"Listen and Understand What I Am Saying": Church-Listening As a Challenge for Non-Native Listeners of English in the United Kingdom

Citation for the original published paper (version of record):
Malmström, H. (2015). "Listen and Understand What I Am Saying": Church-Listening As a Challenge for Non-Native
Listeners of English in the United Kingdom. International Journal of Listening, 29(1): 50-64.
http://dx.doi.org/10.1080/10904018.2014.880928

International Journal of Listening

ISSN: 1090-4018 (Print) 1932-586X (Online) Journal homepage: https://www.tandfonline.com/loi/hijl20

"Listen and Understand What I Am Saying": Church-Listening As a Challenge for Non-Native Listeners of English in the United Kingdom

Hans Malmström

To cite this article: Hans Malmström (2015) "Listen and Understand What I Am Saying": Church-Listening As a Challenge for Non-Native Listeners of English in the United Kingdom, International Journal of Listening, 29:1, 50-64, DOI: 10.1080/10904018.2014.880928

To link to this article: https://doi.org/10.1080/10904018.2014.880928

The Intl. Journal of Listening, 29: 50–64, 2015
Published with license by Taylor & Francis
ISSN: 1090-4018 print / 1932-586X online
DOI: 10.1080/10904018.2014.880928

Routledge
Taylor & Francis Group

"Listen and Understand What I Am Saying": Church-Listening As a Challenge for Non-Native Listeners of English in the United Kingdom

Hans Malmström

Centre for Languages and Literature
Lund University
Department of Applied IT, Division for Language and Communication
Chalmers University of Technology

This article uses computer-assisted analysis to study the listening environment provided by Bible readings and preaching during church services. It focuses on the vocabulary size needed to comprehend 95% and 98% of the running words of the input (lexical coverage levels indicating comprehension in connection with listening) and on the place of infrequent vocabulary in liturgical discourse. The finding that 4,000 words and 7,000 words, respectively, are needed to reach the target levels for lexical coverage suggests that non-native listeners with vocabularies of just a few thousand words may be seriously challenged by church listening.

INTRODUCTION

The vast majority of research to date investigating challenges facing non-native speakers and listeners of English (NNSLs) when listening to English is set in academic or learning-oriented listening contexts (see e.g., Bonk, 2000; Brett, Rothlein, & Hurley, 1996; Chaudron, 1995; Elley, 1989; Ellis, 1994, 1995; Ellis & He, 1999; Ellis, Tanaka, & Yamazaki, 1994; Flowerdew, 1994; Vidal, 2003). In such contexts, the NNSLs are primarily learners; that is, they have a more or less expressed desire to acquire English and/or learn to function in a particular context using English. However, great numbers of NNSLs also face English-listening challenges without an explicit desire to *learn* English.

Church services in Britain provide a case in point; research has shown that such contexts attract large numbers of people for whom the first language is not English (Foner & Alba,

Correspondence concerning this article should be addressed to Hans Malmström, Department of Applied IT, Division for Language and Communication (Johanneberg), Chalmers University of Technology, SE-412 96 Goteborg, Sweden. E-mail: hans.malmstrom@englund.lu.se

2008; Garcia-Munoz & Neuman, 2012). When these people attend a church service, their primary source of motivation is likely not to develop their knowledge of English or their ability to function in an English-speaking environment (that may happen elsewhere); it is much more likely that they come to church for purely religious reasons—they listen to English hoping to *experience* something, for example, as a result of listening to the readings from the Bible and the ensuing exposition of the readings in the sermon. One of the main challenges facing these church-going NNSLs is to *comprehend* the address they are listening to since this is arguably important for their religious experience. Key to overall comprehension is understanding the words used (as evidenced by the vocabulary teaching and learning literature; e.g., Bogaards & Laufer, 2004; Coady & Huckin, 1997; Folse, 2004; Laufer, 1997; Nation, 2001; Schmitt, 2000; Schmitt & McCarthy, 1997). In other words, in order to comprehend what they hear, listeners need to know, or have to be able to guess, the meaning of words they hear.

Research has established that many NNSLs, particularly adults who engage with English for the first time after adolescence, have a limited vocabulary compared with native speakers, sometimes consisting of just a few thousand words (Nation & Waring, 1997). We also know that while contextual knowledge provides some help to listeners trying to understand an unknown word (Nagy, 1997) in order to be able to guess the meaning of a word accurately, a large proportion of the words in the running text need to be understood (cf. the level of "lexical coverage," as conceived by van Zeeland and Schmitt [2012]).

This article is an investigation of the listening environment provided by two key components of the Christian Sunday church service, namely the Bible readings and the sermon that follows the readings. The primary question to be answered in this study is how many English words overall NNSLs need to know in order to reach a lexical coverage point that indicates comprehension of Bible readings and preaching. The study also provides a detailed discussion about the place of infrequent vocabulary (vocabulary beyond the 2,000 most frequent words in English) in liturgical discourse.

BACKGROUND

By drawing on findings from vocabulary acquisition research, this section first proposes a reasonable so-called lexical coverage range needed for comprehension during listening to Bible readings and preaching. This is a necessary first step towards establishing how many words NNSLs need to know. The second part of this section reviews some general claims concerning vocabulary in liturgical discourse.

Large Vocabulary Needed to Understand Spoken English – A Matter of Lexical Coverage

A number of different factors may affect comprehension in connection with a listening activity, for example, background knowledge of the topic (Stahl et al., 1991), individual differences between listeners, and how important the unknown word is (i.e., its relevance in the context; Stahl, Richek, & Vandevier, 1990). However, it has been suggested (Laufer & Sim, 1985; Schmitt, Jiang, & Grabe, 2011) that the most important factor is how many of the running words in the input are known to the listener. Previous research refers to this factor as *text coverage* (e.g., Nation, 2006;

Nation & Waring, 1997; Webb & Rodgers, 2009) but since this study is concerned with aural input only, I prefer the term *lexical coverage*, following Adolphs and Schmitt (2003) and van Zeeland and Schmitt (2012).

Nation (2001, p. 114) suggests that 95% lexical coverage would lead to "reasonable comprehension" in connection with listening but that 98% input coverage "would be better." Nation bases these statements on research in reading comprehension by Hu and Nation (2000), who investigated lexical coverage experimentally. By replacing different proportions of low frequency (actual) words with nonsense words in a fiction text given to non-native speakers, Hu and Nation (2000) measured reading comprehension both through a multiple-choice comprehension test and through written cued recall. At 80% lexical coverage no comprehension was recorded; at 90% some comprehension was recorded for a small minority of the subjects; at 95% better comprehension was established, however, still only with a minority of subjects. The final target for "adequate comprehension" (Nation, 2006, p. 61) was determined to be 98% lexical coverage.

The study by Hu and Nation (2000) and the 98% coverage target they propose has been used as a measure of comparison for several other studies, a small number of which are concerned with English language activities involving listening (rather than reading). Adolphs and Schmitt (2003) used two corpora of spoken English discourse (the Cambridge and Nottingham Corpus of Discourse in English, and the spoken part of the British National Corpus, the BNC) to address a previously well-established assumption that speakers can engage effectively in spoken discourse knowing only around 2,000 words (as proposed by Schonell et al., 1956). Adolphs and Schmitt (2003) found that the 2,000 most frequent words used in the two corpora amounted to about 95% input coverage, thus on the low side if communication is to be successful. Instead they suggested that around 3,000 words are needed to reach 96% lexical coverage and to increase the chances of successful engagement in spoken discourse. Using a slightly different methodology, Nation (2006) was concerned with two types of listening activities: watching a children's movie, and listening to unscripted speech. By comparing scripts from the children's movie *Shrek* and transcripts from a corpus of spontaneous spoken English against frequency lists from the BNC, he concluded that in order to reach the 98% coverage target needed for comprehension, which Nation (2006, p. 79) claims is "ideal," *Shrek* viewers and people listening to unscripted English would need to know between 6,000 and 7,000 words. Finally, Webb and Rodgers (2009) used a similar methodology to Nation (2006) to research the input coverage of movies, asking how many words are needed to understand a movie. They analyzed the scripts of more than 300 American and British movies and showed that if a viewer knows around 3,000 words, this would provide 96.76% coverage (their proposed lower boundary for "adequate comprehension") and that between 6,000 and 7,000 words are needed to reach 98% coverage (their proposed upper boundary for "adequate comprehension").

While the three studies referred to above seem to agree on an "adequacy range" for listening input coverage (95–98%, centering around 98%) and reject coverage numbers in the lower 90% range as insufficient to guarantee comprehension, a recent study (one of just a few actually *testing* lexical coverage and listening comprehension) indicates that less coverage, approximately in the area of 95% coverage, is actually needed and that even at the level of 90% coverage, comprehension is acceptable for a small minority of listeners. Van Zeeland and Schmitt (2012) used the same basic methodology as Hu and Nation (2000) and inserted different proportions of nonsense words into listening passages (four short narratives found online), which were read aloud to the participants. After the listening activity, the participants took a multiple choice test to measure

their comprehension of the story. A crucial result of the study was that listeners "could still comprehend the passages quite well if as much as 5 or even 10% of the words were unknown" (2012, p. 15). Van Zeeland and Schmitt (2012, p. 15) acknowledge that "comprehension of the 98% coverage story was significantly better than that of the 95 and 90% coverage stories." They conclude by saying that "if only very high comprehension will do, then 98% is probably a good coverage target [for listening]. But if less stringent comprehension rates are acceptable, then 90 and 95% seem to enable this" (2012, p. 18).

The difference between 95% and 98% coverage may seem small but is actually substantial when the added "burden" on the listeners is considered; according to van Zeeland and Schmitt (2012, p. 19), the lower target "would reduce estimates of the vocabulary size necessary for listening comprehension from 6,000 to 7,000 words (based on 98% coverage) to around 2,000–3,000 words (based on 95% coverage)." This has important implications for listeners in many contexts of listening, for example, church-listening, the context provided by the present study.

The purpose of this study is to determine how many words NNSLs need to know in order to comprehend Bible readings and preaching during a church service. To decide this, a level of reasonable lexical coverage must first be established. On balance, and in view of the findings presented by van Zeeland and Schmitt (2012) and other studies, it would seem that 95% and 98% lexical coverage are reasonable as possible lower and higher targets. At the level of 98% lexical coverage, one in every 50 words is unknown; at the level of 95% coverage, one in every 20 words is unknown. Both levels appear to indicate comprehension during listening for the majority of listeners in other listening contexts, and indicate a range from "adequate" to "optimal" comprehension (Laufer & Ravenhorst-Kalovski, 2010).

Vocabulary Concerns in Liturgical Discourse

For obvious reasons there are few normative statements in any literature about vocabulary used in Bible readings from a lectionary (reading calendar); it is what it is since the readings are taken from the Bible (and the translation adopted), and the flexibility allowed, at least concerning the Sunday lectionary, appears to be extremely limited. This is unsurprising given the objectives of a set lectionary. According to *The Introduction to the Revised Common Lectionary* (one of the most widely used lectionaries and the source of data used in this study), the lectionary is supposed "to provide whole churches or denominations with a uniform and common pattern of biblical proclamation"; if this objective is to be achieved, deviations (other than those authorized, for example shorter or longer versions of the three texts can be used) cannot be accepted.

A great deal of prescriptive advice has, however, been issued relating to vocabulary use in sermonic discourse, often presented as strong statements about what preachers should or should not do with words during preaching. Broadus and Weatherspoon (1944, p. 244), one of the seminal works in homiletics during the 20th century, say that it is the task of preachers to "render it not merely possible that the people should understand [. . .] but impossible that they should misunderstand (1944, p. 96). More recently, Robinson (2001, p. 186) claims that preachers "must make themselves understood instantly" [and that] no matter how accurately a phrase or word expresses a speaker's meaning, it is worthless if the listeners do not know what it means." Statements like these place considerable demands on preachers' words use. These demands can be addressed both from the point of view of how expansive a vocabulary preachers use in the pulpit and on the basis of the nature of the words themselves.

Buttrick (1987, p. 188) says that a "7,500-word vocabulary will be too large for preaching" and adds that a lexicon of "5,000 words" is appropriate for use in preaching because this is, he says, "the *common shared vocabulary* of a congregation (emphasis in original). Unfortunately Buttrick (1987) does not provide a strong link to research on which to base his claims, and it is therefore difficult to view them as anything but arbitrary numbers; however, the results presented in this paper actually provides some support to Buttrick's claim.

As for the nature of words used in preaching, homiletics caution preachers to avoid "'church chat' that might be the language of theologians and ministers but not the language of most people" (Waznak, 1998, p. 103). Buttrick (1987, p. 194) provides examples of the "religious lingo" which preachers may be tempted to use: "redemption, salvation, sanctification, born again, blessing, justification" and says that "such words may no longer be terribly useful [. . .] they are convenient words in theological discussion [. . .] but alien terminology to most people in a congregation". Preaching without religious words is obviously unthinkable, but if religious words make up too large a proportion of the words used, particularly if they are rare/infrequently used words, this could impact negatively on comprehension. Recent research (Malmström, 2014) has shown that a very small proportion of infrequent words, words which are not among the 2,000 most frequent words in English, (less than 2%) used during preaching are religious words, suggesting that use of religious vocabulary by itself is unlikely to cause listeners any problems. However, little is known about the overall proportion of infrequent words in sermons, or in Bible readings. Use of a too many infrequent words would be "inappropriate" (Craddock, 1985) since NNSLs are unlikely to have come across such words before, and this could add to the listening burden.

MATERIALS AND METHOD

The data for this study come in two parts: sermon manuscripts and Bible extracts in the form of a published Sunday lectionary.

Digital manuscripts of 50 Church of England Easter sermons from 2005 to 2011 (nominally Anglican, but with considerable variations across what is usually referred to as the Anglican spectrum) were obtained by contacting dioceses and parishes in England. The total size of the sermon data set amounts 63,537 running words (tokens).

The data source used for the Bible readings is the most recent edition of the *Revised Common Lectionary for Sundays and Festivals* (RCL; a digital version of the RCL was obtained from the Church of England). The RCL is claimed to be "an international and ecumenical lectionary without rivals" (RCL, Introduction), and with only minor modifications, the RCL is used by a great number of churches around the world, among which is the Church of England. For each Sunday, there is a set lectionary consisting of three readings and a responsorial psalm. The first reading is usually a reading from the Old Testament, followed by a responsorial psalm (also Old Testament), which is sometimes read and sometimes sung. The psalm is then followed by a reading from the New Testament, an epistle text, a section from the Acts, or the Book of Revelation. The Bible reading component of the Sunday/festive liturgy is then concluded by a reading from one of the four gospels. The size of the Bible reading data sample amounts 307,733 words (tokens).

The difference in sample size between the sermon sample and the Bible reading sample is irrelevant for the purposes of the present study. The limitations to the size of the samples can be justified from a natural and a practical standpoint. The sermon data constitute the complete

Anglican section of a research project concerned with linguistic aspects of contemporary sermonic discourse and was readily available. The sermon data were collected and prepared in a deliberate, systematic way, controlling for thematic variation among other things, and includes only one sermon from any one preacher. The Bible reading sample, since it is the complete lectionary for Sundays and festivals, is clearly reflective of readings in the Church of England during the Easter period. Admittedly, in corpus-linguistic terms, the two samples are very small. While small, specialized corpora (or text samples) make generalizations more problematic, they are sometimes used in corpus linguistic research (see, e.g., Hyland, 2008; Hyland & Tse, 2005). It is suggested that the data sizes used in this study provide indicative samples that serve the purposes of the study well.

Before analyzing the sermonic sample and the Bible sample, the data were cleaned where necessary. Many preachers have a practice of including on their sermon manuscripts the complete Bible readings of the day in question. Since this text is not part of the sermon itself it was deleted from the manuscripts (directly quoted text within the sermons was, of course, not deleted).

AntWordProfiler (Anthony, 2013) was used for all analysis of the sermon sample and the Bible reading sample. AntWordProfiler is a computer program that allows you to upload any list(s) of words against which to compare a text sample. In this case, I used 20 lists with words drawn from the British National Corpus and the Corpus of Contemporary American English, with approximately 1,000 words in each list to compare against the sermon sample and the Bible reading sample. All the lists used are frequency based so the first list contains the first thousand most frequently used words in English; the second list contains the second thousand most frequently used words in English, and so on. When a text sample is uploaded, AntWordProfiler will show what proportion of the text contains words from the first thousand-frequency list, the second thousand-frequency list, and so on, and it will also list the words themselves, and their individual frequency. Words which are not among the 20 thousand most frequently used words in English will appear in a separate category called "not in any list." For each level of thousand words, the program also produces a cumulative output indicating the level of lexical coverage so that it is possible to read off how many thousand words are needed to reach a certain level of lexical coverage. In addition to the 20 frequency lists, a list containing approximately 13,000 proper nouns was used. After initial analysis of the "not-in-any-list" category, any Biblical proper nouns that were not part originally part of the proper noun list were added to that list. The inclusion/exclusion of a proper noun list in the analysis makes it possible to calculate different profiles showing the actual proportion of proper nouns in the text sample.

RESULTS AND DISCUSSION

Summary of Results

The investigation indicates that NNSLs need to know 4,000 words (plus proper nouns) to reach 95% lexical coverage and 7,000 words (plus proper nouns) to reach 98% lexical coverage to understand Bible readings and preaching, clearly a challenge to NNSLs with vocabularies of just a couple of thousand words. However, it is suggested that slightly lower levels of vocabulary knowledge may be acceptable because of the discursive link between the readings and the sermon and because of the interpretative and didactic nature of the sermon. Infrequent vocabulary in

liturgical discourse, that is, words beyond the 2,000-level of frequency, is then analyzed in more detail. The analysis suggests that listeners' opportunities for incidental learning through listening exposure are very limited. Even if preachers use an infrequent word approximately once every five seconds during preaching, few of the infrequent words have a wide range, i.e. they do not occur in most of the sermons. The last part of the paper talks about different reasons why NNSLs are likely to encounter infrequent vocabulary in liturgical discourse.

How Many Words Do NNSLs Need to Know?

As Tables 1 and 2 show, the words used in both Bible readings and preaching spread across the 20 frequency lists and beyond (as evidenced by the "not-in-the-list" category). The first row in the tables shows that the first thousand most frequent words cover 83.53% or 86.51% (depending on whether proper nouns are included or not) of the Bible reading data, and 83.82% or 86.90% of the preaching data. When words from the 2,000-level list are added, the cumulative lexical coverage figures for both Bible readings and preaching increase by between 5% and 6 %. Clearly the first and second thousand most frequent words in English account for most of the words found in

TABLE 1
Lexical Coverage in Bible Readings (Revised Common Lectionary for Sundays and Festivals)

Frequency levels	*Cumulative lexical coverage without proper nouns (%)*	*Cumulative lexical coverage including proper nouns (%)*
1,000 level	83.53	86.51
2,000 level	88.71	91.69
3,000 level	90.84	93.82
4,000 level	92.29	95.27[a]
5,000 level	93.46	96.44
6,000 level	94.18	97.16
7,000 level	95.08[a]	98.06[b]
8,000 level	95.41	98.39
9,000 level	95.62	98.60
10,000 level	95.76	98.74
11,000 level	95.82	98.80
12,000 level	95.90	98.88
13,000 level	95.96	98.94
14,000 level	95.98	98.96
15,000 level	96.01	98.99
16,000 level	96.04	99.02
17,000 level	96.06	99.04
18,000 level	96.10	99.08
19,000 level	96.11	99.09
20,000 level	96.11	99.09
(Not in the list)	(3.89%)	(0.9%)
Total	(100.00)	100.00

[a]Reaching 95% lexical coverage.
[b]Reaching 98% lexical coverage.

TABLE 2
Lexical Coverage in Preaching (50 Contemporary Church of England Sermons)

Frequency levels	*Cumulative lexical coverage without proper nouns (%)*	*Cumulative lexical coverage including proper nouns (%)*
1,000 level	83.82	86.90
2,000 level	89.10	92.18
3,000 level	91.91	94.99
4,000 level	93.05	96.13[a]
5,000 level	94.25	97.33
6,000 level	94.89	97.97
7,000 level	95.30[a]	98.38[b]
8,000 level	95.53	98.61
9,000 level	95.67	98.75
10,000 level	95.76	98.84
11,000 level	95.80	98.88
12,000 level	95.84	98.92
13,000 level	95.87	98.95
14,000 level	95.89	98.97
15,000 level	95.91	98.99
16,000 level	95.95	99.03
17,000 level	95.96	99.04
18,000 level	95.97	99.05
19,000 level	0.00	0.00
20,000 level	95.98	99.06
(Not in the list)	(4.02%)	(0.94%)
Total		100.00

[a]Reaching 95% lexical coverage.
[b]Reaching 98% lexical coverage.

Bible readings and preaching. However, it is equally clear that knowledge of only the 2,000 most frequent words is typically not enough for comprehension as the lexical coverage figures at those levels are well below the lower 95% target.

In answer to the research question posed in the Introduction, a vocabulary of the 4,000 most frequent words would provide 95% coverage (on the assumption that proper nouns are unproblematic for comprehension; Nation, 2006), meeting the proposed "target" for adequate comprehension for both Bible readings and preaching. For optimal comprehension, 98% coverage, listeners need to know 7,000 words (including proper nouns) when listening to Bible readings or to someone preaching.

These numbers need, however, be treated with some caution. If a NNSL knows only, for example, 3,000 words, Bible readings and sermons are not necessarily incomprehensible to that person. The proposed levels of vocabulary knowledge needed should not be seen as "thresholds" in a strict sense (van Zeeland & Schmitt, 2012). While it has been widely established that greater vocabulary knowledge (i.e., knowing a higher percentage of the words in some input) leads to better comprehension, adequate comprehension is possible also at lower levels of vocabulary knowledge because other aspects of communication affect comprehension. Good background knowledge of a topic, for example, usually has a boosting effect on comprehension, making guessing at the meaning of unknown words easier than when listeners have little or no background knowledge.

What is more, the discursive link between the readings and the sermon (Barth, 1991; Robinson, 2001) and the often hermeneutical and didactic nature of the sermon will often afford NNSLs additional opportunities to engage with potentially unknown words because a word or concept may be focalized, repeated, elaborated, or re-contextualized as part of an extended communicative event. When a potentially unknown word or phrase is treated in this way during the preaching event, which is by no means standard practice, (e.g., "Let us spend some time reflecting on the *debauchery* in our contemporary society," or "Today, we will talk about the *apocryphal* gospels") infrequent words receive a lot of exposure, and they are "heard" and "explained" and can consequently be learned (Lord, 2010, p. 49). Two words from the Bible readings and their (re-) occurrence in the sermon data provide good examples of this practice; *redemption* (a 6,000-level word) and *leprosy* (a 12,000-level word) both occur with surprisingly high frequency in the sermon data. A closer look at the occurrences reveals that in one of the three sermons where it occurs, *redemption* is highly thematic, and consequently frequent, and in another it is mentioned, and then defined in two different ways. *Leprosy* only occurs in a single sermon, but there it is used to frame a sermon message that repeatedly returns to images of *disease, illness, sickness, cure*, and (incidentally) *redemption*. In other listening contexts, an unknown word may perhaps be used only once, and without any of the "scaffolding" found in the liturgical context, making listening and comprehension more difficult in those contexts.

The indication that church-listeners need to know the 4,000 most frequently used words of English to comprehend Bible readings and preaching (all other things being equal) should alert preachers to take care, at least when they craft their sermons (unless churches opt to use "easy English" Bible readings or alternative translations, little can be done about the Bible texts). If, as Nation and Waring (1997) note, significant numbers of NNSLs have a vocabulary of "much less" than 5,000 words, many NNSLs may be challenged by contemporary church-listening. We do not know what "much less" than 5,000 words refers to—what is clear is that NNSLs who only know around 2,000 words would run into serious church-listening problems. Indeed, several of the lexical coverage studies cited in the Background section of this paper suggest that lexical coverage levels in the low 90% range would mean that *no comprehension* is measurable. Perhaps preachers who know that they are preaching to a primarily immigrant audience could take extra care to use more words which they feel intuitively would be suitable to everyday "familiar conversation" (Waznak, 1998), thus automatically increasing the likelihood of using vocabulary which is known to listeners. It is interesting to compare the reported proportion of high-frequency words (the cumulative percentage of words at the 1,000 and 2,000 levels) of the sermons, which were written to be spoken, with findings reporting the proportion of high-frequency words in spontaneous English conversation. Based on data from two parts of the Wellington Corpus of Spoken English, Nation (2006) presents a lexical coverage figure for high frequency words in "unscripted spoken English" of 90.38%—this is even slightly lower than the coverage figure for the Church of England sermons, 92.18%. This suggests that preachers are heeding the advice from homiletics to use everyday words.

The findings presented here provide another point of comparison with claims from homiletics. Buttrick (1987) argues that a 5,000-word vocabulary is appropriate for use in preaching whereas a 7,500-word vocabulary would be too large. No doubt Buttrick is correct concerning the 7,500 word-vocabulary; this is indeed beyond the comprehensibility level for many NNSLs (and indeed some native speakers and listeners). While the truth of the often-cited claim that everybody in a congregation knows 5,000 words is seriously questionable (a proviso concerning the vocabulary size of the members of the audience has to be added), as an estimate of the

level of words needed to comprehend a sermon, Buttrick's seemingly arbitrary number seems quite reasonable. If a congregation knows the 5,000 most frequently used words in English, congregationers comprehend 97.33% of the words in the sermon, and 96.44% of the words in the readings, assuming that proper nouns are unproblematic. Actually, if church-listeners know the 5,000 most frequently used words, this means that they are closer to optimum comprehension than to adequate comprehension of preaching and readings.

Bible Reading and Preaching Vocabulary Beyond the 2,000 Level – The Place of Infrequent Vocabulary in Liturgical Discourse

All other things being equal, words which are not "everyday" or "conversational," that is, infrequent words beyond the 2,000 level, could cause NNSLs comprehension problems when listening to Bible readings and preaching simply because NNSLs are less likely to have been exposed to infrequent words than frequent words and exposure to an unknown word is crucial for comprehension (Schmitt, 2000; Waring & Takaki, 2003). Incidentally, NNSLs are unlikely to receive a lot of multiple exposures to words beyond the 2,000 level as a result of church-listening. The result of a range analysis (Table 3) of the preaching data indicates that only a small number of infrequent words occur in many different sermons. Only 21 infrequent words (out of a total of nearly 1,900 different infrequent word families) appear in 5 to 10 of the 50 sermons, and only 13 infrequent words appear in 11 to 15 sermons.

TABLE 3
Range Analysis of Words Beyond the 2,000 Word Frequency Level in Church of England Sermons

Words beyond the 2,000 level that appear with a minimum frequency of 1 in at least:	
5–10 sermons:	*11–15 sermons:*
sorrow	resurrection
silence	gospel
symbol	disciple
shadow	tomb
vulnerable	eternal
accuse	Jew
grave	reveal
resurrection	evil
gospel	weep
disciple	crucify
tomb	amen
eternal	transform
Jew	betray
reveal	
evil	
weep	
crucify	
amen	
transform	
betray	
redemption	

The only infrequent word to appear in *all* of the 50 sermons is *resurrection*; this is not surprising given the Easter theme that runs through all the sermons. The finding that so few of the words beyond the 2,000 level appear in a wide range of sermons suggests that the opportunities church-listeners have to gradually learn these words incidentally, merely as a result of multiple exposures, are very limited.

It is difficult to say for certain when a word becomes infrequent enough for it to potentially cause comprehension problems as this is highly listener dependent, but it seems reasonable to claim that anything beyond the 2,000-word level is potentially seriously challenging for NNLSs with a vocabulary of just a couple of thousand words. Tables 4 and 5 contain examples of words at or beyond the 3,000-word frequency level for Bible readings and preaching and give an indication of the kinds of words found at these levels.

As is evident from Table 1, 7.82% of the words found in the preaching sample come from the 3,000-word level or beyond, meaning that preachers use an infrequent word around 94 times in an average 1,200-word sermon (the length of Bible readings vary and an average is not a meaningful measure for that data). Based on a normal speaking rate of 160 words/minute, this would mean that a preacher uses a word like those in Table 5 approximately once every 5 seconds in a 7.5-minute sermon. Considering the transitory nature of spoken discourse (Nation, 2006), where listeners cannot "rewind" and listen again and have little time to digest the input before they encounter another infrequent word, the processing demand placed on NNSLs in the church-listening context is considerable.

It is interesting to speculate about why infrequent words are used to such a high degree in liturgical contexts since they are clearly potential barriers to interpersonal communication (Fry Brown, 2008). It is not as if Bible-reading words like *abhor, debauchery*, and *tempestuous* lack much more frequent synonyms (*hate, dishonesty, intense*); similarly, preachers who are anxious to reach out to everyone and exclude no one could replace preaching words like *affliction, tribulation*, and *extraneous* by their more common everyday equivalents: *sickness, trouble*,

TABLE 4
Examples of Word Families at and Beyond the 3,000 Level in Bible Readings

Frequency lists	*Sample occurrences, minimum frequency 10 tokens, frequency in parentheses*
3,000 level	endure (161), statute (56), scatter (36)
4,000 level	dwell (133), abundant (40), console (22)
5,000 level	deed (88), persecute (49), pasture (28)
6,000 level	perish (66), gracious (40), redemption (34)
7,000 level	dominion (53), reproach (19), vindicate (17)
8,000 level	countenance (30), deceit (30), rebuke (18)
9,000 level	gentile (58), steadfast (41), calamity (14)
10,000 level	sanctify (28), deliverance (19), forbear (13)
11,000 level	defile (19), revile (12), guile (10)
12,000 level	enthrone (25), recompense (10), fornicate (10)
13,000 level	iniquity (46), supplicate (21), gird (17)
14,000 level	licentious (10)
15,000–20,000 level	unleavened (11), accursed (10), reproof (10)

TABLE 5
Examples of Word Families at and Beyond the 3,000 Level in Preaching

Frequency lists	*Sample occurrences, minimum frequency 10 tokens in at least 2 sermons*
3,000 level	reveal (27), presence (20), anxiety (18)
4,000 level	despair (14), oppress (14), grief (14)
5,000 level	tomb (89), sorrow (19), conquer (12)
6,000 level	disciple (89), eternity (11), redeem (10)
7,000 level	baptism (24), bereave (12), humility (10)
8,000 level	resonate (12)
9,000 level	crucify (20), anoint (13), bequeath (11)
10,000 level	supposition (10)
11,000 level and above	n/a

and *irrelevant*. I suggest there are at least three reasons for why low-frequency words are used in church-listening contexts, namely (i) to introduce an element of learnedness; (ii) to guarantee the status and place of "holy speak" in church; and (iii) dependence on translations that faithfully represent the full meaning conveyed by the original biblical text.

Buttrick (1987, p. 88) argues that "the tone of erudition may flatter a preacher, but it will not assist the hearing of the gospel." Using words like those found in Table 5 is arguably dangerous if a preacher is conscious not to come across as learned, so even if it does not hinder comprehension because it may be a minor occurrence, it may affect the listeners' opinion of the preacher. What is more, preachers must be careful not to use words "for show" (Lord, 2010, p. 85). However, findings from qualitative studies on sermons from practicing preachers suggest that this is exactly what happens in some cases. Wallgren-Hemlin (1997) interviewed 45 Protestant preachers in Sweden—the focus of the study was to identify rhetorical strategies used by preachers in different preaching contexts. Several of the preachers interviewed expressed the view the learnedness was not necessarily something bad in a sermon since preachers, after all, must preach in a way that reaches out to different groups of listeners, from different backgrounds. In other words, more "learned" members of an audience will expect, the preachers claimed, a preaching style (and vocabulary to go with it) that matches their intellectual ability.

While it seems fair to associate certain lexical choices with learnedness, lexical choices are also reflective of a "religious style" which is supposed to be fitting for the context, which, in many cases at least, is best described as deferential. Notwithstanding the advice from homiletics and preaching instructions that church language should be like everyday conversational discourse, lest more formal styles alienate church listeners, liturgical discourse "conversants" (the preachers interviewed by Wallgren-Hemlin, 1997) seem to believe that it is important to preserve a speaking style which is a clear signifier of the context of speech, to allow for "holy-speak" in order to preserve certain sanctity of the discourse they are in.

Another reason why low-frequency words are found in liturgical discourse could be a desire by some Bible translators to be maximally faithful to meanings conveyed by the original biblical text. Grudem (2005) talks of a spectrum of Bible translations, ranging from essentially literal to very periphrastic. Certainly, at the literal end of the spectrum, Bible translators are more likely to explore the full potential of the richness of the English language rather than settling for frequently occurring everyday words. Proponents of essentially literal translations, such as Grudem, say that

literal translations (rather than more "dynamic" versions) are important to retain all the aspects of meaning in the original text, not just the main idea, and for that reason translations must sometimes use very specific words, even if they are rare. Grudem explains why the English Standard Version (published in 2001) uses a word like *propitiation* (a word found at the 17,000-word frequency level, thus outside the vocabulary of virtually all NNSLs) instead of a slightly more common synonym such as *soothing, calming* or *conciliation*:

> There is no other English word than 'propitiation' that means 'a sacrifice that bears the wrath of a deity.' But the word 'propitiation' has that meaning. This presents us with two choices. We can simply abandon the word and thereby give up (or make more difficult) the idea of teaching people this important concept (. . .). Or we can retain the word and thereby retain this important concept in the New Testament." (Grudem, 2012, unpaginated)

Thus, as long as Bible translations rely on essentially literal translations, church listeners are likely to be exposed to words like *propitiation* in connection with Bible readings.

CONCLUSIONS

Many (e.g., Robinson, 2001, p. 23) would claim that meaning and understanding in faith-based contexts is primarily not about understanding what individual words mean, but understanding "what the biblical writers mean through their use of words." In other words, just because church-listeners fail to understand the meaning of some individual word(s), this does not necessarily mean that a message at the "level of ideas" is beyond their comprehension. However, if an unknown word is a central word in the Bible story or sermon being heard, then it could cause listener comprehension problems, and important "points" made could be missed.

This article set out to study Bible readings and preaching as "listening environments." The findings that NNSLs need to know 4,000 and 7,000 words, respectively, to reach 95% and 98% lexical coverage suggest that NNSLs with very small vocabularies (smaller than 3,000 words) may be seriously challenged by these church-listening tasks. It should be stressed that the lower target of 95% lexical coverage is by no means a guarantee that all NNSLs will comprehend the input. In view of this situation, and assuming that it is their desire to make church-listening all-inclusive, preachers and governing bodies (at various levels) would be well advised to do what they can to adapt when they draft sermons and decide on lectionary texts for church services that include a predominantly non-native English speaking audience, even if this means that basic communicative concerns (increasing the likelihood of comprehension) would be prioritized over rhetorical and theological concerns governing the particular liturgical context, including a desire to sound "learned" or "holy" or to adhere to literal translations of the Bible.

FUNDING

The research for this paper was funded by a grant from the Swedish National Bank (Riksbankens Jubileumsfond).

REFERENCES

Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, *24*, 425–438. doi:10.1093/applin/24.4.425

Anthony, L. (2013). AntWordProfiler [Software]. Retrieved from http://www.antlab.sci.waseda.ac.jp/software.html

Barth, K. (1991). *Homiletic*. Louisville, KY: John Knox Press.

Bogaards, P., & Laufer, B. (Eds.). (2004). *Vocabulary in a second language: Selection, acquisition, and testing* (Vol. 10). Amsterdam, The Netherlands: John Benjamins.

Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, *14*, 14–31. doi:10.1080/10904018.2000.10499033

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal*, *96*, 415–422. doi:10.1086/461836

Broadus, J. B., & Weatherspoon, J. A. (1944). *On the preparation and delivery of sermons*. New York, NY: Harper & Brothers.

Buttrick, D. (1987). *Homiletic: Moves and structures*. Philadelphia, PA: Fortress Press.

Chaudron, C. (1995). *Academic listening. Department of English as a second language*. Honolulu, HI: University of Hawaii.

Coady, J., & Huckin, T. (Eds.). (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge, England: Cambridge University Press.

Cobb, T. (2013). *Web Vocabprofile*. Retrieved from http://www.lextutor.ca/vp/

Craddock, F. (1985). *Preaching*. Nashville, TN: Abingdon Press.

Drever, E. (2003). *Using semi-structured interviews in small-scale research: A teacher's guide*. Edinburgh, UK: The SCRE Centre.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*, 174–187. doi:10.2307/747863

Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: A review essay. *Applied Language Learning*, *5*, 1–32.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, *16*, 409–441. doi:10.1093/applin/16.4.409

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, *21*, 285–301. doi:10.1017/S0272263199002077

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, *44*, 449–491. doi:10.1111/j.1467-1770.1994.tb01114.x

Folse, K. S. (2004). Myths about teaching and learning second language vocabulary: What recent research says. *TESL Reporter*, *37*, 1–13.

Foner, N., & Alba, R. (2008). Immigrant religion in the U.S. and Western Europe: Bridge or barrier to inclusion? *International Migration Review*, *42*, 360–392. doi:10.1111/j.1747-7379.2008.00128.x

Fry Brown, T. L. (2008). *Delivering the sermon: Voice, body and animation in proclamation*. Minneapolis, MN: Fortress Press.

Flowerdew, J. (1994). *Academic listening: Research perspectives*. Cambridge, England: Cambridge University Press.

García-Mu-oz, T., & Neuman, S. (2012). *Is religiosity of immigrants a bridge or a buffer in the process of integration? A comparative study of Europe and the United States*. Retrieved from http://www.econstor.eu/bitstream/10419/58537/1/715655612.pdf

Grudem, W. (2005). Are only some words of scripture breathed by God: Why plenary inspiration favors essentially literal Bible translation. In W. Grudem, C. J. Collins, V. S. Poythress, L. Ryken, & B. Winter (Eds.), *Translating truth: The case for essentially literal Bible translation* (pp. 19–56). Wheaton, IL: Crossway.

Grudem, W. (2012). *The advantages of the English Standard Version (ESV) translation*. Retrieved from http://www.waynegrudem.com/wp-content/uploads/2013/01/The-advantages-of-the-ESV.pdf

Hu, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge, England: Cambridge University Press.

Laufer, B., & Sim, D. D. (1985). Measuring and explaining the reading threshold needed for English for academic purposes texts. *Foreign Language Annals*, *18*, 405–411. doi:10.1111/j.1944-9720.1985.tb00973.x

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*, 15–30.

Lord, J. L. (2010). *Finding language and imagery: Words for holy speech*. Minneapolis, MN: Fortress Press.

Malmström, H. (2014). "Church chat", academic prose, or conversation? Investigating contemporary sermons as lexical environments. Manuscript under review.

McFadden, K., Barrett, K., & Horst, M. (2009). What's in a television word list? A corpus informed investigation. *Concordia Working Papers in Applied Linguistics*, *2*, 78–98.

Nagy, W. E. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 64–83). Cambridge, England: Cambridge University Press.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9781139524759

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, *63*, 59–82.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge, England: Cambridge University Press.

Robinson, H. W. (2001). *Biblical preaching: The development and delivery of expository messages*. Grand Rapids, MI: Baker Academic.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, England: Cambridge University Press.

Schmitt, N., & McCarthy, M. (Eds.). (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge, England: Cambridge University Press.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*, 26–43. doi:10.1111/j.1540-4781.2011.01146.x

Schonell, F. J., Meddleton, I. J. & Shaw, B. A. (1956). *A study of the oral vocabulary of adults*. Brisbane, Australia: University of Queensland Press.

Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 41. *Journal of Literacy Research*, *23*, 487–508. doi:10.1080/10862969109547755

Stahl, S. A., Richek, M. G., & Vandevier, R. (1990). Learning word meanings through listening: A sixth grade replication. In J. Zutell & S. McCormick (Eds.), *Learning factors/teacher factors: Issues in literacy research, Fortieth yearbook of the National Reading Conference* (pp. 185–192). Chicago, IL: National Reading Conference.

van Zeeland, H., & Schmitt, N. (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, *34*, 457–479. doi:10.1093/applin/ams074

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*, 56–89. doi:10.1093/applin/24.1.56

Wallgren-Hemlin, B. (1997). *Att övertyga från predikstolen* [Persuading from the pulpit]. Nora, Sweden: Bokförlaget Nya Doxa.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader. *Reading in a Foreign Language*, *15*, 130–163.

Waznak, R. (1998). *An introduction to the homily*. Collegeville, PA: The Liturgical Press.

Webb, S., & Rodgers, M. P. (2009). The lexical coverage of movies. *Applied Linguistics*, *30*, 407–427. doi:10.1093/applin/amp010