Knowledge and use of the lexicon in French as a second language

Book

Published Version

Introduction to special issue

Treffers-Daller, J. ORCID: https://orcid.org/0000-0002-6575-6736, Daller, M. H., Malvern, D., Richards, B., Meara, P. and Milton, J., eds. (2008) Knowledge and use of the lexicon in French as a second language. Journal of French Language Studies, 18 (3). Cambridge University Press, Cambridge, pp112. (Special issue 03) Available at https://centaur.reading.ac.uk/29767/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://journals.cambridge.org/action/displayIssue?
decade=2000&jid=JFL&volumeId=18&issueId=03&iid=2370040

Publisher: Cambridge University Press

Introduction: Special issue on knowledge and use of the lexicon in French as a second language

GUEST EDITORS: THE M4 APPLIED LINGUISTICS NETWORK, JEANINE TREFFERS-DALLER AND HELMUT MICHAEL DALLER

University of the West of England, Bristol

DAVID MALVERN AND BRIAN RICHARDS

University of Reading

PAUL MEARA AND JAMES MILTON

Swansea University

INTRODUCTION

This special issue of *JFLS* focuses on what learners know about French words, on how they use that knowledge and on how it can be investigated and assessed.

In many ways, it is a sequel to the special issue on the Acquisition of French as a Second Language edited by Myles and Towell that appeared in *JFLS* in 2004. While articles on the L2 acquisition of the French lexicon have appeared in a variety of journals, including *JFLS*, this special issue (SI) is the first volume which specifically focuses on lexical knowledge and use among learners of French as a second language. The issue is timely, because of the growing importance of vocabulary in the SLA research agenda, but also because research into vocabulary acquisition appears at the top of a list of areas in which teachers of Modern Foreign Languages are most interested (Macaro, 2003: 6).

Over the past few decades interest in research into Second Language Acquisition has grown exponentially, as can be seen in the number of textbooks and handbooks that have appeared in recent years (Ritchie and Bhatia, 1996; Cook, 1996 et seq.; Doughty and Long, 2001; Myles and Mitchell, 2004), as well as in the formulation of new theories of SLA, such as Processability Theory (Pienemann, 1998), Acquisition by Processing Theory (Truscott and Sharwood Smith, 2004) and the Extended Competition Model (MacWhinney, 2005), which complement approaches to SLA that are based on Universal Grammar. As Myles and Mitchell (2004: 91) put it, approaches to SLA which are based on Universal Grammar, have mainly focused on describing and explaining morphosyntactic development in learners, and much less on other aspects of the linguistic system. In the Minimalist Program, however, the differences between languages are seen to be mainly lexical in nature. According to Cook (1998), the Minimalist Program is lexically-driven in that the properties

of lexical items shape the sentence rather than lexical items being slotted into pre-existent structures. As a result, the task the L1 learner faces is mainly one of learning the lexicon (lexical and functional items), which then triggers the setting of universal grammatical parameters. This approach is reflected in the Lexical Learning Hypothesis (Ellis, 1997) according to which vocabulary knowledge is indispensable to acquire grammar (see also Bates and Goodman, 1997).

There are many reasons why understanding how words are learned and used and how we can measure that knowledge is also crucially important for researchers working in Applied Linguistics, in particular for those with an interest in Education. As Milton (this volume) points out, vocabulary knowledge can be quantified in ways that other aspects of language knowledge cannot, which makes this area particularly attractive for the development of indices and measures that can be used in a variety of educational (and clinical) contexts to assess learners' or patients' language profiles.

The importance of vocabulary has been demonstrated in a wide range of studies. It is widely accepted that lexical knowledge is one of the main prerequisites for academic achievement of monolingual and bilingual children (see Daller, 1999; Dickinson and Tabors, 2001). According to Meara and Bell (2001), teachers' judgements of L2 texts appear to be based to a large extent on the type of vocabulary used by the students. In a similar vein, Malvern and Richards (2002) show that teachers' subjective rating of students' range of vocabulary in oral interviews correlated very highly (all values above 0.97) with their judgements about fluency, complexity, content and accuracy.

The language threshold for reading is also largely lexical. Anderson and Freebody (1981) report a high correlation between tests of vocabulary and reading comprehension across a range of studies in first language reading research. Laufer (1992) points to similar results for second language acquisition. She also provides evidence that for text comprehension a vocabulary large enough to provide coverage of 95% of the words in a text is needed. Hu and Nation (2000) even put the boundary for unassisted reading at 98%.

A focus on vocabulary can also provide new insights into the distinction between implicit and explicit learning (Ellis, 1994; Hulstijn, 2003). Recent research has focused on *incidental vocabulary*, i.e. vocabulary that second language learners develop while they are focused on a task other than on learning new words (Gass, 1999). Most scholars agree that except for the first few thousand most common words, L2 vocabulary is predominantly acquired incidentally (cf. Huckin and Coady, 1999). If this is correct, there are important implications for the ways in which vocabulary is offered in language classes and textbooks. The role of explicit versus implicit (or incidental) vocabulary learning is also taken up in two contributions to this special issue (Houten, Bulté, Pierrard and Van Daele; Graham, Richards and Malvern).

Much less is known with respect to the pragmatic and sociolinguistic knowledge about individual items, i.e. information about the ways in which particular words are used dependent on the formality of the conversation, the characteristics of the interlocutors and the topic of the conversation. How L2 learners acquire pragmatic

competence, and how this competence is linked to other levels of linguistic competence is only beginning to be explored (see also Read, 2000 and Nadasdi, Mougeon and Rehner, this volume).

From a theoretical perspective it is important to study the acquisition of a variety of languages as one cannot discover universal principles or processes of Second Language Acquisition if the focus is on one language (English) only. The recent upsurge in interest in French SLA (Myles and Towell, 2004; Prévost and Paradis, 2004; Dewaele, 2005; Ayoun, 2007) is therefore particularly welcome. Each of these volumes makes a very important contribution to our understanding of a range of issues in the acquisition of French, and they all demonstrate the important contribution the analysis of French can make to the development of the SLA research agenda.

This special issue is intended to complement the ones mentioned above, in that it focuses specifically on lexical issues. Most of the contributions to this special issue were first presented in the context of the ESRC seminar series entitled *Models and measures of vocabulary acquisition, knowledge and use: the interface between theory and applications*. This seminar series was jointly organised by researchers of the M4 Applied Linguistics Network (David Malvern and Brian Richards from Reading; Paul Meara and Jim Milton from Swansea and Helmut Daller and Jeanine Treffers-Daller from UWE Bristol) at the three participating institutions between 2006 and 2007.

As Milton (2006) has shown, we know very little about the words that are learnt in French classes or about the vocabulary sizes of learners of French at different levels in the UK or elsewhere, and this volume aims to address this issue. In the first three articles, the focus is on the measurement of lexical richness in speech production. Housen, Bulté, Pierrard and Van Daele make an important contribution to the discussion around the theoretical constructs used in lexical aspects of second language acquisition and how these can be operationalised in empirical studies. In their theoretical framework they distinguish three levels of analysis. At the theoretical level there is the cognitive construct of lexical competence, which consists of different dimensions (width, depth etc.) and a procedural component. As lexical competence cannot be directly observed, they distinguish this from the behavioural construct of lexical proficiency (diversity, sophistication etc.), which then needs to be operationalised in statistical terms in the form of various statistical constructs (TTR, Index of Guiraud etc.). Housen *et al.* subsequently demonstrate how these concepts and operationalisations can be used to analyse lexical development over time in a longitudinal study among Flemish students of French in the Dutch-speaking schools in Brussels. Using a wide variety of measures, the authors show the students progress significantly in terms of lexical diversity, sophistication and productivity over a period of two years (from age 12 until age 14). The authors also call for a greater conceptual clarity as to what lexical sophistication entails. More specifically they come to the conclusion that it is not sufficient to refer to frequency lists alone in operationalisations of lexical sophistication.

This issue is taken up in the second contribution by Tidball and Treffers-Daller who focus on lexical sophistication in the oral productions of British learners of French in Higher Education. They use different operationalisations of the Advanced Guiraud (Daller, Van Hout and Treffers-Daller, 2003) to show that operationalisations based on teacher judgement are superior to those which are solely based on frequency, including a measure based on Vocabprofil, the French version of Laufer and Nation's Lexical Frequency Profile. In addition, they point to the need to address the key role of cognates in vocabulary learning as these often remain outside the scope of measures of vocabulary richness. The study confirms the results of Horst and Collins (2006) that learners often make better use of words in the higher frequency bands, in that they use a larger variety of the first 1000 words (k1), but do not necessarily improve significantly in their use of infrequent words. The authors therefore call for measures that further differentiate between different frequency layers among the k1 group.

In the third study, David presents a detailed analysis of the lexical richness in the speech of 80 British learners in secondary education (Years 9–13). Using a combination of a controlled active and a free active task she shows that lexical diversity as measured on the basis of D (Malvern, Richard, Chipere and Durán, 2004) increases significantly from Year 9 until Year 12 (cross-sectional study), with the largest increase between Years 11 and 12. No significant difference was found between Year 12 and Year 13 in the longitudinal part of the study, which could be due to a test repetition effect. David also demonstrates that the proportion of nouns increases between Years 9 and 10 and thereafter it consistently decreases. This confirms the existence of an early noun bias in the early L2 acquisition of French. The proportion of verb types increases over the different year groups. She finally calls for a further investigation of the effect of tasks (free productive versus controlled productive) and lemmatisation on measures of lexical richness, as comparisons of D-figures between different studies are currently hampered by variation in tasks and lemmatisation methods.

In the fourth study the focus shifts from productive to receptive vocabulary. Milton's study is a follow-up on Milton (2006) in which he studied receptive vocabulary growth among 449 learners across all levels in a British school using a French adaptation of X_lex (Meara and Milton, 2003). In the current study, Milton compares the vocabulary of 21 learners who had studied to 'O' level French between 20 and 50 years ago study with the 24 best students from an equivalent cohort that took GCSE. The comparison shows that the current GCSE students possess only a half to a third of the vocabulary knowledge of learners who took 'O' level in the past. This enormous decline is probably due to pressure to increase the numbers taking and passing the examination, as well as to a reduction in the classroom hours available for learning, and a change in the examination format. Milton also reports about the receptive vocabulary of 66 students taking single and joint honours French at a British university using the same testing methodology. This part of the study shows that students continue to learn words at the same kind of rate that they did at 'A' level; about 500 words per year. This means they

end up with relatively low levels of vocabulary knowledge that make effective communication in the foreign language very difficult. Clearly there are important implications of this study for everyday practice in MFL in schools and universities in the UK.

The fifth study specifically focuses on vocabulary progress in Year 12 in the context of learner strategy training that targeted listening and writing by 150 pupils in 15 English comprehensive schools. Graham, Richards and Malvern analyse the impact of learner strategy training in eleven schools and compare the results with a comparison group of four schools. On the basis of raters' assessments of written picture descriptions, they show that the students' productive vocabulary increases significantly over the two terms. They also used a French adaptation of X_lex (Meara and Milton, 2003) to assess receptive vocabulary. This tool was shown to be a valid measure to distinguish between students with different GCSE grades, but it was more difficult to demonstrate the existence of measurable progress in receptive vocabulary on the basis of this tool. There was an overall gain in listening, writing and productive vocabulary scores, but no group effect. This means that the additional time used on strategy teaching did not negatively affect vocabulary growth.

In the final contribution, the focus is on a dimension of vocabulary knowledge that receptive or productive measures generally ignore, namely socio-stylistic variation. Nadasdi, Mougeon and Rehner's study of lexical variation in the use of Canadian French expressions for *car* demonstrates how difficult it is for learners to get to grips with sociolinguistic variation in French. Among native speakers of Canadian French *auto* is the expression most frequently used variant, with *char*, *machine* and *automobile* being secondary variants whereas *voiture* is virtually absent from their speech. The 41 Canadian immersion students prefer the form *auto*, which may be due to the similarity between this form and the Spanish and Italian translation equivalents, as well as to its frequency in the teaching materials. In addition, they use the hyperformal *voiture*, which is very frequent in the educational input to students, but make no use of the marked informal variants *machine* or *char*. The absence of *char* in the input to the students (in the classroom and the textbooks) is probably a key reason for the students' preferences.

While it is difficult to assess the importance of all six articles for our understanding of lexical dimensions of language proficiency, it is clear that a key aspect of this contribution is a methodological one: the tools and measures used in this study – many of which were developed by members of the M4 Applied Linguistics Network – make it possible to carry out valid and reliable analyses of learner's knowledge across a range of national and educational contexts. The research community will also no doubt find it helpful to see how all tools and measures are evaluated in the different studies.

The studies described here also provide new detailed insights into the development of lexical knowledge among learners in schools and universities. The picture that emerges from these studies is that vocabulary learning is a neglected area in secondary and tertiary education. A further decline in standards can only be

prevented if the information provided here is taken on board by all those who are keen to improve MFL teaching, that is teachers, authors of textbooks and policy makers in different countries, but in particular in the UK, i.e. the Qualifications and Curriculum Authority and the Department for Children, Schools and Families.

An area that deserves further attention from the research community is the construct validity of measures of lexical richness. While lexical knowledge is clearly multidimensional, the frequency of lexical items (types or tokens) is being used more often than anything else in the development of measures of lexical richness. This issue is particularly important for measures which claim to tap into lexical sophistication, as frequency is unlikely to the only determinant of lexical sophistication. A related problem is that valid information about the frequency of lexical items in spoken language remains difficult to obtain due to the fact that oral corpora are much smaller than written corpora. Any measures of lexical richness/sophistication that are based on frequency data from written corpora but are being applied to oral data are likely to run into problems. An important question for future research will also be to establish how cognates and units beyond the single word level such as formulaic sequences (Wray, 2005) can be accounted for in measures of lexical richness. The aim of a special issue such as this can however not be to answer all questions. We hope that the research community will find this special issue has made a useful contribution to answering some of the questions and to pointing out how follow-up studies can contribute to developing this research agenda further.

Address for correspondence:
Jeanine Treffers-Daller
Department of Languages, Linguistics and Area Studies
School of Humanities, Languages and Social Sciences
Faculty of Social Science and Humanities
Frenchay Campus
Coldharbour Lane
Bristol BS16 1QY
UK
e-mail: Jeanine.Treffers-Daller@uwe.ac.uk

REFERENCES

Anderson, R. C. and Freebody, P. (1981). Vocabulary knowledge. In: J. Guthrie (ed.), *Comprehension and Teaching: Research Reviews*. Newark, DE: International Reading Association, pp. 77–117.

Ayoun, D. (2007) (ed.). *French Applied Linguistics*. Benjamins. Language Learning & Language Teaching 16.

Bates, E. and Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. In G. Altmann (ed.), Special issue on the lexicon, *Language and Cognitive Processes*, 12.5/6: 507–586.

Cook, V. (1991; 1996; 2001). *Second Language Learning and Language Teaching*. New York: Hodder Arnold.

Cook, V. (1998). Review of Peter Skehan (1998). *A Cognitive Approach to Learning Language*. Oxford University Press. http://homepage.ntlworld.com/vivian.c/Writings/Reviews/SkehanRev.htm [accessed 16 February 2007]

Daller, H. (1999). *Migration und Mehrsprachigkeit. Der Sprachstand türkischer Rückkehrer aus Deutschland*. Frankfurt am Main, etc.: Peter Lang.

Daller, H., Van Hout, R. and Treffers-Daller, J. (2003). Lexical richness in spontaneous speech of bilinguals. *Applied Linguistics*, 24: 197–222.

Dewaele, J. M. (ed.) (2005). *Focus on French as a Foreign Language*. Clevedon: Multilingual Matters.

Dickinson, O. K. and Tabors, P. O. (eds.) (2001). *Beginning Literacy with Language: Young Children Learning at Home and School*. Baltimore: Paul H. Brookes.

Doughty, C. and Long, M. (2001). (eds). *Handbook of Second Language Acquisition*. Malden, MA: Blackwell.

Ellis N. C. (ed.) (1994). *Implicit and Explicit Learning of Languages*. New York: Academic Press.

Ellis, N. C. (1997). Vocabulary acquisition: Word structure, collocation, word–class, and meaning. In: N. Schmitt and M. McCarthy. (eds.) *Vocabulary: Description Acquisition and Pedagogy*. Cambridge: Cambridge University Press, pp. 122–139.

Gass, S. M. (1999). Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21: 319–333.

Horst, M. and Collins, L. (2006). From *Faible* to strong: How does their vocabulary grow? *The Canadian Modern Language Review*, 63.1: 83–106.

Hu, M. and Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13.1: 403–430.

Huckin, Th and Coady, J. (1999). Incidental vocabulary acquisition in a second language: a review. *Studies in Second Language Acquisition* 21: 181–193.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty and M. H. Long (eds.) *Handbook of second language acquisition*. Malden, MA: Blackwell, pp. 349–381.

Laufer, B. (1992). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability. *Journal of Research in Reading* (UKRA), 15.2: 95–103.

Macaro, E. (2003). *Teaching and learning a second language: a review of recent research*. London: Continuum.

MacWhinney, B. (2005). Extending the Competition Model. *International Journal of Bilingualism*, 9.1: 69–84.

Malvern, D. D. and Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19.1: 85–104.

Malvern, D., Richards, B., Chipere, N. and Durán, P. (2004). *Lexical Richness and Language Development: Quantification and Assessment*. Houndmills, Basingstoke: Palgrave Macmillan.

Meara, P. and Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16.3: 5–19.

Meara, P. and Milton, J. (2003). *X_Lex The Swansea Levels* Test. Newbury: Express.

Milton, J. (2006). Language Lite? Learning French vocabulary in school. *Journal of French Language Studies*, 16: 187–205.

Myles F. and Mitchell R. (eds.) (2004). *Second Language Learning Theories*, second edition. London: Arnold.

Myles, F. and Towell, R. (eds.) (2004). Special Issue on the Acquisition of French as a Foreign Language. *Journal of French Language Studies*, 14.3.

Pienemann, M. (1998). *Language Processing and Second Language Development*. Amsterdam/Philadelphia: Benjamins.

Prévost, Ph. and Paradis, J. (2004). *The Acquisition of French in Different Contexts*. Amsterdam/Philadelphia: Benjamins.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Ritchie, W. C. and Bhatia, T. K. (2005). *Handbook of Second Language Acquisition*. San Diego: Academic Press.

Truscott, J. and Sharwood Smith, M. (2004) Acquisition by processing: a modular perspective on language development. *Bilingualism: Language and Cognition*, 7.1: 1–20.

Wray, A. (2005). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Investigating lexical proficiency development over time – the case of Dutch-speaking learners of French in Brussels

BRAM BULTÉ, ALEX HOUSEN, MICHEL PIERRARD
AND SISKA VAN DAELE

Vrije Universiteit Brussel

ABSTRACT

This article aims (a) to explore the operationalisation and definition of lexical L2 proficiency and related constructs with a view to identifying a set of measures that can adequately capture the dynamics of lexical L2 proficiency development over time, and (b) to shed more light on the development of lexical proficiency in French Foreign Language classes. After a discussion of theoretical, terminological and methodological issues in L2 vocabulary research, we present a longitudinal quantitative study of the lexical development of Dutch-speaking adolescents learning FFL in Dutch-medium schools in Brussels over a three-year period and compare these learners' lexical proficiency in French to native speaker benchmarks.

1. INTRODUCTION

After years of relative neglect, the importance of vocabulary for developing proficiency in a second language (L2) is now generally acknowledged by researchers and theorists (e.g. Hatch and Brown, 1995; Coady and Huckin, 1997; Schmitt and McCarthy, 1997; Nation, 2001; Read, 2000; Malvern, Richards, Chipere and Durán 2004; Bogaards and Laufer, 2004). This importance is also acknowledged by language teachers, who often feel that their students, particularly those in foreign language learning contexts, are not developing their lexicons to levels that would permit them to function adequately in contexts beyond the language classroom (Carter and McCarthy, 1988; Arnaud and Bejoint, 1992; Laufer, 2005; Nurweni and Read, 1999; Ellis, 1994). This sentiment also prevails among teachers of French as a Foreign Language (FFL) in Belgium, the larger language learning context where the present study was conducted (Housen, Janssens and Pierrard, 2002).

However, in order to be able to promote lexical development in FFL classrooms, research first needs to determine the dynamics of the French lexical learning processes in educational contexts, and the factors that impact on these processes. To this end, longitudinal research is needed. Such research, however, is scarce. Most L2 vocabulary studies to date, whether on French or on other L2s, have

adopted a synchronic design, analysing lexical data from L2 learners at one point or over a brief period in time. Such studies can only provide information on aspects of lexical use, processing and representation, but not, or only indirectly, on lexical development over time. Thus, the present study was conducted to (a) explore the operationalisation and definition of lexical L2 proficiency and related constructs with a view to identifying a set of measures that can adequately capture the dynamics of lexical L2 proficiency development over time, and (b) shed more light on the development of lexical proficiency in FFL classes in the specific context of Dutch-medium schools in Brussels, Belgium.

To attain these two objectives we will first consider some basic theoretical, terminological and methodological issues in L2 vocabulary research (section 2). In section 3 the specific context in which the current study was conducted, Dutch-medium schools in Brussels, is outlined. The fourth and fifth sections present a longitudinal study of the lexical development of Dutch-speaking adolescents learning French as a Foreign Language in these Dutch-medium schools. Section 4 outlines the methodology of this study and section 5 presents analyses and results. The final section summarizes and discusses the findings and formulates implications for further research on L2 vocabulary development.

2. INVESTIGATING LEXICAL PROFICIENCY DEVELOPMENT — CONCEPTUAL, TERMINOLOGICAL AND OPERATIONAL ISSUES

A proper understanding of the results of any study of lexical proficiency requires one to be clear about central concepts, terminology and the nature of the measurements used. Unfortunately, terms such as *word*, *lexical proficiency*, *lexical knowledge*, *lexical competence*, *lexical richness*, are used rather loosely and interchangeably in the vocabulary literature, which hinders the interpretation and comparison of individual research findings. Therefore this section is concerned with three things: (a) construct specification (how can key constructs in L2 vocabulary research be further specified), (b) specific lexical assessment metrics used in L2 studies, and (c) the link between the first two.

Figure 1 represents an attempt to classify the different components of lexical competence that are relevant for the present study, and to illustrate how they can be related to lower-order constructs and to the concrete measures and tests commonly used in research on L2 vocabulary.

At the theoretical level of cognitive constructs, we propose to speak of lexical 'competence', which consists of both a declarative component and a procedural component. The declarative component has to do with lexical 'knowledge' as such, which can be further subdivided into the constructs 'size', 'width' and 'depth' of lexical knowledge. Size of knowledge refers to the number of lexical entries in memory (with varying degrees of mastery in terms of production and reception). Width and depth of knowledge both refer to the quality and degree of elaboration of that knowledge, both in terms of intensional relations to other entries in the lexicon (width) and in terms of the relations between the various form and meaning

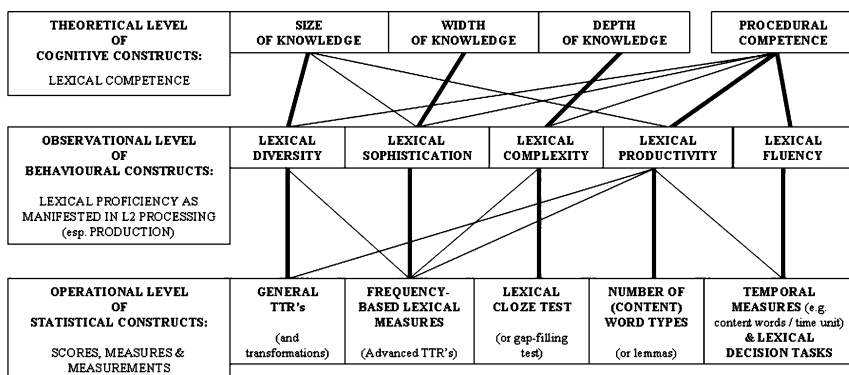| THEORETICAL LEVEL OF COGNITIVE CONSTRUCTS: LEXICAL COMPETENCE	SIZE OF KNOWLEDGE	WIDTH OF KNOWLEDGE	DEPTH OF KNOWLEDGE	PROCEDURAL COMPETENCE
OBSERVATIONAL LEVEL OF BEHAVIOURAL CONSTRUCTS: LEXICAL PROFICIENCY AS MANIFESTED IN L2 PROCESSING (esp. PRODUCTION)	LEXICAL DIVERSITY	LEXICAL SOPHISTICATION	LEXICAL COMPLEXITY	LEXICAL PRODUCTIVITY	LEXICAL FLUENCY
OPERATIONAL LEVEL OF STATISTICAL CONSTRUCTS: SCORES, MEASURES & MEASUREMENTS	GENERAL TTR's (and transformations)	FREQUENCY-BASED LEXICAL MEASURES (Advanced TTR's)	LEXICAL CLOZE TEST (or gap-filling test)	NUMBER OF (CONTENT) WORD TYPES (or lemmas)	TEMPORAL MEASURES (e.g. content words / time unit) & LEXICAL DECISION TASKS

Figure 1. *Analytic framework for investigating L2 lexical proficiency development.*

components of a given entry (depth). The second component, procedural lexical competence, is more a matter of skill and control over knowledge and refers to how strongly linguistic information is stored in lexical memory, which in turn determines how well learners can access, retrieve and encode/decode relevant lexical information in real time.

Lexical competence and its various subcomponents are theoretical constructs which are not open to direct observation or measurement. They are inferred from lower-order constructs which are taken to be the behavioural manifestations of the underlying cognitive constructs in actual L2 performance (production and reception). For the behavioural correlative of lexical competence in actual L2 use we propose the term *lexical proficiency*. As shown in Figure 1, a learner's lexical proficiency can be construed in terms of the lexical diversity, sophistication, complexity, productivity and fluency of his L2 use.

In our terms, lexical diversity (or variation) is the observational correlate of knowledge size and refers to the impression of a learner's lexical proficiency produced by the number of *different* words (or phonologically-orthographical different word forms) which he or she uses. Lexical sophistication refers to the perception of a L2 user's lexical proficiency formed by, among other things, his use of semantically more specific and/or pragmatically more appropriate different words from among a set of related words. In this sense, lexical sophistication implies knowledge of semantic relations (e.g. in terms of hyponoymy, hypernymy, synonymy, antonymy) and, hence, of different but related lexical alternatives for referring to a referent. In these respects, lexical sophistication manifests itself at the macro-level of the lexicon as a whole, making it the nearest correlate of what we termed lexical width. In contrast, lexical complexity manifests itself at the micro-level of the individual *word*. It refers to the impression of someone's lexical proficiency created by, for instance, the ability to comprehend or use not only the prototypical or default semantic, collocational, grammatical or pragmatic aspects of a specific word but also a variety of other, more specific, peripheral and less frequent

properties. In this sense, lexical complexity corresponds to a learner's depth of lexical knowledge. Lexical productivity is a construct at the behavioural–observational level of lexical production which reflects the sheer number of words (tokens) which a speaker uses to complete a given task. For instance, in a given context some learners will describe a given event by using 100 words while others will use 200 words. Finally, the related construct of lexical fluency refers to the speed with which the learner produces or encodes (esp. content) words, as determined by the degree to which the relevant lexical information has been proceduralised.[1] As with lexical competence, more dimensions of lexical proficiency can be distinguished than the five discussed here, but these five already suffice to show that any concrete measure of L2 lexical proficiency is bound to be only partial.

The need for reliable and valid measures for lexical proficiency has long been recognised in SLA research, particularly in the area of L2 production (e.g. Harley, 1995; Malvern *et al.*, 2004; Daller, Milton and Treffers-Daller, 2007). Without such measures several theoretical and practical questions cannot be answered. These questions include determining the number of words 'known' by learners at a particular stage of development (or level of instruction), the rate at which words are acquired, the contribution of factors such as word length, frequency and saliency to the ease/difficulty with which a word is learned, and the relationship between vocabulary size and other aspects of lexical competence (e.g. depth, width, automatisation of vocabulary knowledge). As a result, the number of L2 vocabulary measures and tests has proliferated in recent years, tapping into different dimensions of lexical proficiency. The lower level boxes in Figure 1 are an attempt to classify different types of measures (with an emphasis on *productive* lexical proficiency), and to relate them to the various higher-order constructs discussed earlier. Measures of productive vocabulary size/diversity (*how many words are known?*) are generally based on analyses of oral or written L2 corpora and usually take the form of a general type-token ratio (TTR) and transformations thereof (e.g. Guiraud Index, Über Index), or on an analysis of how fast TTR falls with increasing token count within the language sample (D). These measures have been widely used and discussed in the L2 literature (e.g. Malvern *et al.*, 2004; Van Hout and Vermeer, 2007; Broeder, Extra and Van Hout, 1993; McCarthy and Jarvis, 2007). Other elicitation measures of productive vocabulary size include translation tests and productive vocabulary level tests (Laufer and Nation, 1999). Measures of lexical width/sophistication (*which words are known*) and of depth/complexity *(how well words are known)* focus more narrowly on, respectively, specific or 'advanced' words (e.g. less frequent words) or on specific aspects of the meaning or use of words (cf. the Advanced Guiraud as discussed in Daller *et al.*, 2003 and Tidball and Treffers-Daller, 2007 and this special issue). Each type of measure and each operationalisation has its own inherent

---

[1] Lexical productivity and fluency, as defined here, are not only a matter of the learner's lexical competence (esp. size of knowledge and procedural knowledge) but also of 'external' learner-variables such as task motivation, degree of extraversion, willingness to communicate and personal speech style.

strengths and problems, in terms of validity, reliability, discrimination and feasibility. These strengths and problems are well documented in the literature, and some of these will be taken up in the methodology section. Suffice it here to point out that there is no one-to-one match in Figure 1 between the statistical constructs and the observational-behavioural constructs which these statistical constructs are supposed to measure (nor, for that matter, between the constructs at the observational and theoretical levels). This is indicated by the multiple lines starting from the lower-level boxes in Figure 1. In fact, one could argue that all the quantitative measures listed in Figure 1 say something about all four behavioural constructs distinguished in this framework (which points to a general validity problem) though some measures clearly say more about a given construct than others. This is reflected by the weight of the interconnecting lines.

3. FRENCH AS A FOREIGN LANGUAGE IN BRUSSELS

The French learners whose lexical proficiency is discussed in this article learn French in a specific context, namely Dutch-medium education in Brussels, Belgium. The issue of language learning and teaching in Belgium education has been characterised as paradoxical (Baetens Beardsmore, 1993), and nowhere is this paradox clearer than in Brussels, the officially Dutch-French bilingual but predominantly francophone capital of Belgium. In Brussels, two separate monolingual school systems, one in Dutch and one in French, operate in parallel and independently from each other. The Dutch school system in Brussels was originally established in the 1960s to help maintain the Dutch-speaking minority in Brussels but in recent years an increasing number of Francophone parents and parents from other language backgrounds have sent their children to the Dutch-language schools in the expectation that this will provide their children with the Dutch immersion experience felt necessary to become bilingual in both national languages. This trend has dramatically changed the demographics and the linguistic climate in the Dutch-medium schools in Brussels (henceforth *DuB*) to the extent that French-speaking pupils now begin to outnumber the Dutch-speaking pupils who are becoming a minority in most classrooms. This trend is most salient in kindergarten and primary school, where the number of children from Dutch-speaking families has dropped in 25 years from over 78% to only 14% in 2007. This trend has also now gradually spread to secondary schools, where Dutch-speaking children now represent only 34% of the population (*www.vgc.be*). All this has drastically changed the linguistic balance in Dutch-medium education in Brussels. In most schools French has now become the dominant language in the playground and in the corridors though Dutch is still the only language of instruction (except for in the French-foreign language lessons; cf. below), with little or no provisions for the French-speaking pupils. The general public and policy makers in Belgium believe that the complex multilingual school environment that has thus been created has a profound and overall negative effect on the linguistic and the academic development of the French-speaking and Dutch-speaking pupils alike (see Housen and Pierrard, 2004; Mettewie, 2007; Mettewie,

Housen and Pierrard 2005). The only potentially positive effect, some believe, would be on the Dutch-speaking pupils' proficiency in French which would, with time, perhaps even approximate that of their francophone classmates. Whether, and to what extent this last assumption is correct, is of particular concern for the French-Foreign Language teachers in these Dutch-medium schools. It is this issue which the present study seeks to address, focusing on the lexical proficiency in French of Dutch-speaking pupils in comparison to their francophone peers in the first three years of Dutch-medium secondary education in Brussels.

Reasons for believing that the particular multilingual context in the Dutch-medium schools in Brussels may have positive effects on the lexical development of French of the Dutch-speaking pupils are related to the different sources of exposure to French which these pupils have at their disposal. The first source of exposure for these Dutch-speaking pupils is the French-Foreign Language classroom. Much has been written about the constraints on lexical development imposed by the foreign language classroom context, and the role and efficacy of incidental versus explicit vocabulary learning (e.g. Huckin and Coady, 1999; Gass, 1999; Hunt and Beglar, 2005; Laufer, 2001, 2005; Folse, 2004). Compared to both naturalistic first and second language learners, foreign language (FL) learners often lack an adequate amount of contextualised oral or written input from which they can extract and create the relevant semantic and structural specifications about words and integrate such information in their mental lexicons; consequently, to compensate for FL learners' lack of naturalistic learning opportunities, FL teaching typically resorts to *explicit* instruction.

FFL lessons in Dutch-medium education in Brussels, where Dutch-speaking pupils sit together with their French-speaking peers, start in Year 3 of primary school and are typically taught by non-native and non-specialist teachers for two or three hours a week. Traditional rationalistic-analytic teaching practices prevail in the FFL lessons at primary school, partly because of the spread of one particular textbook *cum* method (Decoo, 2004) which is specially developed for the Dutch-speaking Flemish market and which adheres to the principles of the audio-lingual and direct methods. In this method, vocabulary and grammar receive equal emphasis and are presented explicitly and developed systematically (Housen, 2003). In the first three years of secondary school, French is taught for four hours a week by non-native specialist teachers. Here, teaching syllabi and methods take a more communicative and functional-notional turn but there is also analytic grammar work and a systematic and focused approach to vocabulary building, including the study of decontextualised lexis (e.g. rote memorisation of word lists, sentence-gap filling exercises, translation of Dutch L1 words), the provision of concise definitions of new French words, using dictionaries and consciously inferring vocabulary from context (Kemps, Housen and Pierrard, in press).

In addition to the FFL lessons, the Dutch-speaking pupils in DuB have other contact sources with French which are not normally available in other foreign language teaching contexts. One such additional source is through the informal contacts with their native French-speaking peers in the playground, in the corridors

and often also in class. As mentioned earlier, in many of these Dutch-medium schools, French has become the *de facto* lingua franca for informal pupil interaction. This potentially provides the Dutch-speaking pupils with authentic linguistic role models and rich, varied, and accurate input for learning French. Thirdly, there is the predominantly French-speaking environment outside the Brussels schools, which provides yet another potentially rich source of authentic input and output opportunities for learning French.

In sum, the learning of French by Dutch-speaking pupils in Dutch-medium schools in Brussels can be characterised as a mixture of both *instructed* and *naturalistic* L2 learning, and their lexical development in French as involving both formal and informal learning processes.

4. METHODOLOGY

The study reported here is exploratory in nature and part of a larger research project on the development of L2 proficiency in various educational contexts (e.g. Housen, 2002; Housen, Janssens and Pierrard, 2003; Mettewie *et al.*, 2005; Kemps *et al.*, in press). The present study specifically focuses on the development of lexical diversity, lexical productivity and lexical sophistication of Dutch-speaking pupils learning French in Brussels and compares this development to benchmark data obtained from their French native speaker peers. The central research questions are the following:

1) How does the lexical proficiency of Dutch-speaking L2 learners of French in DuB develop over time in terms of the diversity, sophistication and productivity of their French L2 production?
2) How does the lexical proficiency in French of these Dutch-speaking L2 learners compare to that of their native French-speaking peers?
3) Which measures of lexical proficiency explain the most variance in the Dutch-speaking pupils' lexical proficiency development over time?

4.1 Participants and Data Collection

The French data analysed in this study come from the unplanned oral retellings of the wordless picture story *Frog, Where Are You?* (Mayer, 1969; see also Berman and Slobin, 1994) by 19 Dutch-speaking and 19 French-speaking pupils from ten different Dutch-medium secondary schools in Brussels. These pupils were selected from a larger sample on the basis of their home language (either Dutch or French with their parents) and time spent in Dutch-medium education in Brussels (at least since Year 1 of primary school). At the time of the first data collection, these pupils had all been taught French as a Foreign Language for at least four years for two to four 50-minute sessions a week. In addition, they had had varying amounts of naturalistic exposure to French (cf. section 3).

The progress in French of the Dutch-speaking pupils (the target group) was tracked over a period of two school years, starting when they were in Year 1 of secondary school (age 12) and ending when they were in Year 3 (age 14). They were administered the Frog story task three times, once every school year (Y1, Y2, Y3). The French proficiency of the French-speaking pupils was evaluated only once, in Year 2, to provide a benchmark for the Dutch-speaking pupils.[2] The total corpus for this study thus consists of 76 data sets.

4.2 Data processing

The French oral speech data which the pupils produced during the retell task were recorded and subsequently transcribed, segmented and annotated in CHAT format. Next the transcriptions were analysed with the help of CLAN software (MacWhinney, 2000). As a matter of convenience when reporting research, we use as a unit of analysis the 'word' in the sense of lexical entry (Jiang, 2000). The aim of the analysis was to identify the number and type of words known by the Dutch-speaking L2 learners of French, asking whether they are different from those known by the native French-speaking pupils, with a special focus on their productive lexical knowledge.

To this end, root or base forms (e.g. *grenouille, sauter*) and their related inflected forms (e.g. *grenouilles, saute*) were counted as a single unit (i.e. as different tokens of the same type). Accordingly, all content word tokens in the transcriptions were lemmatised, that is, reduced to their root form or, in the case of inflected verb tokens, to their infinitive form. Finally, all nouns, verbs and adjectives tokens were also coded for part of speech.[3]

4.3 Analytic procedures

4.3.1 Measures

The overall research approach used in our study was to identify a number of properties of learners' oral production while carrying out the narrative task which allowed for quantitative and qualitative characterisations of the Dutch- and French-speaking pupils' lexical proficiency in French on a cross-sectional and a developmental basis. To this end, a total of 42 different quantitative measures were initially calculated, covering the various components of lexical proficiency

---

[2] In spite of the fact that the French-speaking pupils in this study had had their basic education in a second language (Dutch), their lexical proficiency in French was found to be comparable (though not equivalent) to that of monolingual French L1 pupils in Wallonia (Mettewie *et al.*, 2005).

[3] Implicit in our approach is an admittedly simplistic and narrow working definition of lexical proficiency and of what is meant by learning and knowing a word: a learner will be said to have acquired and know an L2 word when he is able to produce its phonological form in a given production task. We acknowledge that this is only part of what it takes to know a word.

Table 1. *Overview of lexical proficiency measures*

Standard measures	Class measures	Frequency-based measures
#-All types (*prod*)	#-Nouns (*prod*)	#-FF1 (*soph, prod*)
#-Content types (*prod*)	#-Verbs (*prod*)	#-FR5 (*soph, prod*)
U-All (*div*)	#-Adjectives (*prod*)	#-FR66 (*soph, prod*)
U-Content (*div*)	G-Nouns/Nouns (*div*)	G-FF1 (*soph, div*)
D-All (*div*)	G-Nouns/All (*div*)	G-FR5 (*soph, div*)
G-All (*div*)	G-Verbs/Verbs (*div*)	G-FR66 (*soph, div*)
G-Content (*div*)	G-Verbs/All (*div*)	
	G-Adjs/Adjs (*div*)	
	G-Adjs/All (*div*)	

distinguished in Figure 1. However, for reasons of space, only 22 of these measures are included in this article. These 22 measures are listed in Table 1, along with an indication of the component of lexical proficiency which they are primarily intended to gauge, viz. lexical diversity, sophistication or productivity. Each of these 22 measures was calculated for the four groups of French data sets in our corpus: the 19 transcriptions of the Dutch-speaking L2 learners in Years 1, 2 and 3, and the 19 transcriptions of the French native speaking pupils.

All measures are either extrinsic frequency measures (indicated in Table 1 by #) or ratio measures (U for Uber Index, G for Guiraud Index), or include the intrinsic frequency distribution of types (like D).

For ease of discussion, we further divided the 22 measures into three categories: 1) 'standard' measures, which simply distinguish between word types and tokens at the level of the lexicon as a whole, 2) lexical 'class' measures, which focus on a particular lexical class (nouns, verbs, adjectives), and 3) 'frequency' measures, which take into account external information about words (i.e. their relative frequency of occurrence) and distinguish between 'basic' versus 'advanced' words.

In our view, lexical proficiency has to do with the knowledge and use of linguistic form units with a specifiable, self-contained semantic-conceptual meaning or, in other words, with semantic *content* words rather than with grammatical *function* words. Knowledge of *function* words essentially pertains to a learner's grammatical competence. The distinction between content and function words is often blurred at the operational level of the quantitative measures that are used in many L2 vocabulary studies. In line with this view, measures of productive lexical proficiency should for the sake of validity preferably be based on an analysis of content words only, rather than on the total number of words used (i.e. content + function words). However, to allow for comparison with other studies, most of which have calculated measures on the basis of the total number of word types and tokens, several measures in this study were computed twice, once for all content and function word types (indicated by 'All' in Table 1) and once for content words only. Moreover, some measures, such as *D*, were difficult to calculate for content words only for practical reasons (e.g. some data sets did not contain the minimum number of 50 content tokens required for the calculation of D). A statistical comparison between the

measures calculated for all words (e.g. G-All, U-All) and those calculated for the content words only (e.g. G-Content; U-Content) revealed that most of these measures were, in fact, highly correlated (with r-values in excess of 0.90).

The set of 'standard' measures include the following seven measures: total number of different word types, total number of different content word types (nouns, adjectives and verbs), Uber's index (U) and Guiraud's index (G) (both calculated for all word types and for content word types only), and finally *D*.

Previous studies suggest that increases in lexical diversity and productivity may be mainly due to lexical growth in one or two content word classes only (e.g. Broeder *et al.*, 1993), though it is not entirely clear which lexical class is most prone to such developmental fluctuations nor what the link is with factors such as task type or the L2 learner's general stage of development. To further explore this issue, we calculated separate measures for the three major lexical content classes (nouns, adjectives, verbs). These 'class' measures are listed in the second column of Table 1. For the ratio measures, only Guiraud's index was calculated,[4] albeit again in two versions for each class, once within the word class and once across all word classes (i.e. relative to text length), using the following formulae:

$$G(N/N) = \frac{noun\ types}{\sqrt{noun\ tokens}} \qquad G(N/All) = \frac{noun\ types}{\sqrt{all\ tokens}}$$

Frequency-based vocabulary measures attempt to capture qualitative aspects of learners' productive lexical proficiency, namely the less frequent and, hence, potentially semantically more specific or more abstract words. This makes these measures potential indicators of what we call lexical sophistication (hence the alternative term 'advanced' measures; Daller *et al.*, 2003). Since there is much discussion about the quality of the word frequency lists on which such measures are based, we again explored different options. First, we used the *Français Fondamental Premier Degré* (FF1) list as discussed by Tidball and Treffers-Daller (2007). This frequency list can be labelled 'extrinsic' in the sense that it was compiled independently from the speech corpus under investigation in this study. The FF1 list contains 1445 'frequent' French words drawn from an established corpus of spoken native French. In addition to the FF1 list, we compiled two additional frequency lists on the basis of our own corpus of French narrative speech data. A first list, called FR66, contains all the types that appear at least once in two-thirds (66%) of the 76 native and non-native speaker transcriptions in our corpus. The FR66 list contains the 43 most frequent types in the entire corpus. The second frequency list is called FR5 and contains all words that appear five times or more in the 19 transcriptions of the French native speaker pupils in our sample. This list includes 149 word types. Both the FR66 and the FR5 frequency lists are 'intrinsic' in the

---

[4] The Uber Index was also calculated but the correlations between the Uber and Guiraud values were sufficiently high (<0.90) to justify the inclusion of only one, the Guiraud Index.

sense that they are compiled from the same speech data to whose analysis they will be applied. The advantage of 'intrinsic' frequency lists is that they are context and task-specific and respect the integrity of the text samples analysed.

All three frequency lists were used to distinguish between 'frequent' and 'rare' words in each of the 76 transcriptions. We tallied not only the number of 'rare' word types but also their relative frequency of occurrence by using the following derivative of the Guiraud Index :

$$GFF1 = \frac{\textit{types not in FF1}}{\sqrt{\textit{all tokens}}}$$

4.3.2 *Statistical analyses*

For each of the four data sets, mean scores were computed for each of the 22 measures. After checking for normality of distribution by means of the Kolmogorov-Smirnov test, a series of repeated measures ANOVA (with 'Year' as within-subjects factor) were performed to investigate how the lexical proficiency in the French oral narrative production data of the Dutch-speaking pupils developed over time. Their development was assessed in more detail by calculating Bonferroni pairwise comparisons (corrected paired sample t-tests) for the different test moments (Year 1–Year 2, Year 2–Year 3 and Year 1–Year 3). Paired sample t-tests were used to compare the scores of the Dutch-speaking target group at the three different test moments with the scores of the French native-speaker benchmark group.

5. RESULTS

The presentation of the results is organised per category of measure ('standard', 'class' and 'frequency'). A first table contains the descriptive statistics (mean scores, standard deviations) for each of the measures for the four data sets (Year 1, 2 and 3 for the Dutch-speaking learners, and the native benchmark group). A second table presents the inferential statistics (RM ANOVA: significance, F-value between brackets; Y1 vs Y2, Y2 vs Y3 and Y1 vs Y3: significance of Bonferroni pairwise comparison). Statistically significant differences (at $p < 0.05$) are marked by three asterisks. The final part of this section presents a correlation matrix for the 22 measures.

5.1 *'Standard' measures*

The repeated measure ANOVA analyses indicate that the factor 'Year' has a significant effect. This means that there is a significant increase in the scores of all seven 'standard' measures of lexical proficiency of the Dutch-speaking pupils over the three observation periods. The pairwise comparisons further indicate that although the increase on all measures is significant over the entire observation period (from Year 1 to Year 3), the increase is most pronounced in the course of the first year (Year 1 to Year 2). The increase on the two frequency measures of

Table 2. *Standard measures – descriptive statistics*

Measure	Year 1	Year 2	Year 3	Benchmark
#–All types	58.53 (19.53)	66.00 (32.39)	72.47 (24.76)	103.53 (15.90)
#–Content types	28.89 (9.98)	35.89 (18.62)	39.79 (14.25)	56.32 (9.45)
U–All	9.33 (1.51)	10.07 (1.58)	10.25 (1.40)	12.55 (0.79)
U–Content	8.13 (1.82)	9.34 (2.14)	9.72 (1.85)	12.18 (1.33)
D–All	15.80 (7.90)	18.39 (8.63)	19.56 (8.18)	33.39 (5.79)
G–All	3.82 (0.83)	4.25 (0.91)	4.37 (0.82)	5.64 (0.40)
G–Content	3.15 (0.74)	3.67 (0.91)	3.87 (0.80)	4.86 (0.45)

Table 3. *Standard measures – inferential statistics*

Measure	RM ANOVA	Pair-wise comparisons			NS vs NNS
		Y1 vs Y2	Y2 vs Y3	Y1 vs Y3	(Y1, Y2, Y3)
#–All types	0.004*** (6.419)	0.371	0.244	0.002***	$p < 0.05$ (3x)
#–Content types	0.000*** (10.431)	0.082	0.239	0.000***	$p < 0.05$ (3x)
U–All	0.000*** (14.915)	0.000***	1.000	0.000***	$p < 0.05$ (3x)
U–Content	0.000*** (13.578)	0.003***	0.914	0.000***	$p < 0.05$ (3x)
D–All	0.001*** (9.040)	0.020***	0.288	0.012***	$p < 0.05$ (3x)
G–All	0.000*** (20.242)	0.000***	0.722	0.000***	$p < 0.05$ (3x)
G–Content	0.000*** (21.807)	0.000***	0.267	0.000***	$p < 0.05$ (3x)

lexical productivity (number of different word types, number of different content word types) is significant neither over the first nor over the second year. The ratio measures of lexical diversity (U and G) and D all show a significant increase over the first year, but not over the second. Finally, the last column in Table 3 shows that the French L1 benchmark group significantly outperforms the Dutch-speaking L2 group on every measure and at each of the three testing moments.

5.2 'Class' measures

Table 4 shows that the verbs are the most frequent word class in our corpus, followed by the nouns and then the adjectives. Adjectives are very much a minor

Table 4. *Class measures – descriptive statistics*

Measure	Year 1	Year 2	Year 3	Benchmark
Nouns				
#–Nouns	11.42	14.95	16.32	23.84
	(3.20)	(7.96)	(5.68)	(4.59)
G (N/N)	1.98	2.26	2.41	3.14
	(0.57)	(0.66)	(0.60)	(0.34)
G (N/All)	0.76	0.96	0.99	1.30
	(0.15)	(0.28)	(0.23)	(0.19)
Verbs				
#–Verbs	15.37	18.16	20.37	29.42
	(6.29)	(9.48)	(7.85)	(5.50)
G (V/V)	2.36	2.69	2.82	3.55
	(0.57)	(0.68)	(0.62)	(0.48)
G (V/All)	1.00	1.17	1.22	1.60
	(0.30)	(0.33)	(0.28)	(0.23)
Adjectives				
#–Adjectives	2.11	2.79	3.11	3.05
	(1.52)	(2.25)	(1.82)	(1.58)
G (A/A)	0.95	1.20	1.27	1.18
	(0.45)	(0.66)	(0.43)	(0.46)
G (A/All)	0.13	0.17	0.19	0.17
	(0.08)	(0.12)	(0.08)	(0.08)

phenomenon, both in the learner and the native speaker benchmark corpus (on average 2 or 3 adjectives per dataset). The statistical analyses of the lexical class measures in Table 5 yields a slightly different picture than that of the standard measures. The ANOVAs indicate that the Dutch-speaking pupils significantly progress over the total observation period on all measures except for one of the two lexical diversity measures for adjectives. However, the pairwise comparisons show that the pupils' overall lexical growth is mainly due to the increase on the lexical diversity measures for the noun and verb classes in the course of the first year (Year 1 to Year 2) only, and that no further significant progress was made on any of the measures over the second year (Year 2 to Year 3). Finally, the paired sample t-tests reveal that at each of the three testing moments the French L1 benchmark group significantly outperforms the Dutch-speaking L2 group on the productivity and diversity measures for nouns and verbs. No significant differences between the learners and the French L1 speakers were revealed for the measures focusing on the adjectives.

5.3 'Frequency' measures

The general trends which emerge from the analysis of the frequency measures, primarily intended to gauge the pupils' more 'advanced' vocabulary, correspond to that of the standard and class measures presented in the previous sections: the Dutch-speaking learners' scores steadily increase in time though they never attain the values of the native-speaker benchmark pupils (Table 5). The ANOVAs indicate

Table 5. *Class measures – inferential statistics*

Measure	RM ANOVA	Pair–wise comparisons			NS vs NNS (Y1, Y2, Y3)
		Y1 vs Y2	Y2 vs Y3	Y1 vs Y3	
Nouns					
#–Nouns	0.001*** (8.257)	0.083	0.856	0.000***	p < 0.05 (3x)
G (N/N)	0.001*** (9.035)	0.038***	0.608	0.001***	p < 0.05 (3x)
G (N/All)	0.000*** (13.170)	0.003***	1.000	0.000***	p < 0.05 (3x)
Verbs					
#–Verbs	0.002*** (7.523)	0.210	0.150	0.005***	p < 0.05 (3x)
G (V/V)	0.000*** (11.754)	0.017***	0.411	0.001***	p < 0.05 (3x)
G (V/All)	0.000*** (10.815)	0.010***	0.768	0.001***	p < 0.05 (3x)
Adjectives					
#–Adjectives	0.040*** (3.523)	0.492	1.000	0.004***	p > 0.05 (3x)
G (A/A)	0.040*** (3.533)	0.151	1.000	0.025***	p > 0.05 (3x)
G (A/All)	0.076 (2.776)	0.481	1.000	0.004***	p > 0.05 (3x)

Table 6. *Frequency-based measures – descriptive statistics*

Measure	Year 1	Year 2	Year 3	Benchmark
#–FF1	3.58 (3.67)	6.11 (6.74)	8.53 (6.92)	17.58 (4.85)
#–FR5	10.63 (6.76)	13.05 (14.47)	14.42 (11.66)	20.89 (7.09)
#–FR66	29.16 (14.99)	34.16 (27.81)	38.52 (21.89)	66.37 (14.61)
G–FF1	0.22 (0.23)	0.35 (0.30)	0.49 (0.32)	0.96 (0.25)
G–FR5	0.67 (0.36)	0.76 (0.54)	0.83 (0.50)	1.12 (0.31)
G–FR66	3.91 (1.15)	4.41 (1.50)	4.80 (1.28)	6.40 (0.60)

that the Dutch pupils' progress on the measures based on the extrinsic FF1 frequency list and the intrinsic FR66 frequency list is significant. The measures on the basis of the FR5-list (which is derived from the French-speaking pupil sub-corpus) does not yield any significant differences. The picture emerging from the pairwise comparisons is less straightforward. Both the number and ratio measure of advanced word types based on the FF1-list significantly increase in the course of the first,

Table 7. *Frequency-based measures – inferential statistics*

		Pair–wise comparisons			
Measure	RM ANOVA	Y1 vs Y2	Y2 vs Y3	Y1 vs Y3	NS vs NNS (Y1, Y2, Y3)
# FF1 types	0.000*** (16.517)	0.030***	0.014***	0.000***	$p < 0.05$ (3x)
# FR5 types	0.134 (2.122)	0.964	0.965	0.124	$p < 0.05$ (3x)
# FR66 types	0.028*** (3.939)	0.707	0.353	0.021***	$p < 0.05$ (3x)
G FF1	0.000*** (25.850)	0.004***	0.007***	0.000***	$p < 0.05$ (3x)
G FR5	0.065 (2.955)	0.662	0.807	0.072	$p < 0.05$ (3x)
G FR66	0.000*** (13.516)	0.006***	0.156	0.000***	$p < 0.05$ (3x)

second and third year. The same measures calculated on the basis of the FR66-list significantly advance only over the first year (Year 1 to Year 2) and over the entire observation period (Year 1 to Year 3). Finally, the last column of Table 5 shows that the French L1 benchmark group once more outperforms the Dutch-speaking learner group on all measures and for all observation periods.

5.4 Correlations

This final section examines the correlations between the different measures from all data sets combined. Table 8 in the appendix shows that all correlations are significant and positive and in some cases very strong. This is the case for the correlation between the number of types and the number of content types (0.98), the correlations between G and U calculated for all words and G and U calculated for content words (respectively 0.96 and 0.95) and for the correlations between the number of noun lemmas (0.92) and verb lemmas (0.95). Especially the 'standard' measures of lexical diversity (G, U and D for all words) are strongly intercorrelated (between 0.95 and 0.99) and they correlate well with the total number of lemmas (0.92 for G, 0.88 for U and 0.92 for D). The lowest correlation coefficients are found for the measures involving adjectives.

5.5 Summary of the results

In summary, results of the statistical analyses reveal certain trends both in relation to the overall development of the L2 learners over time and to the lexical proficiency of the learners and the native speakers compared. A few exceptions aside,[5] the repeated

---

[5] The exceptions in case are G of adjective tokens to all word types, all word types and G based on the FR5 frequency list.

measures ANOVA (with time as within-subjects factor) show that these learners of French progressed significantly on all measures. Bonferroni pairwise comparisons for the different test moments further reveal that the learners' overall development was stronger over the first year (with significant progress on 12 measures evenly distributed over the four dimensions of lexical proficiency they were intended to gauge) than over the second year (learners only progressed for two of the frequency measures, namely the number of non-frequent word types in the FF1 list and the diversity of these 'advanced' word types relative to the text length). Finally, a series of paired sample t-tests shows that except for the 'class' measures of adjective production, the native speakers of French scored consistently and significantly higher than the learners.

6. DISCUSSION AND CONCLUSION

The present study sought to investigate the lexical proficiency in French of the Dutch-speaking pupils vis-à-vis their Francophone classmates in the first three years of Dutch-medium schools in Brussels. A first goal of this study was to explore the operationalisation and definition of lexical L2 proficiency and related constructs with a view to identifying which measures of productive lexical proficiency best capture productive lexical proficiency development in L2 French over time. In future work the model of lexical competence and proficiency proposed in section 2 will be situated within broader models of language proficiency and competence (e.g. Bachman and Palmer, 1996; Hulstijn, 2007). In this article we focused on the strength of different statistical measures of lexical proficiency, particularly in terms of discriminatory power and validity. In this respect the findings of our study suggest that:

1. all selected 'standard', 'class' and 'frequency' measures (except for the measures based on the FF5 list) allow to demonstrate differences between different stages of L2 development as well as between L2 learners and L1 users.
2. most measures but particularly the standard measures of lexical diversity correlate strongly (at least for the –admittedly restricted– data sets in the present study).

These findings imply that the lexical measures calculated here all have longitudinal and cross-sectional discriminatory power. At the same time, however, the observation that all measures are strongly correlated appears to confirm our suspicion that they tap into similar aspects or dimensions of lexical proficiency, which poses problems in terms of construct validity. Particularly the measures proposed here as indicators of lexical sophistication appear to have low construct validity. As already indicated in Table 1, these frequency-based measures do not only measure lexical sophistication (or, rather, aspects thereof) but also other dimensions of lexical proficiency such as diversity and productivity. Moreover, although the frequency-based vocabulary measures can capture some qualitative aspects of learners' productive lexical proficiency (by looking at the less frequent,

more advanced and potentially semantically more specific or more abstract words), lexical sophistication also refers to extensional semantic relations between these words that cannot be captured by referring to frequency lists alone. What is needed then, is greater conceptual clarity as to what lexical sophistication entails and more varied and sophisticated measures of this dimension of lexical proficiency than the ones used in this study.

A second aim of this study was to longitudinally explore the development of French productive lexical proficiency in terms of the diversity, sophistication and productivity of Dutch-speaking pupils vis-à-vis their Francophone classmates in the first three years of secondary education in Dutch-medium schools in Brussels. To this end oral French production data from 19 Dutch-speaking French learners were collected on three occasions at twelve months intervals and compared to the oral production data from 19 native French speakers who attended the same classes. Twenty-two quantitative measures tapping into different components of lexical proficiency were computed and statistically analysed. With respect to the Dutch-speaking pupils' lexical development, the following outcomes appear to be particularly noteworthy:

1. Learners progressed significantly in terms of lexical diversity, sophistication and productivity;
2. The development of lexical proficiency, particularly in terms of lexical productivity and diversity trails off after the first year.

Although these findings clearly indicate an improvement in selected aspects of the French lexical proficiency of the Dutch-speaking pupils over the first year of secondary education, an explanation for the relatively lower scores obtained in the third year is not immediately apparent. Perhaps this last finding can be attributed to methodological aspects of our study. It is not unlikely that after having told the frog story in Year 1 and Year 2 the learners experienced diminished task motivation the third time around and invested less effort in creating a lexically productive and diverse narrative (even though the interviewers were different at each occasion and pretended to be unfamiliar with the story). This point merits further exploration. In future analyses we will include available data on the pupils' attitudinal and motivational dispositions to examine not only individual variation in lexical development but also to what extent the behavioural manifestations of lexical competence are influenced by non-linguistic learner variables. In addition, future work will also include qualitative analyses of the pupils' lexical productions to identify more subtle changes in and differences between the learners' and native speakers' lexical repertoires that may have been overlooked by the quantitative analyses of the present study.

Although at the end of the study the Dutch-speaking pupils had had seven years of FFL instruction and at least nine years of fairly intensive informal contact with French-speaking peers, and although the French-speaking pupils had had their basic education in a second language (Dutch), the findings of this study suggest that the two groups still differ markedly in their French lexical proficiency.

At the end of the observation cycle the native French pupils still outperform the L2 learners on all dimensions investigated except on the measures for lexical productivity and diversity of adjectives. This last finding may again be attributed to methodological issues related to the task design. The fact that adjectives are a minor phenomenon in the French productions of both the Dutch-speaking target group and the French native-speaker benchmark group points to the limitations of the specific data elicitation procedure used in this study. An oral retell task on the basis of a wordless picture story such as the popular Frog story creates too few obligatory or appropriate contexts for using adjectives for any potential differences between the Dutch-speaking pupils and the French native-speaking group to emerge. More generally, this points to the crucial importance, not only of selecting reliable, valid and sensitive proficiency measures, but also of selecting data collection procedures that elicit lexically rich and varied speech.

Methodological limitations notwithstanding, the outcomes of this study unambiguously indicate that even after nine years of intensive and varied contact with French the Dutch-speaking pupils had not yet attained the same level of lexical productivity, diversity and sophistication as their native speaker peers. This observation is obviously not to be taken as a sign of failure on part of the specific pupils, their teachers, the schools or the educational system involved, the more so as the levels of lexical proficiency attained may well be high when compared to FFL contexts elsewhere (a possibility which will be empirically investigated in a follow-up study comparing the French lexical proficiency of Dutch-speaking pupils in Brussels with that of Dutch-speaking FFL learners in monolingual Flanders). Rather, these findings demonstrate that lexical development in a second or foreign language is not always a given, not even in contexts as seemingly conducive to L2 learning as the learning of French in Brussels. More specifically, these findings raise questions about what can be realistically expected from the contribution of formal and informal L2 exposure to lexical development, and of the explicit and incidental learning processes which these different exposure sources are said to elicit. Explicit lexical instruction and learning strategies applied in the FFL lessons in Dutch-medium education in Brussels are ample (e.g. memorisation of word lists, use of dictionaries). Many researchers agree that explicit or form-focused vocabulary teaching, and the explicit learning mechanisms which it elicits, may be necessary, or at least beneficial for learning the core vocabulary in classroom L2 learning, particularly for the learning of basic lexical and semantic knowledge, but informal vocabulary learning should also be encouraged for learning additional vocabulary and for further lexical and semantic development of the words learned through explicit instruction (Huckin and Coady, 1999). Informal vocabulary learning in instructional contexts is often referred to as 'incidental' (cf. Wesche and Paribakht, 1999; Gass, 1999) because it is typically a by-product, not the target, of another meaning-focused communicative activity such as reading, listening and negotiating meaning in the context of authentic communicative native–nonnative interactions. Although incidental vocabulary learning is still not fully understood, it is believed to have clear advantages over explicit instruction and learning, including the

following: (a) it is contextualised, giving the learner a richer sense of a word's use and meaning than can be provided in traditional vocabulary-building exercises, (b) it is more individualised and learner-based because the vocabulary being acquired is dependent on the learner's own selection of communicative topics or reading materials and (c) it occurs through multiple exposures to a word in different contexts (Hucking and Coady, 1999; Gass, 1999). Be that as it may, the findings of this study would still suggest that there are clear limits to the contributions of incidental vocabulary learning, too, and, more generally, that one has to be patient, and realistic in what can be expected from L2 vocabulary development in educational contexts.

Address for correspondence:
Alex Housen
Department of Language and Literature – Centre for Linguistics
Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussels
Belgium
e-mail: Alex.Housen@vub.ac.be

REFERENCES

Arnaud, P. J. L. and Bejoint, H. (1992). *Vocabulary and Applied Linguistics*. London: Macmillan.

Baetens Beardsmore, H. (ed.) (1993). *European Models of Bilingual Education*. Clevedon: Multilingual Matters.

Bachman, L. and Palmer, A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.

Berman, R. A. and Slobin, D. I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bogaards, P. and Laufer, B. (eds.) (2004). *Vocabulary in a Second Language: Selection, Acquisition and Testing*. Amsterdam: Benjamins.

Broeder, P., Extra, G. and van Hout, R. (1993). Richness and variety in the developing lexicon. In: C. Perdue. (ed.), *Adult Language Acquisition: Cross-linguistic Perspectives*, vol. 1. Cambridge: Cambridge University Press, pp. 145–163.

Carter, R. and McCarthy, T. (1988). *Vocabulary and Language Teaching*. New York: Longman.

Coady, J. and Huckin, T. (eds.) (1997). *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.

Daller, H., Milton, J. and Treffers-Daller, J. (2007). Conventions, terminology and overview of the book. In: H. Daller, J. Milton, and J. Treffers-Daller (eds.), pp. 1–32.

Daller, H., Milton, J. and Treffers-Daller, J. (eds.) (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

Daller, H., van Hout, R. and Treffers-Daller, J. (2003). Measuring lexical aspects of oral language proficiency among bilinguals: an analysis of different measurements. *Applied Linguistics*, 24.2: 197–222.

Decoo, W. (2004). *Eventail-Junior 2000*, 1 and 2 (student books, teacher manuals, CD-Roms). Lier: Van In.

Ellis, R. (1994). A theory of instructed second language acquisition. In: N. Ellis (ed.), *Implicit and Explicit Learning of Language*. London: Academic Press, pp. 79–114.

Folse, K. (2004). *Vocabulary Myths: Applying Second Language Research to Classroom Teaching*. Ann Arbor, MI: University of Michigan Press.

Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21: 285–301.

Harley, B. (ed.) (1995). *Lexical issues in language learning*. Amsterdam: John Benjamins.

Hatch, E. and Brown, C. (1995). *Vocabulary, Semantics and Language Education*. Cambridge: Cambridge University Press.

Housen, A. (2002). Processes and Outcomes in the European Schools Model of Multilingual Education. *Bilingual Research Journal*, 26(1): 43–62.

Housen, A. (2003). L'enseignement des langues vivantes dans les écoles primaires en Belgique. In: M. Nikolov and H. Curtain (eds.), *Un apprentissage précoce: les jeunes apprenants et les langues vivantes en Europe et ailleurs*, Strasbourg: Centre Européen pour les Langues Vivantes, Conseil de l'Europe, pp. 153–167.

Housen, A., Janssens, S. and Pierrard, M. (2002). *Frans en Engels als vreemde talen in Vlaamse scholen*. Brussels: VUBPress.

Housen, A., Janssens, S. and Pierrard, M. (2003). *Le français face à l'anglais dans les écoles secondaires en Flandre*. Brussels: Duculot.

Housen, A. and Pierrard, M. (2004). Meertaligheid in het Nederlandstalig secundair onderwijs in Brussel: verslag van een effectenstudie. In A. Housen, P. Van de Craen, and M. Pierrard (eds.), *Taal, Attitudes en Onderwijs in Brussel*. Brussel: VUBPress, pp. 9–34.

Huckin, T. and Coady, J. (1999). Incidental Vocabulary Acquisition in a Second Language: A Review. *Studies in Second Language Acquisition* 21: 181–193.

Hulstijn, J. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal* 91: 663–667.

Hunt, A. and Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, 17(1): 23–59.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics* 21(1), 47–77.

Kemps, N., Housen, A. and Pierrard, P. (in press). La communication en classe de langue: différences de contextes d'enseignement entre le FLE et l'ALE. In: O. Galatanu, M. Pierrard and D. Van Raemdonck (eds.), *Construction du sens et acquisition de la signification linguistique dans l'interaction*. Bern/Brussels: Peter Lang.

Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect*, 16(3): 44–54.

Laufer, B. (2005). Focus on form in second language vocabulary acquisition. In: S. H. Foster-Cohen, M. P. Garcia-Mayo and J. Cenoz (eds.) *EUROSLA Yearbook* 5. Amsterdam: John Benjamins, pp. 223–250.

Laufer, B. and Nation, P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33–51.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Malvern, D., Richards, B. J., Chipere, N. and Durán, P. (2004). *Lexical Diversity and Language Development. Quantification and assessment*. Houndmills, Basingstoke: Palgrave Macmillan.

Mayer, M. (1969). *Frog, Where Are You?* New York: Dial Press.

McCarthy, P. and Jarvis, S. (2007). Vocd: A theoretical and empirical evaluation, *Language Testing*, 24(4): 459–488.

Mettewie, L. (2007). Elèves non–néerlandophones dans l'enseignement néerlandophone à Bruxelles: analyse des répercussions éducatives de la traversée de la frontière linguistique. In: L. Puren and S. Babault (eds.), *L'éducation au-delà des frontières, Apprentissage précoce du néerlandais, apprentissage précoce en néerlandais dans la zone frontalière franco-belge.* Paris: L'Harmattan, pp. 141–178.

Mettewie, L., Housen, A. and Pierrard, M. (2005). Les élèves francophones dans l'enseignement néerlandophone à Bruxelles deviennent-ils bilingues français/néerlandais? In: P. Hiligsmann, L. Beheydt, L. Degand, P. Godin and S. Vanderlinden (eds.), *Les études néerlandaises en France et en Belgique francophone.* Louvain-la-Neuve: Presses Universitaires de Louvain, pp. 165–176.

Nation, P. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nurweni, A. and Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes*, 18(2): 161–175.

Read, J. (2000). *Assessing Vocabulary.* Cambridge: Cambridge University Press.

Schmitt N. and McCarthy M. (eds.) (1997). *Vocabulary: Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press.

Tidball, F. and Treffers-Daller, J. (2007). Exploring measures of vocabulary richness in semi–spontaneous speech of native and non–native speakers of French: a quest for the Holy Grail? In: H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 133–149.

Van Hout, R. and Vermeer, A. (2007). Comparing measures of lexical richness. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 93–115.

Wesche, M. and Paribakht, T. S. (1999). Incidental L2 vocabulary acquisition: Theory, current research, and instructional implications. *Studies in Second Language Acquisition*, 21: 175–335.

APPENDIX

Table 8. *Correlation matrix*

	#-All types	#-Content types	U-All	U-Content	G-All	G-Content	D-All	#-Nouns	#-Verbs	#-Adjectives	G-Nouns/Nouns	G-Nouns/All	G-Verbs/Verbs	G-Verbs/All	G-Adjs/Adjs	G-Adjs/All	#-FF1	#-FR5	#-FR66	G-FF1	G-FR5	G-FR66
#-All types	1	.981	.882	.819	.921	.895	.918	.922	.959	.644	.806	.727	.822	.820	.391	.434	.908	.906	.993	.837	.890	.946
#-Content types	.981	1	.875	.859	.912	.930	.899	.947	.969	.670	.833	.784	.855	.853	.432	.476	.928	.895	.975	.864	.881	.935
U-All	.882	.875	1	.950	.994	.953	.949	.828	.864	.509	.866	.803	.871	.898	.398	.402	.871	.738	.863	.893	.811	.951
U-Content	.819	.859	.950	1	.941	.982	.886	.816	.843	.509	.886	.811	.807	.913	.436	.424	.858	.702	.802	.889	.782	.889
G-All	.921	.912	.994	.941	1	.963	.963	.863	.899	.542	.886	.808	.882	.902	.407	.413	.899	.789	.905	.902	.846	.969
G-Content	.895	.930	.953	.982	.963	1	.916	.882	.910	.576	.897	.844	.915	.924	.459	.458	.904	.783	.880	.907	.836	.938
D-All	.918	.899	.949	.886	.963	.916	1	.864	.874	.542	.846	.784	.831	.837	.337	.394	.914	.807	.913	.899	.842	.922
#-Nouns	.922	.947	.828	.816	.863	.882	.864	1	.852	.554	.905	.907	.731	.716	.312	.362	.902	.804	.911	.849	.780	.877
#-Verbs	.959	.969	.864	.843	.899	.910	.874	.852	1	.602	.740	.660	.911	.921	.402	.405	.896	.885	.958	.835	.882	.919
#-Adjectives	.644	.670	.509	.509	.542	.576	.542	.554	.602	1	.436	.376	.469	.464	.746	.942	.545	.687	.640	.853	.662	.595
G-Nouns/Nouns	.806	.833	.866	.886	.886	.897	.846	.905	.740	.436	1	.937	.673	.707	.292	.316	.840	.663	.783	.826	.705	.840
G-Nouns/All	.727	.784	.803	.811	.808	.844	.784	.907	.660	.376	.937	1	.645	.642	.244	.277	.804	.580	.708	.820	.612	.771
G-Verbs/Verbs	.822	.855	.871	.807	.882	.915	.831	.731	.911	.469	.673	.645	1	.969	.350	.347	.822	.728	.818	.820	.779	.859
G-Verbs/All	.820	.853	.898	.913	.902	.924	.837	.716	.921	.464	.707	.642	.969	1	.400	.352	.827	.726	.815	.836	.791	.872
G-Adjs/Adjs	.391	.432	.398	.436	.407	.459	.337	.312	.402	.746	.292	.244	.350	.400	1	.790	.336	.433	.383	.314	.455	.418
G-Adjs/All	.434	.476	.402	.424	.413	.458	.394	.362	.405	.942	.316	.277	.347	.352	.790	1	.370	.469	.423	.341	.482	.445
#-FF1	.908	.928	.871	.858	.899	.904	.914	.902	.896	.545	.840	.804	.822	.827	.336	.370	1	.818	.910	.975	.823	.891
#-FR5	.906	.895	.738	.702	.789	.783	.807	.804	.885	.687	.663	.580	.728	.726	.433	.469	.818	1	.924	.712	.974	.822
#-FR66	.993	.975	.863	.802	.905	.880	.913	.911	.958	.640	.783	.708	.818	.815	.383	.423	.910	.924	1	.830	.903	.931
G-FF1	.837	.864	.893	.889	.902	.907	.899	.849	.835	.853	.826	.820	.820	.836	.314	.341	.975	.712	.830	1	.755	.875
G-FR5	.890	.881	.811	.782	.846	.836	.842	.780	.882	.662	.705	.612	.779	.791	.455	.482	.823	.974	.903	.755	1	.864
G-FR66	.946	.935	.951	.889	.969	.938	.922	.877	.919	.595	.840	.771	.859	.872	.418	.445	.891	.822	.931	.875	.864	1

Analysing lexical richness in French learner language: what frequency lists and teacher judgements can tell us about basic and advanced words[1]

FRANÇOISE TIDBALL AND JEANINE TREFFERS-DALLER

University of the West of England, Bristol

ABSTRACT

In this paper we study different aspects of lexical richness in narratives of British learners of French. In particular we focus on different ways of measuring lexical sophistication. We compare the power of three different operationalisations of the Advanced Guiraud (AG) (Daller, van Hout and Treffers-Daller, 2003): one based on teacher judgement, one on 'le français fondamental 1er degré' and one on frequency of lexical items. The results show that teacher judgement is a highly reliable tool for assessing lexical sophistication. The AG based on teacher judgements is better able to discriminate between the groups than the other operationalisations. It also works better than Vocabprofil (the French version of Laufer and Nation's (1995) Lexical Frequency Profile).

1. INTRODUCTION

In this paper we aim to come to a better understanding of a particular aspect of lexical richness, namely lexical sophistication, in learner language of British learners of French. As is well-known, lexical richness is a multidimensional feature of written or spoken language. Read (2000: 200) distinguishes four dimensions of lexical richness, and one of these is lexical sophistication, which he defines as 'the use of technical terms and jargon as well as the kind of uncommon words that allow writers to express their meanings in a precise and sophisticated manner'. The key question is of course how to operationalise what counts as a sophisticated word or expression. Many measures of lexical richness are based on the assumption that the key factor behind the difficulty of a lexical item is its frequency. Laufer and

Nation's (1995) Lexical Frequency Profile, for example, is based on the assumption that frequent words are easier than infrequent words. This view is shared by Vermeer (2000) and Meara and Bell (2001). Similarly, Malvern, Richards, Chipere and Durán (2004: 3) define lexical sophistication as the appropriate use of low frequency vocabulary items.

The question is, however, whether frequency is the only dimension that counts. The psycholinguistic literature shows that the cognate status of items is also an important factor in processing. Cognate items, i.e., translation pairs in which the words are similar in sound and spelling are processed faster than non-cognates (Van Hell and De Groot, 1998: 193) confirming that, as one would predict, cognates are easier than non-cognates.[2] For British learners of French therefore, many infrequent items are easy, because the French and English translation equivalents are cognates, e.g. French *détester* 'to detest', which is infrequent but probably highly transparent to learners.

Support for the fact that cognates play an important role in L2 acquisition comes from Laufer and Paribakht (1998) who demonstrate that French-speaking ESL students obtain higher scores on a test of English controlled active vocabulary than learners with no French because of the large number of French-English cognates. In a similar vein, Horst and Collins (2006) in their study of the longitudinal development of French learners of English in Canada show that learners initially prefer certain low frequency items such as *respond* over the high frequency alternative *answer*, because the former is a cognate of the French translation equivalent *répondre*.

Experienced teachers, in putting together textbooks or other material for learners will therefore not rely *solely* on the frequency of lexical items, but they will use additional criteria, such as the cognate status of items, in judging which words learners need. It is interesting to find out to what extent teacher judgements can help us to get a better understanding of the lexical richness of learner language.

We assume that measures of lexical sophistication which involve teacher judgement are better than those that are solely based on frequency. More specifically, the key hypothesis of the current article is that measures of lexical richness which are based on a basic vocabulary list which is derived from teacher judgements should be better able to discriminate between groups than measures that are based on a basic vocabulary list which consists of the most frequent words or on a traditional basic vocabulary list such as *Français Fondamental Premier Degré* (FF1). The latter is based on a variety of criteria and contains several items that are now out of date (see Tidball and Treffers-Daller, 2007).

We test the key hypothesis of this article by investigating how different operationalisations of the concept of basic vocabulary affect the power of a measure of lexical sophistication which was recently proposed by Daller, Van Hout and

---

[2] We have adopted the psycholinguistic definition of cognates given in Van Hell and De Groot (1998: 193) rather than a historical linguistic definition of cognates because many speakers will not know whether the words in French and English are derived from the same source.

Treffers-Daller (2003), the Advanced Guiraud (AG) (see below for a description). The current paper is a follow-up to Tidball and Treffers-Daller (2007) in which we found that the AG was less able to discriminate between groups than measures of lexical diversity that do not make use of external criteria such as a basic vocabulary list. The choice of a different basic vocabulary can possibly improve the performance of the AG.

We will compare different operationalisations of the concept of basic vocabulary by calculating AG in a variety of ways. These results will subsequently be compared with those obtained with the help of Vocabprofil, the French version of Nation's Range programme which gives a Lexical Frequency Profile of texts (Laufer and Nation, 1995). We will also establish how these measures compare with measures of lexical diversity, which do not make reference to any external criteria or lists, such as the Index of Guiraud (Guiraud, 1954) and *D* (Malvern *et al.*, 2004).

Before going into the details of the current study, we will first present different ways of measuring lexical sophistication (section 2), a brief appraisal of recent work on word frequencies in French and how this relates to the operationalisation of the concept of a basic vocabulary (section 3). In section 4 we present the methodology of the current study, and section 5 gives an overview of the results. Section 6 offers a discussion of the results and a conclusion.

2. MEASURING LEXICAL SOPHISTICATION

We agree with Meara and Bell (2001) that it is important to assess the quality of vocabulary used by L2 learners by making reference to external criteria, such as basic vocabulary lists or frequency lists of lexical items, if one wants to gain a better understanding of lexical richness. As Meara and Bell's (2001: 6) now rather famous examples (1) to (3) show, measures of diversity which are based on distribution of types and tokens in a text will produce the same result for each of these examples.

(1)   The man saw the woman
(2)   The bishop observed the actress
(3)   The magistrate sentenced the burglar

These three sentences are however quite different in the quality of the vocabulary used, as the words in (1) are less difficult (and more frequent) than those in (2) and (3). As Malvern *et al.* (2004: 124) notice, the dimensions of diversity and rarity (sophistication) are of course linked, because 'over a longer stretch of language, diversity can only increase by the inclusion of additional different words, and the more these increase, the more any additional word types will tend to be rare'. The question is therefore whether the development of lexical resources in L1 or L2 learning is due to an increase in the number of low frequency words, or whether the children or students make better use of a wider range of higher frequency words.

Hayes and Ahrens (1988; in Malvern *et al.*, 2004) and Laufer (1998), found that the percentage of low frequency vocabulary did not increase in the spoken or (free

active) written data of their informants. Recently, Horst and Collins (2006) have shown that 11- and 12-year-old francophone learners of English in Québec do not use a higher number of low frequency words after 400 hours of tuition, but a larger variety of high frequency words (up to k1 layer), and they draw less upon cognates (see above). This illustrates most clearly that other factors, such as the cognate status of items, in addition to frequency, play an important role in lexical development, and that frequency bands to which vocabulary items belong do not always provide a good indication of students' progress.

In the present study we use the Advanced Guiraud (AG), as proposed by Daller *et al.* (2003), to measure the differences in lexical sophistication in the speech of British learners of French and a French native speaker control group. The AG is derived from the Index of Guiraud (Guiraud, 1954), which is the ratio of types (V) over the square root of tokens (N) as expressed in the following formula (V/$\sqrt{}$N). The AG is ratio of advanced types over the square root of the tokens (V$_{adv}$/$\sqrt{}$N). For its calculation, one needs to distinguish basic and advanced vocabulary and in this paper we do this in three different ways; with the help of a) the traditional *français fondamental premier degré* (FF1); b) a list of basic words based on teacher judgements; c) a list of basic words derived from the Corpaix frequency list (Véronis, 2000).

As teacher judgements of the difficulty of words were a reliable tool in measuring lexical richness among Turkish-German bilinguals (Daller *et al.*, 2003), the second operationalisation is based on teacher judgements. A list of the frequency of words in a corpus of spoken French, the Corpaix oral frequency list (Véronis, 2000), forms the basis of the third operationalisation. We wanted to find out whether frequency lists are able to successfully capture words which intuition tells us belong to basic vocabulary. We believe with Gougenheim, Rivenc, Michéa and Sauvageot (1964: 138) that 'Ils [les mots concrets] semblent se dérober à la statistique' ('concrete words seem to escape statistics'). A comparison of basic vocabulary lists based on frequency with those based on teacher judgement may well be able to shed new light on the validity of this claim.

We compare different operationalisations of the AG with a well-known measure of lexical diversity, *D* (Malvern *et al.*, 2004), which represents the single parameter of a mathematical function that models the falling TTR curve (see also Jarvis, 2002 and McCarthy and Jarvis, 2007 for an appraisal of this measure). The different measures can give us an indication to what extent the students from the three groups differ from each other in the quantity and/or in the quality of the vocabulary they use.

Finally, we compare these results with those obtained with the help of the frequency bands in Vocabprofil. The output of the programme is a Lexical Frequency Profile (Laufer and Nation, 1995), which gives the frequency of words according to the following four frequency layers: the list of the most frequent 1000 word families (K1), the second 1000 (K2), the Academic Word List (AWL) and words that do not appear on the other lists (NOL). Laufer (1995) shows that a condensed version of the LFP, which distinguishes between the basic 2000 words and the 'beyond 2000' words can also be used to measure lexical richness across different levels of proficiency.

The frequency data on which Vocabprofil is based are derived from a written corpus (see below), but Ovtcharov, Cobb and Halter (2006) claim that Vocabprofil can be used to analyse vocabulary in oral data. It would be useful to know to what extent the profiles for oral and written data as produced by Vocabprofil differ but the authors do not offer such a comparison. They do, however, show that the profiles of advanced Canadian learners of French are not significantly different from those of Beeching's corpus of oral data from French native speakers, which is also freely available on the internet. Vocabprofil differs from the LFP in that the third frequency layer (K3) contains words which occur at a frequency of 2001–3000 in the corpus, French having no equivalent to the Academic Word list (Cobb and Horst, 2001).

For the purposes of the current article it is important to note that Beeching's corpus contains a very high proportion of NOL words, namely 10.87%, and the same is true for the learners in Ovtcharov *et al.*'s study: their scores for the NOL category range from 4.02% for learners at the lowest level to 8.71% for learners in the top group. As the percentages are so high, it is likely that the NOL category does not only contain exceptionally rare words, but also many words that may be frequent in spoken language but not in written language, and which therefore do not occur in the written corpus on which the frequency profiles are based. Whether or not this is the case in our data will be investigated below.

3. BASIC VOCABULARIES AND WORD FREQUENCIES IN FRENCH

Until recently, the only existing basic vocabulary list was *Le français fondamental premier degré* (Gougenheim, 1959) which has been widely used as a reference in many studies on vocabulary. *Le français fondamental premier degré* (FF1) is largely based on an oral corpus. This list is not solely based on frequency: to the most frequent words were added 'available words', i.e. common words (e.g. *fourchette* 'fork', *chocolat* 'chocolate', *autobus* 'bus') which did not appear in the corpus because they are topic-specific and therefore have a lower frequency, but were frequently mentioned in additional surveys on specific themes, or were deemed essential for teaching French as a foreign language.

Given the importance of lexical frequency in language processing, a selection of the most frequent words is an obvious alternative to FF1. We chose to use the Corpaix frequency list for oral French (Véronis, 2000) as this frequency list of 4,592 tokens is based on oral data, and it is freely available on the internet (in unlemmatised form). The corpus of one million words from which the list was derived is based on 36 hours of recordings of interviews held in real-life situations, collected over 20 years at the Université de Provence (now part of the DELIC team).

Because the list is drawn from a relatively limited corpus, some contexts are clearly represented more than others. A few examples can illustrate that there are some unexpected results in this list. In the corpus *orthographe* 'spelling' occurs 235 times, and it has a higher frequency than *finir* 'to finish', regarder 'to look at', and *bactérie* 'bacteria', which occur 144 times. It is also surprising that *fromage* 'cheese'

and *pipette* 'pipette' occur with the same frequency (16) and that *fromage* 'cheese', which appears in FF1, does not feature in the first 1000 words of Corpaix (see appendix for more examples).

In addition we used the frequency profiles that can be obtained with Vocabprofil. The frequency information on which the programme is based stems from a corpus of 50 million words (Verlinde and Selva, 2001) from two newspapers *Le Monde* (France) and *Le Soir (*French-speaking part of Belgium).

It is doubtful whether information about frequency of lexical items in written texts can be used for an analysis of oral data, because of the discrepancies between spoken and written French, but we felt it was interesting to see whether this tool is able to uncover the differences in lexical richness between our three groups, what percentage of the students' tokens belongs in the category NOL (not-on-lists) and whether the lexical frequency profiles are better able to discriminate between the groups in this study.

4. METHODS

The participants were two groups of British undergraduates studying French as part of a Languages Degree at the University of the West of England (UWE), Bristol –21 level 1 (first year), 20 level 3 (final year) – and a control group of 23 native French speakers, also students at UWE. All students undertook the same task under the same conditions: they were asked individually to record their description of two picture stories presented as cartoon strips of six pictures each (Plauen, [1952] 1996]). The corpus contains 23,332 tokens (January 2008).

The general language proficiency of each participant was measured by means of a French C-test which provided a useful external criterion against which the different measures could be validated. This test was highly reliable (Cronbach's alpha = 0.96, 6 items).

The data were transcribed and coded in CHAT, lemmatised and analysed using CLAN (MacWhinney, 2000). More details on the informants, the C-test, the transcription and the lemmatisation are given in Tidball and Treffers-Daller (2007).

For this project we operationalised the concept *basic vocabulary* in three different ways. First of all we used a list based on frequency, availability and judgement (FF1); second, a list based on oral frequency (Corpaix); and third, an intuition-based list (judgements of teachers). We will present each briefly here.

We used FF1 as our first operationalisation, even though this list is rather old and contains several items which relate to rural life in France, such as *charrue* 'plough' and *moisson* 'harvest', which are probably no longer part of the basic vocabularies of speakers living in cities.

Our second list is based on the non-lemmatised oral frequency Corpaix list (Véronis, 2000). This list contains many elements that are typical for spoken data, such as the interjections *euh* (20,897), *ben* (2,936) or *pff* (298). It also unfortunately splits up words that contain an apostrophe, such as *aujourd'hui* 'today' into two words, giving a frequency of 261 for each part.

Homographs constitute another issue: *voler* 'to steal/to rob' (in our data) also means 'to fly' (a bird/ aeroplane), *vol* 'theft/robbery' or 'flight'. The frequency lists on which Vocabprofil is based (see below) do not differentiate between the two meanings and the frequency rank of the word is therefore not entirely meaningful. FF1, on the other hand, gives the two different meanings of *voler* under two different entries. VocabProfil lists *vol* in K1 (first thousand words), *voler* in K2 (1100–2000) and *voleur* 'thief/robber' is NOL (beyond the first 3000 words). The latter is listed in FF1.

As we wanted to compare the results based on the Corpaix list to those based on FF1 (which only contains lemmas), we needed to lemmatise this list. The lemmatisation gave us a frequency list of 2767 types. The methodology followed for the lemmatisation can be found in Tidball and Treffers-Daller (2007).

For our third basic vocabulary list we used the judgement of three experienced tutors of French, two of whom were French native speakers, and one was a bilingual who had grown up with English and French. They were given a list of all 932 types produced by our learners and asked to rank them on a scale on 1 to 7, according to how basic or advanced they judged them to be, with 1 being the most basic and 7 the most advanced. A reliability analysis showed the raters' judgements correlated almost perfectly with each other (Cronbach's Alpha $= 0.943$ (N $=$ 3). Two weeks later we asked the tutors to give us a second judgement of a random sample of 10% of these judgements, which enabled us to carry out a test-retest reliability analysis. The scores given to each item by individual judges in the first and the second rounds correlated strongly and significantly with each other for the first two judges ($r = 0.88$ and 0.84) and significantly but less strongly for the third judge ($r = 0.56$). We also calculated Cohen's kappa to establish to what extent raters agree on what constitutes a basic and a non-basic word in the two rounds. Agreement turned out to be substantial for raters one and two ($k = 0.624$ and 0.601; $p < 0.001$) but only fair for the third rater ($k = 0.252$; $p < 0.001$). The latter was therefore excluded from further calculations.

We defined our basic vocabulary as follows: First, we totalled the scores given by the two remaining raters. Then we selected all words which obtained total scores in the lower quartile (i.e. scores of 4 or less out of a possible 14) for our basic vocabulary list. This gave us a list of 246 basic words.

In order to make a comparison between the different operationalisations possible, we used the Corpaix frequency list to create three different basic vocabulary lists: the first one (Corpaix 246) contained the same number of words as the judges' file (246 words), the second one (Corpaix 1378) contained the same number of words as FF1 (1378), and the third (Corpaix 2000) corresponded to Laufer's Beyond 2000 measure (i.e. it contained the 2000 most frequent words in the list).

5. RESULTS

In this section we first present the results of the C-test, to show how the language proficiency of the three groups differs on a measure that is independent of the story

telling task. We then discuss to what extent the different operationalisations of the concept *basic vocabulary* overlap (section 5.2) and in section 5.3 we will present the results of the analysis of lexical sophistication using different measures based on those basic vocabularies.

5.1 *The C-test*

The C-test results demonstrate that there are significant differences between the French proficiency of the two learner groups and the native speakers (ANOVA, F (df 2,61) = 105.371. $p < 0.001$). The Tukey post hoc test shows that all groups are significantly different from each other. This information is important as one would expect that measures of vocabulary richness should be able to demonstrate the existence of such a clear difference between the learner groups and between learners and native speakers. The power of the C-test to discriminate between groups turned out to be very high as can be seen in the *Eta²* of .776 (see section 5.3 for more details on Eta squared).

5.2 *The overlaps between the basic vocabularies*

Before going into a discussion of the different measurements of lexical sophistication, it is interesting to see to what extent the different basic vocabularies overlap. No two lists, even those drawn from very large corpora of similar origin, will overlap completely. Comparisons of the first 1000 words of three existing frequency lists derived from different large French literary corpora, of which the *Trésor de la langue Française* (TLF) (INALF, 1971) is one, showed that they had 80% of words in common, whereas *le Français fondamental* had 65% in common with TLF (Picoche, 1993).

We have used CLAN to compare the content of FF1 with two other operationalisations: the Corpaix oral frequency list and the basic vocabulary list which is based on the judgements of the teachers. As FF1 contains 1378 words, we compared the first 1378 words of the Corpaix oral frequency list with FF1, and found that 725 words (52.6%) of FF1 are also found in the first 1378 words of the Corpaix frequency list. The judges' file shares 236 words (95.9% of the 246 words it contains) with FF1.

Subsequently, we entered all different operationalisations into Vocabprofil, to find out what percentage of the words in FF1, the Corpaix list and the judges' file belongs in the different frequency bands distinguished by Vocabprofil. The results of these analyses can be found in Table 1. It shows that FF1 and the first 1378 words of Corpaix have roughly similar profiles, with approximately 60% K1 words, whilst the first 2000 words of Corpaix contains almost 50% K1 words. The judges' file and the first 246 words in Corpaix contain a far larger proportion of K1 words: respectively 89 and 94%.

The percentage of words that do not appear in any list is very high for FF1 (21.9%) and Corpaix (14.44%) and this is probably due to the fact that Corpaix

Table 1. *Percentage of words in each Vocabprofil frequency band in each corpus*

Vocabprofil Bands	FF1 (1378 words)	Corpaix first 1378 words	Corpaix first 2000 words	Judges (246 words)	Corpaix first 246 words
K1 Words (1 to 1000)	60.12	59.29	49.10	89.0	93.57
K2 Words (1001 to 2000)	14.15	21.19	23.0	6.46	3.61
K3 Words (2001 to 3000)	3.84	5.08	6.6	0.76	0.40
NOL (not in list of first 3000 words)	21.90	14.44	21.3	3.80	2.41

contains many elements that are frequent in spoken language but which are not found in the written corpora.

Examples of words from our corpus which are NOL – apart from the many interjections mentioned in section 4 – are nouns such as *chapeau* 'hat' and *voleur* 'thief', an adjective such as *gentil* 'nice' and verbs such as *nager* 'to swim' and *repartir* 'to set off again'. The problem, from our perspective, is that Vocabprofil puts these very common words (all of which are in FF1) in the same category as *bousculer* 'knock down' and *canne* 'walking stick', which are highly specific and very infrequent, and which are not found in FF1 or Corpaix.[3] This illustrates the difficulty of using a written frequency list for the analysis of oral data. It is very unlikely that these words would all be classified in the same category if Vocabprofil was based on an oral frequency list.

5.3 *Measures of vocabulary richness*

Table 2 gives an overview of the results obtained for our different measures of vocabulary richness, including list-free measures. All measures show that there are significant differences between the groups in the vocabulary used. The smallest basic vocabularies – defined by the judges or the Corpaix 246 list – were successful at differentiating between all groups. The AG (judges) and the AG (Corpaix 246) yielded significant differences between the two learner groups (level 1 and level 3) but the AG (FF1)[4] and AG (Corpaix 1378) or AG (Corpaix 2000) did not.

In order to find out which measure is best able to discriminate between groups, we calculated Eta². Eta squared is the percentage of the variance in the dependent variable that can be accounted for by the independent variable (i.e. group membership in this case). As Table 2 shows, the AG (judges) obtains a

---

[3] Corpaix does list CANNES, but this is presumably the city, and not the plural form of *canne* 'walking stick'.

[4] In Tidball and Treffers-Daller (2007), the results for the AG (FF1) were marginally significant, but in the current study, with four more informants, this was no longer the case. As two of these additional learners had C-test scores well below the Mean of 75, this may be an indication that these learners are rather weak. The person with the lowest C-test score also had an exceptionally low score for advanced words, namely 3, which is more typical of the level 1 group.

Table 2. *Mean scores on different measures of vocabulary richness and results of a one-way ANOVA/Tukey post hoc*

	Level 1 (N = 21)	Level 3 (N = 20)	Native speakers (N = 23)	F-value, df (2,61)	Tukey post hoc 1–3	1–NS	3–NS	Eta²
Guiraud	4.29	5.28	6.27	57.0	**	**	**	.651
D (lemmat.)	18.78	26.53	34.87	58.9	**	**	**	.659
AG (FF1)	0.30	0.54	1.21	32.4	–	**	**	.515
AG (judges)	0.08	0.10	1.50	58.7	*	**	**	.658
AG (Corpaix, 246)	1.68	2.56	3.35	47.8	*	**	*	.610
AG (Corpaix, 1378)	0.48	0.69	1.18	36.8	–	**	*	.547
AG (Corpaix, 2000)	0.30	0.45	0.87	31.2	–	**	*	.505

* = p < 0.05; ** p < 0.001.

Table 3. *Percentage of the tokens belonging to different frequency layers (Vocabprofil) for all three groups (One-way ANOVA/Tukey post hoc)*

	1–1000 (K1)	1001–2000 (K2)	2001–3000 (K3)	NOL (not on lists)
Level 1 (N = 19)	92.77	4.47	.69	2.07
Level 3 (N = 20)	90.87	4.61	.15	4.37
Native speakers (N = 25)	88.83	4.99	.916	5.63
F (2, 61)	19.03	0.741	1.80	31.58
P (Tukey)	<0.001 (all groups sign. different)	n.s.	n.s.	<0.001 (all groups sign. different)
Eta2	0.384	0.024	0.056	0.509

higher *Eta²* than the Index of Guiraud. The *Eta²* for the AG (judges) and D are virtually identical. The AG (judges) compares very positively with the other operationalisations of the AG, including its closest ally, the AG (Corpaix 246). This clearly shows that using teacher judgement is a better way to obtain a basic vocabulary list than using frequency data.

We also submitted the data to Vocabprofil, to find out to what extent the frequency layers as distinguished in Vocabprofil could help to distinguish our three groups (see Table 3).

The results given in Table 3 show that, as one might expect, the native speakers make less use of words belonging to the highest frequency layer than the level 3 learners, and the latter use fewer words from the highest frequency layer than level 1 learners. The percentages are comparable to those of Beeching's corpus, although a smaller proportion of the words used by Beeching's informants belonged to the K1 category (83.99%), and more of their words are NOL (12.07%). This is probably

Table 4. *Diferent keywords from the stories in Vocabprofil, FF1, Corpaix and the teachers' judgements*

Keywords	Vocabprofil	FF1	Rank order in Corpaix	Teacher judgements (out of 14)
banque	K1	no (FF2)	926	4
bâton	NOL	yes	2362	5
canne	NOL	no	no	7
chien	K2	yes	649	3
coup	K1	yes	224	6
criminel	K2	no	no	10
lac	K1	yes	2434	6
vol	K1	no	1438	8
voler	K2	yes	1480	8
voleur	NOL	yes	no	7

due to the fact that our informants did not produce free speech but they all narrated the same two stories, which limits the choices for the informants. Of the words used by Beeching's informants, 3.94% belonged to the K2 layer and 1.2% to the K3 layer, and these percentages are very similar to those for our informants.

All three groups differ significantly from each other in their use of K1 words and also in their use of NOL words (see ANOVA/Tukey post hoc in Table 3). It is interesting that the groups do not differ from each other with respect to the K2 and K3 layers of Vocabprofil. While level 3 students use more words from the K2 layer than level 1 students, these differences are too small to become significant. As for the K3 layer, the students seem to use even fewer words of this frequency layer than the level 1 students.

Table 3 shows the effect size of the measurement of the vocabulary used at each frequency layer distinguished by Vocabprofil. It clearly demonstrates that the choice of the words which are *not* in the frequency list is the most powerful indicator of the differences between the groups, but the *Eta²* of the scores obtained on the basis of Vocabprofil are clearly lower than those obtained with the help of the basic vocabulary lists (see Table 2).

The *Eta2* obtained for the C-test outshines all the results found for the lexical richness measures, as the C-test obtained an *Eta²* of .776. A simple C-test may therefore well be a more effective way to distinguish language proficiency levels among learner groups.

The use of cognates by learners also conveys important information that needs to be taken into account in analyses of vocabulary richness. In our data, for example, speakers from all groups describe the thief in the second story most often as a *voleur*, but the learners' second most popular word for this character is the cognate *criminel* 'criminal', which is not used at all by the native speakers, even though it can be used as a noun in standard French. This word belongs to the K2 layer in Vocabprofil, and it is not listed in the Corpaix oral frequency list at all. Students who use this word would get higher scores on Vocabprofil or on the AG (corpaix) as these measures are exclusively based on frequency (see Tables 4 and 5). The use of cognates is

Table 5. *Examples of keywords from the stories and their allocation to different frequency layers in Vocabprofil*

K1	K2	K3	NOL
banque	chien	amuser	bâton
vol	voler	(se) promener	voleur
aller	content		chapeau
retrouver	enlever		nager
lancer	rapporter		bouche
autant	emmener		braqueur
pouvoir	ramener		bousculer
se rendre compte			désarroi

however not necessarily an indication that the speaker possesses a rich vocabulary. Rather, it shows that the speaker knows how to strategically exploit similarities between languages in telling a story. How we can account for the strategic use of cognates in the context of studies on vocabulary richness therefore deserves to be investigated further.

6. DISCUSSION AND CONCLUSION

The results presented in section 5 clearly show that the power of a list-based measure such as AG depends on the way in which researchers operationalise the concept of basic vocabulary (see also Daller and Xue 2007, who make a similar point). We found that an operationalisation based on teacher judgements was more powerful than different operationalisations based on frequency, in that the AG (judges) was better able to discriminate between the groups. It is possible though that the performance of the AG based on frequency data can be improved if alternative frequency lists are being used. The Corpaix frequency list may not have been ideal for the current purposes, as it was drawn from a relatively small corpus. Alternative frequency lists based on the CRFP or *Lexique* could be used in future studies on this topic.

The results of the Vocabprofil analyses show that the students make better use of a range of relatively easy words. The learners used more NOL words at level 3, such as *gentil* 'nice' and *chapeau* 'hat', which are common in spoken language but happen not to occur in the written corpora on which Vocabprofil is based. Our results confirm those of Horst and Collins (2006) whose learners did not use a higher number of low frequency words after 400 hours of tuition either but a larger variety of high frequency words which belong to the k1 layer. As many researchers have found that the percentage of K2 words (and beyond) is very low in learner language, it may be important to further differentiate between different frequency layers among the k1 group. In this paper we have shown that using a *small* basic vocabulary (n = 246 words) in calculations of the AG works better than using a large basic vocabulary (n = 1378 or n = 2000). In Tidball and Treffers-Daller (in preparation) we illustrate this further by focusing on motion verbs. While the level

1 learners prefer to use basic deictic verbs (*aller* 'go' and *venir* 'come') over path verbs such as *entrer* 'to enter', all of which belong in the K1 layer of the vocabulary frequency lists, level 3 learners increasingly use *entrer* to express the same motion event, which shows they have progressed in comparison with the level 1 learners. The percentage of low frequency words is therefore for these learners possibly a less suitable indicator of the differences in the lexical richness of their speech.

The analysis of our data with Vocabprofil provided interesting information about the frequency layers of the vocabulary used by our learners, but the large number of words in the NOL category is worrying in that this category contains both very rare words such as *bousculer* 'to knock over' and highly frequent items which are characteristic of spoken rather than written language. It would therefore be very useful for the research community if a new version of Vocabprofil could be created which is based on frequency data for oral language.

Finally, a preliminary analysis of the use of *criminel* by the learners indicates that cognates play an important role in L2 vocabulary acquisition, which confirms the results of Laufer and Paribakht (1998) and Horst and Collins (2006). The strategic use learners make of cognates is an area that deserves further attention in future studies of vocabulary richness.

Address for correspondence:
Françoise Tidball
Jeanine Treffers-Daller
Department of Languages, Linguistics and Area Studies
School of Humanities, Languages and Social Sciences
Faculty of Social Science and Humanities
Frenchay Campus
Coldharbour Lane
Bristol BS16 1QY
e-mail: Francoise.Tidball@uwe.ac.uk
*   Jeanine.Treffers-Daller@uwe.ac.uk*

REFERENCES

Cobb, T. and Horst, M. (2001). Is there an academic word list in French? In: P. Bogaards and B. Laufer (eds.), *Vocabulary in a Second Language*. Benjamins, pp. 15–38.

Daller, H. and Huijuan, Xue (2007). Lexical richness and the oral proficiency of Chinese EFL students. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

Daller, H., Van Hout, R. and Treffers-Daller, J. (2003). Lexical richness in spontaneous speech of bilinguals. *Applied Linguistics* 24: 197–222.

Goodfellow, R., Jones, G. and Lamy, M.-N. (2002). Assessing learners' writing using Lexical Frequency Profile. *ReCALL* 14: 129–142.

Gougenheim, G. (1959). *Le Français fondamental 1ᵉʳ degré*. Ministère de l'Education Nationale. Paris: Institut National de Recherche et de Documentation Pédagogiques.

Gougenheim, G., Rivenc, P., Michéa, R. et Sauvageot, A. (1964). *L'élaboration du français fondamental*. Paris: Didier.

Guiraud, P. (1954). *Les caractéristiques du vocabulaire*. Paris: Presses Universitaires de France.

Hayes, D. P. and Ahrens, M. G. (1988). Vocabulary simplification for children: a special case of 'motherese'. *Journal of Child Language* 15: 395–410.

Horst, M. and Collins, L. (2006). From *Faible* to strong: how does their vocabulary grow? *The Canadian Modern Language Review*, 63(1): 83–106.

INALF (1971). *Dictionnaire des fréquences du Trésor de la Langue Française*. Paris: Didier.

Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19: 57–84.

Laufer, B. (1995). Beyond 2000. A measure of productive lexicon in a second language. In: L. Eubank, L. Selinker and M. Sharwood Smith (eds.). *The Current State of Interlanguage. Studies in Honor of William E. Rutherford*. Amsterdam/Philadelphia: John Benjamins, 265–272.

Laufer, B. (1998). The development of active and passive vocabulary in a second language: same or different? *Applied Linguistics* 16: 307–22.

Laufer, B. and Nation, P. (1995). Vocabulary size & use: Lexical richness in L2 written productions. *Applied Linguistics* 16(3), 307–322.

Laufer, B. and Paribakht, T. S. (1998). The relationship between active and passive vocabularies: effects of language learning context. *Language Learning* 48(3), 365–391.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Lawrence Erlbaum.

Malvern, D. D., Richards, B. J., Chipere, N. and Durán, P. (2004). *Lexical Diversity and Language Development: Quantification and Assessment*. Houndmills, Basingstoke: Palgrave Macmillan.

McCarthy and Jarvis (2007). VOCD a theoretical and empirical investigation. *Language Testing*, 24(4): 459–488.

Meara, P. and Bell, H. (2001). P_Lex: a simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect* 16: 5–19.

Ovtcharov, V., Cobb, T. and Halter, R. (2006). La richesse lexicale des productions orales: mesure fiable du niveau de compétence langagière. *The Canadian Modern Language ReviewLa revue canadienne des langues vivantes*. 63(1): 107–125.

Picoche, J. (1993). *Didactique du vocabulaire*. Paris: Nathan.

Plauen, E. O. [1952] 1996. *Vater und Sohn*, Band 2. Ravensburger Taschenbuch.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Tidball, F. and Treffers-Daller, J. (2007). Exploring measures of vocabulary richness in semi-spontaneous French speech: A quest for the Holy Grail? In: H. Daller, J. Milton and J. Treffers-Daller (eds.) (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press, 133–149.

Tidball, F. and Treffers-Daller, J. (In preparation). Variability in the expression of motion events by British learners of French: the role of transfer and simplification.

Van Hell, J. G. and de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition* 1(3), 193–211.

Verlinde, S. and Selva, T. (2001). Corpus-based versus intuition-based lexicography: Defining a word list for a French learner's dictionary. In: P. Rayson, A. Wilson, T. McEnery, A. Hardie and S. Khoja (eds.). *Proceedings of the Corpus Linguistics 2001*

Conference. Lancaster: Lancaster University, University Centre for Computer Corpus Research on Language, pp. 594–598.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing* 17, 65–83.

Véronis, J. (2000). *Fréquence des mots en français parlé*. http://www.up.univ-mrs.fr/ ∼ veronis [accessed 1 October 2007]

WEBSITES

The Range programme (Paul Nation): http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx

Vocabprofil: http://www.lextutor.ca/vp/fr/

Beeching's corpus of spoken French: http://www.uwe.ac.uk/hlss/llas/iclru/corpus.pdf

DELIC team (Description Linguistique Informatisée sur Corpus): http://sites.univ-provence.fr/delic/

Véronis' frequency lists: http://sites.univ-provence.fr/veronis/donnees/index.html

A developmental perspective on productive lexical knowledge in L2 oral interlanguage[1]

ANNABELLE DAVID

Newcastle University

ABSTRACT

This article reports on productive vocabulary development by instructed British learners of French over a five-year period (from age 13 to 18). Lexical diversity development was investigated through a semi-guided oral picture-based task. Results show that the students' lexical diversity (as measured by D) did significantly improve throughout the five years showing little sign of slower periods. Overall more noun types were observed than verb types in the composition of the lexicon throughout the study but with a consistent decrease in its proportion after Year 10. Further results using the Limiting Relative Diversity measure indicate that learners vary their use of nouns to a much larger extent than verbs. The discussion focuses on the noun-bias hypothesis and the use of different elicitation tasks.

INTRODUCTION

The aim of this paper is to describe and analyse the development of lexical diversity (the range of vocabulary used in a text or transcript) as well as the type of vocabulary used (from different grammatical categories) during semi-spontaneous oral production amongst instructed learners of French through cross-sectional and longitudinal data representing a five-year learning period. A brief background is first provided, describing the current literature in L2 vocabulary and outlining the need for the current study. Then, the data used as well as the unit of analysis and methods of analysis are presented. Finally, the details and results of the current study are laid out and discussed.

Vocabulary acquisition is one of the key basic aspects of language learning and current models of language competence all give a central role to the lexicon, including both Chomskyan minimalist accounts and connectionist and emergentist accounts (see Collentine, 2004 for discussion). Speakers continue to learn new words well into adulthood (Hall, Paus and Smith, 1993) even though other aspects of language might have been fully acquired or have fossilised (e.g. grammar). Vocabulary is also one of the key predictors of school success (Verhallen and Schoonen, 1998).

There are very few studies in the current literature about L2 lexical development. L1 acquisition research, on the other hand, has produced a much larger number of studies. L1 researchers have focused, amongst other aspects, on profiling the developing lexicon of young children and overall, cross-linguistically, the same developmental lexical patterns have been described by the literature. Children acquire nouns first, verbs second and closed-class items later (see David, 2004, for a cross-linguistic discussion). However, there are differences in noun/verb productions across languages (Childers and Tomasello, 2006). Recent research has provided evidence, against the noun-first prediction, showing that verbs can be learnt as early as nouns (e.g. Bassano, 2000) and that certain languages are less noun-biased. Tardif, Shatz and Naigles (1997) found that Mandarin-speaking children produced more verbs than English- or Spanish-speaking children at earlier stages, for example.

As far as L2 acquisition is concerned, there is a growing literature on the lexical development of English as a second language (see Nation, 2001). Key issues addressed by the literature include how words might be organised in the mind (e.g. Kroll and Tokowicz, 2001), how the lexicon might relate to other aspects of language (e.g. Clark, 1993) and how to measure lexical knowledge (e.g. Meara and Milton, 2003). Studies dealing with French are rarer. However, this is an area which is growing strongly and the present issue is proof. Lexical development amongst instructed second language learners of French in the UK educational setting has been the focus of a few studies. In a cross-sectional study, Milton (2006) used a test measuring receptive vocabulary knowledge (X-Lex) amongst secondary school students. He concludes that an A-level student (aged 18 and at the end of his/her secondary education) knows on average 2000 words (after learning French for seven years). He found that GCSE learners (aged around 16 years old), have receptive knowledge of about 850 words while final year graduates know about 3300 words, according to the measure used. He also noticed a slower period of lexical growth between the second and fourth year of teaching. Milton and Meara (1998) report on a comparative study of French foreign language learners in Britain and EFL learners in Germany and Greece, finding a passive vocabulary of about 800 French words on average after four years of study; a learning rate of some 200 words per year. All learners in their study appeared to gain between three and four words per contact hour on average, in line with other studies reviewed by Milton and Meara. These studies used vocabulary tests that assessed receptive knowledge as this is easiest to elicit (Fitzpatrick, 2007). But Waring (1997 cited in Daller, Milton

and Treffers-Daller, 2007) claims that productive vocabulary is only about 50% of receptive vocabulary. As a result, even though measuring receptive vocabulary is interesting for other purposes and it is linked to productive vocabulary (Laufer and Paribakht, 1998), it is not an accurate indication of a learner's productive abilities.

Currently we have few independent descriptions of vocabulary use during semi-spontaneous oral production amongst instructed foreign language learners. We do not have a detailed picture of learners' access to their L2 lexicon when faced with an unprepared oral task with an unknown interlocutor, a context frequently adopted to represent communicative competence. Such descriptions have the potential to inform teaching, at both a macro-level (syllabus design, assessment criteria) and micro-level (classroom practice). Nation (2007) and Read (2000) also propose that we need to measure vocabulary in use (e.g. having a conversation about a holiday) to gain a complete picture of the learners' vocabulary. This type of task is interesting to see how varied a learner's vocabulary is when the learner is carrying out an activity whose apparent sole purpose is not to assess vocabulary knowledge. Nation (2007) points out that if a learner is asked to write a piece on (or to talk about) his latest holidays, s/he is unlikely to be aware that lexical use will be the focus of the assessment. In contrast when a learner answers a yes/no vocabulary test, s/he knows that vocabulary is the focus of the researcher's agenda. This issue is linked to incidental vocabulary learning (Gass, 1999). Learning lexical items while this is not the target of the activity (i.e. incidental learning) is a well-documented phenomenon. Producing lexical items when it is not the focus of the task could be referred to as incidental productions. Gass (1999: 322) suggests that learners are more likely to learn words incidentally. Similarly, we could argue that incidental productions are more likely to be richer/larger than intentional productions. Although the topic of this study is not to compare results that different tasks could provide, the oral task that will be used in this research calls upon incidental productions.

In terms of vocabulary in use, there has been considerable attention given to assessing learners' use of vocabulary in writing (see Malvern, Richards, Chipere and Durán, 2004, and see Fitzpatrick, 2007 for discussion). Nevertheless, little attention has been given to oral production. One of the reasons for the lack of studies is that it remains difficult to assess precisely the breadth and depth of the lexical knowledge of language users. Tidball and Treffers-Daller (2007) explore different measures of vocabulary richness in L2 university-level learners of French through a cartoon-based story telling task. They showed significant differences between learners at different levels using several different measures as well as correlations between lexical scores and general proficiency scores. They claim that the different measures used (or at least those that did demonstrate a difference between levels) are all valid. These are: D, Guiraud index and Guiraud Advanced (D and the index of Guiraud are further described in the methodology section of this paper). In addition, D and Guiraud did correlate strongly with each other. Other studies based on oral semi-spontaneous tasks for L2 learners exist but these are not dealing with learners of French (e.g. Daller and Xue, 2007). This is supported by Tidball and

Treffers-Daller (2007) who say that very few measures of lexical diversity/richness have been tried out on learners of French.

As illustrated above, a number of recent studies have focused on measurement and assessment methods (e.g. Eyckmans, van de Velde, van Hout and Boers, 2007; Brown, 2003). Few L2 studies have focused on aspects similar to those present in the L1 literature. As mentioned earlier, one of the focuses of L1 research has been the nature of the developing lexicon. The L2 developing lexicon is much more of a mystery to researchers. In her longitudinal study about the development of narrative abilities in French, Myles (2003) concludes, that the learners (aged between 12 and 15) who were the best at telling a detailed story were also those with the richest vocabulary and at the more advanced syntactic stage. She also claims that lexical chunks and nouns appear first whilst verbs (produced outside of lexical chunks) come later. However, this claim is not backed up by any lexical analyses (other than the number of verbs used). Hence the development of the nature of the L2 lexicon remains under-studied.

AIM AND RESEARCH QUESTIONS

The aim of the study is to provide a developmental profile of the lexicon in secondary school students learning French by focusing on the kind of oral productive French vocabulary that is under-represented in the current literature. Through the use of corpus data, the data will span five different school years.

In view of the present aim and the background literature, we identified the following research questions which this paper will attempt to answer:

- Does learners' productive lexical diversity increase significantly over the course of five years of instruction?
- What is the nature of their developing productive lexicon? Are nouns and verbs developing at the same rate? Or do verbs appear later?

METHODS

Below we highlight the methods used to answer the previous research questions.

The corpus used for the current study

The data are taken from the French Learner Language oral corpora (FLLOC). This dataset (constructed by Myles, Mitchell and their research teams (e.g. Rule, Marsden, Myles and Mitchell, 2003)) is publicly available to the research community (http://www.flloc.soton.ac.uk/). The corpus contains digital audio files, related transcripts formatted using the CHILDES software and conventions and files tagged for parts of speech (see Rule, 2004; Myles, 2005, 2007 and Myles and Mitchell, 2004 for discussions of the issues relating to transcription and analysis of oral L2 data). The corpus holds data from a series of cross-sectional studies from British

students learning French (the youngest are in their first year of secondary school and the oldest are in their final undergraduate year at university) and native speakers. These students perform a range of semi-spontaneous oral tasks on a one-to-one basis with a researcher. These oral tasks can be used to investigate a range of issues: from the emergence of aspects of morphosyntax to lexical development, learners' use of formulaic language, or aspects of discourse development (both monologic and dialogic). Here, we concentrate on lexical development.

Participants

This study describes and compares lexical richness as measured during an oral semi-guided conversation amongst learners of French in Years 9, 10, 11, 12 and 13 of the British school system. Students are aged 13 to 14 in Year 9 and 17 to 18 in Year 13. There are 20 learners in each year group. Most of the data is cross-sectional with the exception of Year 12 and 13. Those learners were tested in Year 12 and once again, a year later, in Year 13. Therefore, the total number of different learners is 80. In Year 9 (their third year of classroom learning) students will have received about 150 hours of instruction. By Year 13, they will have had a maximum of approximately 600 hours. The learners were all tested between December and March (so half-way through the school year). The learners are from different state schools in the U.K. Year 12 and 13 learners are from the North East of England and Years 9 to 11 are from the South (the area in and around Southampton).

The oral task

All participants carried out the same task which involved a conversation about a set of six different photos including questions relating to past, current and future activities. This task takes the form of a one-to-one semi-structured interview in French between individual learners and members of the research team. The task is in two parts. In the first part, the learners are shown two separate sets of stimulus photographs representing young people doing various activities and they have to find out as much information as they can about the young people shown in the pictures, the location, and so on, by asking questions. This task is therefore referred to as the Photos task. In the second part, the researcher asks the learner a range of questions about their current interests, their family life and (for the most advanced learners) their past holidays, and their plans for the future. For this second part, the photos only serve as a starting point to the conversation but the discussion is not solely based on the photos. This task is a combination of what Laufer and Paribakht (1998) call controlled active and free active tasks. It is controlled as the learners are told to ask questions about the picture. Consequently, they are limited to the context of the pictures. The second part, however, is an active one as they are more or less free to talk about whatever they like within the remits of the researcher's vague questions.

Unit of analysis

It is essential to define the unit of analysis that is used in this paper. According to Richards and Malvern (2007), this is one of the most crucial decisions to be taken by a researcher investigating vocabulary diversity. The 'quick and dirty' method (Richards and Malvern, 2007: 88) adopted by many researchers who do not define what counts as a word is not sufficient. In terms of productive vocabulary use, it has been shown by Vermeer (2004) that the lemma is the most valid unit of counting. Lemma will be defined here in the morphological sense of the word, as the canonical form of a word or lexeme. For example, in English, the lemma *go* represents the inflected forms *go*, *goes*, *going*, *went* and *gone*. Lemmas are especially significant in highly inflected languages such as French. There are several reasons for using lemmas as our unit of analysis. Firstly, lemmas will allow us to minimise transcription inconsistencies regarding inflected forms in particular. For example, learners often mispronounce forms such *un/une* or *petit/petite*. These mispronunciations mean that the transcribers have to make decisions as to the best way of coding them. By using lemmas as the unit of counting, the problem will be eliminated. Secondly, using lemmas rather than words means that the data and transcripts have to be carefully prepared and reliability is, therefore, increased. Finally, using lemmas as the unit of counting allows, for example, for different collocations and grammatical constructions to be counted. Nation (2007) stresses that if a researcher was to use another classification, word families for example, these aspects of the lexicon would be masked. Wrong units of analysis could also lead to overestimation or underestimation of the learners' lexicon. An example would be a form-based word family where lexical items such as *famille*, *familier*, *familiarité*, etc would all be counted as one.

Consequently, different morphological forms of the same stem e.g. *regardez* and *regarde* or *le*, *la* and *les* were counted as one lemma. Words that have different meanings (e.g. *avocat* = avocado or lawyer/solicitor) or different grammatical functions depending on the context were also counted as one single lemma. For example, the word *que* can be a conjunction or a relative pronoun. No difference was made between those two uses and *que* was counted as one lemma. Derivational morphological forms (such as *apprendre* and *apprenant*) were counted as different lemmas though. We did not take into account grammatical inaccuracies as the aim of the paper relates to vocabulary and not grammatical development. Fillers (*ah*, *euh*), imitations of the researcher's utterances or words as well as words in languages other than French were excluded. Proper nouns (e.g. of geographical areas and people's names) were excluded from the analyses.

*P34: et ils font euh du scuba?
*JUL: ils font de la plongée.
*P34: de la plongée ou.

In the example above, *euh* was excluded (filler pause) as well as *de la plongée* as the learner is imitating the researchers' utterance and did not appear to have prior

knowledge of the expression. In cases of repetitions and retracings (repetitions with corrections), only final repairs were counted.

*P47: quel a faire leur [//] le garçon?

In the example above, *leur* was excluded and only *le* was counted.

Measure used

As outlined before, there is a growing literature (see Daller, Milton and Treffers-Daller, 2007) on the quest for the best measure in terms assessing a learners' lexical development and in particular, lexical diversity. This paper focuses on one measure in particular: D to assess lexical diversity in the production of the learners described above.

Probably the most common measure used, TTR is based on the ratio of different words (Types) to the total number of words (Tokens). This is known as the Type-Token Ratio (TTR). However, there is well-reported controversy over this measure (see e.g. discussion in Daller *et al.*, 2007), as it does not account for the fact that the longer someone speaks (or writes) for, the less varied their language is likely to be, thus misrepresenting some learners' lexical richness. The texts (or transcripts) in our corpus are of very varied length as students tend to speak more as their proficiency increases. Thus, comparing students using TTR would prove unreliable. The Guiraud (1954) index is the ratio of types by the square root of tokens (Types/$\sqrt{}$tokens). It is one of the alternatives to TTR put forward to minimise the impact of text length. The introduction of the square root compensates all samples by multiplying TTR by $\sqrt{N}$ (square root of the number of word tokens). However, this measure does not change the issue of the text-length dependency (see Malvern *et al.*, (2004) for a discussion).

Consequently, the measure used here will be D. D is a measure of lexical diversity created to avoid the inherent flaws in raw TTR and other mathematically related measure (like Guiraud). The approach taken is based on an analysis of the probability of new vocabulary being introduced into longer samples of speech (or writing). D uses random sampling of tokens in plotting the curve of TTR against the increasing token curve (see McKee, Malvern and Richards, 2000 for a more detailed description of the program). D has three main advantages: it is not text-length dependent; it uses all of the data in a single text (transcript); and it is more informative than TTR 'as it is based on the TTR versus token curve calculated from data for the transcript as a whole, rather than a particular TTR value on it' (MacWhinney, 2000). The measure has been validated across a wide range of language learners (Malvern *et al.*, 2004). D has been integrated within CLAN (Computerised Language Analysis program available through CHILDES at http://childes.psy.cmu.edu/) and is computable through the VOCD program. Recent criticisms of D have emerged. In particular, McCarthy and Jarvis (2007) claim that D is affected by text length. However, they conclude by saying that D remains a very useful measure and that, even if researchers need to use it

Table 1. *Post hoc test (Tukey) for types and tokens*

	Types	Tokens
Year 9 and 10	.129	.571
Year 10 and 11	.757	.919
Year 11 and 12	.032*	.292
Year 12 and 13	.983	.939

*The mean difference is significant at the .05 level.

with caution, 'D is undoubtedly a better performer than most alternative indices' (McCarthy and Jarvis, 2007: 480).

Comparisons of different lexical categories

The present study includes comparisons of different lexical categories and the use of different types of words (nouns, verbs, adjectives, etc). To do so, type/type ratios are used (see Malvern *et al.*, 2004, for a discussion of different type/type ratios). We also calculated the Limiting Relative Diversity (LRD) measure, as proposed by Malvern *et al.* (2004). This is a type-type ratio that enables the study of the ratio of one category of words over another (e.g. nouns over verbs). The formula, in the case of verbs/nouns ratio, is: $LRD = \sqrt{D(verbs)/D(nouns)}$. This allows the researcher to examine the diversity of one word class compared to another word class and as this measure is based on D, it is not a function of text length. One drawback of this measure is that it only works if the sample of each word class is larger than 50. Therefore, beginner learners who produce less than 50 verb tokens would not be able to be included in this measure.

RESULTS

In this section we will first describe the overall development of the productive lexicon of the learners. Secondly, we will present an analysis of the types of words produced.

Overall developmental trend

The first developmental trend worth mentioning is the increase of the number of tokens in the productions of the learners. Figure 1 shows the apparent ever increasing number of types and tokens from Year 9 until Year 13.

Performing the same semi-spontaneous task, learners do appear to produce more and more types and tokens. This might not be surprising as far as the number of tokens is concerned. But it is not as evident for the number of types. An ANOVA highlights the fact that there are significant differences between the year groups for tokens $F (4, 95) = 8.475$, $p < 0.001$ and types $F (4, 95) = 17.216$, $p < 0.001$. Further details of the post hoc test are found in Table 1.

Table 2 shows the mean scores and standard deviations for D per year group. It demonstrates that the five groups differ from each other in predictable ways:
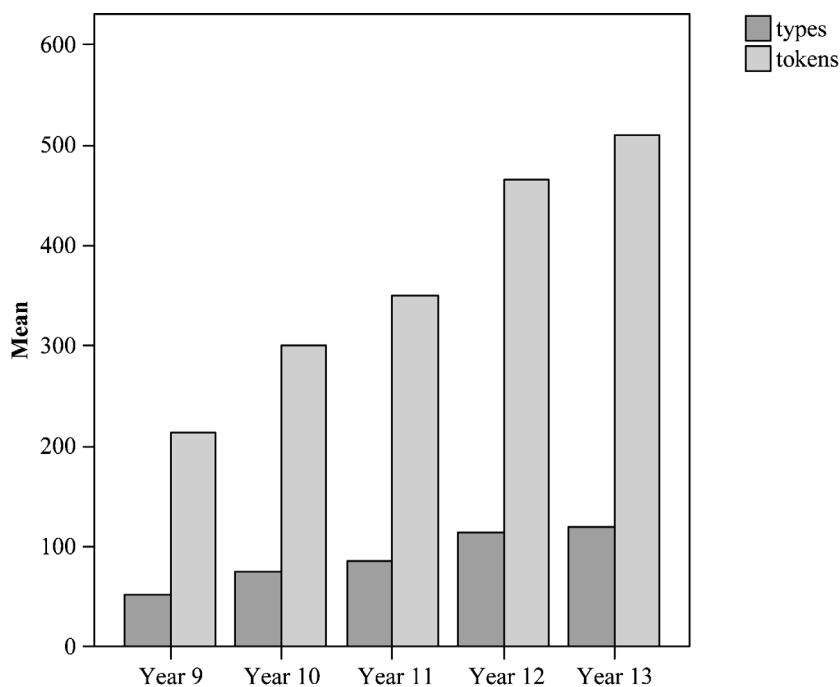
Figure 1. *Mean number of types (number of different lemmas) and tokens (total number of lemmas) produced.*

Table 2. *Mean scores and standard deviation for D per year group*

		D
Year 9	**Mean**	**15.62**
	Std. Deviation	5.43
Year 10	**Mean**	**19.88**
	Std. Deviation	4.87
Year 11	**Mean**	**23.01**
	Std. Deviation	5.54
Year 12	**Mean**	**28.41**
	Std. Deviation	5.87
Year 13	**Mean**	**26.39**
	Std. Deviation	5.46

productive lexical diversity increases between each year group. T-tests reveal that there are significant differences between Years 9 and 10 ($p < 0.05$), Years 10 and 11 ($p < 0.1$) and Years 11 and 12 ($p < 0.05$). There are no significant differences between Years 12 and 13.

The developmental thread emerging from D is further illustrated with Figure 2. One aspect of that chart that needs to be explained is the apparent decrease of D
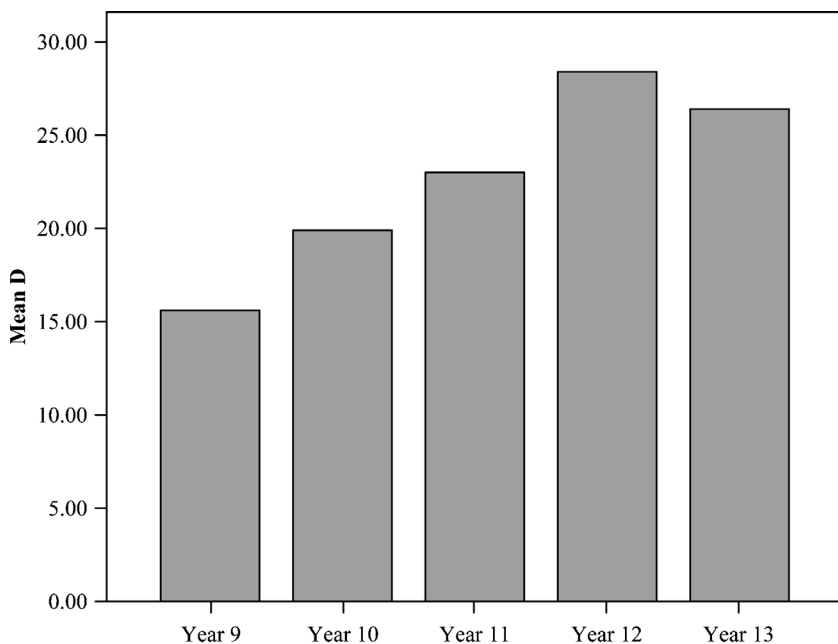
Figure 2. *Mean D score per year group.*

between Year 12 and 13. First of all, the decrease is not statistically significant. But one reason for it might be that the students who did the task in Year 12 were the same who were asked to do it again in Year 13. It would appear then that these students might not have been trying as hard as in the previous year. This part of the data collection process was longitudinal and not cross-sectional. Therefore, there could be a task effect there.[2]

The following example illustrates the level of vocabulary and language a Year 13 student produced during the oral task. In this example, the student recalls her Christmas holidays with very little prompting from the researcher. She uses adverbs, adjectives, conjunctions to coordinate her story as well as a range of verbs to describe different events.

*P31: ehm j' ai passé les vacances avec ehm ma famille.
*P31: ehm <c' était> [/] c' était absol(ument) [//] ab(solument) [//] absolument super.
*P31: euh le matin euh j'ai visité ma grand_mère.
*P31: et nous avons ouvrir des cadeaux et euh tous les choses euh pour la famille.
*P31: et nous avons <mangé le> [/] mangé le déjeuner là aussi.

[2] Or as one reviewer suggested, 'informant fatigue' where the learner behaves differently in response to the same task being put forward again (manifested by a lack of motivation in some cases).

P31: ehm j' ai passé ehm le soir encore avec ma famille avec ma mère ehm et mon frère.
P31: et euh <ça c' est> [//] ça c' était tout pour < le jour> [/] euh le jour de Noël.
P31: mais ehm le jour après je [//] ehm mon frère est venu en Angleterre parce qu'il [/] il [/] euh il habite en Irlande.
P31: alors il [/] il [/] il est venu en Angleterre.
P31: et je [/] euh je passer le jour après <avec il> [//] avec lui.

The D values (as shown in Table 2) are relatively low compared with those found by Malvern *et al.* (2004). They found that students taking their oral examination at the end of Year 11 had mean D values of 56.9 (Malvern *et al.*, 2004: 102). The students in the current study have a mean D value of 23.01 in Year 11 and 28.41 in Year 12. However, the difference in scores is most likely due to the fact that the present study used lemmas as the unit of analysis and Malvern *et al.* counted inflected forms as different forms. The present lemmatised D values are higher, however, than those found by Tidball and Treffers-Daller (2007) when they report mean values of 18.78 for level 1 students (i.e. first year university students of French with A level qualifications). One factor that could explain this difference is the nature of the task undertaken. Tidball and Treffers-Daller gave their students a story-telling task based on cartoon strips. As pointed out by Laufer and Paribakht (1998), free active and controlled active tasks can yield different results. The nature of story-based tasks (which we could class as controlled active) means that students repeat certain words as they attempt to describe things happening to the same characters, for example, thus limiting their active vocabulary. We are of course aware that it is not possible to compare absolute values of D as the elicitation materials were different in all three studies discussed here. The comparisons are used here purely to illustrate the importance of the elicitation material and of the effect of lemmatisation on D-values.

Types of words used[3]

We have established that lexical diversity develops more or less constantly from Year 9 to Year 13. Our next step is to find out the composition of the productive lexicon.

We chose to run type-type ratios. We followed a method proposed by Kauschke and Hofmeister (2002) where types belonging to each word class are analyzed as a proportion of all word types.

Figure 3 shows the mean percentage of noun and verb types for each year group. What this shows is that, in the overall increasing vocabulary, the proportion of nouns increases between Years 9 and 10 and thereafter it consistently decreases. This means that, after Year 10, a growing proportion of the words that are acquired

---

[3] In this section of the results, lemmas are still used as the unit of analysis but they have been classified according to their grammatical category in the given context. Hence, the word *grand*, for example, could be counted as either an adjective or a noun depending on the context.
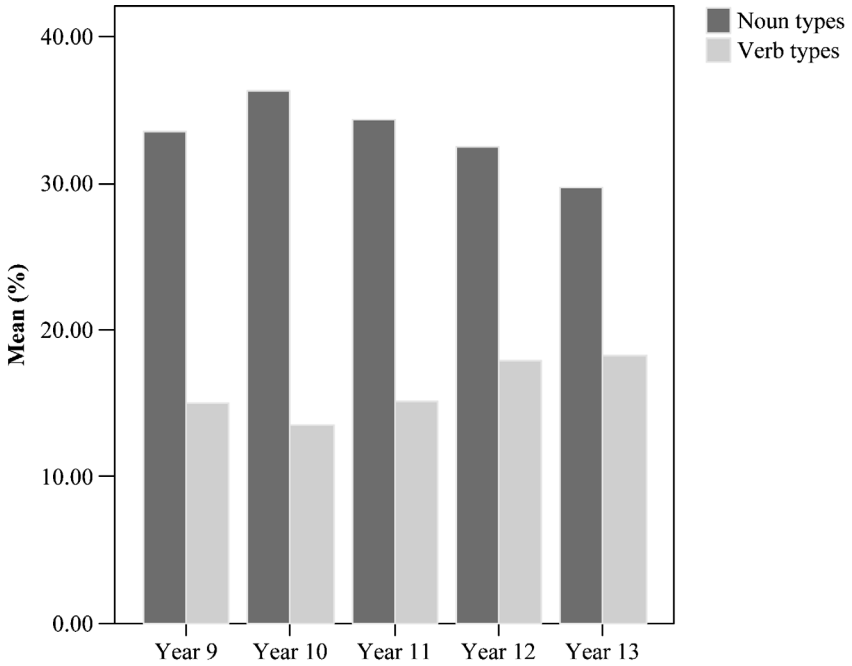
Figure 3. *Mean noun types and verb types as a proportion of all word types per year group.*

are not nouns anymore. On the other hand, the proportion of verbs has the opposite effect: learners do appear to have an increasing larger proportion of verb types in their lexicon from Year 10 onwards. Between Year 9 and 10, the proportion of verbs decreases. It is possible that as well as the increasing proportion of verbs, other categories (e.g. adjectives) increase too. This should be the subject of further analyses. When comparing means, an analysis of variance shows that the difference in noun type proportions is statistically significant across the different year groups ($F_{(4, 95)} = 5.749$, $p < 0.001$). There is also a strong negative correlation between the proportion of noun types and verb types ($r = -0.315$, $p = 0.001$). This means that as the proportion of noun types decreases the proportion of verb types increases. Overall, however, there are always more nouns used than verbs, confirming an early noun-bias in the early stages of lexical production. This appears to be in line with L1 acquisition data (see Kern, 2007, for data on French, and Caselli, Casadio and Bates, 1999, for cross-linguistic comparisons). Below are examples of two learner's production in Years 9 and 13 respectively representing an apparent move from a more 'nouny'-type production to a more advanced production.

Example 1:
P11: oh euh # nom de garçon?
FLO: c'est David.

P11: nom de fille?
FLO: elle s'appelle Lisa.
P11: um quel âge le garçon?

Example 2:
P24: ehm pourquoi est ce qu'elles euh sont allées là?
VIV: elles travaillent comme bénévoles oui.
P24: ehm qu'est ce qu'elles ont fait là?
VIV: alors ici ehm elles ont plongé.
P24: ehm qu'est ce qu'elles veulent faire à l' avenir?

Table 3. *Mean, standard deviation, minimum and maximum Limiting Relative Diversity (LRD) for each year group*

	N	Minimum	Mean	Maximum	Std. Deviation
Year 9	3	.27	.3533	.51	.13577
Year 10	9	.23	.3456	.58	.10990
Year 11	10	.28	.3430	.41	.04423
Year 12	13	.24	.3323	.55	.09084
Year 13	13	.25	.3523	.50	.07316

From Example 1, we can see that student P11 (in Year 9) makes no use of verbs and simply uses nouns to ask questions. In contrast, student P24 (in Year 13) does use verbs in the right places and in morphologically and syntactically complex utterances. It is important to note that type–type ratios are dependent on text length. Subsequently, it is possible that this result is influenced by the growing mean number of types and tokens produced by the learners across the year groups.

Secondly, we use the Limiting Relative Diversity measure (LRD) for verbs over nouns to allow us to compare the diversity of nouns and the diversity of verbs within the productive lexicon of the learners.

Overall, as Table 3 shows, LRD indicates, for all year groups, that the noun category is (token for token) more diverse than the verb category (overall mean ratio verbs/nouns: 0.345). Furthermore, no trend is evident and no statistical difference is present across the groups ($F_{(4, 43)} = 0.101$, $p = 0.981$). This means that the relative diversity of the two categories remains stable. However, as evidenced in Table 3, LRD could only be calculated for a small number of students as most did not produce enough (lemma) tokens (a minimum of 50 is required by the program). So, the lack of a pattern could be due to simply the lack of data. Alternatively, it could simply be that new verbs and nouns are developed in the same way as those learnt previously and that relative diversity of the two categories remains stable.

DISCUSSION AND CONCLUSION

Previous studies have shown that the overall (receptive/passive) vocabulary of French instructed learners increases (e.g. Milton, 2006). This study shows that

second language learners' lexical diversity increases in a rather constant positive trend from near-beginner stages through to more advanced stages. Their ability to use that increasing knowledge observed by others also increases. The results (as calculated by D) show clear progressive improvement in lexical diversity and range. It could be said that there is a 'slow down' in or around Years 10 and 11 as the difference found between those two groups was only marginally significant (at 10% level while others were at 5%). However, this 'slow down' in lexical diversity is not observed at the same time as that observed by Milton (2006), for example. He finds a slower growth in receptive vocabulary acquisition between the second and fourth year of teaching (i.e. Years 8 and 10). Clearly this will have to be further investigated with a study looking at both receptive and productive vocabularies. The largest increase in our data was found between Years 11 and 12. But this is to be expected as Year 12 learners are those who have opted to further their learning of French to an advanced level. There is no observed difference between Years 12 and 13. This could be due to the test repetition effect or it could be an indication of a ceiling effect on that particular task. This will need to be further analysed with comparisons with native speakers' productions performing the same task.

As far as the composition of the lexicon is concerned, the two measures used to assess the use of verbs and nouns enable us to present a more detailed picture of the lexical diversity of learners. Throughout the study period, more nouns types are used than verb types. This noun-bias is more pronounced in the earlier period of study (between Years 9 and 10). The noun-bias between Years 12 and 13 is not due to the increase of the proportion of verb types as this remains largely constant. This suggests that learners are using more types of words which are neither nouns nor verbs. Further analyses taking into account more parts-of-speech need to be carried out. In spite of the greater rate of increase in verb types, the relative diversity of verbs to nouns remains stable. This indicates that, in this task, learners vary the nouns they use to a larger extent than their verbs. It has been suggested in the literature that nouns were acquired first because the nouns children use label concrete, individual and enduring objects (e.g. Gentner, 1982). However, this is not an argument which is completely valid as far as L2 learners are concerned. Myles (1995, 2004) suggests that verbs take longer to be acquired, and therefore used, as they require more processing. Knowing a verb involves knowing its argument structure: knowing about what kind of complements/subjects they require. This is often too complex for beginner learners. There is little morphology on French nouns (compared to verbs) which may make them easier to process and the use. An alternative option would be that learners are mostly taught nouns in earlier stages, at least. A further study controlling input would be useful to test this hypothesis.

The picture of productive lexical development we have provided is far from being complete. It represents only one aspect of vocabulary (productive) and it is limited by the semantic fields relating to the pictures. As with any measure of productive lexical knowledge, there is no indication of the limits of knowledge of our informants. It would be interesting to compare the production of learners in light of the

distinction Laufer and Paribakht (1998) make (i.e. controlled active versus free active). A different type of elicitation task, more controlled (e.g. story-telling) might help in that respect. This study has highlighted the issue of comparisons based on the same measure (D or any other measure of lexical diversity) using different task types. This is a topic which warrants further research in the future.

Address for correspondence:
Annabelle David
School of Modern Languages
Old Library Building
Newcastle University
Newcastle upon Tyne
NE1 7RU
UK
e-mail: annabelle.david@ncl.ac.uk

REFERENCES

Bassano, D. (2000). Early development of nouns and verbs in French: exploring the interface between lexicon and grammar. *Journal of Child Language*, 27: 521–559.

Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20.1: 1–25.

Caselli, C., Casadio, P. and Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, 26: 69–111.

Childers, J. B. and Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In: K. Hirsh-Pasek and R. Golinkoff (eds.), *Action Meets Word: How Children Learn Verbs*. Oxford: Oxford University Press.

Clark, E. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26.2: 227–248.

Daller, H., Milton, J. and Treffers-Daller, J. (eds.). (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

Daller, H. and Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 150–164.

David, A. (2004). *The developing bilingual lexicon*. Unpublished PhD thesis, Newcastle University.

David, A. (2007). *Do story (re-)telling tasks really tell us something about lexical development?* Paper presented at the Models and Concepts: practical needs and theoretical approaches in modelling and measuring vocabulary knowledge (ESRC seminar series), Swansea, UK, 6–7 July 2007.

Eyckmans, J., van de Velde, H., van Hout, R. and Boers, F. (2007). Learners' response behaviour in yes/no vocabulary tests. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 116–132.

Fitzpatrick, T. (2007). Productive vocabulary tests and the search for concurrent validity. In H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 116–132.

Gass, S. (1999). Discussion. Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21: 319–333.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj, II (ed.), *Language Development: Vol. 2. Language, Thought, and Culture*. Hillsdale, NJ: Erlbaum, pp. 301–334.

Guiraud, P. (1954). *Les Caractéristiques statistiques du vocabulaire*. Paris: Presses Universitaires de France.

Hall, J., Paus, C. and Smith, J. (1993). Metacognitive and other knowledge about the mental lexicon: do we know how many words we know? *Applied Linguistics*, 14.2:189–206.

Kauschke, C. and Hofmeister, C. (2002). Early lexical development in German: a study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language*, 29: 735–757.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, 27.3: 227–250.

Kroll, J. and Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. Nicol (ed.), *One Mind Two Languages: Bilingual Language Processing*. Cambridge, MA: Blackwell, pp. 49–71.

Laufer, B. and Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning contexts. *Language Learning*, 48.3: 365–391.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk, 3rd edn, vol. 2. Mahwah, NJ: Erlbaum.

Malvern, D., Richards, B., Chipere, N. and Durán, P. (2004). *Lexical Diversity and Language Development: Quantification and Assessment*. Basingstoke: Palgrave Macmillan.

McCarthy, P. and Jarvis, S. (2007). Vocd: a theoretical and empirical evaluation. *Language Testing*, 24.4: 459–488.

McKee, G., Malvern, D. and Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15.3: 323–337.

Meara, P. and Milton, J. (2003). *X_Lex The Swansea Levels Test*. Newbury: Express.

Milton, J. (2006). Language lite? Learning French vocabulary in school. *Journal of French Language Studies*, 16.2, 187–205.

Milton, J. and Meara, P. (1998). Are the British really bad at learning foreign languages? *Language Learning Journal*, 18: 68–76.

Myles, F. (1995). Interaction between linguistic theory and language processing in SLA. *Second Language Research*, 11.3: 235–266.

Myles, F. (2003). The early development of L2 narratives: a longitudinal study. *Marges Linguistiques*, 5: 40–55.

Myles, F. (2004). From data to theory: the over-representation of linguistic knowledge in SLA. In R. Towell and R. Hawkins (eds.), *Empirical evidence and theories of representation in current research in Second Language Acquisition*: Special Issue of Transactions of the Philological Society, pp. 139–168.

Myles, F. (2005). Interlanguage corpora and second language acquisition research. *Second Language Research*, 21.4: 373–391.

Myles, F. (2007). Using electronic corpora in SLA research. In: D. Ayoun (ed.), *Handbook of French Applied Linguistics*. Amsterdam and Philadelphia: John Benjamins, pp. 377–400.

Myles, F. and Mitchell, R. (2004). Using information technology to support empirical SLA research. *Journal of Applied Linguistics*, 1.2: 169–196.

Nation, I. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, P. (2007). Fundamental issues in modelling and assessing vocabulary knowledge. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 33–43.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Richards, B. and Malvern, D. (2007). Validity and threats to the validity of vocabulary measurement. In H. Daller, J. Milton and J. Treffers-Daller (eds.), pp. 79–92.

Rule, S. (2004). French interlanguage oral corpora: recent developments. *Journal of French Language Studies*, 14: 343–356.

Rule, S., Marsden, E., Myles, F. and Mitchell, R. (2003). Constructing a database of French interlanguage oral corpora. In: D. Archer, R. Rayson, E. Wilson and T. McEnery (eds.), *Proceedings of the Corpus Linguistics 2003 Conference*, Vol. 16. University of Lancaster: UCREL Technical Papers, pp. 669–677.

Tardif, T., Shatz, M. and Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian and Mandarin. *Journal of Child Language*, 24: 535–565.

Tidball, F. and Treffers-Daller, J. (2007). Exploring measures of vocabulary richness in semi-spontaneous French speech. In H. Daller, J. Milton and J. Treffers-Daller (eds), pp. 133–149.

Verhallen, M. and Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics*, 19.4: 452–470.

Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In: P. Bogaards and B. Laufer (eds.), *Vocabulary in a Second Language*. Amsterdam and Philadelphia: John Benjamins, pp. 173–189.

Waring, R. (1997) A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata: The Occasional Papers at Notre Dame Seishin University*, 94–114.

French vocabulary breadth among learners in the British school and university system: comparing knowledge over time

JAMES MILTON

Swansea University

ABSTRACT

This paper investigates the scale of vocabulary learning that appears to be common among British foreign language learners of French up to university graduation. These results are compared with historical data and the knowledge learners possessed in previous generations. This is done by comparing current students' vocabulary knowledge with students who studied more than 20 years ago. These comparisons provide very powerful evidence in support of claims made to the Qualifications and Curriculum Authority of a progressive decline in the knowledge of learners and the standard of school examinations over a period of decades.

1. INTRODUCTION AND BACKGROUND

It is hard to overstate the importance of learning vocabulary in the process of learning a foreign language. Long and Richards (2007: xii) report that it is 'the core component of all the language skills'. Wilkins (1972: 111) emphasises its importance by noticing that, 'without grammar very little can be conveyed, without vocabulary *nothing* can be conveyed'. It is surprising, then, that we know so little about this aspect of language knowledge in British schools where the learning of a foreign language remains mandatory. A recent renaissance in the study of vocabulary learning has been restricted, largely, to the knowledge and progress of EFL learners. Macaro (2003: 251) comments that it is a feature of research in foreign languages generally that very little literature exists which has a European, let a lone a British, perspective and which systematically investigates foreign languages other than English.

The collection of data on vocabulary knowledge and vocabulary size ought to be a highly useful exercise in the British, or any other, system for a number of reasons. Firstly, it allows teachers and learners to chart their progress in an important area of knowledge over the course of learning. Learning a foreign language is probably the largest and most academically taxing undertaking any student engages in at school and it is often difficult to detect real progress in knowledge or ability. But vocabulary is an area where recent innovations in testing method can make even relatively small

incremental gains measurable. Being able to see progress, where it occurs, can be a reassurance and a motivation to all. Secondly, vocabulary knowledge can be countable in a way that other aspects of language knowledge are not. A mark of 8 out of 10 for an essay in French, for example, does not imply that the learner has twice the knowledge or ability of another learner who scores 4 out of 10. These are grades rather than scores and it is often hard to explain definitively how one grade of this kind relates to another. But a learner who knows 2000 words in a foreign language can be argued, rather more persuasively, to have twice the knowledge of someone with only 1000 words. Measuring the vocabulary knowledge of learners can help give a much better impression of the scale of learning which is taking place than is possible with other measures of language proficiency. Thirdly, and connected with the last point, it allows comparisons between learners, between groups of learners, and between systems to be made rather more systematically. Foreign language learning, and the learning of French in particular, is a subject area which is under criticism for a decline in standards (for example, Qualifications and Curriculum Authority, 2002; Milton, 2006) but these concerns are hard to quantify or demonstrate explicitly. Despite these concerns the official line, at least from schools' ministers, is that there is no decline in standards and that the increasing numbers who take and pass exams, for example at 'A' level, is the result of improved teaching (Adonis, 2005: 2). If the techniques and systems we have for teaching foreign languages really are getting better then we ought to see this in learners' vocabulary knowledge and any differences ought to be quantifiable.

The broad purpose of this paper, therefore, is to report on the knowledge of learners of French in the British educational system, and to investigate in more detail how these scores compare with historical data and the knowledge of learners from previous eras. It will test learners who studied French more than a generation ago, to ascertain their vocabulary knowledge. A more considered conclusion can be reached as to whether the vocabulary knowledge of learners has changed over recent years.

2 . LEVELS OF VOCABULARY KNOWLEDGE IN BRITISH SCHOOLS

A study of 449 learners across all levels in a British school (Milton, 2006) gave the results reported in Table 1.

These scores, which are estimates of the vocabulary knowledge of the 5000 most frequent words in French, suggest that learners reach GCSE after five years of study with about 850 of French on average, and take 'A' level two years later with just under 2000 words. Students, on average, gain about 170 words per year up to GCSE, and about 535 words per year between GCSE and 'A' level. When time in class is taken into account it emerges that learners progress at about 2.7 words per contact hour up to GCSE and at about 4.1 words per hour during the course of 'A' level study. These results are broadly similar to Milton and Meara's (1998) results which suggest that learners in a variety of language settings learn

Table 1. *Vocabulary knowledge in a British school, divided by year of study (Milton, 2006, 192)*

year	learners	number	hours/year	mean	max	min	Sd
1	School year 7	80	78	311	900	0	193
2	School year 8	78	58.5	411	1150	0	228
3	School year 9	106	58.5	460	1150	0	261
4	School year 10	57	78	592	1350	50	263
5	School year 11	49	78	852	1800	0	440
6	School year 12	10	175.5	1555	2000	1100	349
7	School year 13	69	117	1930	3100	650	475

about three to four words per classroom hour. They are also, broadly, confirmed by recent replications (for example, David, forthcoming.).

While these scores take us only to the end of formal school study, they appear to provide a cogent picture of the way language develops. Vocabulary size tends generally to increase over the course of study as might be expected, and learners taking 'A' level, for example, have greater vocabulary knowledge than those taking GCSE exams, and these learners have greater knowledge than beginners at the end of only one year of study. Milton (2006: 193) also reports a statistically significant relationship between vocabulary size and 'A' level grade; the larger a learner's vocabulary size the more likely they are to get a higher grade.

This study has been extended to include 66 students taking single and joint honours French at a British university using the same testing methodology. Students in their first, second and final year were tested in October at the very start of the academic year and again in June at the end of the academic year's study. Students taking the intercalary year abroad, however, could not be tested. While not a complete sample of students on these courses, it is thought that the results are representative of learners at this university, and a similar study in another British university by David (forthcoming) produces almost identical scores. The scores the university level students obtained are reported in Table 2 and when their year final scores are added to those of the students in the Milton (2006) school study, the progress it implies over the whole of foreign language learning in the British educational system can be seen in Diagram 1.

It appears that foreign language learners at university continue to learn vocabulary within the 5000 word band at the same kind of rate that they did at 'A' level; about 500 words per year. If it were not for the intercalary year progress would be completely, and rather impressively, regular from GCSE onwards.

Table 2. *Vocabulary knowledge in a British university, divided by year of study*

level	n	mean entry	mean exit	mean change	SD
1	29	1950	2555	605	678
2	20	2367	2930	563	828
4	17	2832	3326	504	579
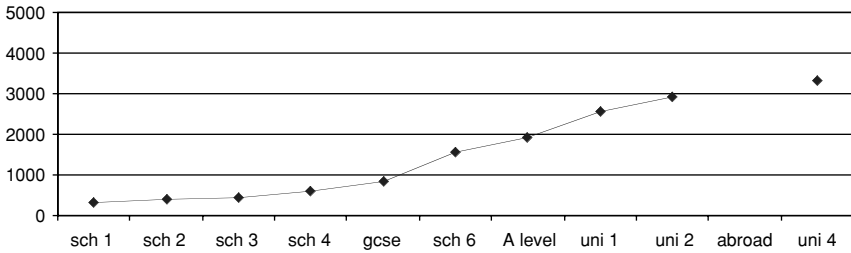
Figure 1. *Year-end mean vocabulary scores at a British school and university.*

The scores suggest some features that deserve closer investigation and explanation. Firstly, it appears that little vocabulary is learned during years 2, 3 and 4 of study although it is not immediately apparent why this should be so. If progress can be sizeable in year 1 and both sizeable and regular after year 4, why should this hiatus in learning occur? It appears that learners have been held back in this period. Secondly, it is hard to discern vocabulary gains during the year abroad. These figures suggest that learners enter their final year with approximately the same level of vocabulary knowledge that they had on completion of their second year, a feature also apparent in David (forthcoming) data. It is almost as though the intercalary year does not exist and learners progress after it much as they progressed before. This probably is not a ceiling effect. Learners are not approaching the maximum score of 5000 and native speakers and genuinely able French speakers, university lecturers for example, routinely score 5000 or very close.

In addition, the scores appear quite modest in scale given the years devoted to learning including, in the case of university graduates, a whole academic year spent in France. It raises the question whether school learners have the scale of vocabulary knowledge appropriate for communication in French at the CEFR's A2, B1 and B2 levels in the formal exam system. Meaningful communication in anything but the most contrived and predictable circumstances must be difficult with so few lexical resources. Learners at these levels who study English as a foreign language tend to have much larger vocabularies when tested by the same methodology. It raises too, the question of what level of competence we expect of our foreign language graduates when they appear to be far away from native-like levels of knowledge of frequent and important vocabulary.

These, then, are the levels of vocabulary knowledge current learners display. What evidence is there that standards have changed or that learners in the past knew more and were more competent?

3. HAS THERE BEEN A DECLINE IN STANDARDS OF FRENCH AT SCHOOL?

There are some reasons for thinking that French in particular has suffered a decline in the standard of the milestone exams at age 16 and 18, and a reduction in what we expect learners to gain during their language study. Reports from the Qualifications and Curriculum Authority concerning French GCSE, for example, report that 'there was evidence that the standard declined between 1986, 1991 and 1996. This was not confined to accuracy, but also included linguistic range and critical response'. (Qualifications and Curriculum Authority, 2002: 1). It will be appreciated that in a subject where formal exams rely very heavily on the subjective evaluation of essays and oral interviews by many different markers, and for a quality as difficult to measure as critical response, it can appear hard to be certain of such changes, still less to quantify the scale of any change.

Milton (2004) points to a further problem in trying to gauge the maintenance of standards over time. Prior to 1988, it seems, the numbers of learners taking formal exams, the number of schools offering these exams, and the nature of the exams themselves, had remained fairly constant over time which allowed markers to norm reference their assessments. The introduction of GCSE changed both the number and nature of the students taking the age 16 exam, and the nature of the exam itself. Norm referencing was no longer possible and it is far from clear that criterion referencing has satisfactorily carried out the function of moderating the level of exams. Milton's (2004) paper points to the disproportionate increase in pass rates after the introduction of GCSE, and in the proportion of grade A awards in particular, which suggest grade inflation. The same paper analyses the lexical sophistication of the French texts used in 'O' level and GCSE over a 20-year period and reports a decline in lexical difficulty in these texts after the introduction of GCSE. These conclusions suggest both a decline in the knowledge and ability of learners, and in the standard of the age 16 exam. While is hard to find a change in teaching methodology which could produce an increase of 250% in the number of grade A awards within two years of the introduction of GCSE, a reduction in the difficulty of the exam could certainly produce this change.

The current levels of vocabulary knowledge reported in the previous section add to this picture. A number of studies, now rather old, review the scale of vocabulary learning before and just after the Second World War and in every case they report that learners had higher rates of vocabulary acquisition than are measurable today leading to the conclusion that '. . . the average French learner [today] learns less vocabulary than was the case two generations ago' (Milton, 2006: 202–203). There are methodological differences between the tests administered to current learners and historical learners, however, and while the size of the difference in their

vocabulary sizes means that this conclusion is probably valid, it is hard to estimate how large that difference really is.

4. MEASUREMENTS OF VOCABULARY SIZE IN BRITAIN BEFORE GCSE

Studies reporting vocabulary knowledge among foreign language students in the British education system are few in number. Some of them are now distinctly venerable. Robson (1934) examined the first year of French learning in 16 schools of varying kinds: elementary, central, secondary, convent, high and grammar schools. The nature and range of teaching, and the volume of learning which resulted, appears very diverse. Classroom time varied from 25 to 130 hours, vocabulary exposure varied from 212 to 1112 words and passive vocabulary knowledge at the end of study varied from 132 to 665 words. Burns (1951) studied first year learners of French in British Grammar Schools. The mean estimated productive vocabulary was 555 words among these learners, and the mean passive receptive vocabulary was 718 words at the end of the academic year (p. 44).

Erigna (1974) estimated that learners with five or six years of French, presumably those who had completed 'O' level and who might have progressed to their first year of 'A' level in the exam system of the time, possessed receptive vocabularies of 2500 to 3500 words with productive vocabularies of 1000 to 1500 words. It is not clear how these figures have been arrived at. Nonetheless, these particular estimates sit comfortably with the anecdotal evidence from the literature of language learners who took 'O' and 'A' level in the 1960s and 1970s and learners who have been tested for this study. These learners report that they expected to learn about 4000 words for 'O' level and double this to succeed at 'A' level. These kinds of estimates also sit well with learner recollections of how they managed to learn these volumes of vocabulary. Learners recall being given 60 words a week as part of the routine of foreign language study (Fine, 2005: 7). Sixty words a week, every week over several years of study could certainly result in a vocabulary approaching 4000 words after five years even if some of the words were repeated and others forgotten. If the learners reported by Robson and Burns were to continue to learn at the rate they did in their first year, then, they also could have acquired the several thousand vocabulary items which Erigna reports. These, historical studies, are consistent with each other, therefore, and appear to tell a believable story about the scale of vocabulary learning at that time.

Another feature of Erigna's study (1974: 260) is that it reports on the time spent learning French in school. In the first three years of French study the learners receive 280 hours of classroom instruction; about 50% more classroom input than current learners receive in the same period (see Table 1). Again, this figure is very similar to learners' own recollections where a 40 minute French class almost every day at school before 1980 is reported.

In all these articles the volumes of vocabulary being learned appear substantially larger than the recent studies suggest. However, differences in the testing method make this difficult to confirm unequivocally and make it difficult to estimate the

scale of the difference. This difficulty is not easy to overcome unless all the learners take the same test so that scores can be directly compared.

5. EXTENDING THE COMPARISON USING THE SAME TESTS

To address this problem, the specific objectives of this study are:

- to test the vocabulary size of a sample of learners who studied French before GCSE and who took 'O' level around the time of Erigna's study, and
- to compare the scores they obtain with those gained by modern learners taking GCSE and other exams.

6. METHODOLOGY FOR ESTIMATING VOCABULARY SIZE

Since the 1930s the methodology for selecting the vocabulary to teach has been, at least in principle, relatively systematic. Writers generally attempt to incorporate into foreign language learning materials the most frequent words in the language. The most frequent words in a language provide most coverage of a text and are, arguably, the most useful as a consequence. Robson (1934: 265) draws explicit attention to *Longman's Modern Method French Book* where the contents are derived in this way, and modern textbooks continue to identify frequency criteria in their explanations of vocabulary choice (e.g. Rixon, 1990: 5). If this principle were rigorously applied then it would make testing relatively straightforward; a test of the most frequent words in the language would produce a good estimate of the learners' knowledge. However, books also need subject and thematic organisation if they are to be usable, and strict frequency-based selection does not allow this. Some vocabulary must, as a consequence, fall outside the most frequent lexical bands and this material can be chosen idiosyncratically which makes a single test for all users more difficult to construct.

To overcome these differences in the vocabulary the learners are exposed to, investigators who examine vocabulary size have traditionally worked with beginners only and have based vocabulary estimates on tests of the contents of the books the learners use (for example, Robson, 1934 and Burns, 1951). While these tests ought to have a strong frequency bias, it is unclear precisely how well these scores can compare with each other or with modern tests which are more strictly derived from frequency data and are designed to estimate knowledge of learners from a wide variety of learning backgrounds and at a variety of levels of knowledge and ability.

The test used to estimate vocabulary size in Milton (2006), for example, is an adaptation of the *X_Lex* vocabulary levels test (Meara and Milton, 2003), an EFL vocabulary size test which tests knowledge of each of the first five, 1000 word frequency bands, estimates overall knowledge of this vocabulary, and gives an overall score of words known out of 5000. The French version is based on Baudot's (1992) frequency lists. Since the most frequent 5000 words in Baudot's word lists provide

Please look at these words. Some of these words are real French words and some are invented but are made to look like real words. Please tick the words that you know or can use. Here is an example.

☑ chien

☐ de ☐ distance ☐ abattre ☐ absurde ☐ achevé ☐ manchir

Figure 2. *Illustration of test format.*

fractionally under 95% coverage of the whole of his 1.2 million word corpus, this suggests that the results of this test might compare quite well with older tests. The older tests, and the books they were derived from, would have to contain very large volumes of infrequent and largely un-useful vocabulary to produce substantially different results. Nonetheless, since the content of these tests can no longer be checked this is still a potential source of difference and ideally learners should take the same test so results can be directly compared. The 'O' level takers in this study have therefore taken the French version of *X-Lex* like the learners in Milton (2006).

The French version of *X-Lex* is a Yes/No test which presents learners with 120 words, one by one. Learners have to indicate whether they know each word. There are 20 randomly selected words from each 1000 word frequency band and a further 20 pseudo words which are designed to look like words in French but are not real French words. The number of Yes responses to these pseudo words allows the score on the real words to be adjusted for guessing and overestimation of knowledge. Each Yes response to a real word scores 50 on this test, and each Yes response scores –250, allowing the adjusted estimate to be made. As in the Milton (2006) study testees with guessing rates above 20% were excluded. The test was delivered in a paper and pencil (15 testees) and computer versions (6 testees). Adamopoulou (2000) suggests in practice the scores produced by these two difference forms of the test vary scarcely at all. The test instructions, an example, and the first line of test items in the pencil and paper version of the test are shown in Diagram 2.

7. THE SUBJECTS IN THIS STUDY

Test data was collected from 21 learners who had studied to 'O' level as part of their school study but who had not pursued their French, or used the language, subsequently. These learners varied in age from forty to over sixty. No regard was

Table 3. *Vocabulary knowledge among learners who took 'O' level*

number	mean	max	min	Sd
21	2188.1	4800	850	1259.75

paid to the grade these testees had received at 'O' level and they varied from a Grade 1, the top grade, to a Grade 9, which was the greatest degree of graded failure possible. Testees were excluded if they had continued their study of French, say at evening class, or had lived for any length of time in France, although the occasional holiday in France was accepted. These criteria ruled out a large number of 'O' level takers and the remaining sample is opportunistic to a degree, but it should allow some estimate to be made of the vocabulary gained and retained by learners who studied to 'O' level more than twenty years ago.

It must be recalled that the proportion of the age group who took 'O' level French in the 1950s, 1960s and 1970s was much smaller than the proportion who now take GCSE. One of the reasons for making the change to GCSE was to increase the numbers taking milestone exams at age 16 and leaving school without qualifications. The 'O' level takers will have had a different learning experience. The majority of them had studied Latin at school and often another foreign language, and went on to study at university. The majority had also gone on to professional careers in areas such as education, the civil service and the social services. By no means all followed this career path, however, and 8 of the 21 testees had not gone to university, had remained housewives or had taken jobs such as bank cashier or travelling salesmen.

8. RESULTS

It is not immediately apparent what scores might be expected from the 'O' level group. It might be thought in the light of the studies discussed above, that these learners will produce larger scores on the vocabulary test than current learners, reflecting the greater vocabulary knowledge which was reported at the time. However, in all cases, considerable time has passed since the language was formally studied or even used and some attrition of word knowledge must have taken place. It is not entirely clear, therefore, what results may emerge. The results of the learners who took 'O' level are summarised in Table 3.

Even after at least two decades of not using or learning French, and in some cases nearly fifty years of non-use, learners recall on average over 2000 words. There must be some attrition here. Levels at the time of taking their French language 'O' levels might reasonably be expected to have been higher. The highest scorer on the test, a lady now in her sixties, does not appear to have a freak test score but really did learn large numbers of vocabulary lists, lists of bird names and trees for example, many of which she can still recall.

9. DISCUSSION

9.1 Comparing 'O' level and GCSE vocabulary knowledge

At the outset of this paper it was suggested that the levels of French vocabulary known by current learners might be compared with the knowledge of learners who studied several decades ago in order to ascertain whether knowledge appears to have changed over time. The results obtained from the group which took 'O' level all those years ago suggest that passive vocabulary knowledge in a foreign language can be remarkably robust and that large volumes of words can be retained over decades even when the language is not being used. The mean score obtained is close to the bottom level of vocabulary knowledge Erigna (1974) suggests for learners at this level. This difference can be accounted for in two ways. One is that Erigna's figures must have included some vocabulary outside the 5000 word range and this will not appear in the current estimate. But this may not be a huge amount because of the weight of the second reason. Notwithstanding the way this vocabulary has survived, there must have been some attrition over time and learners will have known more when they actually took the exam. It seems to me, therefore, that these figures support the idea that the estimates of vocabulary knowledge which were made some 30 years ago are consistent with those made using current methodology. Current results can also be compared directly with historical vocabulary size estimates despite the change in testing method. What does this tell us about the way French vocabulary knowledge in schools has changed over time?

At one level the conclusion to be drawn is very clear. Learners who took 'O' level all those years ago had greater vocabulary knowledge than equivalent learners taking the equivalent age 16 exam now. Their scores suggest they still have on average more than double, in fact closer to treble, the vocabulary knowledge of GCSE takers. Allowing for attrition over time the difference is probably even greater than this and Erigna's figures, which suggest they might have had vocabularies quadruple the size of modern learners, appear entirely believable. The importance of this cannot be under-estimated. This is not a change in abstract knowledge irrelevant to the communicative goals of modern language teaching, it implies that modern learners will be much weaker and less successful at every aspect of language, communication and language use. Learners cannot be as communicatively able with less than 1000 words of French as they will be with more than 2000 or 3000 of the most frequent words. If the objective in introducing and running the GCSE exam was to maintain the standard of the age 16 exam, so that users have a qualification we can all recognise and can understand, then GCSE has signally failed to achieve this. The vocabulary knowledge required for 'O' level in the 1970s is much closer to that required for 'A' level now, and the mean score of 'O' level takers, even after decades of non-use, is still higher than the current 'A' level mean.

It is worth asking the question whether a decline in the standard of the age 16 exam is unequivocally a bad thing. One of the features of GCSE is that it has brought into the sphere of the exam a much greater proportion of the age range,

Table 4. *Comparison of 'O' level group vocabulary scores with an equivalent GCSE cohort*

	number	mean	max	min	Sd
'O' level	21	2188.1	4800	850	1259.75
GCSE	24	1304.16	1900	950	273.83

and many more school students now know some French, and have a qualification, than was the case 30 years ago. Probably this is a good thing. But this change has been bought at a price and that price appears not to be a reduction in the level of the age 16 exam to the standard where increased numbers of learners stand a chance of completing it. This would seem to me a price well worth paying provided this change in the overall standard of the exam has not compromised the ability of the most able learners to achieve their full potential. The question arises, therefore, whether the best performing students in the modern cohort, the ones who 30 years ago would have taken 'O' level, are attaining the levels of knowledge that Erigna reports.

Figures from WJEC exams (Milton, 2004: 6) suggest that after the change from 'O' level to GCSE approximately double the number students took the age 16 exam in French. The best performing 24 learners in the Year 5 GCSE group in Table 1 have been separated out and the mean scores recalculated to allow the 'O' level group to be compared with an equivalent cohort from the modern group. The results are presented in Table 4.

These figures suggest that even most able learners taking GCSE know much less vocabulary than equivalent learners did on taking 'O level. Even the best performing GCSE students have scores below the 'O' level takers' mean score. Partly this must be the result of the reduction in the number of hours available for learning French. The figures produced by Erigna (1974) as well as the recollections of learners confirm that this has occurred. Partly too, it will be the product of the way GCSE sets a low threshold of knowledge for completion, and teachers will inevitably teach to this threshold rather than press the best learners to achieve their fullest potential. The strikingly low figure for standard deviation among GCSE learners in Table 4, suggests to me that the most able current learners are capped in some way. These figures suggest very strongly that not only is the standard of the average student taking GCSE much lower than that of the average student who took 'O' level, but that the standard of the very best GCSE students fail to match that of the old 'O' level. A lower standard at the outset of study and up to age 16 must put pressure on the remainder of the French teaching system, for example at university where students are expected to study in France and through the medium of French.

9.2 Progress at University and the Intercalary Year

If the levels of vocabulary knowledge among learners have declined so much over recent decades, how has this change come about? It is possible to provide a

Table 5. *Vocabulary gain per hour of study in a British university*

year	number	mean gain	contact hours	mean gain per contact hour	contact SD	module hours	mean gain per module hour	module SD
8	29	605	88	6.88	7.7	400	1.51	1.7
9	20	563	66	8.52	12.56	200	2.81	4.1
11	17	503	66	8.07	9.37	200	2.51	2.9

convincing argument that the comparatively slower progress in learning vocabulary which current school students make is the result of a decline in the hours available for study. At university, however, learners continue to learn at about the same annual rate even though much more time is, in theory, available to them. How can this be explained? Most of this time will be spent studying through English, however, and even literature modules are likely to be tackled through translations. I suspect that part of the answer lies, again, with the hours available for classroom language teaching. University students have a number of timetabled, classroom hours which are explicitly devoted to language development including the introduction and learning of new vocabulary. But in addition, there are a large number of hours, 200 hours per language learning module, which are intended for learners to read, to write assignments and to study independently. How much of this second group of study hours are actually devoted to learning French is uncertain and very probably will vary enormously from one individual to another. Depending on which numbers of hours is taken as the basis of calculation, progress can vary considerably. Table 5 shows these separate calculations.

If students are devoting every hour that the modules make available to them in formal French study, then the progress they are making is small at around one to two words per hour. These figures are comparable with the rate of progress in Years 2, 3 and 4 of secondary school. In reality, some of this time at least will be taken up with testing and other activities rather than purely in vocabulary learning. Nonetheless, there is evidence that non-classroom activities such as reading in a foreign language or listening to foreign language songs, can lead to very rapid vocabulary growth when properly structured (for example, Horst and Meara, 1999), and there seems little evidence that students are taking best advantage of the time available to them to advance and develop an essential area of their foreign language knowledge. When progress is recalculated so that only classroom hours are used, the figures which emerge are quite believable for able learners, and remember that university level learners should comprise only highly able and successful language learners. Figures of six to eight words per contact hour are reported for able learners elsewhere (for example, Vassiliu, 2001). My suspicion is that learners either are not using the time available for learning outside of the classroom or, if they are, then they are not taking good advantage of the time, perhaps because they do not know how to.

A second reason why many students may make only modest progress is the focus of the classroom teaching which is available to them. It appears that disproportionate
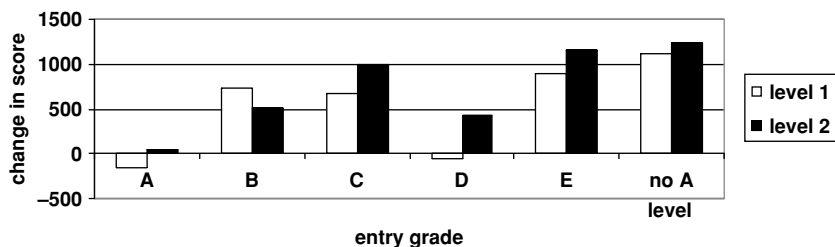
Figure 3. *Lexical growth at university and 'A' level grade.*

effort is devoted to the least able students at the expense of the able. This can be seen when the progress of the university level 1 and level 2 students is subdivided by the grade they received at 'A' level as in Diagram 3.

Students who enter with a grade A, and consequently have a comparatively large vocabulary, make little or no progress. Students who have a grade E or no 'A' level at all make very impressive progress, on average about 1000 words in the year. My suspicion is that classroom materials and teaching effort are pitched at the level of the least able, giving the more able little opportunity to learn. Where students are expected at the end of the second year at university to spend a year independently in France it is, perhaps, understandable why teachers should focus their effort this way.

If progress at university in Britain appears perplexingly slow, the effect of the intercalary year is even less easy to understand. It is hard to see any vocabulary gain among the students in this study and certainly nothing like the very large gains reported in Milton and Meara (1995) who studied European learners on an equivalent exchange at a British university. This really needs investigation in much more depth, and probably with a 10,000 word test to give a better gauge of knowledge. These learners seem to be well below the levels needed for native like fluency and accuracy and it is strange that they do not improve, given the chance to live abroad, when European learners in Britain do improve. There appear to be some important differences in the experience of the intercalary year between the two groups.

The European learners in Britain studied by Milton and Meara (1995) were taking degree level courses with native speakers, and were obliged to take lectures, write assessed essays and pass exams for credit like every other student in the university. They continued to take formal language classes and took the Cambridge Proficiency in English exam at the end of their year abroad. Even if they did not mix as expected with native-speaking students, their social life with other overseas students appears to have been conducted with English as the *lingua franca*. They have had huge exposure to English, therefore, have been put under pressure to progress in English, and been given language support at an appropriate level to help them achieve this. British students in France may have a very different experience.

It is not clear that they have to complete modules in France successfully to gain the credits they need at home, and given their level of language they would find it hard to do this. It is not clear that the British learners in France have formal French language classes nor the formal external language accreditation that their European counterparts sought. Nor do the British students in France, unless on a work placement, develop social patterns that involve the use of French. Socialising with other students frequently involved, as in Britain, mixing with other international students and the use of English, rather than French, as the general *lingua franca*. It would be wrong to disparage the potential value of time spent in the country of the language you are learning but the kind of large-scale progress which might reasonably expected from such an experience is missing in this group of learners. This should be a concern. The nature of the experience and the benefits derived from it require more and closer investigation. It seems quite likely that the nature of the student experience needs to be changed quite significantly if learners are to make the kind of progress which traditionally we have expected of them and any change will have to take into account the very modest levels of vocabulary and language knowledge that these learners now possess.

10. CONCLUSIONS

The introduction to this paper raised concerns that the standards of knowledge had declined among learners at British schools and taking formal exams in French. The examination of current and historical measurements of vocabulary size confirms that these concerns are not misplaced. The use of vocabulary size allows a comparison to be made and even allows a judgement to be made about the scale of this decline which appears enormous. Learners taking GCSE under the current system have only a half to a third of the vocabulary knowledge of learners who took 'O' level between 20 and 50 years ago. This lends credence to the figures provided by Erigna (1974) and others which suggest that learners taking 'O' level about 30 years ago had vocabularies three or four times larger than current learners. It is hard to believe the claims that the increases in the numbers passing GCSE and gaining a grade A are the product of improved teaching, when the knowledge displayed by learners has diminished this dramatically. Even the best learners taking GCSE have only a fraction of the knowledge of learners taking age 16 exams historically. Because vocabulary knowledge is so central to language performance generally this must mean that learners are much less able in French generally than they were before the introduction of GCSE. It is not hard to suggest the reasons for this decline. Pressure to increase the numbers taking and passing the exam, coupled with a reduction in the classroom hours available for learning, and a change in the exam format, have probably made retaining continuity in standards impossible.

This reduction in standards has had implications for the rest of the language teaching system and for the learners themselves. If 16-year-olds know less on taking GCSE, then progress to 'A' level becomes much harder. There is evidence here that the standard of 'A' level has likewise been compromised as a result. The

mean vocabulary size of 'O' level takers in this study is larger than the mean score for 'A' level students now entering university to study French. This in turn has consequences for the study of French at university. It appears that the efforts of French teachers are devoted to learners now entering the system with very low levels of French, and intermediate level students make little progress as a result. The cohort which completes two years of university study has levels of vocabulary knowledge which must make university study abroad through the medium of French incredibly difficult if not impossible. The effects of the year abroad have been little studied but this investigation suggests that, linguistically, learners may take little from the experience. Graduates in French now have a comparatively modest knowledge of the most frequent French vocabulary and the mean scores they produce fall within the vocabulary size range suggested by Erigna for 'O' level takers 30 years ago. This suggests that the goals of university study in a foreign language should be reassessed with native-like proficiency no longer being routinely achieved by the majority of graduates.

A final conclusion which emerges from this investigation is how potentially useful vocabulary measures can be in the process of measuring knowledge and comparing the standards of learners and exams over time. This is an area where intelligent discussion is bedevilled by the absence of objective and quantifiable information. This study has shown, however, that it is possible to make a quantified and objective measure of learners' knowledge and to draw meaningful conclusions from this information. It would surely be beneficial if learners, schools and examining authorities made greater use of these kinds of technique.

Address for correspondence:
James Milton
School of Arts
Swansea University
Singleton Park
Swansea
SA2 8PP
Wales
UK
e-mail: j.l.milton@swansea.ac.uk

REFERENCES

Adamopoulou, R. (2000). *The Equivalence of a Computer-delivered and a Paper-delivered Yes/No Vocabulary Test.* Unpublished MA Thesis; University of Wales Swansea.

Adonis, A. (2005). Speech by Andrew Adonis, Parliamentary Under-Secretary of State for Schools, Canterbury Christ Church University College, Wednesday 17 August 2005 at: http://www.dfes.gov.uk/speeches/media/documents/alevels.doc, accessed 16 January 2008.

Baudot, J. (1992). *Fréquences d'utilisation des mots en français écrit contemporain.* Montréal: Presses de l'Université de Montréal.

Burns, G. (1951). An investigation into the extent of first-year vocabulary in French in boys' grammar school. *British Journal of Educational Psychology* 21, 36–44.

David, A. (forthcoming 2008). Vocabulary breadth in French L2 learners. *Language Learning Journal*, Winter 2008.

Erigna, D. (1974). Enseigner, c'est choisir: vocabulaire-verwerving. *Levende Talen*, 306, 260–267.

Fine, A. (2005). My favourite lesson. *The Education Guardian*, 7 June 2005: 7.

Horst, M. and Meara, P. M. (1999). Test of a model for predicting second language lexical growth through reading. *The Canadian Modern Language Review/La Revue canadiennne des langues vivantes*, 56.2: 308–328.

Long, M. H. and Richards, J. C. (2007). Series Editors' Preface. In: H. Daller, J. Milton, and J. Treffers-Daller (eds.), *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press, pp xii–xiii.

Macaro, E. (2003). *Teaching and Learning a Second Language: A guide to recent research and its applications*. London: Continuum.

Meara, P. and Milton, J. (2003). *X_Lex, The Swansea Levels Test*. Newbury: Express.

Milton, J. (2004). Comparing the lexical difficulty of French reading comprehension exam texts. *Language Learning Journal*, 30: 5–11.

Milton, J. (2006). Language Lite: Learning French vocabulary in school. *Journal of French Language Studies*, 16: 187–205.

Milton, J. and Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL Review of Applied Linguistics*, 107/108: 17–34.

Milton, J. and Meara, P. (1998). Are the British really bad at learning foreign languages? *Language Learning Journal* 18, 68–76.

Qualifications and Curriculum Authority (2002) *Five yearly subject reviews: conclusions – French*, at: http://www.qca.org.uk/nq/subjects/conclusions_french.asp (accessed 06/07/03).

Rixon, S. (1990). *Tip Top Teacher's Book 1*. London: Macmillan.

Robson, W. F. (1934). The vocabulary burden in the first year of French. *British Journal of Educational Psychology*, 4, 264–293.

Vassiliu, P. (2001). *Lexical Input and Uptake in the Low Level EFL Classroom*. Unpublished PhD dissertation; University of Wales Swansea.

Wilkins, D. A. (1972) *Linguistics in Language Teaching*. London: Arnold.

Progress in learning French vocabulary in a one-year advanced course at school[1]

SUZANNE GRAHAM, BRIAN RICHARDS
AND DAVID MALVERN

University of Reading

ABSTRACT

This article reports on vocabulary learning in French during Year 12 by 150 pupils in 15 English comprehensive schools. Vocabulary progress over two school terms was investigated in the context of learner strategy training that targeted listening and writing. This intervention was carried out in eleven schools and results compared with a comparison group of four schools. Recognition vocabulary and writing vocabulary measures taken during the students' course are compared with listening and writing scores, and examination results. Results suggest that interventions which target listening and writing skills do not lead to significant gains in vocabulary through incidental learning and the deployment of vocabulary in writing tasks. However, it also appears that extra time and effort devoted to developing listening and writing is not at the expense of vocabulary learning.

1. INTRODUCTION

The decline in the numbers of pupils studying French beyond the age of 16 in England has been well documented (CILT, 2005; Fisher, 2001; Graham, 1997, 2002, 2004). Reasons for this are complex, but important factors are the difficulty learners experience, both at the General Certificate of Secondary Education (GCSE) and in the first year of post-16 study, and their perceived lack of progress when faced by the more stringent linguistic demands of the Advanced Level syllabus. The transition from Year 11 (GCSE) entails a steep increase in difficulty, arising partly from exposure to large amounts of new and more complex vocabulary as students move from a basic transactional and survival-based syllabus to more extended and abstract texts and topics. There are also indications that learners have insufficient

awareness of how to make progress (Graham, 2004, 2006). In one of the few large-scale studies conducted into students' perceptions of language-learning post 16 in England, Graham (1997) found that many students reported difficulties learning the large amounts of new lexis they encountered from the start of their A-level course. In addition, she found that students used a narrow range of strategies for learning vocabulary, which mainly consisted of fairly shallow strategies such as list-making and 'look-cover-write-check'.

Graham's (1997) study does not, however, attempt to chart the development of learners' vocabulary knowledge during advanced level study. Indeed, in comparison with what we know about how learners of English develop vocabulary, little research has been conducted into learners of French in schools in England in general and at A-level in particular. Milton (2006a) represents probably the only study to address this question. Using data drawn from X_Lex (Meara and Milton, 2003), a vocabulary recognition test, he estimates that pupils learn on average 170 words per year in the first five years of secondary schooling (Years 7 to 11). More progress seems to occur in Year 12, however, where the mean number of words known rises from 852 in Year 11 to 1,555. Milton (2006a) also reports a strong relationship between vocabulary levels and the A-Level grade learners achieve in French, which underlines the importance of learners' vocabulary levels for overall progress.

Milton's (2006a) study concludes that the rate of vocabulary growth among learners of French in England is slow when compared with learners of other languages in different countries, even when the smaller number of lessons typically received by pupils in English schools is taken into account. Reasons for this are unclear, although Milton suggests that an analysis of the textbook contents used by English learners might provide insights.

Milton's research looks at vocabulary development under 'normal' classroom conditions, which presumably involve a combination of explicit and incidental vocabulary learning. Schmitt (2000: 116) defines the former as 'the focused study of words', and the latter as learning 'through exposure when one's attention is focused on the use of language, rather than the learning itself' (e.g. through reading or listening). That both explicit and implicit modes play some role in vocabulary acquisition is generally acknowledged (e.g. Min, 2008). Nevertheless, the relative importance of these two modes of learning has been the subject of controversy (Laufer, 2006), with a recent gulf emerging between research evidence and semi-official advice (see Milton's, 2006b, review of the Harris and Snow, 2004, book on vocabulary building published by The National Centre for Languages, London).

A number of studies have been conducted into both processes. In terms of explicit learning, one area of research has been how learning can be enhanced through word-focused activities such as sentence construction, sentence completion and using target vocabulary in essays (Laufer, 2006). In addition, studies of instruction in vocabulary learning strategies have provided generally positive results (e.g. Burgos-Kohler, 1991, for learners of Spanish in a US university, and Lawson and Hogben, 1998, for secondary school learners of Italian in Australia. See also Nyikos and Fan, 2007, for a review of vocabulary intervention studies).

Regarding incidental learning, there is a sizeable body of research (reviewed in Pigada and Schmitt, 2006, and in Swanborn and de Glopper, 1999) which suggests that vocabulary growth can occur through extensive reading, although gains tend to be small. As both Pigada and Schmitt (2006) and Swanborn and de Glopper (1999) found, there is a lack of clarity surrounding such research, in terms of differences in the measurement instruments used and exactly what factors have led to the vocabulary gain. Is it frequency of occurrence of words in the text, for example? If so, to what extent is this dependent on textual support mechanisms such as marginal glosses or dictionary use? (see, for example, Hulstijn, Hollander and Greidanus, 1996). Few studies, however, have looked at incidental L2 vocabulary acquisition through listening. Vidal (2003) located only three such studies in non-laboratory settings, ranging from studies of young beginners (Schouten-van Parreren, 1989) to university-level Japanese learners of English (Toya, 1992), all of whom made gains in vocabulary through listening activities. Vidal herself conducted a study of vocabulary gains among ESP students in a Spanish university, to see whether vocabulary was gained through listening to academic lectures. Before watching and listening to video-taped lectures, learners were tested on their knowledge of 36 technical, academic and low-frequency words contained in the lectures. This test was then repeated after listening, and again as a delayed post-test. Results showed significant gains in knowledge of the selected items at post-test, but this declined at delayed post-test.

Clearly, incidental learning of L2 vocabulary *can* occur through listening and reading. Nevertheless, it is also true that second language learners fail to learn words they encounter (Hulstijn et al., 1996) and that incidental learning will be influenced by factors that impact on the learner's ability to infer the meaning of unknown words such as the ratio of known to unknown words (Laufer, 1989, 1992). Of particular relevance here seems to be the 'quality of information processing' (Laufer and Hulstijn, 2001: 12) involved in the task through which incidental vocabulary learning may occur. Interventions that seek to improve learners' performance in listening or reading skills, and which typically involve learners in greater depth of processing as learners consciously apply strategies to tasks, would therefore be likely to have an additional benefit for their vocabulary acquisition. We located only one published study (Fraser, 1999), however, that looked at this by-product of strategy instruction, and that was in reading. Fraser investigated the impact of instruction in 'lexical processing strategy' (LPS) use on vocabulary learning among eight university-level Francophone learners of English. Participants received training over a period of two months in three forms of LPS: ignore the word and continue reading; consult (a dictionary or another person); and infer (via linguistic or contextual clues). Data were collected on eight occasions, with four measurement points to determine the extent of LPS use: a baseline period; after the metacognitive strategy training; after the focused language instruction; and one month after the treatment. Learners' use of LPS was ascertained through a think-aloud procedure, whereby they identified unknown words in a text they were asked to read and then explained how they had dealt with these unknown words. A week after this

procedure, each learner was presented with ten of the words identified as unknown and asked to indicate whether they knew them or not. The mean 'retention rate' (Fraser, 1999: 238) was 28%, although there was much individual variation and a standard deviation of 12%. Fraser argues that this retention rate is higher than reported in previous studies of incidental vocabulary learning through reading. She also reports that learners' rate of ignoring unknown words decreased and their rate of success when making inferences increased. Thus the instruction seemed to have improved the underlying conditions for vocabulary learning. As there was no comparison group, however, and the number of participants small, the results need to be interpreted cautiously.

The limited amount of research in this area thus suggests that the potential impact on vocabulary acquisition of strategy interventions in other skill areas requires further exploration. One possible argument might be that the whole process of strategy awareness raising and a focus on alleviating students' problems in listening and writing would have beneficial spin-offs. In relation to listening tasks, these could result from increased incidental vocabulary learning through mechanisms such as including improved word segmentation, identification and comprehension. In addition, students might benefit from multiple spaced repetition in listening texts (Nyikos and Fan, 2007) and the recycling of previously encountered words known only superficially. The use of writing strategies might benefit vocabulary through, among other things, deeper processing, more efficient use of feedback, redrafting, and vocabulary-focused planning and monitoring. On the other hand, a counter-argument might be that strategy training would not benefit lexical learning because a tight focus on other skill areas such as listening and writing would be at the expense of curricular time devoted to vocabulary learning and promotion of vocabulary-specific strategies.

The study described here forms part of a project to evaluate the impact of strategy instruction on Year 12 learners' performance in these two skills. Results show a clear-cut positive effect of the intervention for listening (Graham and Macaro, forthcoming, 2008). The benefits for writing were less clear, however, with a much smaller effect (Macaro, Graham, Richards, Spelman-Miller and Vanderplank, 2006). Against this background, and in view of the paucity of studies looking at vocabulary development as a by-product of strategy instruction in other skill areas, we identified the following questions specific to vocabulary:

1. In view of the increased vocabulary demands post-GCSE and Milton's (2006a) findings of a vocabulary 'spurt' in Year 12 (see above) together with the difficulty of finding reliable indices of vocabulary growth over short periods (Tonkyn, 2006), can this progress be measured over two school terms in students' writing as well as their recognition vocabulary?
2. What student variables at pre-test predict post-test scores and progress in vocabulary over this period?
3. If vocabulary progress is measurable (Q1), how much progress is made compared with listening and writing?

4. What is the effect of strategy intervention that targets listening and writing? There are three possibilities:
    a. there are incidental benefits to vocabulary;
    b. the extra attention paid to listening and writing is at the expense of other areas of development such as vocabulary learning;
    c. there is no effect on vocabulary.

2. METHOD

2.1 Participants

The Year 12 students were aged 16 to 17 and were in the first year of post-compulsory education, having elected to continue with French following their GCSE. These participants were preparing for the Advanced Subsidiary (AS) examination at the end of Year 12, with the option of continuing their studies into Year 13 and sitting the Advanced Level (A2) examination. Typically they had already studied French for five years, receiving between 400 and 600 hours of instruction. At the outset, a total of 150 pupils (120 females, 30 males) took part. By the time of the post-test, however, the number of participants had reduced, and inevitably there was some absenteeism on days when data were collected, leaving 107 pupils who completed both pre- and post-tests for listening, writing and vocabulary. Such a high attrition rate is not unusual at this level and reflects a number of factors including pupils moving to different schools or colleges, changes of subject choice and, above all, the high drop-out rates for languages.

Background data on grades indicate that the 150 participants were high achievers in their GCSE French examinations at the end of the compulsory phase of their education: 42.2% obtained an A* grade (the highest), 39.1% an A, 12.5% B, 5.5% C and 0.8% (one student) was awarded a D. This reflects the typical profile of students in AS level classes.

2.2 The schools

Students attended 15 comprehensive schools. The comparison group consisted of four schools, with the remaining eleven schools allocated to one of two treatment groups (see below). We sought to obtain a stratified sample with matched pairs, i.e. allocation of schools was conducted in such a way as to obtain three groups that were as well-matched as possible for type, location and make-up of school (e.g. general level of pupil achievement). Randomised allocation was neither ethically nor logistically possible within the framework of a state school setting. The comparison schools were not located in the same counties as the treatment schools.

2.3 Strategy instruction

Schools receiving strategy instruction were placed in one of two groups: high scaffolding (HSG, five schools) or low scaffolding (LSG, six schools). Both groups

received two initial one-hour sessions from researchers, in which modelling of selected strategies for listening and writing took place. Over the following six months, class teachers then led further modelling and practice activities (five main activities were provided for teachers in each skill) in normal class time. Additionally, students were encouraged to use the strategies that had been introduced whenever they were engaged in writing and listening tasks through the use of strategy prompt sheets and record sheets. Detailed instruction notes were provided for teachers as well as briefing meetings to guide them through the implementation of the strategy instruction. For the HSG only, scaffolding of strategy use was provided in the form of additional awareness-raising and reflection about strategy use in the initial researcher-led sessions, a diary in which to record reflections on strategy use, and feedback from researchers both on these diary entries and on the strategy record sheets that learners submitted along with the language tasks that they accompanied (see Graham and Macaro, forthcoming, 2008). Students in the comparison group received no strategy instruction but simply followed their normal French classes.

Selection of strategies was based on the problems in strategy use exhibited by a different, but comparable, sample of students. For listening, these included: poor use of prediction and inferencing; lack of monitoring; and difficulties with identifying familiar words and word boundaries, the latter being a particular problem for English-speaking learners of French. Materials for developing effective use of prediction, inferencing and monitoring, and clusters of strategies were therefore created, along with those aimed at improving students' perception skills. For writing, students had exhibited problems at the 'formulation' stage of composing, i.e, the point at which they wanted to turn their ideas into French. Instruction materials therefore focused on formulation strategies (e.g., re-combining or restructuring known phrases) but also included planning (e.g. 'mind-mapping' or 'brainstorming' of known French that fitted the task requirements), monitoring and using feedback. For details and examples of materials in both skills, see Graham and Macaro (2007).

2.4 Assessing students

At the beginning of their AS course (pre-test) and after two school terms (post-test), students' listening was assessed through a written recall protocol after they had listened to four short passages on the theme of holidays. A different set of passages was used each time, with the difficulty level held constant in terms of length, percentage of unknown words, speech rate and judgements of a group of students who had listened to passages during piloting. Students were instructed to write down in English everything they had understood in each passage. Responses were written during the two hearings of each passage.

Recall protocols were scored by two raters, who used a banded rating score (four bands) to assess how many idea units had been recalled (in the form of words or phrases) across all four passages. There was a high level of agreement for the total

scores of the two raters: 0.95 at Time 1, at 0.96 at Time 2 (Pearson correlations). Differences in scores were resolved by discussion.

Students' performance in writing was assessed through a narrative writing task. They were given a six-picture narrative and asked to write a past-tense account of approximately 200 words in 30 minutes. A different but comparable set of pictures was used at each time point. Consulting a bilingual dictionary was allowed because bilingual dictionary use was included in the strategy training. Scoring was conducted using a six-dimensional analytical marking scheme adapted from Jacobs, Zingraf, Wormuth, Hartfiel, and Hughey (1981) and Weir (1993): Content (max. = 20), Organisation (20), Local coherence (20), Vocabulary (15), Grammar (15), Mechanics (10). Organisation includes the control of genre conventions, Local Coherence covers cohesion and the development and integration of ideas, and Mechanics focuses on spelling and punctuation. Each dimension was divided into bands of marks with descriptors for each band. For vocabulary, these were: 'inadequate vocabulary, basically translation (0–2)'; 'frequent lexical inappropriacies, circumlocution, and/or repetition (3–7)'; 'some lexical inappropriacies and/or circumlocution (8–12)'; 'almost no inadequacies in vocabulary for the task, effective range of vocabulary and appropriate register (13–15)'. Scoring was carried out independently by two expert raters and interrater reliability across the six categories ranged from 0.69 to 0.77 (Pearson correlations). Discrepancies were resolved by negotiation. While the reliability of writing assessment is often improved by including more than one task, the pressures of working in busy school contexts meant that this was out of the question. Fortunately, this constraint had no adverse effect on the total writing score whose reliability (Cronbach's apha) was 0.957 at pre-test and 0.955 at post-test.

In addition to vocabulary production in narratives, students' receptive vocabulary was assessed using X_Lex (Version 2.02), the Swansea Vocabulary Levels Test (Meara and Milton, 2003). This is a computerised 'yes/no test' that asks respondents to indicate whether they know the meaning of a series of 120 words that appear on screen. One hundred of these words are real words taken from five frequency bands (1K to 5K) based on Baudot's (1992) frequency count. In order to control for false positives resulting from guessing, overconfidence or cheating, 20 of the items are non-existent but plausible words that follow the phonotactic rules of French, for example 'clabrer' or 'muce'. An adjusted score is calculated that takes account of guessing (see Milton, 2006a). A pilot study (Richards and Malvern, 2007) had shown that X_Lex was appropriate for Year 12 pupils and that their results across the five bands were sensitive to word frequency. In the intervention study reported here, students did two parallel forms of the test both at pre- and post-test. The intention was to use the second set of results because we feared that students would overstate their knowledge at their first attempt before realising how heavily this would be penalised in the adjusted score.

We have referred above to 'vocabulary production' and 'receptive vocabulary'. As Read (2000:154–157) has pointed out, however, there is much confusion about the distinctions between receptive and productive vocabulary. It is important,

Table 1. *Descriptive statistics for the students' four attempts at X_Lex (maximum possible score = 5,000) and internal consistency of scores across frequency bands (Cronbach's alpha)*

	N	Mean	SD	Min.	Max.	Alpha
Pre-test 1st attempt	138	2281.9	674.3	550	4300	.903
Pre-test 2nd attempt	130	2297.3	839.7	100	4550	.923
Post-test 1st attempt	115	2700.0	697.5	1050	4600	.907
Post-test 2nd attempt	112	2723.7	719.7	100	4350	.896

therefore, to note that our two measures actually address subsets of receptive and productive knowledge. Using Read's terminology, these entail *recognition* of pre-selected, decontextualised L2 words and ratings of contextualised *use* of L2 vocabulary in writing constrained only by the content of the picture stimuli.

3. RESULTS

3.1 *Progress on vocabulary measures (the three groups combined)*

This section addresses the first research question by examining whether, for the whole sample of students, their progress on both vocabulary indices is measurable. In doing so, we will need to consider the relative validity of the students' two attempts at X_Lex on each occasion.

It will be recalled that students had two attempts at X_Lex in order to benefit from a predicted practice effect. In practice, however, some students lost motivation on their second attempt, with several achieving surprisingly low scores and others who failed to finish (see Ns and minimum scores in Table 1). Table 1 shows the results for the four attempts including Cronbach's alpha coefficients assessed from the raw totals from each of the five frequency bands as an indicator of internal consistency. These all indicate high reliability. Tests of normality indicate that the first three sets of results were normally distributed with zero skew but the second post-test attempt was strongly negatively skewed (skew = −0.558, s.e. skew = 0.228, z = −2.45, p < 0.05) and there were two outliers with low scores. Parallel forms reliability was therefore estimated using Spearman rank order correlations. These were also satisfactory, though less impressive: 0.636 (N = 130) for the two pre-test attempts and 0.687 (N = 112) for the two post-test attempts.

A comparison of the average scores for the two attempts in Table 1 using the Wilcoxon signed ranks tests shows that, contrary to expectations, there was no advantage for the second attempt at either time point (pre-test: z(129) = 0.469; post-test: z(111) = 0.169; ps > 0.05). However, on both attempts, the students performed substantially better at the second time point (first attempt: z(107) = 5.34, p < 0.001; second attempt: z(99) = 4.65, p < 0.001), thus demonstrating measurable progress on receptive vocabulary over the period of the study.

Because of concerns over loss of motivation and students failing to finish their second attempt, the analyses that follow include only the data from their first

Table 2. *Descriptive statistics for ratings of vocabulary in writing at pre-test and post-test (maximum possible score = 15)*

	N	Mean	SD	Min.	Max.
Pre-test	139	7.79	3.03	0	14
Post-test	121	9.58	2.53	3	15

Table 3. *Intercorrelations between vocabulary, listening and writing variables at both time points*

	1	2	3	4
Pre-test				
1. X_Lex	–	.455	.410	.441
2. Vocabulary in writing		–	.942	.603
3. Total writing score			–	.601
4. Total listening score				–
Post-test				
1. X_Lex	–	.342	.327	.340
2. Vocabulary in writing		–	.933	.561
3. Total writing score			–	.544
4. Total listening score				–

Note. All ps < .001.

attempt at pre-test and post-test. This has the additional advantage of allowing the use of parametric statistics.

Raters' scores for vocabulary in students' writing at pre-test and post-test are shown in Table 2. Both variables are normally distributed and a paired samples t-test shows gains over the course of the study (t(112) = 5.75, p < 0.001, Eta² = 0.228).

Substantial progress can thus be reliably demonstrated for both productive and receptive vocabulary. For the latter, however, analysis of the two successive attempts suggests that students tend to demonstrate a fatigue effect rather than a practice effect.

3.2 Inter-relationships between variables at each time point

In this section we explore the relationships among the vocabulary, listening and writing variables and participants' success at GCSE as a precursor to identifying factors that predict success in vocabulary learning (Research Question 2). For the continuous variables we examine Pearson correlations at both time points. All are highly significant (Table 3), with inter-relationships being slightly weaker at the second point. Productive and receptive vocabulary are related, but of particular interest are the strong correlations between writing vocabulary and the total writing score (0.942 and 0.933), as well as the moderately strong correlations between writing vocabulary and listening (0.603 and 0.561) and between listening and writing scores (0.601 and 0.544). It must be remembered, of course, that the

writing scores were not independent of vocabulary production in the sense that 15% of the total marks allocated were assigned to rating of vocabulary. A new variable was therefore created at both time points which excluded the vocabulary ratings from the total writing scores. While this does not ensure total independence because of possible halo effects between different dimensions that contribute to the total, it does provide mathematical independence. Correlations between the vocabulary rating and these new, more independent, writing variables were 0.919 at pre-test and 0.905 at post-test, consistent with a highly important contribution of vocabulary to success in writing at both time points.

When using examination results in educational research it is common to convert the grades into points and to treat the resulting scale as ordinal or even interval level of measurement (see, for example, Croll, 1995). The skewed distribution of GCSE grades in our data, however, makes this inappropriate and the restricted range makes even an analysis based on rank orders questionable. Participants were therefore divided into a high grade and low grade group. In practice this meant that the high group consisted entirely of those whose grade was A* while all other grades were allocated to the low group. Univariate ANOVAs were then carried out to determine relationships with the continuous variables that entered into the correlational analysis above. At pre-test all differences were substantial and highly reliable (ps < 0.001): for X_Lex means were 2083.8 for the low group and 2576.5 for the high group ($F(1,120) = 17.64$, $Eta^2 = 0.128$); for writing vocabulary the means were 6.32 and 10.06 respectively ($F(1,121) = 76.20$, $Eta^2 = 0.386$); for the total writing score the means were 40.55 versus 65.24 ($F(1,121) = 82.59$, $Eta^2 = 0.406$); and for listening they were 14.54 versus 22.98 ($F(1,121) = 48.69$, $Eta^2 = 0.287$). As can be seen from the Eta^2 values the largest effect sizes are for the two writing variables and the smallest is for the receptive vocabulary test.

3.3 Predictors of success in vocabulary learning

Having examined inter-relationships at each time point, the next step was to discover which variables predicted success in receptive and productive vocabulary after two terms and the amount of progress made (Research Question 2). Predicting success and predicting progress are, of course, entirely different procedures. Correlations between pre-test and post-test variables will predict *success*, but may only tell us that students who did well at Time 1 also did well at Time 2. Of more interest is *progress* and this needs to be assessed by the *gain* made from Time 1 to Time 2. The measurement of gains is problematic, however. Simple or 'crude' gain scores (Time 2−Time 1) or percentage gain scores (($T2−T1$)/$T1 \times 100$) are not independent of Time 1 scores, and tend to be negatively correlated with them (Barnes, Gutfreund, Satterly and Wells, 1983). A solution is to use residual gain scores (O'Connor, 1972), that is to say the difference between actual and predicted scores obtained from the regression of the post-test scores on the pre-test scores. By definition, these are independent of Time 1 scores. Unfortunately, residual gains are affected by similar problems to other gain scores, as their reliability is not only

Table 4. *Pearson correlations between pre-test variables and success and progress in receptive and productive vocabulary*

Pre-test measures	Post-test measures			
	X_Lex score	X_Lex residual gains	Writing vocabulary score	Writing vocabulary residual gains
X_Lex	.489***	–[a]	.424***	.225*
Vocabulary in writing	.371***	.217*	.640***	–[a]
Total writing score	.317***	.180	.641***	.062
Total listening score	.426***	.304**	.574***	.339***

Note. [a]Correlations between residual gain scores and the independent variable from which they were calculated are always zero.
*p < .05. **p < .01. ***p < .001.

a function of the reliability of both the Time 1 and Time 2 scores but also of the correlation between them. The *stronger* the correlation between Time 1 and Time 2 the *lower* the reliability (see Ross, 1998). Ideally, therefore, our vocabulary scores need to be significantly correlated between the two occasions in order to justify the calculation of residuals, but not strongly enough to impair reliability. As can be seen from Table 4, this optimum level of correlation was the case for both X_Lex and writing vocabulary, and so residual gains were computed for these two variables.

Pearson correlations were computed between pre-test and post-test continuous variables (Table 4). For GCSE grade, ANOVAs were conducted with GCSE group as the independent variable. As can be seen, all the pre-test variables are highly significant predictors of post–test receptive and productive vocabulary (all ps < 0.001), although predictors for writing vocabulary tend to be rather stronger than those for X_Lex. Correlations with residual gains indicate how strongly the pre-test variables predict later success *over and above* what would be predicted by the students' original pre-test status on that measure. These associations can therefore be expected to be weaker than those discussed above, and this is indeed the case. Nevertheless, both receptive and productive vocabulary at the first time point is significantly associated with later vocabulary gains, as is pre-test listening. However, there is no relationship between pre-test total writing scores and gains for vocabulary. This can be accounted for by the high correlation at pre-test between writing and writing vocabulary: in other words, writing does not account for a significant amount of additional variance beyond that already explained by pre-test writing vocabulary.

Results of univariate ANOVAs with GCSE group (upper versus lower) as the independent variable follow a similar pattern: GCSE predicts post-test X_Lex scores for receptive vocabulary $(F_{(1,100)} = 6.23,\ p = 0.014,\ Eta^2 = 0.059)$ and, much more strongly, productive vocabulary in writing $(F_{(1,106)} = 47.47, p < 0.001, Eta^2 = 0.309)$. GCSE also weakly predicts gain scores for writing vocabulary $(F_{(1,102)} = 5.66, p = 0.019, Eta^2 = 0.053)$ but not for X_Lex $(p > 0.05)$.

3.4 Comparing progress within and between groups

To compare how much progress the students made on vocabulary, listening and writing (Research Question 3), a series of repeated measures ANOVAs was conducted comparing pre-test and post-test scores. In order to remove any confounding effects of the intervention, these were performed on the comparison group only. Significant progress could be detected on all variables except receptive vocabulary: X_Lex ($F(1,41) = 3.51$, $p = 0.068$), writing vocabulary ($F(1,45) = 9.24$, $p < 0.001$, $Eta^2 = 0.170$), listening ($F(1,37) = 42.85$, $p < 0.001$, $Eta^2 = 0.537$), writing ($F(1,44) = 5.66$, $p = 0.022$, $Eta^2 = 0.114$). Although we are not able to show that effect sizes differ significantly from each other, it is clear that by far the largest increase is for listening. Progress on writing and written vocabulary is similar with a slight advantage for writing. No effect size is reported for receptive vocabulary as there was no significant difference between pre-test and post-test for this group.

In order to test whether the intervention programme had focused on listening and writing to the detriment of vocabulary or, conversely, whether it had provided positive spin-offs for vocabulary (Research Question 4), an ANCOVA was conducted with the group as the independent variable, pre-test receptive vocabulary as a covariate, and post-test receptive vocabulary as the dependent variable. This was not statistically significant ($F(2,104) = 1.31$, $p = 0.275$). A corresponding analysis for writing was also carried out, again with group as the independent variable, but entering pre-test writing vocabulary as the covariate and post-test writing vocabulary as the dependent variable. There was no significant effect of group ($F(2,110), = 0.556$, $p = 0.569$), indicating that the intervention had neither positive nor negative effects on students' vocabulary.

4. DISCUSSION

The research reported above investigated four research questions about the measurement, characteristics and extent of development of vocabulary proficiency over a period of two terms in English comprehensive schools. The first question considered whether students' progress could be measured reliably over such a short period. Even though Year 12 is a period of increased vocabulary growth compared with Years 7 to 11 (Milton, 2006a), it was not clear that measures of general vocabulary proficiency, particularly the X_Lex test of receptive vocabulary, would be sensitive to developments across 15 schools using different textbooks and different examination syllabuses whose varied range of topics and linguistic content lead to different expectations concerning vocabulary. Nevertheless, for our measures of both receptive and productive vocabulary, reliable scores were obtained from which significant progress could be measured even over a relatively short period.

One potential measurement problem had been the issue with yes/no tests of how to prevent respondents from wrongly claiming to know the meaning of a word. X_Lex attempts to control for this by interspersing the stimuli with highly plausible nonsense words and heavily penalising those who claim to know them. For each

such 'error', the raw score out of 5000 is adjusted downwards by 250 points. Thus a student who really did know all the genuine words but also claimed to know all 20 nonsense words would receive a raw score of 5000 and an adjusted score of zero. We were concerned that students prone to guessing or risk-taking would fail to realise the impact of lack of caution and obtain scores that greatly underestimated their knowledge. We therefore allowed two attempts on each occasion in order that students could benefit from practice. In the event, this provided no advantage and second attempt scores were no better than for the first attempt. If anything, some students became bored and demotivated. Nevertheless, concerns remain about the effect of test-taking style or test-taker personality on the validity of such tests, and this is an area that needs further investigation.

The second research question addressed predictors of success and progress in vocabulary. Firstly, however, concurrent intercorrelations at pre-test and post-test between measures of vocabulary, listening and writing were examined. These were all substantial and highly reliable, as are their relationships with GCSE grades at pre-test. What was particularly striking, however, were the very high correlations (0.94 and 0.93) between productive vocabulary and writing at both time points, providing evidence of the central role of vocabulary in the writing process at this stage.

This research question made an important distinction between success *after* two terms and progress, that is, gains, *over* two terms and asked which student pre-test variables predicted these for vocabulary. Both pre-test vocabulary measures and pre-test writing and listening scores were highly significant predictors of post-test vocabulary, as were GCSE grades (particularly of productive vocabulary). Our measure of progress used residual gain scores to obtain an index for each vocabulary measure that reflected progress over and above what would be predicted from the pre-test score on that measure. As would be expected, these were predicted less strongly, but were still significantly related to pre-test productive vocabulary, X_Lex and, especially, listening. GCSE grade was also a significant predictor of productive vocabulary.

With regard to our third research question concerning the relative amount of progress made in vocabulary, listening and writing, this analysis was carried out on the comparison group only in order to control for the effects of the intervention programme. It was clear that students made progress in listening, writing and writing vocabulary. The result for X_Lex, however, was not significant. It seems to be the case that it is more difficult to make measurable progress on receptive vocabulary, particularly when compared with listening which had the largest effect size. It may well be that the skill element of listening undergoes rapid development in Year 12, even for students who do not receive strategy training, and this has a large impact on test scores. Writing, on the other hand, which offers more opportunity for reflection, planning and self-direction may progress more gradually. With regard to receptive vocabulary versus vocabulary in writing, it seems likely that students would be able to demonstrate progress more easily on the latter when they are able to choose the lexical items they will use rather than relying on any chance

correspondence between the vocabulary they have learnt and the items contained in X_Lex's dictionary.

Finally, we considered the three possible effects on vocabulary learning of an intervention that addressed listening and writing skills. The first possibility was that students' vocabulary in the intervention groups would benefit indirectly from the attention to listening and writing strategies, possibly through incidental learning from more effective listening, and practice and depth of processing in writing. The second possibility was that the intervention could harm lexical development by diverting attention from it, and the third possible result was that the intervention made no difference. Although we had expected vocabulary to benefit from the experimental programme, analyses found no difference between the groups. It may be, as Walters (2006: 238) argues, that the effects of strategy instruction on vocabulary development take time to emerge fully, perhaps because of the 'incremental nature of incidental vocabulary development'. It is also possible that strategy instruction that had focused more sharply on word identification or inferencing might have led to greater vocabulary gains in the intervention groups. These are questions that need to be explored more fully in future research. Nevertheless, the fact that the strategy instruction had a positive impact on students' listening skills, while not jeopardising their vocabulary development suggests that strategy instruction is a fruitful avenue to explore in terms of pedagogy and improving learners' attainment in French at lower-intermediate level.

Address for correspondence:
Brian Richards
University of Reading
Institute of Education
Bulmershe Court, Earley
Reading RG6 1HY
UK
e-mail: b.j.richards@reading.ac.uk

REFERENCES

Barnes, S., Gutfreund, M., Satterly, D. and Wells, C. G. (1983). Characteristics of adult speech which predict children's language development. *Journal of Child Language*, 10.1: 65–84.

Baudot, J. (1992). *Fréquence d'utilisation des mots en français écrit contemporain*. Montréal: Presses de l'Université de Montréal.

Burgos-Kohler, N. (1991). The effect of a selected group of language learning strategies upon language development. Unpublished PhD dissertation, University of Texas at Austin.

CILT. (2005). *Languages yearbook 2005*. London: The National Centre for Languages.

Croll, P. (1995). Early linguistic attainment, family background and performance in 16+ examinations. *Educational Studies*, 21: 13–28.

Fisher, L. (2001). Modern foreign languages recruitment post-16: The pupils' perspective. *Language Learning Journal*, 23: 33–40.

Fraser, C. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21.2: 225–241.

Graham, S. (1997). *Effective Language Learning. Positive Strategies for Advanced Level Language Learning.* Clevedon: Multilingual Matters.

Graham, S. (2002). Experiences of learning French: a snapshot at Years 11, 12 and 13. *Language Learning Journal*, 25: 15–20.

Graham, S. (2004). Giving up on modern foreign languages? Students' perceptions of learning French. *The Modern Language Journal*, 88.2: 171–191.

Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34.2: 165–182.

Graham, S. and Macaro, E. (2007). Designing Year 12 strategy training in listening and writing: from theory to practice. *Language Learning Journal*, 35.2: 153–173.

Graham, S. and Macaro, E. (forthcoming, 2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*.

Harris, V. and Snow, S. (2004). *Doing it for Themselves: Focus on Learning Strategies and Vocabulary Building.* London: CILT, the National Centre for Languages.

Hulstijn, J., Hollander, M. and Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80: 327–339.

Jacobs, H., Zingraf, S., Wormuth, D., Hartfiel, V. and Hughey, J. (1981). *Testing ESL Composition: a Practical Approach.* Rowley, MA: Newbury House.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In: C. Lauren and M. Nordman (eds.), *Special Language: From Humans to Thinking Machines.* Clevedon: Multilingual Matters, pp. 316–323.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In: P. J. L. Arnaud and H. Bejoint (eds.), *Vocabulary and Applied Linguistics.* London: Macmillan, pp. 126–132.

Laufer, B. (2006). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59: 567–587.

Laufer, B. and Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22.1: 1–26.

Lawson, M. and Hogben, D. (1998). Learning and recall of foreign-language vocabulary: effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, 8.2: 179–194.

Macaro, E., Graham, S., Richards, B. J., Spelman-Miller, K. and Vanderplank, R. (2006). Strategy training in year 12 French: listening and writing. Final report to the Economic and Social Research Council (Project RES-000-23-324).

Meara, P. M., Milton, J. L. (2003). X_Lex: Swansea Vocabulary Levels Test (Version 2.02) [Computer software]. Swansea: Lognostics.

Milton, J. (2006a). Language lite? Learning French vocabulary in school. *Journal of French Language Studies*, 16: 187–205.

Milton, J. (2006b). Review of K. S. Folse, *Vocabulary Myths: Applying Second Language Research to Classroom Teaching.* Ann Arbour: The University of Michigan Press, 2004; and V. Harris and D. Snow. *Doing it for Themselves: Focus on Learning Strategies and Vocabulary Building.* London: CILT, 2004. *Language Learning Journal*, 33: 75–77.

Min, H.-T. (2008). EFL vocabulary acquisition and retention: reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58.1: 73–115.

Nyikos, N. and Fan, M. (2007). A review of vocabulary learning strategies: focus on language proficiency and learner voice. In: A. D. Cohen and E. Macaro (eds.), *Language Learner Strategies: Thirty Years of Research and Practice*. Oxford: Oxford University Press, pp.251–273.

O'Connor, E. F. (1972). Extending classical test theory to the measurement of change. *Review of Educational Research*, 42, 73–97.

Pigada, M. and Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18.1: 1–28. Retrieved December 7, from http://nflrc.hawaii.edu/rfl.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Richards, B. J. and Malvern, D. D. (2007). Validity and threats to the validity of vocabulary assessment. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press, pp. 79–92.

Ross, S. J. (1998). *Measuring Gain in Language Programs: Theory and Research*. Sidney: National Centre for English Language Teaching and Research, Macquarie University.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.

Schouten-van Parreren, C. (1989). Vocabulary learning through reading: which conditions should be met when presenting words in texts? *AILA Review*, 6: 75–85.

Swanborn, M. S. L. and de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Review of Educational Research*, 69.3: 261–285.

Tonkyn, A. (2006). *Towards progress-sensitive measures of L2 speaking proficiency: a case study of instructed learners*. Unpublished PhD Dissertation, University of Reading.

Toya, M. (1992). *Form of explanation in modification of listening input in L2 vocabulary learning*. Unpublished MA in ESL thesis, Department of ESL, University of Hawai'i, Honolulu.

Vidal, K. (2003). Academic listening: a source of vocabulary acquisition? *Applied Linguistics*, 24.1: 56–89.

Walters, J. M. (2006). *Beyond exhortation: training students of English as a second language to infer meaning from context*. Unpublished PhD thesis, University of Nottingham, UK.

Weir, C. J. (1993). *Understanding and Developing Language Tests*. Hemel Hempstead: Prentice Hall.

Factors driving lexical variation in L2 French: A variationist study of automobile, auto, voiture, char and machine[1]

TERRY NADASDI

University of Alberta

RAYMOND MOUGEON

York University, Toronto

KATHERINE REHNER

University of Toronto at Mississauga

ABSTRACT

Our paper examines lexical variation in the spoken French of second language learners and focuses on words referring to the notion of 'automobile' (i.e., *automobile*, *auto*, *voiture*, *char* and *machine*). Results reveal that while students do follow the native speaker pattern of using the neutral variant *auto* in most instances, they diverge from native speakers by making no use of the vernacular form *char* and relatively high use of the prestige variant *voiture*. The principal external factors that influence variant choice are students' home language and the representation of variants in the input to which students are exposed.

1. INTRODUCTION

Our study presents a sociolinguistic analysis of lexical variation in the spoken French of 41 French immersion students in Ontario, Canada. Specifically, it focuses on words that in Canadian French refer to the notion 'car' (e.g.: *auto*, *char*, etc.). While a large body of research now exists on grammatical variation in the spontaneous speech of French immersion students (see Mougeon, Rehner and Nadasdi, 2004) relatively little is known about the distribution of lexical variants in such speech. To date, only one such study (Nadasdi and McKinnie, 2003) has considered this issue. Studies of lexical variation in spoken Canadian French are also rare. This reflects in part the fact that some sociolinguists have expressed reservations about extending the concept of the sociolinguistic variable beyond phonology (e.g.,

---

[1] The research reported upon in the present article was made possible in part through a research grant awarded to Raymond Mougeon and Terry Nadasdi by the Social Sciences and Humanities Research Council of Canada from 1996 to 2002.

Lavandera, 1978). However, other sociolinguists have argued that non phonological variables, notably lexical ones, can be the object of sociolinguistic investigation with due consideration given to the methodological principle of the semantic equivalence of variants, see Armstrong, (2001); Lodge (2004); Sankoff (1988) and especially Chambers (1999, 2000, 2002, 2004). Chambers has notably examined lexical variation in Canadian English from the dual perspective of sociolinguistics and geolinguistics. Another reason for the scarcity of variationist studies of lexical variation in spoken Canadian French is the fact that the corpora of semi-directed taped interviews used by variationist sociolinguists yield only a limited number of lexical variables that are frequent enough to lend themselves to statistical analyses of variance.

As a backdrop to the present study, we first present an overview of previous studies of sociolinguistic variation in immersion French and make several hypotheses concerning the use of the variants that make up the current variable. In addition to this review, we will review previous studies having examined the use of variants meaning 'car' in the speech of Canadian francophones in order to arrive at an L1 benchmark for the present study.

The main questions addressed in this paper are: a) do the French immersion students use all of the variants that mean 'car' in Canadian French; b) does their frequency of use of such variants differ from that of native speakers of Canadian French (e.g., do they under-use the marked informal variant *char*, do they overuse the hyper-formal variant *voiture*); c) does their use of the variants under study reflect the treatment of these variants in the educational input of the students' (i.e., French immersion teachers' speech and French language arts teaching materials) – specifically, we examine the frequency of the variants, the information regarding their socio-stylistic status, and any special activities targeting their use; and d) what are the linguistic and extra-linguistic factors that correlate with the students' variant choice (e.g., lexical priming by the interviewer, the students' extra-curricular exposure to Canadian French, the students' home language).

2 . METHODOLOGY

The 41 speakers in the present study were Grade 9 and 12 students (ages 14–17) taken from Mougeon and Nadasdi's 1996 corpus of immersion French. The data they have provided come from semi-directed interviews similar to those used in previous sociolinguistic research on variation in L1 Canadian French. All speakers were enrolled in programmes where 50% of courses were taken in French from Grades 5–8, followed by 20% in high school. Though these students speak a variety of languages in the home, none of them learned French as a first language and all come from homes where neither parent is a native speaker of French and where French was not spoken in the household. As such, these students' use of and exposure to French has been largely limited to the school setting. Although these students were not from French-speaking homes, they were by no means all from unilingual Anglophone homes. In fact, 51% of our subjects came from homes where

a language other than English was used to varying degrees. Of these students, 39% came from homes where Italian or Spanish were spoken and the rest were from non-Romance language homes. There was approximately the same number of Grade 9 and 12 students, more females than males, and over half were from middle-class families with all but one of the remaining being from lower-middle-class families. The majority of the students have received between 26–37% of their schooling through the medium of French. Most students never use the spoken French media; however, there were more Grade 12 students than Grade 9s who did so occasionally. The Grade 12 students had also spent more time in Francophone environments and with Francophone families than was the case for the Grade 9 students. These stays in a Francophone environment or with a Francophone family took place, for the most part, in Quebec. The average duration of these stays is 17 days.

In addition to the student corpus, we use two corpora in our research that provide data on the type of educational input to which students have been exposed. This data takes two forms. First, we have examined Allen, Cummins, Harley and Swain's 1987 corpus of French immersion teachers' speech to gain insight into the range of variants used by teachers when speaking in the classroom.[2] Second, we have examined in their entirety two series of materials, one called *Portes ouvertes sur notre pays* which included series 1A and B (Roy Nicolet and Jean-Côté, 1994) and 3A and B (Le Dorze and Morin, 1994), which was used in the school district where we gathered our student corpus, and the other called *Capsules* (Deslauriers and Gagnon, 1995, 1997), which was used in the Toronto District School Board. In our results section, we consider patterns of convergence and divergence between students' use of the 'car' variable and the availability of the variants in the input to which they are exposed.

3 . PREVIOUS RESEARCH

3.1 Sociolinguistic variation in the speech of French immersion students

Over ten years of research on our French immersion corpus has brought to light a variety of trends concerning sociolinguistic variation in the speech of French immersion students. This research has focused on sociolinguistic variables that include variants that can be placed at various points on a socio-stylistic continuum ranging from marked informal variants to hyperformal ones, discussed notably in Mougeon *et al.* (2004). Our socio-stylistic categorisation of variants is based on the findings of research on the effect of speaker social class and sex on variant choice and on variant frequency in corpora of spoken Quebec French or Ontario spoken French gathered via semi-directed taped interviews (see Mougeon, Nadasdi

---

[2] This corpus includes recordings made in the classroom of seven Grade 3 and 6 French immersion teachers from the Greater Toronto Area and Ottawa while teaching French or other subjects.

and Rehner, 2003).[3] It includes the following categories; i) marked informal variants (used by working class and/or male speakers markedly more often than by speakers from the other social classes and the other sex); ii) mildly-marked informal variants (gradient or weak correlations with social class and sex and high discursive frequency); iii) neutral variants (no correlation with speaker social class and sex); iv) formal variants (used by upper middle class and/or female speakers markedly more often than by speakers from the other social classes and the other sex) and v) hyperformal variants (used only by speakers at the top of the social ladder and low discursive frequency).

The above-mentioned research has revealed that French immersion students:

a) never, or rarely make use of marked informal variants (e.g., *(ça) fait (que)* with the meaning 'therefore', cf. Rehner and Mougeon, 2003).

b) use mildly-marked informal variants at levels of frequency well below FL1 norms (e.g. *ne* deletion, cf. Rehner and Mougeon, 1999)

c) tend to overuse formal and hyper-formal variants (e.g. subject pronouns *nous*, cf. Rehner *et al.*, 2003)

d) use neutral variants in a manner that reflects the systemic properties of these forms (e.g. the under-use of morphologically complex forms such as *falloir + que* + subjunctive, cf. Lealess, 2005)

e) tend to use variants at levels similar to those found in their educational input (e.g. conjunctions *alors* and *donc* to express the notion of consequence, cf. Rehner and Mougeon, 2003)

f) follow some, but not all linguistic constraints on variation found in L1 French (e.g. they respect only one of the linguistic constraints on choice of future tense variants attested in L1 French, cf. Nadasdi, Mougeon and Rehner, 2003)

g) follow some, but not all social constraints on variation in L1 French (e.g. female and middle-class students use hyperformal subject pronouns *nous* more often than male and lower-middle-class students, cf. Rehner, Mougeon and Nadasdi, 2003)

h) tend to use variants that have a home-language counterpart (e.g. students who speak Italian at home use the variant *alors* exclusively (modeled on *allora*) to express the notion of consequence, cf. Rehner and Mougeon (2003).

After having considered the distribution of the variants for the current variable in L1 Canadian French, we will make hypotheses concerning the use of these variants by French immersion students, drawing on the previous findings outlined above.

Concerning the precise topic of lexical variation in the speech of French immersion students, we note that two lexical variables have been previously studied: a) the alternation between verbs with that mean 'dwell', i.e., *habiter, vivre, rester,*

---

[3] For some variants we also use the findings on the effect of topic (in)formality. Such findings are scarce though, since studies of variation in spoken Canadian French have tended to neglect the investigation of the influence of topic on variant frequency.

demeurer and b) nouns used to refer to remunerated work, i.e., *travail, job, ouvrage, emploi, poste, situation* (cf. Nadasdi and McKinnie, 2003).

Results reveal that French immersion students express the notion of 'dwell' with either *vivre*, a socio-stylistically neutral variant in spoken Quebec French and *habiter* (a formal variant in spoken Quebec French). Interestingly they do not use *demeurer*, which is also a marked formal variant in spoken CF and they do not use the verb *rester* with the meaning of 'dwell' (a marked informal variant in spoken Quebec French). As for lexical items referring to remunerated work, it was determined that, here too, French immersion students make use of only a subset of the variants found in spoken Quebec French (namely *travail*, *emploi* and *job*). In keeping with findings regarding the 'dwell' variable, the immersion students make no use whatsoever of the marked informal variant *ouvrage* and also never use the formal variant *poste*. For both variables, there is evidence that immersion students' use of lexical variants strongly reflects the input to which they are exposed. For example, like the students, French immersion teachers and the authors of the French language arts textbooks do not make use of variants *rester* and *ouvrage*. Another relevant finding is that lexical priming exercises a strong influence on the use of variants used by French immersion students, i.e. the use of a given variant in the interviewer's question strongly promotes the use of the same form in the learner's response. One final noteworthy finding from previous research on lexical variation in French immersion students' discourse is that some of the students in our corpus prefer variants that have an equivalent in their home language. This was evidenced, for example, by the fact that the use of the form *travail* is greatest among individuals who speak Spanish and Italian in the home since both of these languages have a cognate form (cf. Nadasdi and McKinnie, 2003).

3.2 Lexical variants meaning 'car' in L1 Canadian French

Let us turn now to variationist studies that have examined nouns that mean 'car' in the speech of native speakers of Canadian French. Only two such studies exist, namely Martel's (1984) study of lexical variables (including 'car') in the variety of French spoken in Sherbrooke (in the Eastern Townships of Quebec) and Nadasdi, Mougeon and Rehner (2004) in the varieties of French spoken in Hawkesbury, Cornwall, North Bay and Pembroke, Ontario.

Given that the immersion students focused upon in our research have had interactions with L1 speakers of Canadian French mostly in Quebec, Martel's study is especially relevant for comparative purposes, unfortunately it is based on a corpus gathered in the early 1970s. As for Nadasdi *et al.*'s (2004) study, it is based on the Mougeon and Beniak corpus gathered in 1978. To obtain a more recent picture of the frequency of the variants under study we present below the findings of a special analysis of variant frequency in the corpus of Ontario French gathered in 2005 by Mougeon, Nadasdi and Rehner in the same localities where Mougeon and Beniak collected their corpus. The relevance of the Ontario French corpora lies in the fact that the francophone communities where they were gathered are located

Table 1. *Variants meaning 'car' in Canadian French*

Variants	Quebec spoken French (1971–73)		Ontario spoken French (1978)[4]		Ontario spoken French (2005)[5]	
	N	%	N	%	N	%
auto	154	42%	196	62%	406	66%
char	84	23%	85	27%	171	23%
machine	71	19%	0	0%	3	0.5%
automobile	51	14%	28	9%	8	1.5%
voiture	8	2%	7	2%	51	9%

near the Quebec/Ontario border and are the result of Quebecois immigration. It is also noteworthy that the speakers of Ontario French who provided these corpora are also of the same age as the immersion students under study.

The two studies mentioned above identified five variants. Examples of these variants, taken from the Mougeon and Beniak and Mougeon, Nadasdi and Rehner corpora, are presented below:

1) *j'étais en-dessous du **char***
2) *c'est à peu près une demi—heure de **voiture** d'ici*
3) *un' **auto**[6] pouvait te durer au moins 10 ans*
4) *peut-être j'm'achèterais un'* automobile
5) *ça va pour des 'spark plug' de ma **machine** des bougies pour ma machine*

The overall frequencies of variant usage are presented in Table 1.

Table 1 reveals that the most frequent word used to mean 'car' in Quebec and Ontario spoken French is *auto*. In the spoken French of Sherbrooke *char, machine* and *automobile* can be characterised as secondary variants, while *voiture* is clearly a marginal variant. The Franco-Ontarian corpora yield similar results: *auto* is the base variant, *char* is a secondary variant, used approximately one quarter of the time, *voiture*, which was quite rare in the 1978 corpus has undergone a modest rise in the 2005 corpus and both *automobile* and *machine* are marginal. The rarity of *machine* in the corpora of Ontario French (corpora gathered among adolescent speakers) is in keeping with Martel's findings. *Machine* is primarily used by the speakers who are 50 years old or older and is hardly used by the younger speakers (20–29 years old) in the Sherbrooke corpus.

3.3 Effect of linguistic constraints on variant choice

In Martel's (1984) study, no linguistic constraints were examined for this variable. However, Nadasdi *et al.* (2004) identified two linguistic factors influencing variant

[4] These data are from Nadasdi *et al.* (2004).
[5] These data are from Mougeon, Nadasdi and Rehner's 2005 corpus of Ontario spoken French.
[6] When *un* and *une* are used before vowel-initial words in informal speech their realizations tend to merge (primarily as a result of vowel denazalization in *un* and vowel laxing in *une*).

choice, namely interviewer priming and preceding element. Lexical priming by the interviewer is illustrated in 5), taken from the Mougeon and Beniak corpus:

Lexical Priming
5) Interviewer:
ah l'école garde l'auto?
Response:
oui elle garde l'auto

Nadasdi *et al.* (2004) found that when the interviewer asked a question containing one of the variants, this variant was used categorically in the response.

Concerning preceding element, three contexts were examined: a) preceded by a preposition; b) preceded by an adjective; and c) preceded by a determiner. These three contexts are presented below (taken from the Mougeon and Beniak corpus):

Preposition
6) *comment je me rends là? en auto*
7) *non en en char*

Adjective
8) *bien j'aimerais avoir une belle auto*
9) *il était le plus beau char au monde*

Determiner
10) *il y a un' auto stationné là-bas*
11) *il avait un char*

Note, that the effect of context was examined only in relation to *auto* and *char*. The other two variants, namely *automobile* and *voiture* were not frequent enough to lend themselves to this type of analysis. Consideration of preceding contexts led Nadasdi *et al.* (2004) to the conclusion that, unlike *char*, *auto* is favoured in contexts where the preceding element is a preposition. The explanation provided for this pattern is that expressions like *en auto* are partially lexicalised (especially more so than *en char*).

3.4 *Effect of social constraints on variant choice*

Martel (1984) made several observations concerning the effect of social factors on variant choice. He found that *char*, which he describes as a vernacular variant, is favoured by working class speakers and is used only marginally by those in the professional class. On the other hand, *auto* is used by speakers from all social groups, and is particularly frequent with speakers under 30 years of age. The form *automobile* is found, in particular, in the speech of middle-aged speakers. As for the variant *machine* it is also labelled as vernacular by Martel. Further, it is associated with older speakers and is marginal in the speech of younger speakers. Finally, Martel describes *voiture* as a prestige variant found almost exclusively in the speech of the upper-middle class.

Nadasdi *et al.* (2004) arrived at the following results concerning social constraints on the variable in Ontario spoken French. First, female students make greater use

of the variant *auto* than do males, while males show a stronger preference for *char* than do female speakers. Concerning social class, no discernible pattern obtains for *char*. However, *auto* is found more often in the speech of middle-class speakers than working class ones. These results suggest that *auto* may be more marked socially as a formal variant than *char* as an informal variant. Concerning variant *voiture*, it was too infrequent in their data to draw conclusions about its distribution according to sex or social class.

One particularly interesting result reported by Nadasdi *et al*. (2004) concerns the distribution of the variant *char* according to the extent to which speakers used French and English in daily interactions. In particular, they found that speakers who used French and English about as frequently (semi-restricted speakers in the Mougeon and Beniak corpus, see Mougeon and Beniak, 1991) showed a preference for *char*. The interpretation they offer for this correlation is that it can be attributed to the fact that the balanced bilinguals have relatively high contact with vernacular French, the source of *char*, and English which possesses the near cognate *car*. The hypothesis that French/English bilinguals may show a preference for variant *char* due to its semantic and morphological similarity receives indirect support from the findings of Quebecois lexicologists. They found that the French word *char* acquired the meaning of English word *car* in Quebec French around the beginning of the twentieth century, i.e. at the time when cars started to be mass-produced in North-America, see Poirier (1998). In other words, according to this author use of *char* to mean an 'automobile' is a case of semantic extension induced by contact with English.

In sum, while speakers of Quebec French use *auto* as the base variant whose socio-stylistic status is relatively neutral, in the speech of Franco-Ontarian adolescents this word is more aptly categorised as a formal variant. Both Quebec and Ontario speakers make very rare use of *automobile* and even rarer use of *voiture* although the latter variant may have undergone a relative rise recently (as indicated by the 2005 corpus of Ontario spoken French). The rarity of *voiture* in the Mougeon and Beniak 1978 corpus made it impossible to determine its socio-stylistic status. In the Sherbrooke corpus however, it was clearly a hyperformal variant. As for *char* this variant can be viewed as marked informal (although evidence for this in the Ontario corpus is less clear-cut). Finally, balanced bilinguals show a predilection for this latter variant, a result that may be attributed in part to the existence of *car* in their other language.

4 . HYPOTHESES

We have seen that L1 speakers of Canadian French make use of five variants to refer to the notion of 'car' in their spoken discourse. Furthermore, our review of previous studies of the variable has revealed that the social status of the variants spans the full stylistic continuum. At the vernacular end of this range, one finds variants *char* and *machine*, at the other end one finds the variants *automobile* and *voiture*, while the variant *auto* holds an intermediary position between the two ends. We have

Table 2. *Percentage of variants in Quebec spoken French and in the educational input of immersion students*

Variants	Spoken Quebec French		Immersion Teachers		Materials: Text		Materials: Dialogues, etc.	
	N	%	N	%	N	%	N	%
char	84	23	0	0	0	0	0	
machine	71	19	0	0	0	0	0	
auto	154	42	1	—	2	18	6	75
automobile	51	14	0	0	6	64	0	0
voiture	8	2	2	—	3	28	2	25

also noted that the variable correlates with two linguistic factors and a variety of social factors. These findings, coupled with our review of previous research on the French immersion corpus led us to a number of hypotheses regarding the use of variants by immersion students:

a) it is unlikely that immersion speakers will make use of the vernacular variant *machine* since previous research shows that marked informal variants are all but absent from immersion speakers speech and since, during their brief stays in Quebec, they are unlikely to have had many interactions with older speakers of Quebec French, the vanishing users of this variant.
b) it is possible that immersion students will make use of the vernacular variant *char*, given its formal similarities with the English equivalent *car* and the fact that it ranks second in frequency behind *auto* in Quebec spoken French.

Before finishing with our hypotheses, let us briefly examine the representation of these variants in the input to which French immersion students are exposed. Information regarding the availability of variants in this input is provided in Table 2.

The first observation to be made is that none of the different kinds of input make use of the marked informal variants *char* and *machine*. Furthermore, the frequency of variant *voiture* in the educational input of the students is higher than in spoken Quebec or Ontario French (including the most frequent 2005 corpus). It should be noted, however, that the teacher classroom speech corpus contained too few instances of the variable to arrive at firm conclusions. While the frequency of *voiture* in the teaching materials is higher than in spoken Canadian French the disparity is not as extreme as what we observed with other hyperformal variants (e.g., consequential conjunction *donc* see Nadasdi *et al*. 2005). In fact in the dialogic parts of the materials the frequency of *auto* is not too far removed from that found in corpora of spoken Canadian French. In the textual parts of the materials, however, *automobile* is clearly the preferred variant, a finding which suggests that this particular word is associated with formal written Canadian French.

Consideration of the input to which students are exposed led us to further hypotheses, regarding the neutral variant *auto* and the formal ones, namely *voiture*

Table 3. *General distribution of variants among immersion students*

Variants	Ns	Percentage	Dispersion[†]
machine	0/92	0%	0% (0/41)
char	0/92	0%	0% (0/41)
auto	68/92	74%	68% (28/41)
automobile	5/92	5%	5% (2/41)
voiture	19/92	21%	20% (8/41)

[†]Dispersion rate: number of speakers who used a given variant at least once divided by the total number of speakers.

and *automobile*:

c)   the immersion students will make frequent use of the variant *auto* since it is abundant in the input to which they are exposed (it is found in all input sources) and is not structurally complex;

d)   the immersion students will make use of the hyperformal variant *voiture* and of the formal variant *automobile* since both variants are found in their input and *automobile* has an English cognate; however, in line with their educational input, the frequency of the formal variants should be lower than that of *auto*.

Regarding correlations between variants and the linguistic and social factors that have been found in L1 Canadian French, it is likely that some, but not all of these factors will condition variant usage.

5. RESULTS

5.1 *Overall distribution of variants*

The general distribution of the five variants in the immersion students' spoken French speech is presented in Table 3.

As we can see, neither marked informal variant (*char* and *machine*) was used by speakers in the French immersion corpus. This finding clearly confirms hypotheses a) and b) mentioned in the preceding section. It is also in keeping with the general trends discussed above in section 3.1 and is reminiscent of the finding that the vernacular **lexical** variants *rester* 'live' and *ouvrage* 'job' are absent from the speech of French immersion speakers. While the absence of *char* is not surprising, it should be pointed out that it constitutes a clear divergence between L1 and L2 speakers since the former speakers use it in approximately one quarter of occurrences in the same situation of communication, namely a semi-directed taped interview. Note also that we had considered the possibility that *char* might have some currency in the speech of immersion speakers since it is formally similar to the English equivalent *car* and that a correlation between the use of English and this variant was found in Nadasdi *et al.*'s (2004) study of the same variable in Ontario French. However, it would appear that significant exposure to the variant is a pre-condition to its usage. As we saw in our examination of variants in the educations input, no model for the use of *char* is found in the input to which immersion speakers are exposed. It may

374

Table 4. *Effects of linguistic constraints on* auto *versus* voiture

Factor Groups	*Auto* (N)	*Voiture* (N)	*Auto* (%)	*Voiture* (%)	Total (N)	Factor Effect on *Auto*
Priming						
By *auto*	12	0	100	0	12	knock out
By *voiture*	0	1	0	100	0	
Unprimed	56	18	76	24	74	
Preceding Element						
Adjective	12	2	86	14	14	N.S.
Preposition	14	2	87.5	12.5	16	
Determiner	42	15	74	26	57	
TOTAL	68	19	78	22	87	
Significance = .01 Input = .89						

also be surmised that although half of the students have stayed with Francophone families in Quebec, that such stays were too short (17 days on average) to provide the frequency of exposure to *char* necessary for them to learn this variant.

The primary variant used by students is *auto*, found in 68% of occurrences. Furthermore, this variant is widely dispersed and is used at least once by almost 70% of speakers in the corpus. These rates of frequency, which are in line with those found in the dialogues included in the teaching materials, confirm hypothesis c) above. The next common form found is the hyperformal variant *voiture* which is used approximately 20% of the time by students, a rate which is similar to that found in the teaching materials and which confirms our hypothesis that the frequency of *voiture* would be lower than that of *auto*. The third variant found in the immersion corpus is *automobile*. Our results, however, suggest that this is a marginal variant since it is only found in 5% of occurrences and is only used by two speakers. Thus it would seem that the frequent use of *automobile* in the textual parts of the teaching materials has not, so to speak, rubbed off on the students' speech.

In order to examine the effect of linguistic and social factors on the variable, we have performed regression analysis of the two primary variants, namely *auto* and *voiture* and have considered the two linguistic constraints that were found to influence variant selection in L1 French, namely lexical priming and the type of preceding element. Results concerning the effect of these two linguistic constraints are presented in Table 4.

5.2 *Linguistic factors conditioning the variable in French immersion*

Table 4 reveals that, as was the case for Canadian Francophones, priming is highly relevant for the use of lexical variants in the spoken French of immersion students. Thirteen occurrences of the variable were primed, and each case gave rise to the primed variant (12 times for *auto* and once for *voiture*). While the frequency of lexical variants in spontaneous speech is relatively modest, the general trend that

emerges is identical to that of native speakers: a variant used in the interviewer's speech to elicit speech on a given topic is always used in the speaker's response. This result is perhaps not surprising since lexical priming was selected as a highly significant factor in Nadasdi and McKinnie's (2003) study of verbs of residency. This result constitutes further evidence that immersion students' variant use in a semi-directed interview is strongly influenced by variants used by the interviewer.

Let us now consider the role of the preceding element. The reader will recall that when the effect of preceding element was considered for L1 speakers, it produced a robust association. In particular, it was noted that the presence of a preceding preposition strongly favored the use of *auto* and disfavored the variant *char*. The explanation given for this distribution was that there is a strong lexical association between *en* and *auto* such that this combination has the status of a fixed lexical expression. However, no such association is found for the immersion students, suggesting that no such lexicalization has taken place for them. One should bear in mind, however, that the association found between *en* and *auto* in L1 Canadian French resulted from a comparison with *en char*. The reader will recall that in the 2005 corpus of Ontario spoken French, we found 51 tokens of *voiture* representing 8% of all tokens of the variable. Preliminary analysis of these tokens and of the tokens of *char* and *auto* suggests that while *voiture* may be somewhat on the rise in Ontario spoken French, *en* still exerts the favourable effect on *auto* found in the Mougeon and Beniak corpus but does not favour *voiture* or *char*. Specifically, in the 2005 corpus, after *en*, we found 48 tokens of *auto*, and only 5 tokens of *char* and 5 tokens of *voiture*. Thus we have an indication that even though the immersion students do not use *char* they have not mastered the association between *en* and *auto* observable in Canadian spoken French.

5.3 Social factors

Results concerning the effect of social factors on variant usage are presented in Table 5.

The first observation we can make is that there is a clear correlation between speakers' home language and variant use. In particular, speakers from a Romance language background (five students who speak Italian and one who speaks Spanish), display a greater preference for the variant *auto* than do speakers from other language background groups. While the general rate of *auto* usage is 68%, it climbs to 93% when the student also speaks a Romance language at home. We attribute this correlation to the fact that the form *auto* is highly reminiscent of the Italian words *auto* and *automobile* and the Spanish words *auto* and *automóvil*. This result is in line with previous research on verbs meaning 'work' since here too, home language was selected as a significant factor with Romance speakers showing a preference for *travail* versus *emploi* (both Italian and Spanish have a cognate for the former). That said, the words *auto*, *automobile* and *automóvil* are not the primary variants used by Italian and Spanish speakers, who in fact make more frequent use respectively of variants *macchina* and *carro* and variant *coche*. Still, based on the correlation we have

Table 5. *Effects of extra-linguistic constraints on* auto *versus* voiture

Factor Groups	*Auto* (N)	*Voiture* (N)	*Auto* (%)	*Voiture* (%)	Total (N)	Factor Effect on *Auto*
Home Language						
Romance	14	1	93	7	15	.97
English	26	11	70	30	37	.27
Other	28	7	80	20	35	.45
Fr. Environment						
0 hours	17	1	94	6	18	.75
1 − 6 days	11	7	61	39	18	.07
7 − 20 days	33	2	94	6	35	.86
Over 3 weeks	7	9	44	56	16	.11
Sex						
Female	55	12	82	18	67	N.S.
Male	13	7	65	35	20	
Grade						
9	32	4	89	11	36	N.S.
12	36	15	71	29	51	
French Media						
Never	39	5	89	11	44	N.S.
Occasional	29	14	67	33	43	
Social Class						
Middle	43	12	78	22	55	N.S.
Upper– Working	22	7	76	24	29	
TOTAL	68	19	78	22	87	

Significance = .01 Input = .89

documented, we surmise that the Romance speakers' preference for *auto* reflects the fact that *auto* and its morphologically non reduced variants have wider currency in Italian and Spanish than in English.

Let us now consider exposure to French outside the classroom. When previous research has found stays in a French environment to correlate linearly with variants, the pattern has generally been that the greater the length of stay, the more use of informal variants one finds. The explanation offered for this pattern is that greater exposure in a Francophone environment would provide students with more opportunities to hear non-standard variants not found the classroom input (e.g. *ne* deletion, cf. Rehner and Mougeon, 1999, *on* for *nous* Rehner, Mougeon and Nadasdi, 2003).

We saw above that the limited stays that the students made in Quebec were not enough of a factor to induce learning of the informal variant *char*. The lack of an effect of amount of exposure to L1 French in the present study also suggests that such limited stays have not been enough of a factor to sensitise them to the rarity of hyperformal *voiture* in Quebec spoken discourse.

Results for sex and social class suggest that students are unaware of the hyperformal status of *voiture* in Canadian French since no correlations exist between the variable and these categories. That said, previous research on the immersion

Table 6. *Comparison of variants in all corpora*

Variants	Quebec spoken French	Immersion Teachers	Materials: Text	Materials: Dialogues, etc.	Immersion Students
char	23%	0%	0%	0%	0%
machine	19%	0%	0%	0%	0%
auto	42%	33.5%	21%	75%	74%
automobile	14%	0%	55%	0%	5%
voiture	2%	63.5%	24%	25%	21%

corpus has shown that this is not always the case. For example, middle-class students insert the hyperformal negative particle *ne* more often than their working-class counterparts and female students use hyperformal *nous* more often than informal *on*. Still, it may be that the relatively low frequency of lexical variables prevents students from gleaning information about their sociolinguistic status. Evidence for this claim is also found in Nadasdi and McKinnie's 2003 study of lexical variation in immersion French since neither sex nor social class was found to correlate with the lexical variables they examined.

Finally, as with past studies, French-media consumption does not influence variant choice (cf. Mougeon, Nadasdi and Rehner, 2002).

6. DISCUSSION

An overall comparison of the distribution of variants in Quebec spoken French, pedagogical input and the speech of French immersion students is presented in Table 6.

Several observations can be made regarding this comparative data. First, the French immersion students make no use of the marked informal variants *char* and *machine*. This result is probably unavoidable since they are never exposed to these forms through the classroom input and are unlikely to have been intensively exposed to them in their limited contacts with Francophones outside of the classroom. It is true that like Francophones (and unlike the teachers!), *auto* is the dominant variant used by immersion students. However, their use of this form surpasses what is found in native speaker discourse. Still, this is a socio-stylistically neutral variant in Quebec French and is not necessarily an undesirable outcome. In other words, it does not cause their speech to appear non-native like. On the other hand, the students' use of the hyperformal variant *voiture* in close to a quarter of the contexts for the variable is a notable divergence from L1 Canadian Francophone norms (even from those of contemporary adolescent speakers in Ontario). Like the French immersion students, the teaching materials use *voiture* more often than native speakers do! It is perhaps understandable that texts meant to represent written French make significant use of the *voiture* variant. However, its overuse in materials designed to replicate spoken dialogues is in need of revision. We therefore suggest that, along the lines of Nadasdi, Mougeon and Rehner (2005) the educational materials be

modified to familiarise students with informal variants like *char* and that they make less frequent use of forms like *voiture* that are rare in spoken Canadian French. This would not be difficult to accomplish since it does not involve the learning of a complex grammatical rule. It is simply a matter of raising their awareness of informal lexical variants and downplaying the prominence of *voiture* in the classroom input. This is particularly necessary for the teachers' discourse to ensure that students are presented with a sociolinguistically accurate model for variant usage. It should also be noted that such a proposal is in keeping with the Ontario Ministry of Education (2000) guidelines for the teaching of French in high school immersion programmes which underscore the need for students to develop familiarity with varieties of Canadian French.

In the case of *automobile* the infrequency of this variant in the students' speech is likely the result of several factors: 1) infrequency of this variant in the teachers' speech; 2) absence of this variant in the dialogues included in the teaching materials; and 3) the possibility that the students equate this form with the English word *automobile*, which is marked in spoken English, and therefore avoid the use of its French equivalent.

Finally, the noteworthy finding that the immersion students strongly prefer *auto*, in spite of its limited use in classroom teacher speech deserves an explanation. Here again, it is possible to invoke the influence of inter–systemic factors. Specifically, the morpheme *auto-* is widely used in English in the same semantic field, even though it does not function as a noun (e.g., *the auto-industry*, *auto-workers*, *automotive*) and this may have triggered a process of convergence towards the French variant *auto* on the part of the immersion students. That said, it is also plausible to hypothesise that such convergence may have been reinforced by the fact that, as we have seen, in the dialogues included in the teaching materials *auto* is clearly preferred over *voiture*.

As regards factors that correlate with the variable, we have noted that like native speakers, lexical priming is a key factor in predicting variant use of the French immersion students. That said, the immersion student do not follow the constraint involving the preceding element found in L1 research, namely a strong association between *auto* and the preposition *en*. The only social factor that was found to correlate with variation was home language and that the effect of this factor was higher rates of *auto* use on the part of speakers of a Romance language. We note that in our corpus all of our Romance language speakers speak either Italian or Spanish which have the related words *automobile* and *automóvil*.

Address for correspondence:
Terry Nasdasi
Department of Linguistics
University of Alberta
Edmonton, Alberta Canada
T6G 2E6
e-mail: terry.nadasdi@ualberta.ca

REFERENCES

Allen, P., Cummins, J., Harley, B. and Swain, M. (1987). *Development of Bilingual Proficiency Project*. Toronto: OISE, University of Toronto.

Armstrong, N. (2001). *Social and Stylistic Variation in Spoken French. A Comparative Approach*. Amsterdam/Philadelphia: Benjamins.

Chambers, J. K. (1999). Converging features in the Englishes of North America. *Cuadernos de Filología Inglesa*, 8: 117–127.

Chambers, J. K. (2000). Region and language variation. *English World-Wide*, 21: 1–31.

Chambers, J. K. (2002). Dialect Topography of Canada. Macro-sociolinguistic regional surveys of Golden Horseshoe 1991, Golden Horseshoe 2001, Ottawa Valley, Quebec City, Montreal, New Brunswick, Eastern Townships and Greater Vancouver, as well as adjacent U.S. regions (upstate NY, Maine, Vermont and Washington): http://www.chass.utoronto.ca/~chambers/dialect_topography

Chambers, J. K. (2004). 'Canadian Dainty': the rise and decline of Briticisms in Canada. In: R. Hickey (ed.), *The Legacy of Colonial English: A Study of Transported Dialects*. Cambridge: Cambridge University Press, pp. 224–241.

Deslauriers, L. and Gagnon, N. (1995). *Capsules. Manuels 5A, 5B*. Montreal: Modulo.

Deslauriers, L. and Gagnon, N. (1997). *Capsules. Manuels 6A, 6B*. Montreal: Modulo.

Lavandera, B. (1978). Where does the sociolinguistic variable stop? *Language and Society*, 7: 171–182.

Lealess, A. (2005). En français, il faut qu'on parle bien: Assessing native-like proficiency in L2 French. MA thesis, University of Ottawa.

Le Dorze, P. and Morin, N. (1994). *Portes ouvertes sur notre pays. Manuel 3A, 3B*. Montreal: Guérin.

Lodge, R. A. (2004). *A Sociolinguistic History of Parisian French*. Cambridge: Cambridge University Press.

Martel, P. (1984). Les variantes lexicales sont-elles sociolinguistiquement intéressantes? *Sociolinguistique des langues romanes. Actes du XVIIème Congrès International de Linguistique et Philologie Romanes*. Aix-en-Provence: Université de Provence, pp. 181–193.

Mougeon, R. and Beniak, É. (1991). *Linguistic Consequences of Language Contact and Restriction: The Case of French in Ontario, Canada*. Oxford: Oxford University Press.

Mougeon, R., Rehner, K. and Nadasdi, T. (2004). The Learning of Spoken French Variation by Immersion Students from Toronto, Canada. *Journal of Sociolinguistics*, 8.3: 408–432.

Mougeon, R., Nadasdi, T. and Rehner, K. (2002). État de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE. In: J.-M. Dewaele and R. Mougeon (eds.), L'acquisition de la variation par les apprenants du français langue seconde, Special issue of *Acquisition et Interaction en Langue Étrangère*, 17: 7–50.

Mougeon, R., Nadasdi, T. and Rehner, K. (to appear). The sociolinguistic competence of French immersion students. Amsterdam: John Benjamins.

Nadasdi, T. Mougeon, R. and Rehner, K. (2005). Learning to speak everyday (Canadian) French. *Canadian Modern Language Review*, 61.4: 543–561.

Nadasdi, T., Mougeon, R. and Rehner, K. (2004). Expression de la notion de «véhicule automobile» dans le parler des adolescents francophones de l'Ontario. *Francophonies d'Amérique*, 17: 91–106.

Nadasdi, T., Mougeon, R. and Rehner, K. (2003). Emploi du futur dans le français parlé des élèves d'immersion française. *Journal of French Language Studies*, 13.2: 195–219.

Nadasdi, T. and McKinnie, M. (2003). Living and working in Immersion French. *Journal of French Language Studies*, 13.1: 47–61.

Poirier, C. (1998) (ed.). Dictionnaire historique du français québécois. Les Presses de l'Université Laval.

Rehner, K. and Mougeon, R. (1999). Variation in the spoken French of immersion students: To *ne* or not to *ne*, that is the sociolinguistic question. *Canadian Modern Language Review*, 56.1: 124–154.

Rehner, K. and Mougeon, R. (2003). The effect of educational input on the development of sociolinguistic competence by French immersion students: The case of expressions of consequence in spoken French. *Journal of Educational Thought*, 37.3: 259–281.

Rehner, K., Mougeon, R. and Nadasdi, T. (2003). The learning of sociolinguistic variation by advanced FSL learners: The case of *nous* versus *on* in immersion French. *Studies in Second Language Acquisition*, 25: 127–156.

Roy Nicolet, L. and Jean-Côté, M. (1994). *Portes ouvertes sur notre pays. Manuel 1A, 1B.* Montreal: Guérin.

Sankoff, D. (1988). Sociolinguistics and syntactic variation. In: F. Newmeyer (ed.), *Language: The Socio-cultural Context*. Cambridge: Cambridge University Press, pp. 140–161.

Note aux auteurs

Les propositions d'articles et la correspondance doivent être envoyées à:

Professor Florence Myles
School of Modern Languages
Newcastle University
Old Library Building
Newcastle upon Tyne NE1 7RU
England UK
E-mail: jfls@newcastle.ac.uk

Les publications soumises pour revue et les manuscrits de comptes-rendus doivent être envoyés à:

Dr David Hornsby
Department of French
School of European Culture and Languages
Cornwallis North West
University of Kent at Canterbury
Kent CT2 7NF
England UK
E-mail: d.c.hornsby@kent.ac.uk

Les langues du *JFLS* sont l'anglais et le français. Les auteurs sont invités à soumettre articles et comptes-rendus dans leur langue première. Aucun article présenté au *JFLS* n'aura ni été soumis simultanément pour une évaluation à une autre revue scientifique, ni fait l'objet d'une publication préalable ailleurs.

COPYRIGHT

Les auteurs doivent obtenir eux-mêmes les autorisations de citation lorsqu'elles sont soumises à copyright. Il sera demandé aux auteurs dévolution de copyright au profit de Cambridge University Press, sous des conditions à préciser.

PRÉPARATION DES MANUSCRITS

1. Les propositions d'articles doivent être soumises soit en **cinq** exemplaires, soit sous forme électronique (disquette IBM PC compatible, logiciel Microsoft Word ou WordPerfect) accompagnée d'un imprimé rigoureusement conforme à la version informatisée. (La voie électronique permet une évaluation internationale plus rapide.) **La longueur des articles est habituellement comprise entre 6000 et 8000 mots**. L'auteur indiquera à la fin le nombre approximatif de mots. Les propositions de revues doivent être soumises en **deux** exemplaires papier, ou sous forme électronique. On se bornera au nombre de mots spécifié au préalable. Les manuscrits soumis, tant d'articles que de revues, sont définitifs et les épreuves ne doivent pas donner lieu à des révisions d'auteur.

2. Établis sur papier A4 (ou de format équivalent), les manuscrits auront double interligne et les marges ne seront pas inférieures à 4 cm (1.5 pouces).

3. Les manuscrits d'articles seront accompagnés d'un résumé d'environ 100 mots présentant le contenu conceptuel de la contribution. Le résumé en double interligne sera présenté sur feuille séparée.

4. Pour les articles comme pour les revues, tous les paragraphes sauf le premier commencent par un retrait. Il n'y a pas d'espace particulier entre les paragraphes. Les paragraphes ne sont pas numérotés.

5. Les citations sont introduites par des guillemets simples, les double guillemets n'étant utilisés que pour les citations à l'intérieur de citations. Les citations de plus de 60 mots environ seront disjointes du texte et décalées de la marge gauche. Elles sont présentées sans guillemet.

6. Les tableaux, figures et graphiques, à l'exclusion des simples règles, exemples et formules, seront présentés chacun sur feuille séparée. Ils seront numérotés en continu de façon appropriée 'Figure 2', 'Tableau 4'. Dans le texte, on y renverra par ce numéro et jamais par une formule du genre 'voir la figure suivante'.

7. Pour indiquer les caractères **gras**, *italiques* ou les PETITES MAJUSCULES, il est acceptable d'utiliser les fonctions du logiciel, pourvu que le recours à celles-ci soit systématique. Autrement, l'indication d'italique sera notée par un soulignement simple, le gras par un soulignement ondulé, les petites majuscules par un soulignement double. On ne mélangera pas les deux systèmes.

8. Autant que possible, les transcriptions phonétiques seront conformes à l'Alphabet Phonétique International (API/IPA). Elles n'excéderont pas en étroitesse ni en précision ce qui est absolument nécessaire à l'analyse présentée.

9. On évitera les allusions sexistes et autrement blessantes.

10. Lorsqu'un article a été accepté, il est souhaitable que l'auteur fournisse une copie de la version finale sous forme informatique (IBM PC compatible) en plus de la version finale papier. Les indications relatives au logiciel utilisé seront également fournies (Microsoft Word, WordPerfect). L'éditeur se réserve la possibilité de faire établir lui-même le manuscrit si la disquette fournie se révèle inadéquate.

RÉFÉRENCES

1. **On évitera, autant que faire se peut, les notes infrapaginales.**

2. Pour les citations et références, on indiquera dans le corps du texte le nom de l'auteur, la date de publication et les numéros de pages entre parenthèses (Dupont, 2002: 56–58). Lorsque cela est nécessaire, on distinguera les publications d'un même auteur et d'une même année par des lettres minuscules (a, b, c) placées après la date. Les références multiples seront ordonnées par date et séparées par des point-virgules. Pour les co-auteurs, on utilisera 'et/and' et non '&'.

3. La liste complète des références citées, présentée en ordre alphabétique de nom d'auteur et en double interligne, sera jointe à la fin de l'article. On indiquera le lieu de publication ainsi que l'editeur. Les titres de livres, articles et revues seront complets et développés conformément aux exemples ci-dessous:

Adam, J.-M. (1999). *Linguistique textuelle: des genres de discours aux textes*. Paris: Nathan.

Balibar-Mrabti, A. (dir.) (1997). *La Synonymie.* = *Langue Française* 128. Paris: Larousse.

Basset, P., Amelot, A., Vaissière, J. et Roubeau, B. (2001). Nasal airflow in French spontaneous speech. *Journal of the International Phonetic Association*, 31: 87–99.

Bentolila, F. (1988). Autour du verbe. *La Linguistique*, 24.1: 3–141.

Bilger, M. (2001). Retour sur 'le futur' dans les corpus de français parlé. *Recherches sur le Français Parlé*, 16: 177–188.

Brugmann, K. (1906). *Grundriss der vergleichenden Grammatik der indo-germanischen Sprachen,* 2ième éd., vol. 2, 1ère partie. Leipzig: Trubner.

Johnson, K. (2000). Grammaticalization chains and French complex-inversion. Dans: S. N. Dworkin et D. Wanner (dir.), *New Approaches to Old Problems*. Amsterdam: Benjamins, pp. 183–198.

Kaye, A. S. (2003). Compte-rendu de: I. Fónagy, *Languages Within Language. Language*, 79: 652.

Weidenbusch, W. (1993). *Funktionen der Präfigierung: präpositionale Elemente in der Wortbildung des Französischen*. (Beiheft zur *Zeitschrift für Romanische Philologie* 247). Tübingen: Niemeyer.

Notes for Contributors

Contributions and correspondence should be sent to:

Professor Florence Myles
School of Modern Languages
Newcastle University
Old Library Building
Newcastle upon Tyne NE1 7RU
England UK
E-mail: jfls@newcastle.ac.uk

Material for review and review copy should be sent to:

Dr David Hornsby
Department of French
School of European Culture and Languages
Cornwallis North West
University of Kent at Canterbury
Kent CT2 7NF
England UK
E-mail: d.c.hornsby@kent.ac.uk

The languages of the journal are French and English. Contributors should submit articles and reviews in their first language. Submission of an article is taken to imply that it has not been previously published, and has not been submitted for publication elsewhere.

COPYRIGHT

Authors are responsible for obtaining necessary permissions to quote copyright material. They will be requested to assign their copyright, on certain conditions, to Cambridge University Press.

PREPARATION OF MANUSCRIPTS

1. Typescripts of articles should be submitted either in **five** paper copies or in electronic format (IBM PC compatible diskette, using Microsoft Word or WordPerfect software), together with one print-out identical in all respects to the electronic version. (Electronic submission speeds up the international refereeing process.) **Articles should normally be between 6000 and 8000 words**. Authors should indicate the approximate number of words at the end of their text. Typescripts of reviews should be submitted in **two** copies, or in electronic format. Reviews should not exceed the number of words suggested in each case. Typescripts, whether or articles or reviews, should be definitive, and author's revisions should not be made at proof stage.

2. All copy must be double-spaced throughout on A4 paper (or equivalent), with margins of at least 4 cm (1.5 inches).

3. The typescript of an article should be accompanied by an abstract of about 100 words summarising its conceptual content. It should be typed (double-spaced) on a separate sheet of paper.

4. In both articles and reviews, all paragraphs except the first are indented. There is no blank line between paragraphs. Paragraphs are not individually numbered.

5. Quotations should be marked by single inverted commas, with double inverted commas reserved for quotations within quotations. Quotations longer than about 60 words should be set off from the text and indented from the left-hand margin, without inverted commas.

6. Tables, charts and diagrams (though not simple rules, examples or formulae) must each appear on a separate sheet, and must be labelled 'Figure 2', 'Table 4', as appropriate. In-text references should use the same number, and never an expression like 'see the following diagram'.

7. **Boldface**, *italic type* and SMALL CAPITALS may be indicated by using software functions, provided that this is done systematically. Otherwise, italics should be indicated by a single underline, bold by wavy underlining, and small capitals by a double underline. The two systems must not be mixed.

8. Phonetic transcriptions should, wherever possible, make use of the symbols of the International Phonetic Alphabet, and should be no 'narrower' than is absolutely necessary for the purpose.

9. Sexist and other offensive language should be avoided.

10. When an article has been accepted for publication, all authors are asked to send a copy of the final version on computer disk (IBM PC compatible) together with the hard copy typescript, giving details of the word processing software used (Microsoft Word or WordPerfect). However, the publisher reserves the right to typeset material by conventional means if an author's disk proves unsatisfactory.

CITATIONS

1. Footnotes should as far as possible be avoided.

2. Citations in the body of the text should give author's surname, date of publication and page number in parentheses, as in (Dupont, 2002: 56–58). Where a second or subsequent work by a particular author in the same year is cited, references should be distinguished by letters (a, b, c) placed after the date. Multiple references should be ordered chronologically and separated by semicolons. Co-authors should be indicated by 'and/et' and not by '&'.

3. The complete list of references cited, arranged alphabetically by first author's surname should be typed double-spaced at the end of the article. Give place of publication and the publisher. Titles of books, articles and journals should be given in full. References will normally match one of the following patterns:

Adam, J.-M. (1999). *Linguistique textuelle: des genres de discours aux textes.* Paris: Nathan.

Balibar-Mrabti, A. (ed.) (1997). *La Synonymie.* = *Langue Française* 128. Paris: Larousse.

Basset, P., Amelot, A., Vaissière, J. and Roubeau, B. (2001). Nasal airflow in French spontaneous speech. *Journal of the International Phonetic Association*, 31: 87–99.

Bentolila, F. (1988). Autour du verbe. *La Linguistique*, 24.1: 3–141.

Bilger, M. (2001). Retour sur 'le futur' dans les corpus de français parlé. *Recherches sur le Français Parlé*, 16: 177–188.

Brugmann, K. (1906). *Grundriss der vergleichenden Grammatik der indo-germanischen Sprachen*, 2nd edn, vol. 2, part 1. Leipzig: Trubner.

Johnson, K. (2000). Grammaticalization chains and French complex-inversion. In: S. N. Dworkin and D. Wanner (eds), *New Approaches to Old Problems*. Amsterdam: Benjamins, pp. 183–198.

Kaye, A. S. (2003). Review of: I. Fónagy, *Languages Within Language. Language*, 79: 652.
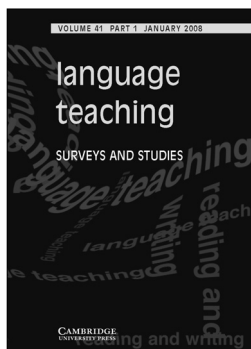
Weidenbusch, W. (1993). *Funktionen der Präfigierung: präpositionale Elemente in der Wortbildung des Französischen*. (Beiheft zur *Zeitschrift für Romanische Philologie* 247). Tübingen: Niemeyer.

language teaching

SURVEYS AND STUDIES

VOLUME 41 PART 1 JANUARY 2008

language teaching

SURVEYS AND STUDIES

CAMBRIDGE
UNIVERSITY PRESS

Editors
Ewa Jaworska, University of Essex, UK
Graeme Porte, University of Granada, Spain

Language Teaching is a well-established and unique research resource for both professionals and students in the broad area of second and foreign language teaching and learning. Volume 41 (2008) marks a turning point in the journal's history. While specially commissioned articles continue to survey up-to-date research on specific key topics, new strands of articles review recent second language acquisition and teaching research conducted in individual countries and also research on the teaching of languages other than English. There is a special series of plenary and keynote speeches from recent international conferences as well as annual reports from research groups around the world, including those from IATEFL, AAAL and BAAL. A further section is dedicated exclusively to empirical research papers which follow a replication approach in their methodology.

Price information is available at:

journals.cambridge.org/lta

Free email alerts
**Keep up-to-date with new material –
sign up at *journals.cambridge.org/register***