *Sprache* (language T-score) and *Seiten* (sides; (T-score 3.6)also point towards conditions for successful integration. The latter occurs in the pattern (*Integration erfordert beide/zwei Seiten* (integration requires) two/both sides; *beide* (both) and *zwei* (two) co-occur with *Seiten* 25 times, of which 19 times one left of the collocate. *Einbahnstraße* occurs three positions to the right of integration to 88% in the fixed phrase *Integration ist keine Einbahnstraße* (integration is not a one-way road) – reflecting a perceived need of the state to facilitate Integration as well as a need on part of immigrants to make an effort to integrate.

Another group of collocates seems to generally reflect the thematic connection between integration and immigration; *Migration* (T-score 12.5), *Zuwanderung* (T-score 9.8), *Einwanderung* (T-score 6.4) – all of which translate ‘immigration’ – and *Assimilation* (T-score 6.3) co-occur with *Integration* mostly in coordinating patterns: Taken together, these collocates in 74% of their co-occurrence with *Integration* appear two positions left or right of *Integration*; *Integration und* (and) *Migration*, *Integration statt* (instead of) *Assimilation*, *Einwanderung und Integration*.

Finally, a last thematic category can be seen in a number of collocates that are indicative of the politicization of integration in Germany, i.e., political institutions and measures responsible for triggering and implementing Integration: *Aktionsplan* (action plan; T-score 3.6), *Staatsministerin* (Minister of State; T-score 4.4), *Sachverständigenrat* (advisory council; (T-score 4.8), *Maria* (T-score 4) and *Böhmer* (T-score 5.9), first and last name of the Minister of State responsible for Migration, Refugees and Integration between 2005 and 2013.

5.4 Collocations of intégration

There are no additional thematic categories emerging from the French corpus. Out of the eight categories within the German corpus, six are well represented: Firstly those that indicate

the connection between integration as a sub-topic of immigration as well as secondly, those that indicate the politicization of integration. Third, collocations include labels for subjects or target groups of indication and fourth, indicators of metacommunication. Fifth, the French corpus also features collocations relating to domains of integration and finally, collocations that describe the state or progress of integration. Two of the thematic categories that emerge in the German corpus, efforts made to achieve integration as well as requirements for integration, are not within the 100 strongest collocates from the French corpus. A table that lists all content word collocations of *intégration* with their translation and T-score values and which indicates the thematic categories to which they were allocated can be found in the appendix 2.

The most frequent collocate of *intégration* is *immigration* (T-score 25.9), indicating the theme of the articles. This result might not be very surprising since it was used as a search-word, but, still, the very high T-score indicates a close discursive relationship between the two words, on top of thematic associations. To explore this relationship, a statistical measure that delinearizes the syntagmatic axis the data proves a good starting point, but in order to arrive at a deeper analysis of the use of the word, we have to not only consider the frequency of co-occurrence, but also consider syntagmatic patterns of co-occurrence by looking at concordance lines. The word *immigration* appears with sound regularity (40.5%) coordinated or juxtaposed with *intégration*: *immigration et de l'intégration* ('immigration and of integration', 106 occurrences out of 607 occurrences of *immigration* to the left of *intégration*, i.e., 17.4%), *l'immigration et l'intégration* ('immigration and integration', 56 times, 9.2%), *l'immigration et à l'intégration* ('immigration and to integration', 27 times, 4.4%), *l'immigration, l'intégration* and *immigration, intégration* ('immigration, integration', respectively 33 and 25 times, 5.4% and 4.1%).

This “immigration and integration” pattern is itself embedded in patterns related to policies and institutions such as *ministre/ministère de l’immigration, de l’intégration (et de l’identité nationale)* (Immigration Minister/Ministry of immigration, integration (and national identity), 114 occurrences), *ministre/ministère de l’intégration*, 61 occurrences), *loi relative à l’immigration et à l’intégration* or *projet de loi relatif à l’immigration et à l’intégration* (law/bill on immigration and integration, 21 and 25 occurrences), *Office français de l’Immigration et de l’Intégration* (39 occurrences). Overall, those patterns account for 260 out of the 693 occurrences of the collocate *immigration* (i.e., one third) and were classified under the semantic category ‘policies and institutions’. Thus, the strong collocate *immigration* is relevant to two categories, ‘thematic indicators’ and ‘policies and institutions’.

This latter category includes *politique* (‘policy’ or ‘political’), the lexical item with the second highest T-score (17.8). The following collocates are part of labels for policies or institutions: *haut* (T-score 15.8) and *conseil* (T-score 15.7), which stand for *Haut Conseil à l’intégration* (‘Higher Council for Integration’, 3% of all occurrences of *conseil*), *contrat* (‘contract’; T-score 15.3) and *accueil* (‘welcome’; T-score 11.4) as in *contrat d’accueil et d’intégration* (literally ‘welcome and integration contract’, 14,7% of the occurrences of *contrat*), *nationale* (‘national’; T-score 13.6), *identité* (‘identity’; T-score 13.2) and *ministre* (‘minister’; T-score 10.7) as in *ministre de l’immigration, de l’intégration et de l’identité nationale* (‘minister for immigration, integration and national identity’, 6,6% of the occurrences of *ministre*).

The collocate *politique* is ambiguous and can be understood either as *policy* (noun) or as *political* (adjective). The patterns and their frequency provide disambiguation. Half of the occurrences of *politique* are embedded in the pattern *politique(s) d’intégration* (integration policy(ies), 201 occurrences of 402, 50%) and only 28 rely on the adjective (‘political

integration’, 6.9%). *Politique* therefore can be seen to belong to the category of ‘policies and institutions’.

A series of adjectives with high T-score values constitute the category ‘domain of integration’; e.g., *européenne* (‘European’, T-score 12.1, *intégration européenne* 111 occurrences), *sociale* (‘social’, T-score 9.1, *intégration sociale* 46 occurrences), *républicaine* (‘republican’, T-score 8.6, *intégration républicaine*, 67 occurrences), or *française* (‘French’ adj., T-score 10.6, 58 occurrences of *intégration à la française*). These collocates do not all refer to immigration, especially *européenne*. However, this thematic heterogeneity is decreasing overtime (see section 6.1).

The noun *intégration* is derived from a verb, so that we find collocates indicating groups integrating, mostly appearing in stable patterns based on a noun phrase. The frequency of occurrence of the collocate in the pattern is as high as 85.4%. Examples are: *intégration de/des* + nouns such as *jeunes* (35 of 72 occurrences, 48.6% of the occurrences of *jeunes* appearing in the pattern), *étrangers* (55 of 108 occurrences, 50.9%), *populations* (53 of 62 occurrences, 85.4%), *immigrants* (36 of 61 occurrences, 59%), *immigrés* (98 of 249 occurrences, 39.3%) (‘integration of young people, foreigners, population, immigrants, people having immigrated’). To underline the necessity of resorting to patterns, the case of *nouveaux* (‘new’, plural) is enlightening. This collocate has a T-score value of 7.2 with *intégration*, but this measure does not give any explanation of the reason behind this collocation. In more than half of its occurrences (55.3%), this collocate appears in the pattern *intégration des nouveaux (arrivants, venus)* (integration of new immigrants, of newcomers), which suggests to classify it in the ‘subjects of integration’ category.

Another linguistic characteristic of the noun *intégration* is its aspectual dimension, stemming from its verbal origin. This is well represented by collocates indicating integration is a process, pointing at the process (*favoriser* ‘to promote’, T-score 8.4, *faciliter* ‘to

faciliter' T-score 7.1), at the end of the process as a success (*réussie*, 'successful', T-score 10) or as a failure (*échec*, 'failure', T-score 9.6). The two words *problème(s)* and *difficultés* ('problem(s)'; T-score 8.8 for the plural form and 7.9 for singular, 'difficulties'; T-score 7.2), can express opposite meanings through different patterns, either pointing at immigrants supposedly having *problème(s) d'intégration* / *difficultés d'intégration* (problem/difficulties to integrate, respectively 69 of 148 occurrences (46.6%) of *problème(s)* and 18 out of 54 (33.3%) of *difficultés*), or causing *problèmes de l'intégration* and *difficultés de l'intégration* (problem/difficulty caused by integration, respectively 13 occurrences out of 67 (19.4%) and 19 of 54 (35.1%)). Both patterns are related to integration as a process. The last strong collocation referring to integration as a process occurs with the collocate *modèle* ('reference model', T-score 14). The pattern *modèle (français, républicain) d'intégration* (French/republican reference model for integration) accounts for 173 of the 200 occurrences of *modèle* (86.5%). The discursive meaning of the verb *fait* (3rd pers. *faire*, to do), is related to integration as a process as it appears in patterns such as *l'intégration se fait/ ne se fait pas* (integration is realised/not realized, respectively 11.6% and 5.1% of the occurrences of *fait*), contributing to express a judgment on the state of integration.

These judgments about the state of integration are expressed in a debate where integration is being discussed, which is attested by collocates such as *question* and *matière*. Both appear in patterns indicating their metacommunicative function: *question de l'/son/leur intégration* ('question of (his/her/their) integration', 45 occurrences out of 76, 59.2%) and *en matière d'intégration* or *en matière de contrat d'intégration / de politique d'intégration* ('concerning integration', 'concerning the integration contract, integration policy', 53 occurrences out of 57, 92.9%).

It appears that the category containing the more and the strongest collocations is 'policies and institutions', which needs to be interpreted with regard to the historical context

outlined above, especially the role played by the HCI, the successive legislation between 2003 and 2006 and the legislation concerning the ‘integration contract’.

5.5 Comparison

Firstly, it is interesting to note that there are more content words among the collocates for German Integration than for French integration; 87 compared to 43, respectively. By the same token, there are considerably more function words in French (57) compared to German (13).. To some extent this might be due to differences in morphology and syntax. In French, words are combined through prepositions and determiners, whereas German exploits compounding (see 6.2 below). However, the discussion below in 6.1 will show that there were more function words in German initially, but as the frequency of *Integration* increases, there are fewer function words and more content words among the first 100 collocates.

Secondly, the German data show distinct patterns of co-occurrence of *Integration* and its collocates, pointing towards an established discourse featuring repeated claims, such as *Bildung ist der Schlüssel zur Integration* (education is the key to integration) or *Integration ist keine Einbahnstraße* (integration is not a one way road). While collocates in the French corpus also tend to occur within certain patterns, some of these patterns seem less distinct and less specific than in German.

Thirdly, the collocates suggest the presence of two additional semantic categories in German; efforts (towards integration), involving the notion of efforts made by either government and/or immigrants and the measuring/assessment of such measures as well as requirements for integration, demonstrating concern over what needs to be done in order to oblige or to help immigrants to become valuable members of the German society.

Finally, even though integration has been politicised and legislated in Germany and even though there are related collocations in German, lexis relating to political institutions

collocate less strongly with integration than in French. In German, indicators of metacommunication feature more saliently among the first 100 collocates than in French, i.e. there are more such collocates with stronger collocation values in German than in French, indicating a distinct awareness about integration as a current issue under debate.

6 Indicators of discursive salience

6.1 Collocates and their development over time

Firstly, indicators of the discursive salience of integration can be seen in metacommunicative collocates such as *question*, *matière*, *Thema*, *Debatte*, and *Fragen*, which illustrate a degree of self-reflective awareness about integration as part of an ongoing debate about migration in both countries.

Secondly, we looked at the development of collocation profiles of the key word both in French and German over time. Along with the increasing frequency of the use of *Integration* and *intégration* in both corpora, we observe an increase in the number of collocations. There is also a qualitative change in the collocations in that there is an increase in content words over function words, fewer topic indicators, and, for the German corpus, more metacommunication and more evaluations regarding the state of integration over time.

In the French corpus, the development of collocations over time was analyzed based on a sub-corpus consisting of the year before the rise (2001), the overall peak in relative frequency (2004) and when the absolute frequency levels out (2009). As before, we computed collocates of *intégration* within a span of five words before and five words after the keyword. We took into account only the content words within the fifty strongest collocates. The number of collocates (types) relates to the frequency of the DKW in that it rises when the frequency of the word reaches its highest point. *Intégration* has 1742 collocates in 2004, while there are only 637 in 2001 and 777 in 2009. The number of

content words rises slightly: only 11 out of 50 collocates are content words in 2001, as opposed to 17 in 2004 and 18 in 2009.

Through the years, words referring directly to immigration become more numerous and with stronger links to integration. In 2001, only *immigration* (T-score 2.5) refers explicitly to it. In 2004, it is joined by *immigrés* ('people having immigrated'; T-score 5.9). In 2009, the thematic collocates are more diverse: in addition to those appearing in 2004, there is *immigrants* ('immigrants'; T-score 2.4) and *musulmans* ('Muslims'; T-score 2.6). In parallel, *européenne* ('European') disappears: it is present in 2001 (T-score 3.2) and absent in 2009. This is interpreted as a stabilization of the association of *intégration* with the theme of immigration rather than with Europe. It is confirmed by a rise in T-score values. For instance, *immigration* starts at 2.5, rises to 6.6 in 2004 and to 7.2 in 2009.

This phenomenon needs to be considered in regard to another one, the rise of collocates referring to French institutions. Many labels emerge: in 2004, *Haut* (7.4), *Conseil* (7.2), for Higher Council for Integration, *contrat* (5.9) and *accueil* (4.5) for the Integration Contract; in 2009, *identité* (5.4), *nationale* (5.4) and *ministre* (2.2) appearing in the label *Ministre de l'immigration, de l'intégration et de l'identité nationale*. The name of the minister is also in the list (*Besson*, 2.8).

Words denoting negative or positive outlook appear quite irregularly: *échec* ('failure') is more frequent between 2001 and 2004, passing from 8 to 18 occurrences, with an increasing T-score (2.8 to 4.2), but disappears from the first 50 collocates in 2009. Its occurrences might be related to the peak of frequency of *intégration* in 2004.

To conclude, we can say that the number of collocates increases through the sub-corpora, and that T-score values increase as well. This can be interpreted as a densification of the collocations revolving around *intégration*. Collocations are more numerous and stronger, which means that the word progressively becomes a key-word and a node in the discourse.

In German, we compared the collocations of the time span of low frequency 1998-2000, before the use of *Integration* gathered momentum and compared it to the year 2010 which yields the highest frequency. Increasing frequency goes along with a diversification of the collocation environment, so that there are 102 collocations in 1998-2000 and 346 in 2010. A closer look at the 50 strongest collocations in 1998 yields nine content words with strong collocation values which generally situate *Integration* in the context of migrants and immigration: *Zuwanderer, Migranten, Ausländern, Zuwanderung, Ausländer, Einwanderung, Migration, Einwanderern, Einwanderer* (immigrants/ immigration/ foreigners). In 2010, there are still eight such thematically situating words among the 50 strongest collocations, but more and other content words show stronger collocation values at that point.

In 1998-2000, there are 18 function words (articles, prepositions, connectors) among the 50 strongest collocations, and there are only two metacommunicative collocations, *Thema, Themen* (topic/topics), which indicate that *Integration* is an issue under debate. In this time span, there are no words among the 50 strongest collocates that evaluate the ‘state of integration’. In 2010, there are only 12 function words among the 50 strongest collocations. Discursive salience is now metalinguistically indicated by more collocates that appear with high collocation values: *Thema* (topic), *Debatte* (debate), *Sachen*, (*in Sachen Integration* – concerning/with regard to integration), *Diskussion* (discussion), *Thesen* (theses), *Themen* (topics).

In contrast to 1998-2000, the 2010 list of the 50 strongest collocates contains evaluating lexis concerning the state of integration: *bessere /besseren besser* (better), *gelingen /gelingen /gelingt* (successful, accomplished/success/succeeds). The collocation *verweigert* (refused) is new on this list, indicating the increasing concern about groups of people who supposedly refused to integrate despite political measures promoting integration.

6.2 Integration and intégration as the basis for compounds and neologisms

A third point to make regarding discursive salience is the occurrence of compounds and neologisms. Compounding is of particular interest in German because compounds can replace what would have to be expressed with an object or relative clause in English. Integration needs a junction element 's' to compound with other words, and it is interesting to see that Integration occurs 688.98 PMW in the corpus, but Integrations* occurs even more often, 714.81 PMW, which means that nominal compounds including *Integration* are more frequent than the noun by itself. The years of peak in usage (2001, 2006 and 2010) are the same for both noun and compounds. The most frequent compounds point to the politicization, institutionalization and legislation of integration in the German context:

- *Integrationspol** (121.55 PMW) (regarding politics of integration, both possible as noun: *Integrationspolitik* and adjective: *integrationspolitisch*)
- *Integrationsbeauftragte* (59.25 PMW) (state commissioner for integration),
- *Integrationskurs* (56.05 PMW) (integration course)¹

A larger group of less frequent compounds indicate concern over the perceived readiness of immigrants to integrate.

- *Integrationswill** (15.18 PMW) (willingness to integrate, both possible as noun: *Integrationwille* and adjective: *integrationswillig*)
- *Integrationsprobleme* (28.62 PMW) (problems with integration)
- *Integrationsverw** (13.85 PMW) (either *Integrationsverweigerung*, refusal to integrate or *Integrationsverweigerer*, those who refuse to integrate)
- *Integrationsbereit** (12.78 PMW) (readiness to integrate, both possible as noun: *Integrationsbereitschaft* and adjective: *integrationsbereit*)

¹ Classes provided with the aid of state funding that have become increasingly obligatory since the introduction of integration politics for new immigrants and/or those applying for citizenship if not born in Germany, comprising of a language as well as a civic education component.

- *Integrationsfähig** (10.92 PMW) (ability to integrate, both possible as noun:

Integrationsfähigkeit or adjective: *integrationsfähig*)

These findings support the above collocation analysis, especially concerning necessary efforts and requirements for integration. However, they somewhat strengthen the category of ‘politics and institutions’.

Compounding is less common in French, where the favored way to express noun determination uses the preposition ‘of’ (*de*). There are few compounds in the corpus, most notably *non-intégration* (8 occurrences) and *l’intégration-assimilation* (4 occurrences) and a few cases of morphological neology such as *intégrationniste(s)* (advocating integration, 12 occurrences), *intégratrice* (integrating, 19), *hyperintégrés* (hyper-integrated, 2) or *intégrabilité* (integrability, 1). The word *désintégrés* (disintegrated, 4) usually refers to a concrete broken item and is semantically extended to the breaking up of relationships between immigrants and society. Even if numbers are very low, these examples echo judgments over the degree of integration achieved by immigrants.

7 Conclusion

Our corpus assisted approach allows us to establish that *Integration* and *intégration* are keywords in both French and German migration discourses and that their usage is indicative of main developments and events pertinent to the public debates about migration and immigrants in both countries. They peak at certain times when the discourse salience increases, prompted not least by specific events and longer period of high salience that lead to an increased density of content words among the collocation and stabilizing patterns of co-occurrence and, therefore, discourse contexts.

The analysis of these two key words exhibits similarities and differences in usage. In German, *Integration* occurs more often than *intégration* in French, and it yields a higher

number of content words among the collocations. However, despite the extensive politicization and legislation around this issue in Germany, collocates referring to concrete political measures are more prominent in the French discourse.

Other words from the same discourse that could be potentially considered as keywords also warrant attention, especially considering the area of tension between integration, assimilation and multiculturalism as well as different word forms, e.g., the verbs *intégrer/integrieren*, *assimiler/assimilieren*, the adjective *multikulturell/multiculturel* and the noun *Multikulturalismus/multiculturalisme*.

A comparative, discourse and context oriented project questions the notion of lexical equivalence. While we select lexical items for comparison on the basis of their surface equivalence, we then explore in depth to what extent they are used in a similar way in discourse contexts and whether there are differences. Differences in usage to some extent undermine the idea of lexical equivalence across languages and confirm the role of discourse contexts as a determining factor for lexical semantics. Reference to research to establish the historical and political context of the discourses is also needed to gauge whether or not we compare words (that may or may not be formally equivalent) that hold a similar floor in the related discourse. We started from our insight into debates about migration in both countries and purpose-built a corpus that needed to focus on migration; we had to argue from the discursive context into the corpus and out of the corpus into the discursive context so that our study was not primarily determined by one or the other.

Acknowledgements

We would like to express our gratitude to Andreas Blätte, University of Duisburg-Essen, for his help with providing the German corpus as well as to Serge Fleury, Université Paris 3-Sorbonne Nouvelle, for his help with preparing the French corpus. We would also like to

thank Sylvia Jaworska, University of Reading, and, once again, Serge Fleury, for their helpful comments on earlier versions of this article.

References

- Anthony, L. (2011). AntConc (Computer Software Version 3.4.3.0). Tokyo: Waseda University.
<http://www.antlab.sci.waseda.ac.jp/>, 16/12/2014.
- Aprile, S. and S. Dufoix (2009). *Les Mots de l'Immigration*. Paris: Belin.
- Baker, P., C. Gabrielatos, M. KhosraviNik, M. Krzyzanowski, T. McEnery and R. Wodak (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society* 19(3), 273–306.
- Baker, P. (2004). Querying keywords: Questions of Difference, Frequency and Sense in Keyword Analysis. *Journal of English Linguistics*, 32(4), 346–359.
- Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics* 4(2), 197–226.
- Baker, P (2006). *Using Corpora in Discourse Analysis*. London, New York: Continuum.
- Baker, Paul, Gabrielatos, C. & McEnery, T. (2013). *Discourse analysis and media attitudes: the representation of Islam in the British press*. Cambridge: Cambridge University Press.
- Barats, C. (1994). *Intégration et immigration dans les discours de François Mitterrand (1981 -1990)*, thèse de science politiques, Université Paris 1.
- Bednarek, M., & Caple, H. (2012). *News Discourse*. London, New York : Continuum.
- Bennett, T., Grossber, L., & Morris, M. (2005). *New Keywords. A Revised Vocabulary of Culture and Society*. Malden, MA; Oxford, Carlton: Blackwell.
- Benson, R. (2013). *Shaping Immigration News. A French-American Comparison*. New York: Cambridge University Press.
- Bondi, M., & Scott, M. (Eds.). (2010). *Keyness in Texts*. Amsterdam, Philadelphia: John Benjamins.

- Bonnafous, S. (1991). *L'immigration prise aux mots, Les immigrés dans la presse au tournant des années 80*. Paris, éditions Kimé.
- Brunner, O., Conze, W., Kosellek, R. (1972ff). *Geschichtliche Grundbegriffe. Historisches Lexikon zur politisch-sozialen Sprache in Deutschland*. Stuttgart: Klett Cotta.
- Bubenhofer, N. (2013). Quantitativ informierte qualitative Diskursanalyse. Korpuslinguistische Zugänge zu Einzeltexten und Serien. In K.S. Roth & C. Spiegel (Eds.), *Perspektiven einer angewandten Diskurslinguistik* (103–129). Berlin: Akademie-Verlag.
- Cruse, A. (2000). Aspects of the Micro-structure of Word Meanings. In Ravin, Y., & Leacock, C. (Eds.), *Polysemy. Theoretical and Computational Approaches* (0-51). Oxford: Oxford University Press.
- Eitz, T. (2010). Das missglückte Wort. Bundeszentrale für politische Bildung, *Themen*.
http://www.bpb.de/themen/A3M5EN,0,Das_missgl%FCckte_Wort.html (18.11.2014).
- Göktürk, D. Gramling, D. & Kaes, A. (2007). *Germany in Transit. Nation and Migration 1955-2005*. Berkley, London: University of California Press.
- Green, S. (2004). *The Politics of Exclusion. Institutions and Immigration Policy in Contemporary Germany*. Manchester: Manchester University Press.
- Hart, C. (2010). *Critical Discourse Analysis and Cognitive Science. New Perspectives on Immigration Discourse*, London: Palgrave Macmillan.
- Hardie, A. (2012). CQPweb – combining power, flexibility in a corpus analysis tool. *International Journal of Corpus Linguistics* 17(3): 380-409.
- Herbert, U. (2003). *Geschichte der Ausländerpolitik in Deutschland. Saisonarbeiter, Zwangsarbeiter, Geatarbeiter, Flüchtlinge*. Bonn: Bundeszentrale für politische Bildung.
- Jung, M., Niehr, T., & Böke, K. (2000). *Ausländer und Migranten im Spiegel der Presse. Ein diskurshistorisches Wörterbuch zur Einwanderung seit 1945*. Wiesbaden: Westdeutscher Verlag.
- Mahlberg, M. (2007). Lexical items in discourse: identifying local textual functions of sustainable development. In Hoey, M., Mahlberg, M., Stubbs, M., Teubert, W. (Eds.), *Text, Discourse and Corpora. Theory and Analysis* (191–218). London: Continuum.

- Mautner, G. (2009). Checks and Balances. How Corpus Linguistics can Contribute to CDA. In Wodak, R., & Meyer, M. (Eds.), *Methods of Critical Discourse Analysis* (122–143). London: Sage.
- Née, E., & Veniard, M. (2012). Lexical Discourse Analysis: Toward a Revival Using a Semantic Approach? *Langage et Société* 140, 15-28, http://www.cairn-int.info/resume.php?ID_ARTICLE=E_LS_140_0015 (4.12.2014).
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse. Theory and practice in corpus-assisted discourse studies*. Amsterdam, Philadelphia: John Benjamins.
- Schröter, M. (2013). Die kontrastive Analyse politischer Diskurse. Skizze, Verkomplizierung und Ausblick. In Kilian, J., & Niehr, T. (Eds.), *Politik als sprachlich gebundenes Wissen. Politische Sprache im lebenslangen Lernen und politischen Handeln* (91-105). Bremen: Hempen.
- Schröter, M., & Storjohann P. (2015). Patterns of Discourse Semantics. A Corpus-Assisted Study of *Financial Crisis* in British Newspaper Discourse in 2009. *Pragmatics and Society* 6(1), 43-66.
- Sinclair, J. (2004). *Trust the text. Language, corpus and discourse*. London: Routledge.
- Sinclair, J. (2007). Meaning in the framework of corpus linguistics. In Teubert, W., & Krishnamurthy, R. (Eds.), *Corpus linguistics*, vol 1, (182-196). Routledge, London.
- Stubbs, M. (2001). *Words and phrases*. Oxford: Blackwell.
- Stubbs, M. (2010). Three concepts of keywords. In Bondi, M., & Scott, M. (Eds.), *Keyness in Texts*, (21-42). Amsterdam, Philadelphia: John Benjamins.
- Taylor, C. (2014). Discourse keywords of migration: Community and community in UK and Italian newspapers. Paper presented at the 10th European Social Science History Conference, Vienna, April 2014.
- Teubert, W., & Čermáková, A. (2007). *Corpus Linguistics. A Short Introduction*. London: Continuum.
- Teubert, W. (2010). *Meaning, Discourse and Society*. Cambridge: Cambridge University Press.
- Vollmer, B. (2014). *Policy Discourses on Irregular Migration in Germany and the United Kingdom*. Basingstoke: Palgrave Macmillan.

Wierzbicka, A. (1997). *Understanding Cultures through their Key Words. English, Russian, Polish, German, and Japanese*. Oxford: Oxford University Press.

Wierzbicka, A. (2006). *English: Meaning and Culture*. Oxford: Oxford University Press.

Wierzbicka, A. (2010). *Experience, Evidence, and Sense. The Hidden Cultural Legacy of English*. Oxford: Oxford University Press.

Williams, R. (1976). *Keywords. A vocabulary of culture and society*. London: Fontana.

Wodak, R. & Boukala, S. (2015). (Supra)National Identity and Language: Rethinking National and European Migration Policies and the Linguistic Integration of Migrants. *Annual Review of Applied Linguistics* 35, 253-273.

Appendix 1 Collocations of Integration in German

Thematic category	Collocation	Translation	T-score value (rounded)
Indicators of thematic environment	Migration	Migration	12.5
	Zuwanderung	immigration	9.8
	Einwanderung	immigration	6.4
	Assimilation	assimilation	6.3
	Integration	integration	5.6
	Islam	Islam	3.7
Subjects of integration	Zuwanderern/ Zuwanderer	immigrants (nom + gen)	10/6.5
	Migranten /Migrantinnen	migrants (m/f)	13.8/4.1
	Ausländern /Ausländer	foreigners (nom + gen)	8.8/5.9
	Einwanderern/Einwanderer,	immigrants (nom + gen)	8.4/5.5
	lebenden	living	7.2
	Türken	Turks	5.2
	Zugewanderten	(past participle; those who	3.8
muslimischen/muslimischer	immigrated)		
	Muslim (adj., nom. + gen)	3.5/3.5	
Destination or domain of integration	Gesellschaft	Society	5.8
	Arbeitsmarkt	labour market	5.6
	gesellschaftliche	social	4.8
	soziale	social	4.8
	Einbürgerung	naturalisation	4.0
kulturelle	cultural	3.4	
Indicators of metacommunication	Thema	topic	11.3
	Debatte	debate	8.0
	Sachen	(concerning/with regard to)	6.8
	Themen	topics	5.0
	fordert	demands (<i>reported speech</i>)	4.8
	Diskussion	discussion	4.5
	bedeutet	means (verb 3rd p sg)	4.0
	Thesen	theses	4.0
	Sarrazin	(name of book author)	3.8
	Fragen	questions	3.7
Bedeutung	meaning	3.6	
State of integration	Gelungene	successful/accomplished	8.3

	bessere Probleme erfolgreiche gelingen gelingt besser besseren gescheitert funktioniert gelungener Beispiele verbessern mangelnde gelungenen gescheiterten funktionieren	better (adj) problems successful succeeding succeeds better (adv) better (adj, dat) failed works (~works well, succeeds) successful/accomplished examples improve lacking/not good enough successful/accomplished (gen) failed to work (~to succeed)	8.0 7.2 5.8 5.6 5.3 5.0 4.8 4.8 4.6 4.3 3.9 3.9 3.7 3.5 3.4 3.4
Efforts made and measured	Schlüssel fördern Engagement Beitrag Stand Bereitschaft Aufgabe Anstrengungen erschwert Förderung braucht Bemühungen	key promote, foster engagement contribution state of affairs readiness task efforts make difficult promotion/support needs (verb) efforts	7.1 6.7 4.8 4.6 4.0 4.0 4.0 4.0 3.8 3.8 3.7 3.5
Policies	Stiftungen Böhmer Sachverständigenrat/Sachverständigenrats/ Sachverständigenrates Staatsministerin Maria SVR Aktionsplan Minister	foundation family name of minister of state for integration affairs Advisory board/panel of experts, different case endings minister of state (f) First name of minister of state for integration affairs Acronym of Sachverständigenrat action plan Minister	6.1 5.9 4.8 /4.3 /3.8 4.4 4.0 3.7 3.6 3.4
Requirements for integration	Bildung Voraussetzung Seiten Sprache Einbahnstraße	Education requirement sides language one-way-road	5.5 5.1 3.6 3.6 4.0

Appendix 2 Collocations of *intégration* in French

Category	Relevant collocation and rank	Translation	T-score value (rounded)
Indicators of thematic environment	immigration	immigration	25.9
Policies and institutions	Politique /Politiques Haut/Conseil/hci Contrat/accueil	Policy or political ('political integration') Haut Conseil à l'intégration contract (mandatory for every immigrant who enters France)	17.8 15.8/15.7/7.8 15.3/11.4

	Nationale/Identité/Ministre/ Ministère	nationale, identité, minister and ministère refer to the ministry/minister of immigration, integration and national identity	13.6/13.2/10.7/ 7.2
Destination or domain of integration	Immigration européenne Française France pays Français sociale Républicaine société	immigration European French France Country French social Republican Society	25.9 12.1 10.6 10.1 9.5 9.2 9.1 8.6 7.7
Subjects of integration	immigrés étrangers jeunes populations immigrants	immigrants foreigners young people people immigrants	15.5 10.1 8.0 7.8 7.6
State of integration / Integration as a process	modèle réussie échec processus problèmes favoriser fait (verb) Problème Difficulties faciliter	model successful failure process problems promote do problem difficulty facilitate	14.0 10.0 9.6 9.2 8.8 8.4 8.0 7.9 7.2 7.1
Indicators of metacommunication	question matière	question concerning	8.3 7.3
Efforts made and measured	n/a	n/a	
Requirements for integration	n/a	n/a	