Context Effects in Spoken Word Recognition of English and German by Native and Non-native Listeners

by

Robert Albert Felty

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Linguistics & Germanic Languages & Literatures)
in The University of Michigan
2007

Doctoral Committee:
- Professor Patrice Speeter Beddor, Co-Chair
- Assistant Professor José R. Benkí, Co-Chair, Michigan State University
- Professor Robert L. Kyes
- Assistant Professor Andries W. Coetzee

If we knew what we were doing, it wouldn't be called research

— Albert Einstein

To my loving wife, Clare Jessamyn Dibble

Acknowledgments

I would like to thank the members of my dissertation committee for all of their help and support during my time at U-M:

- José Benkí for giving me the opportunity to work in his lab for two years, without which experience I most likely would not have come upon this topic. I am also grateful for his constant encouragement.
- Pam Beddor, for all her phonetic insights, administrative and professional knowledge, and for her attention to detail, which have greatly improved my dissertation.
- Andries Coetzee for his constant encouragement and enthusiasm for my research.
- Bob Kyes, for his teaching and mentoring early in my graduate career (I took a total of five classes with Bob), and also for his encouragement.

Two of the four experiments in my dissertation were carried out at the University of Konstanz in Germany. I would like to thank all the members of the Linguistics department there, who showed me great hospitality while I was there collecting data in the Spring of 2006. In particular I would like to thank:

- Prof. Dr. Aditi Lahiri
- Dr. Mathias Scharinger
- Verena Felder
- Frank Zimmerer

I would also like to thank all the members of the Phonetics-Phonology Discussion Group past and present, who gave me innumerable insights during several of the presentations of my dissertation work in progress. In addition, I feel very fortunate to have been a member of both the German and Linguistics departments, both of whose faculty, staff, and graduate students have made my years studying at the University of Michigan an enjoyable experience.

In addition to support from colleagues, there are also many friends and relatives who have greatly helped me throughout my education:

- Fran Felty, my mother, for being the best mom I have ever had, for forcing me to do things that I didn't want to at the time, but am now happy that I did, for allowing me to be my own person, for always being willing to help me when I needed it, and for making so many wonderful pies to fulfill my sweet tooth.
- Harold Felty, my father, for being the best dad I have ever had, for taking the time to be involved in my life, for helping me practice my German, for teaching me about insurance, finances, and linguistics, and for his meticulous proofreading of this and other papers I have written.

- Sean Gallagher, my longest friend, who has stood by me in good times and bad alike. He always helps keep me grounded when my thoughts start to stray too far into the clouds.
- Brian Grekowicz, who first encouraged me to think critically and question many of the assumptions which many hold as fact. He has always challenged me to be a better person.
- Luis Sanchez, who has always supported me in both academic and extracurricular activities, and who has always provided me with some healthy competition.
- Danny Reeves, who (except for my wife) is probably the closest friend I made while in Ann Arbor. Not only did he expose me to new thoughts on politics and feminism, he also introduced me to new sports such as inline-skating and stair-climbing. I will always remember hiking with him and his family in Yosemite. Last but not least, he gave me the LaTeX class file for the dissertation (though I ended up writing my own improved version), and immeasurable help learning the wonderful program.
- Alex C. de Baca, my friend and best roommate ever. Alex was a great source of inspiration, as well as an incredible resource for computer programming help. Without him, I would not know nearly as much as I do about programming, which has greatly improved the analysis of my dissertation. We also had some great times listening to Afro-Cuban jazz and debating politics.
- Grinnellians — those who I knew during my years at Grinnell College, as well as alumni I have met here in Ann Arbor. My experience at Grinnell was very special, and I attribute that in large part to the sort of people that Grinnell attracts.
- Clare, my wife, who basically just makes me happy. We both went through stages in the dissertation process where we simply wanted to quit, and the other one was there for encouragement. It is hard to imagine doing this without her support.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

Abstract

Spoken word recognition involves integrating acoustic/auditory information extracted from the signal with linguistic knowledge, including sentential and discourse context, as well as the frequency of the words in the signal, and the similarity of target words to other words in the mental lexicon. Recent research on visual word recognition has shown that morphology may also affect lexical access, and that the effects of morphology on lexical access may be language-specific. This study investigates the effect of morphology on spoken word recognition using two languages which share many phonological characteristics but differ in key aspects of morphological structure.

Four separate experiments investigated open-set spoken word recognition in noise using English and German disyllabic words and nonwords, testing both native and non-native listeners of each language. Results from native listeners showed facilitatory effects of lexical status and lexical frequency, as well as inhibitory effects of neighborhood density, consistent with previous studies using English CVC stimuli. In addition, the results showed a processing advantage for monomorphemic words over bimorphemic words, indicating that morphology also has an influence on spoken word recognition. The processing advantage of monomorphemes was greater for native listeners of German than of English, which is taken as evidence that the morphological structure of the language plays a key role in the influence of morphology on spoken word recognition. Results from non-native listener experiments were largely consistent with the native listener results, suggesting that non-native listeners are sensitive to the same context effects as native listeners, although the size of the context effects were generally somewhat smaller for non-native listeners, suggesting that the amount of exposure to a language can also affect processing.

No current models of spoken word recognition can account for all of the effects found in this study. Full storage models cannot account for effects of morphology, while morphological decomposition models cannot account for neighborhood density effects. Therefore, a revised version of the Neighborhood Activation Model (Luce & Pisoni, 1998) of spoken word recognition is proposed which posits that words are stored whole in the lexicon, and that in addition to orthographic, phonological, semantic, and frequency information, lexical entries also contain morphological information.

Chapter 1

Introduction

THE abundant research on lexical access in the last 30 years has greatly improved our understanding of spoken word recognition. Spoken word recognition is a complex process of integrating acoustic, lexical, and grammatical information. Unlike hearing a completely foreign language, in which the only information available to the listener is the acoustic signal, listeners perceiving utterances in a known language have a wealth of additional information stored in long term memory to aid them in spoken word recognition. A common way to view this process is as a matching process, whereby listeners match acoustic/auditory information with words they already know. Words that closely match the acoustic information are activated in the brain, and if the activation reaches a certain threshold, then a decision is made. Both the speed and accuracy of this process have been shown to be affected by several context effects, including:

1. lexical frequency — high frequency words are processed more quickly and accurately than low frequency words (e.g. Broadbent, 1967; Taft, 1979)
2. neighborhood density — words that are highly similar to other words are processed more slowly and less accurately than words that have a low degree of similarity (e.g. Luce, 1986; Luce & Pisoni, 1998; Benkí, 2003a; Imai, Walley, & Flege, 2005)
3. morphology — processing of multi-morphemic words involves activation of the constituent morphemes of the word, which can create a processing disadvantage for multi-morphemic words (Taft & Forster, 1975; Laine, Vainio, & Hyönä, 1999; Lehtonen, Vorobyev, Hugdahl, Tuokkola, & Laine, 2006; but see also Andrews, 1986; McClelland & Patterson, 2002), and words can prime morphologically related words (Marslen-Wilson, 2001; Meunier & Longtin, in press)

The role of morphology in lexical access has received much attention, with particular attention to the storage and retrieval of multi-morphemic words. Proponents of associative models of lexical access hypothesize that morphologically complex words are stored whole and accessed directly (e.g. Rumelhart & McClelland, 1986; Plaut & Gonnerman, 2000; Hahn & Nakisa, 2000), while combinatorial models hypothesize that morphemes are stored separately and combined during lexical access (e.g. Pinker & Prince, 1988; Marcus, Brinkman, Clahsen, Wiese, & Pinker, 1995; Clahsen et al., 2001). Associative models thus predict that monomorphemic and bimorphemic words should be processed in the same way, while combinatorial models predict that monomorphemic words should show a processing advantage.

A growing body of research also suggests that cross-linguistic differences in morphological structure can have a profound impact on the ways in which morphology affects lexical access. Though the majority of research on lexical access has been concentrated on only a few languages (mostly English and Dutch), this trend has begun to change in recent years, with several new studies investigating lexical access in diverse languages such as Finnish (Vannest, Bertram, Järvikivi, & Niemi, 2002), Arabic (Boudelaa & Marslen-Wilson, 2000), Polish (Reid, 2001), and Chinese (Zhou & Marslen-Wilson, 1995, 2000). These studies have shown that languages with rich morphology tend to exhibit more morphological effects on lexical access than languages with relatively sparse morphology, which points to the importance of further cross-linguistic research for adequate formation and testing of models of lexical access.

While many researchers agree that words that differ in morphological structure are often processed differently, the cause of these differences is still under debate. It is possible that these processing differences are reflective of underlying differences in the structure of the mental lexicon of these speakers, or that processing differences arise simply from the distributional properties of the language during on-line processing. Assuming that the structure of the mental lexicon is to some extent language-dependent, non-native speakers might carry over some of the properties of their native mental lexicon when learning a foreign language. While several studies have shown that second language learners are sensitive to lexical frequency and neighborhood density in a second language (L2) (Bradlow & Pisoni, 1999; Imai et al., 2005), the effect of morphology on L2 word recognition has only recently been addressed (Hahne, Mueller, & Clahsen, 2006).

Another important question in lexical access is whether humans process visual and aural language in the same way. Although several studies have concluded that readers convert spelling to phonemes before lexical access (Rubenstein, Garfield, & Millikan, 1970; Rubenstein, Lewis, & Rubenstein, 1971; Pexman, Lupker, & Reggin, 2002; Sparrow & Miellet, 2002; but see also Forster & Chambers, 1973; Forster & Shen, 1996, for negative evidence, and Frost, 1998; Harm & Seidenberg, 2004, for hybrid views), this is not always the case, especially in languages with fairly ambiguous orthographies, such as English. In addition, the temporal nature of visual and spoken word recognition differ greatly, in that written words (especially high-frequency words) can be processed as wholes. In contrast, the stimulus in spoken word recognition is a continuous signal that unfolds over time. When processing words with suffixes, one might predict that suffixes could have a greater influence on lexical access in visual tasks as opposed to aural tasks, due to this temporal processing. While several studies using visual tasks have found evidence of morphological decomposition in German (Clahsen, 1999; Clahsen et al., 2001; Sonnenstuhl & Huth, 2002), it is still unclear whether these effects will also be found using auditory tasks. In addition, previous research in spoken word recognition using open response tasks has been limited to monosyllabic words (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Olsen, Tasell, & Speaks, 1997; Luce, 1986; Luce & Pisoni, 1998; Benkí, 2003a). It is not known whether context effects on spoken word recognition of multisyllabic words are the same as for monosyllabic words.

From this overview several research questions arise:

- Are monomorphemic and bimorphemic words processed in the same way, as associative models predict, or are bimorphemic words decomposed into their constituent morphemes before lexical access, as combinatorial models propose?
- What role does morphology play in spoken word recognition, and how do phonetic and morphological effects interact in lexical access?

- To what extent are context effects in lexical access dependent on the morphological structure of the language?
- Do cross-linguistic differences in the mental lexicon carry over to learning a second language?
- Do previously found effects of lexical frequency and neighborhood density in monosyllabic words extend to disyllabic words?

This study seeks to address these questions by providing new experimental results from four separate experiments investigating spoken word recognition in noise using English and German disyllabic words and nonwords, conducted with both native and non-native listeners. Although one can draw some conclusions from comparing studies using a variety of tasks and languages, it is more reliable to directly compare results from experiments differing in as few variables as possible. In order to make a direct cross-linguistic comparison of effects of morphology in lexical access, two morphologically divergent languages that share many phonological properties have been chosen to address these questions: English and German. English provides a good base, since the great majority of spoken word recognition and lexical access research has used English. German is an ideal language to compare morphological effects with English, since German has a much richer inflectional morphology than English.

The organization of the dissertation is as follows. In Chapter 2, an overview of relevant research is provided. Chapter 3 presents an explanation of the basic design of the current study. Chapters 4–7 contain the specific methods and results of each experiment. A general discussion of all four experiments and final conclusions are given in Chapter 8.

Chapter 2

Background

THE study of lexical access investigates how word recognition is affected by the mental lexicon. According to Balota & Chumbley (1984: 341) , the notion that speech processing requires access to lexical representations involves three assumptions:

> (a) Lexical access involves some matching of the features extracted from the stimulus to an internal representation of words: (b) word frequency determines the availability of lexical representations either by ordering them or by affecting their thresholds: (c) higher order semantic information for a word presented in isolation becomes available only after lexical access has taken place.

Almost all models of lexical access rely on these three assumptions. These three assumptions also situate the field of lexical access within linguistics. Research on lexical access seeks to discover how language is processed and how the mental lexicon is arranged, which can also have an impact on our general understanding of grammatical knowledge and language.

Since the inception of modern research on lexical access in the 1950's,[1] several context effects have consistently been found to influence how humans process speech. The context effects in question here are effects of context from the lexicon, as opposed to phonological, syntactic, or discourse context. Lexical context refers to the fact that words are not processed in isolation. Word recognition is typically viewed as a matching process by which an acoustic input activates words in the mental lexicon. Words which are phonologically, orthographically, morphologically, or semantically related to the input are also activated. In much the same way that sentence processing is affected by syntactic context or speech perception by phonetic context, word recognition is affected by lexical context. One of the earliest and most robust findings of research in word recognition was that lexical frequency has a strong influence on lexical access. Repeated research has shown that high-frequency words elicit quicker and more accurate responses than low-frequency words in a large variety of experimental conditions (e.g. Broadbent, 1967; Taft, 1979; Benkí, 2003a). Another factor that has been reliably shown to affect lexical access is neighborhood density. Neighborhood density is a metric of similarity, roughly defined as the degree to which a word is similar to others (in phonological or orthographical characteristics). Words, for which there are many similar words

---

[1] though it is also appropriate to note that Bagley (1900–1901) was undertaking very similar research

are said to be in dense neighborhoods, whereas words for which there are few similar words are said to be in sparse neighborhoods. In contrast to lexical frequency, which facilitates the activation of a word in the brain, neighborhood density has been found to inhibit activation in word recognition (e.g. Luce, 1986; Luce & Pisoni, 1998; Benkí, 2003a; Imai et al., 2005).

2.1 Models of lexical access

A number of different theoretical models of lexical access exist, most of which address one particular aspect of lexical access. One fundamental distinction is whether the lexicon is being accessed in comprehension or production; since the present study is investigating comprehension, only comprehension models will be addressed. A second fundamental distinction is whether the model focuses on phonological, morphological, and/or semantic levels of representation. The debate over morphological processing centers around whether words or morphemes are stored in the lexicon, and consequently, whether comprehension of morphologically complex words involves rule-based processes. While several specific models have been proposed to account for processing of morphologically complex words, these models are frequently grouped into two categories — associative and combinatorial models (e.g. Clahsen et al., 2001).

2.1.1 Associative Models

The key defining trait of associative models of lexical access is that they posit that all words are stored whole in the mental lexicon, including both monomorphemic and multimorphemic words. The assumption that words are stored whole in the lexicon leads to the prediction that, all other factors being equal, morphologically simple and complex words should be processed in the same manner. In practice, it is nearly impossible to construct an experiment in which monomorphemic and multimorphemic words differ only in morphology, and not in phonological structure, semantic similarity, lexical frequency, neighborhood density, phonotactic probability, or other linguistic traits. In fact, several researchers have convincingly argued that any processing differences between mono- and multimorphemic words are due to differences in phonological structure or semantic content, and not a result of differences in morphological structure (Ramscar, 2002; Baayen & Martin, 2005).

A variety of associative models have been proposed, the most prominent being schema-based models (e.g. Bybee, 1995, 2001) and connectionist models (e.g. Rumelhart & McClelland, 1986; McClelland & Elman, 1986; Hahn & Nakisa, 2000). Connectionist models have generally received the most attention in the lexical access literature, in part because they provide bold, quantitative predictions about how listeners process and acquire words, which are contradictory to traditional linguistic theories of morphology. Connectionist models such as TRACE (McClelland & Elman, 1986) have been very successful in accurately predicting the effects of lexical frequency on lexical access, which is one of the most consistent and wide-spread context effects on lexical access. Connectionist models have also had some success in modeling morphological effects in lexical access without explicit morphological representations (Rumelhart & McClelland, 1986; MacWhinney & Leinbach, 1991; Seidenberg & McClelland, 1989; Plunkett & Marchman, 1991; Hahn & Nakisa, 2000), though these studies were also met with considerable criticism (Pinker

& Prince, 1988; Marcus et al., 1995; Clahsen, 1999; Albright & Hayes, 2003). Many of the studies supporting associative models of lexical access have either only used monomorphemic stimuli, and concentrated on explaining effects of lexical frequency or neighborhood density, or they have explained processing differences between mono- and multimorphemic words by appealing to differences in phonology or semantics (Ramscar, 2002; Baayen & Martin, 2005). However, several studies have found processing differences between mono- and multimorphemic words even when phonological structure and semantics were highly controlled (e.g. Roelofs, 1996; Gumnior, Boelte, & Zwitserlood, 2006), which would seem to pose problems for associative models as currently implemented.

2.1.2 Combinatorial Models

In contrast to associative models, combinatorial models, also known as morphological decomposition models, hypothesize that only word stems are stored in the mental lexicon, and that affixes are either combined with stems (in word production) or stripped off of multimorphemic words (in word recognition). Combinatorial models predict a processing advantage of monomorphemic words over multimorphemic words, under the assumption that affix stripping (or combining) requires additional processing. Such processing advantages have been found in a number of experiments (e.g. Taft & Forster, 1975; Taft, 1988; Gürel, 1999), though, as noted above, some claim that such processing advantages are largely due to phonological or semantic rather than morphological differences.

Researchers working on combinatorial models realized that wholly combinatorial models cannot account for a number of phenomena, most notably irregular forms. For example, there is no rule or generalization to capture the fact that the past tense of *go* is *went*. Therefore, most researchers using combinatorial models posit a dual-route mechanism of lexical access (e.g. Clahsen, 1999; Clahsen et al., 2001; Marcus et al., 1995), a rule-based route that accounts for regular forms, and a direct-access route that accounts for irregular forms. Such dual mechanism approaches can account for effects of regular morphology as well as high-frequency irregular forms, but, as with wholly combinatorial models, they do not make specific predictions as to the influence of lexical frequency on lexical access, and make no predictions whatsoever as to the influence of neighborhood density.

2.1.3 Summary of lexical access models

This brief discussion of models of lexical access has shown that both associative and combinatorial models have had a fair amount of success in explaining effects of context on lexical access, but that each class of model fails to account for all context effects. Associative modelshave succeeded in accurately predicting effects of lexical frequency and neighborhood density on lexical access, but have not always been able to account for morphological effects. In contrast, combinatorial models have successfully predicted effects of morphology on lexical access, but have only marginally addressed effects of frequency, and have not addressed the effects of neighborhood density at all. The present study further tests the predictions of associative and combinatorial models by investigating effects of morphology, lexical frequency, and neighborhood density, while attempting to control for phonological and semantic effects. Specific models of lexical access are discussed in greater detail in §8.3, and an alternative model is proposed, which seeks to address all three of these context effects.

2.2 Cross-linguistic research in Lexical Access

While most research on lexical access has been done with English, interest in investigating issues of lexical access with other languages is increasing. Marslen-Wilson (2001), in an overview of cross-linguistic research in his laboratory, reports differences across Polish, Arabic, English, and Chinese in terms of how morphology is processed and represented in the lexicon. Results from English show that complex words such as *darkness* are represented by their constituent morphemes, and are combined during lexical access. The findings of Marslen-Wilson and colleagues also show stem-priming for English, whereby the stem in *darkness* and *darkly* primes *dark*. This is not the case for semantically opaque words such as *department*, which does not prime *depart*. Results from Polish show even more such combinatorial effects, including affix priming (e.g. *kotek/ogródek* 'a little cat' / 'a little garden'), in which the diminutive affix in the prime facilitates perception of the target), and suffix interference (e.g. *pis-anie/pis-arz* 'writing'/'writer'), in which no facilitation is found in pairs which share stems, but differ in suffixes. They also find evidence for morphological decomposition for Arabic words, which, like other Semitic languages, have a three-consonant morphological root, leading them to conclude that root priming in Arabic parallels stem priming in other languages. In contrast to English and Polish stem priming however, they do find evidence for root priming even for semantically opaque words. Chinese has virtually no inflectional or derivational morphology, and is therefore also a key language to study cross-linguistic differences in morphological processing. The only aspect of Chinese morphology which could possibly show effects of morphological decomposition is compounding, which is very productive in Mandarin Chinese, with bimorphemic compounds accounting for up to 70% of all word forms in the language. Marslen-Wilson and colleagues find no evidence for morphological decomposition in Mandarin compounds, however, much like English. The cross-linguistic differences that Marslen-Wilson find suggest that experimental evidence supporting either combinatorial or associative models of lexical access may be highly dependent on the language studied.

Vannest et al. (2002) also find similarly various results in a comparison of English and Finnish derivational morphology. Since previous research on Finnish inflectional morphology has shown support for combinatorial-like processing (e.g. Laine et al., 1999), Vannest et al. (2002) hypothesize that Finnish will show more evidence for morphological decomposition with derivational morphology than for English. However, they find exactly the opposite result, which they account for in terms of the lexical-statistical properties of the two languages. Whereas most derivational affixes in English combine with monomorphemes, most words with derivational affixes in Finnish also contain inflectional affixes. They hypothesize that words with derivational affixes are stored separately in Finnish in order to decrease the amount of morphological processing that the Finnish speaker must compute.

Especially relevant to the current investigation, several studies on lexical access in German have been published. As mentioned in §2.1, Marcus et al. (1995) reported evidence for a default plural rule in German using nonword rating tasks. Clahsen (1999), in a review of collaborative research on morphological effects on lexical access, summarizes the German evidence as showing:

Table 2.1 Morphological markedness (from Clahsen et al., 2001)

-e	-s	-m
	[-PL]	[-PL]
-	[-FEM]	[-FEM]
-	[-MASC]	-
[-OBL]	[-OBL]	[+OBL]
-	-	[+DAT]

Table 2.2 Example of materials from Clahsen et al. (2001). Numbers given are raw frequency counts from the CELEX database Baayen & Rijn (1993). Stem gives the lemma frequency for each word, and -m and -s list the wordform frequency with the given suffiix.

-m dominant adjectives				-s dominant adjectives			
	Stem	-m	-s		Stem	-m	-s
ruhig	838	51	13	rein	783	14	38

(1) frequency effects for irregular verbs, but not regular verbs in lexical decision; (2) full priming of regularly inflected verbs, but only partial priming of irregularly inflected verbs; and (3) differences in brain response (ERP) to regularization of noun plurals as opposed to irregularization of noun plurals, and likewise for the past participles of verbs.

Differences in response latencies to inflected adjectives in lexical decision and cross-modal priming tasks have also been reported by Clahsen et al. (2001). They propose that of the five possible adjective endings in German (-r, -n, -m, -s, -e — see also Table 3.1), some endings are more marked than others (based on proposals by Bierwisch (1967); Zwicky (1986); Blevins (1995, 2000); Wunderlich (1997)), and thus should show a difference in processing time. Under their model, the endings -e, -s, and -m have the representations shown in Table 2.1.

Clahsen et al. (2001) argue that -m is the most marked, because it is positively specified for dative and oblique, whereas -e and -s are negatively specified. They hypothesize that adjectives inflected with -m will take longer to process than those with -s or -e. To test this, they selected 81 stimuli from the CELEX database (Baayen & Rijn, 1993) that were matched for the lexical frequency of the lemma (dictionary entry) but differed in frequency of the word form. One example is shown in Table 2.2. Clahsen et al. (2001) used this set of materials in a lexical decision task (LDT) and a cross-modal priming task.

The associative and combinatorial models make different predictions on the speed of processing of these materials. A combinatorial approach predicts that adjectives inflected with -m will be processed more slowly, because they are more marked. This is in contrast to the associative model, which predicts that the word forms with lower frequency would be processed slower (*ruhiges* and *reinem*). The results from both the LDT and the cross-modal priming task found longer reaction times to adjectives with *-m* than with *-s*, in support of the combinatorial model. However, as they note on page 518, the ending *-s* occurs approximately twice as often as *-m* overall. Thus, the finding could be due to overall frequency of the endings rather than morphological markedness. Nevertheless, even if the results are due to frequency and not morphological markedness, the results still show that German speakers are sensitive to these morphological differences.

Cross-linguistic differences in lexical access also extend beyond morphology. Current research by Benkí, Myers, & Nearey (in preparation) on Taiwanese Mandarin has found no effect of lexical status (words vs. nonwords), which has been one of the most robust and consistent findings in the research on lexical access. They posit that this could be due to the very restricted syllable structure of Mandarin.

Such results emphasize the need for lexical access research on a variety of languages, in order to determine what kinds of effects are language-specific, and which effects may be more general. In particular, cross-linguistic effects of morphology in spoken word recognition (as opposed to visual word recognition) have yet to be examined. A controlled study of cross-linguistic morphological effects should compare languages which share many phonological properties, yet morphologically diverse, as does the present study.

2.3 Lexical Access by non-native speakers

While cross-linguistic perception and second language (L2) perception have been studied for quite some time now (see Strange, 1995 for an overview), researchers have only recently begun to investigate lexical access and spoken word recognition in bilinguals and L2 speakers. One of the fundamental concepts in second language acquisition (SLA) is language transfer. The method of contrastive analysis (Lado, 1957) aimed to predict which grammatical features are difficult for learners to acquire by comparing the grammars of the L1 and L2 in question. While later research showed that contrastive analysis cannot account for several important results from SLA research (Corder, 1967), language transfer continues to be a relatively good predictor of selected aspects of L2 acquisition, particularly in the domains of phonetics and phonology. While many studies have shown that non-native listeners have difficulty discriminating between phones that are not contrastive in their native language (e.g. Best, 1995; Beddor, Harnsberger, & Lindemann, 2002; Flege, 1993), few studies have examined effects of lexical context on non-native word recognition. Crucial questions for research on lexical access by non-native speakers include: (1) Do non-native speakers access both L1 and L2 simultaneously?, in which case language transfer effects would be found; and (2) Is the structure of the L2 lexicon the same as the L1 lexicon, i.e. do context effects such as lexical frequency, neighborhood density, and morphology affect the responses of non-native speakers in the same way as those of native speakers?

One issue in L2 word recognition is the effects of vocabulary size on lexical competition. It is widely assumed that non-native speakers have a smaller vocabulary than native speakers. This difference could lead to reduced effects of lexical competition. Weber & Cutler's (2004) test of this hypothesis using several eye-tracking experiments with English and Dutch listeners led them to conclude that non-native listeners have additional sources of lexical competition compared to native listeners, including: (1) competition from other L2 words that would not be competitors for L1 speakers, and (2) competition from the L1. The first conclusion comes from the result that words such as *ballot box* and *belly button* were not disambiguated by the Dutch listeners until after the /l/, presumably because /æ/ and /ɛ/ are often confused by Dutch listeners. Weber & Cutler (2004) also found that L1 words such as *kist* /kɪst/ 'chest' were in competition with L2 words such as *kitten*, though they did not find L2 words being activated using the same task with L1 stimuli. This is in contrast to the study of Marian & Spivey (1999), who found competition in both directions using Russian and English. Weber and Cutler attributed this difference to the fact that Marian and Spivey's participants were living in the L2 environment, while Weber and Cutler's participants were living in the L1 environment. While Weber & Cutler's (2004) results do show that non-native listeners have additional sources of lexical competition, their results do not address the question of whether global effects of lexical competition such as neighborhood density differ between native and non-native listeners.

Recent work on L2 lexical access in German has shown that non-native speakers are affected by many of the same lexical and grammatical properties of German as L1 speakers. Hahne et al. (2006) applied the techniques of Marcus et al. (1995) and Clahsen (1999) to learners of German. They performed two nonword production tasks (similar to Marcus et al., 1995) as well as ERP experiments to investigate differences in irregular and regular noun plurals and past participles of verbs. They found that the non-native speakers (L1 = Russian) behaved almost identically to the native speakers in producing past participles, rating regularizations as more natural than irregularizations. The L2 speakers also patterned similarly to the L1 speakers in rating noun plurals, though the difference in ratings between regular and irregular plurals was not as great as for L1 speakers. The results from the ERP experiments were similar, with online processing of past participles more similar to L1 speakers than the processing of noun plurals. They suggest that L2 learners acquire the German noun plural system later than the verbal system, since the plural system is more complicated.

Though they did not directly investigate lexical access, a recent study by Cutler, Weber, Smits, & Cooper (2004) showed that Dutch listeners performed slightly worse than English listeners in a speech-in-noise test of English CV and VC syllables at all signal-to-noise ratios (S/N). The lack of interaction between S/N and language background is in contrast to earlier studies that have suggested that the gap between first language (L1) and L2 performance may increase with the amount of noise present. Cutler et al. (2004) interpreted the results to mean that there is a greater phonetic processing load in general for the L2 speakers. If their interpretation is correct, this could also have implications for lexical access in L2 speakers, especially in auditory tasks such as the one used in the present study. If non-native listeners' overall perceptual accuracy is lower than native listeners, then non-native listeners are forced to rely more on lexical information to fill in the missing acoustic information in a spoken word recognition task.

Two studies to date have investigated such an interaction between lexical access and phonetics. Bradlow & Pisoni (1999) investigated talker- and item-related effects of spoken word recognition in noise with native and non-native listeners. Previous studies have shown that listeners are sensitive to talker-specific information, specifically that listeners are better at perceiving familiar talkers than unfamiliar talkers (Mullennix, Pisoni, & Martin, 1989; Bradlow & Pisoni, 1994; Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998; Goldinger, 2003). Bradlow & Pisoni (1999) found that these effects are largely the same for native and non-native listeners. However, they found that lexical effects differ between L1 and L2 listeners. Their materials included "easy" and "hard" words — the "easy" words had high lexical frequency and sparse neighborhoods, whereas the "hard" words had low lexical frequency and were in dense neighborhoods. Native listeners showed a small (4.3%) difference in recognition rate between easy and hard words, the non-native listeners exhibited a much larger (25.2%) difference. They interpret the results as evidence that loss of fine-grained phonetic detail (due to the noise in the stimuli) affects lexical access of non-native listeners more than native listeners.

One drawback of the Bradlow & Pisoni (1999) study is that neighborhood density and lexical frequency covaried, making it impossible to determine whether the differences between the “easy” and “hard” words were due to frequency, density, or some combination thereof. Imai et al. (2005) addressed this shortcoming by comparing spoken word recognition scores from three groups of listeners who heard native-accented and Spanish-accented English. The listeners consisted of native English speakers, and two groups of native Spanish speakers, separated into low- and high-proficiency groups. Listeners heard English words (mixed with multi-talker babble) which differed according to lexical frequency and neighborhood density in a 2x2 design. The L1 listeners scored consistently higher on the native-accented speech than the L2 listeners, while the L2 listeners scored better on the Spanish-accented for words in dense neighborhoods, but not sparse neighborhoods. No significant effect of word frequency was found, but a significant effect of word familiarity was found, which also interacted with neighborhood density and accent, in that neighborhood density caused a large effect for high familiarity Spanish-accented stimuli, but no effect for low familiarity Spanish-accented stimuli. Imai et al.’s results suggest that low-level phonetic differences can affect more global effects of lexical access, and that this effect also depends upon the proficiency level of the L2 listener.

In addition to previous research on L2 lexical access, there is a growing body of research on lexical access in bilinguals which is also relevant. The major focus in this line of research has been to address the question of whether bilinguals have one lexicon containing information from multiple languages, or separate lexicons for each language they know. This question can also be thought of as a difference between simultaneous activation of both languages versus activation of only one language. Inhibitory effects are generally seen as evidence of simultaneous activation. In addition to questions of simultaneous activation, the bilingual literature has also addressed other effects such as lexical frequency and lexical competition. Pallier, Colome, & Sebastian-Galles (2001) used a repetition priming task, in which participants make auditory lexical decisions on a list of items, some of which are repeated; response times to repeated words are generally lower than the first presentation of the word. They tested Spanish-Catalan bilinguals (half of whom where Spanish-dominant, and half of whom were Catalan-dominant) on Catalan words which included minimal pairs that had a phonemic contrast shared by both languages (e.g. /p b/) as well as pairs which only contrasted in Catalan (e.g. /e ɛ/). The Spanish-dominant bilinguals exhibited a repetition effect for the words with Catalan specific contrasts, while the Catalan-dominant bilinguals did not. In other words, the Spanish-dominant listeners had interference from their Spanish phonology, similar to the effect that Weber & Cutler (2004) reported for Dutch speakers listening to English.

2.4 Summary

This brief discussion of research on lexical access and spoken word recognition has identified some of the key findings in previous research and also highlighted important gaps in the literature, some of which this study addresses. In particular, the theoretical predictions of associative and combinatorial models should be further tested using an auditory task with a cross-linguistic design, in order to investigate the influence of stimulus presentation and language structure on morphological processing. In addition, models of spoken word recognition which have only been verified using monosyllabic stimuli need to be tested using multisyllabic stimuli. Finally, lexical access research with non-native speakers can provide additional information about the structure of the lexicon. The present study will address all of these issues.

Chapter 3

Experimental Design

This chapter presents an overview of the design of the four experiments used in this study. A brief explanation of previous related experimental procedures is given, followed by a summary of the tasks used in the present experiments. The method of analysis and predictions for all four experiments are also given.

3.1 Experimental procedures

Research in lexical access has used a variety of different experimental apparati to investigate how the lexicon is accessed when processing speech. All experiments in lexical access can be said to have following four components, which can be combined in a number of ways.

1. stimulus presentation method
    (a) aural presentation
    (b) visual presentation
2. measurement method
    (a) behavioral measures, e.g. reaction time and accuracy
    (b) neurological measures, e.g. electro-encephalography (EEG) and magnetic resonance imaging (MRI)
3. task (only a partial list)
    (a) lexical decision, in which the participant is asked to respond whether the target word is a real word or a nonword
    (b) naming, in which the participant is asked to speak aloud the target word as quickly as possible
    (c) rating tasks, in which the participant is asked to rate the target word along a particular dimension, e.g. how familiar the word is
    (d) open response word recognition, in which participants hear a stimulus and are asked to record (orthographically, or auditorily) their response.
4. priming
    (a) no priming

(b) form priming, in which a phonologically related word is presented shortly before the target word

(c) semantic priming, in which a semantically related word is presented shortly before the target word

All of the possible options for each component have their own advantages and disadvantages, depending upon what exactly is being studied. One of the main goals of the current study is to test whether effects of morphology found in studies using visual tasks also applies to aural tasks. Though there is a fair amount of evidence that readers convert spelling to phonemes before lexical access (Rubenstein et al., 1970, 1971; Pexman et al., 2002; Sparrow & Miellet, 2002; but see also Forster & Chambers, 1973; Forster & Shen, 1996, for negative evidence, and Frost, 1998; Harm & Seidenberg, 2004, for hybrid views), this is not always the case, especially in languages with fairly ambiguous orthographies, such as English. In addition, the temporal nature of visual and spoken word recognition differ greatly, in that written words (especially high-frequency words), can be processed as wholes — that is, all of the letters of a word can be seen simultaneously. In spoken word recognition, the stimulus is a continuous signal which unfolds over time. Given an auditory stimulus which is revealed over time, it is possible that more weight may be given to the beginning of words than to the end, which has been suggested by Marslen-Wilson & Zwitserlood (1989). When processing words with suffixes, one might predict that suffixes could have a greater influence on lexical access in visual tasks as opposed to aural tasks.

Several different types of tasks can be used in spoken word recognition. One of the most frequently used tasks is the lexical decision task (LDT), in which participants are asked to decide whether a stimulus is a word or not. Some have criticized use of the LDT, in that it over-emphasizes frequency effects, and that the cognitive demands it places on participants are quite different from the demands of other tasks (Balota & Chumbley, 1984). In fact, recent studies by Vitevitch (2006) (using LDT) and Altieri (2006) (using a naming task), have found opposite effects of clustering coefficient (basically the number of neighbors which are neighbors of each other). Another disadvantage of LDT and other measures based on response time is that no information about the activation of competing words is given. That is, when listeners make errors, what types of errors do they make? Open-response tasks allow researchers to investigate the types of errors that listeners make. For the present study, an error analysis provides insight into whether listeners' misperceptions are morphologically and/or phonologically related to the target word, and what role frequency has on the types of errors that listeners make.

Most listeners' performance reaches near 100% accuracy in an ordinary open-response spoken word recognition paradigm, which does not reveal much about the types of misperceptions they make. To avoid these ceiling effects, the difficulty of open-response tasks must be increased in some way. One of the most common ways to do this is to degrade the acoustic signal, either through filtering, additive noise, or reducing the signal strength (i.e. reducing the volume of the signal). For the present study, signal-dependent noise was chosen as the method of signal degradation (Schroeder, 1968). This method has several advantages over other methods of signal degradation. Unlike filtering or additive broadband noise, signal-dependent noise can be added to the stimuli during the experiment, which is a practical advantage. More importantly however, the signal-to-noise ratio for signal dependent noise is calculated on a sample per sample basis, with the result that all parts of the signal are masked equally, as opposed to broadband noise, in which quieter segments (such as consonants) are masked more than louder segments (such as vowels). Filtering and additive broadband noise rely on average amplitude of the signal, and thus mask quieter segments more than louder segments. Signal-dependent noise also has the advantage over reducing signal strength, in that it does not require measuring the hearing threshold of listeners prior to the experiment.

Open-response paradigms also have the advantage that phonetic, as well as phonological, neighborhood effects can be investigated. Most research on neighborhood density effects look at phonological neighborhood. That is, the phonemes of a particular word are compared with the phonemes of all the other words in a database, and words differing in only one phoneme are counted as neighbors (see e.g. Newman, Sawusch, & Luce, 1997). This makes the assumption that *cat* and *pat* are as likely to be confused with one another as *cat* and *mat*. This assumption is not valid though, as Miller & Niceley (1955) clearly showed that confusion among consonants is systematic and involves only limited errors. Using an open-response paradigm, the probability of confusing any phoneme with another phoneme can be used to calculate a measure of phonetic neighborhood density (for further discussion see Luce & Pisoni, 1998).

Although tasks incorporating priming can be used to investigate effects of morphological similarity on lexical access (e.g. Marslen-Wilson, 2001), it is unclear whether effects of priming are reflective of the way in which words are stored in the mental lexicon, or whether they are reflective of on-line processing. For example, both schema models (Bybee, 1995, 2001) and connectionist models (Rumelhart & McClelland, 1986; McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000) could predict that morphologically related words could prime one another, since such words are usually also phonologically and semantically related; however, these models would not predict that bimorphemic and monomorphemic words presented in isolation would be treated differently. Since the present study wishes to address the different predictions of these models, effects of priming were not included.

3.2 Basic Design

The present study seeks to investigate the role of context effects in spoken word recognition, in particular the role of morphology. Context effects such as lexical status, lexical frequency, and neighborhood density have been shown to play a role in spoken word recognition (Luce, 1986; Luce & Pisoni, 1998; Benkí, 2003a), but the role of morphology in spoken word recognition has not been widely investigated, and to my knowledge, no studies have been undertaken exploring the effects of morphology in open response spoken word recognition. This study attempts to bridge that gap by adding morphology to the list of context effects to be studied in spoken word recognition research. The design of the present study is largely inspired by Clahsen et al. (2001), but differs in several key ways. The major difference is the type of task. Clahsen et al. (2001) used a lexical decision task and a cross-modal priming task, both of which required the participants to make a lexical decision, which is known to be more sensitive (perhaps over-sensitive) to frequency effects than many other tasks (Balota & Chumbley, 1984). In addition, their materials were not balanced for word length, as is clear from the example in Table 2.2, which has been shown to be a relevant factor in lexical access (Frisch, Large, & Pisoni, 2000). Finally, all of the target words in both experiments used by Clahsen et al. (2001) were presented visually. There are therefore several reasons to question if similar results will be found using an auditory-based task. The word materials in the present study include adjectives, nouns, and verbs, and all have the same syllabic structure (CVCCVC), so chosen because it is a fairly common syllable structure for both English and German words, and allows for the inclusion of bimorphemic and monomorphemic words (and nonwords). The task for the current study is a speech-in-noise task, which allows one to address some questions that other methods cannot. By looking at confusions of both words and nonwords, it should be clear to what degree the perception is being influenced by acoustics and also lexical factors.

English and German are well-suited for investigating cross-linguistic influences of morphology in spoken word recognition, since they are phonologically quite similar, yet morphologically quite different. Both English and German are Germanic languages, with similarly-sized phonological inventories, including a high degree of overlapping phonemes. English has 23 syllable initial consonants, 21 syllable final consonants (counting affricates as phonemes), and 15 stressed vowels (including diphthongs). German has 21 syllable initial consonants, 14 syllable final consonants (counting affricates as phonemes), and 18 stressed vowels (including diphthongs) (International Phonetic Association, 1999). German and English also have similar phonotactics: both languages allow consonant clusters in syllable onsets and codas, though German does exhibit final obstruent devoicing, resulting in a lower number of syllable-final consonants.

Though phonologically similar, English and German are quite different morphologically. Modern English has lost most of the inflectional morphology that Old English had, and is now restricted to five inflectional suffixes, *-s* (plural), *-s* (possessive), *-s* (third person singular), *-ed* (past tense of regular verbs), and *-ing* (progressive aspect of verbs). In contrast, German has a fairly rich morphology, including an adjective inflection system that indicates case, gender, and number. It is a synthetic system (i.e. one ending encodes all three morphological categories, as opposed to agglutinative languages such as Turkish), yet not all forms are distinct. That is, some endings are homophonous. In addition, German has so-called ‘strong’ and ‘weak’ endings – the strong endings are used for adjectives that do not follow a determiner or demonstrative; the weak endings are used with strongly inflected determiners. The German adjective inflection is displayed in Table 3.1.

Table 3.1 German adjective declension

	singular			plural
	masc.	fem.	neut.	
		strong declension		
nom.	r	e	s	e
acc.	n	e	s	e
dat.	m	r	m	n
gen.	s	r	s	r
		weak declension		
nom.	e	e	e	n
acc.	n	e	e	n
dat.	n	n	n	n
gen.	n	n	n	n

The inflectional system of German has some unique properties which make it an ideal language to study the interactions of morphology, lexical access, and phonetics. As can be seen from Table 3.1, some of the adjective endings occur much more often in the paradigm than others, with -n occurring most often. In addition, the endings contain some phonemes which are more confusable than others — /m/ and /n/ are known to be highly confusable, especially in syllable final position (e.g.

Benkí, 2003b), whereas /s/ is much more salient. While previous studies have investigated the interaction of morphology with frequency effects, to my knowledge no study has investigated the interaction of morphology and phonetics. One possible prediction is that the /m n/ pair is perceptually more distinct in bimorphemes than monomorphemes due to a greater functional load. On the other hand, the opposite result (that the /m n/ pair is more distinct in monomorphemes) could be due to a difference in uniqueness points between the monomorphemic and bimorphemic words, or due to semantic factors.

Four separate experiments were carried out. As mentioned in Chapter 1, no other studies have used an open response spoken word recognition task with disyllabic words. Therefore Experiment One uses English CVCCVC words and nonwords with native speakers of English as listeners. These results are used as a baseline to determine the size of the various context effects, and as an estimate of the sample size required for further experiments. Experiment Two consists of German words and nonwords presented to German-speaking listeners. This experiment explores the first two research goals, general vs. language-specific results, and associative vs. combinatorial models of lexical access. The third and fourth experiments use the same stimuli and experimental design as the first two, except with non-native listeners. In Experiment Three, native speakers of English with intermediate fluency of German heard the same German stimuli presented in Experiment Two; in Experiment Four, native speakers of German with intermediate fluency of English heard the same English stimuli presented in Experiment One. The non-native listener experiments further address the structure of the lexicon, testing whether any possible cross-linguistic differences in the structure of the lexicon are carried over to the processing of non-native languages.

3.3 Analysis

3.3.1 The *j*-factor model

The primary method of analysis in this study is the j-factor model of Boothroyd & Nittrouer (1988). The j-factor model allows for a more detailed analysis than a traditional analysis based on percent correct. Although it is possible to discern differences between listeners' responses to words and nonwords using a percent correct method, there are several shortcomings with this method. One shortcoming is that percent correct differences vary with different signal-to-noise ratios. As Boothroyd & Nittrouer (1988: 102) note, one cannot assume that a difference between 70% for words and 50% for nonwords is equivalent to the difference between 40% and 20%. This is exacerbated as one approaches either 0 or 100%. Furthermore, if one wishes to measure the context effects in spoken word recognition one needs a measure which is reliable independent of context. The j-factor model is one such measure. The j-factor model provides a measure of the number of independent units in a stimulus. The units under investigation in this study are phonemes, but it is also possible to carry out a j-factor analysis using other units such as syllables or features. The probability of correctly identifying a given word (or nonword) can be calculated as the product of the probabilities of its constituent phonemes.

$$p_w = p_{C1} p_{V1} p_{C2} p_{C3} p_{V2} p_{C4} \tag{3.1}$$

where p_w is the probability of correctly identifying a word (or nonword). Assuming that the constituent phonemes in a stimulus are perceived statistically independently of each other, (3.1) can be rewritten as:

$$p_w = p_p^n \tag{3.2}$$

where n is the number of phonemes, and p_p is the geometric mean of the recognition probabilities of each constituent phoneme. Following Fletcher (1953), Boothroyd & Nittrouer allow for violation of the assumption that phonemes are perceived independently of one another, by positing that

$$p_w = p_p^j \tag{3.3}$$

where $1 \leq j \leq n$.[1] Rewriting 3.3, the quantity j can be empirically determined from confusion matrices by:

$$j = \frac{log(p_w)}{log(p_p)} \tag{3.4}$$

A value of $j = n$ implies that phonemes are perceived independently of one another, while a value of $j = 1$ implies that correct recognition of one phoneme is sufficient to correctly recognize the whole stimulus.

Although the j-factor model assumes that the phoneme is the basic unit of speech perception, this remains an empirical question. Other models of speech perception have proposed different basic units, ranging from features (Stevens, 1989; Stevens & Blumstein, 1981) to whole words as in some exemplar-based models (Johnson, 1997). If the phoneme assumption holds true, then the j-score for an n-phoneme word should be equal to n. This is what has been found for nonwords in several studies, using different speakers, listeners, materials, and types of masking (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Olsen et al., 1997; Benkí, 2003a). However, for CVC words, all of these studies found j-scores of approximately 2.5, indicating an effect of lexical status. (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Benkí, 2003a; Olsen et al., 1997). Nearey (2001, 2004, in press) has interpreted these results as due to response bias. Various different explanations have been given for effects of the lexicon in spoken word recognition. Ganong (1980) interpreted a tendency for listeners to select the word rather than the nonword in a *dask — task* continuum as a boundary shift, implying that the underlying perceptual mechanisms were altered. Later research has shown that these effects can also be accounted for by response bias. In other words, the effect that Ganong (1980) found should not be attributed to psychoacoustic processes, but rather to lexical access processes. In summary, for CVC syllables, a j-score of 2.5 for words can be interpreted as a bias towards words, while a j-score of 1 would imply that words are being perceived as wholes.

In order to better illustrate the j-factor model, several hypothetical and actual examples are discussed next. One key aspect of the j-factor model is that it relies on averaged results. A j-factor analysis (like most other analytic techniques) cannot be performed on a single trial. In order to perform a j-factor analysis, the results must be averaged either over subjects or over items. Averaging over items provides more reliable results, since this also means that results are averaged across phonemes as well, and therefore the influence from individual items is diminished.

Consider an experiment involving 5 listeners and 2000 stimuli, each 6 phonemes long. Half

---

[1] Actually, contrary to what Boothroyd & Nittrouer propose, it is possible to find $j > n$. Examples of situations where $j > n$ are given later in this section.

	Listener				
	1	2	3	4	5
	Nonwords				
p_p	.610	.701	.802	.900	.949
p_w	.051	.117	.261	.530	.738
	Words				
p_p	.1	.3	.5	.7	.9
p_w	.1	.3	.5	.7	.9

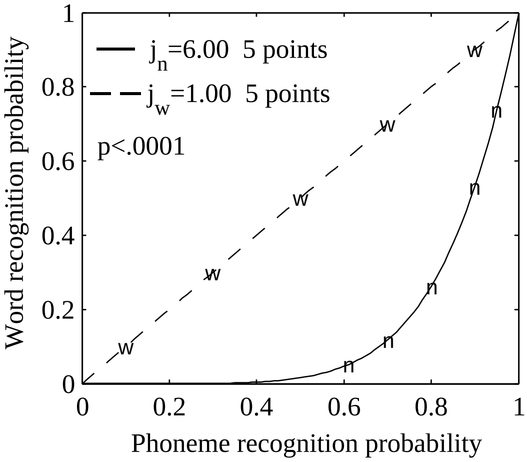


Figure 3.1 Hypothetical j-factor results. The data show hypothetical results for five subjects who heard 1000 words and 1000 nonwords. The plot shows curves representing $p_w = p_p^j$ for the mean j of each group

of the stimuli are nonwords, and half are very frequent words. Each listener hears the stimuli at different masking levels (the masking could be any sort of noise, a filter, or played at a very low volume). Hypothetical extreme results are presented for this experiment in Figure 3.1. Even though the overall levels of performance as measured by p_p and p_w span a large range, the j-score is nearly identical for each listener. The difference between the performance on the words and nonwords is due to the fact that p_p is the geometric mean of the average recognition probability of each of the six phonemes. That is, there are many ways to arrive at a given p_p. A p_p score of .5 could be the result of correctly perceiving all 6 phonemes for half of the words, and perceiving no phonemes correctly for the other half (this results in a j-score of 1, as illustrated by listener 3 in the word condition). This is the all or nothing case. The other extreme is that the listener correctly perceives most of the phonemes of each individual stimulus, but regularly misses one or two, resulting in a relatively high p_p, but a relatively low p_w. This is the case for listener 3 in the nonword condition.

One frequently asked question about the j-factor model concerns the possible j-factor values, especially how to interpret $j > n$. The previous example showed that the minimum possible j-score is 1, since p_w can never exceed p_p. There is no theoretical upper bound for the j-score, as shown in Equation 3.5, but in practice it is uncommon for a subjects analysis to return a j-score higher than n, and in fact such a result is somewhat difficult to interpret. In an items analysis, there are specific instances in which a j-score higher than n is interpretable.

$$\frac{\lim_{p_w \to 0} log(p_w)}{\lim_{p_p \to 1} log(p_p)} = \infty \tag{3.5}$$

Consider a hypothetical example of an items analysis (averaging over subjects) from an experiment containing the English words *hot* and *hut*. The raw CELEX frequency of *hot* is 2498 (log frequency per million = 2.14) and *hut* has a frequency of 396 (log frequency per million = 1.34). Hypothetical spoken word recognition results for *hot* and *hut* from 100 listeners are shown in Table 3.2. In this example, there is a clear bias against *hut*. In spite of being able to perceive the initial and final consonant of *hut* and *hot* equally well, listeners respond with *hot* more often than *hut* due to its higher frequency.

Table 3.2 Items analysis of j-factor results for hypothetical example

	p_{C1}	p_V	p_{C2}	p_p	p_w	j
hot	.9	.9	.9	.90	.8	2.12
hut	.9	.2	.9	.54	.1	3.74

Table 3.3 Items analysis of j-factor results for real example. The errors listed here are type errors, not token errors. That is, some of the errors were given as responses by more than one participant.

Item	freq	dens	p_p	p_w	j	p_{C1}	p_{V1}	p_{C2}	p_{C3}	p_{V2}	p_{C4}	errors
hosted	1	1.11	.71	.1	6.74	.13	.97	1	1	1	1	posted, coasted, hasted, toasted
chances	2.5	4.91	.92	.8	2.67	.83	1	1	.97	.87	.87	chancing, cancers, cancer, Candice

One final real example clearly exhibits what sort of responses evoke a particular j-score. The following example consists of data collected from Experiment One. In the case of *hosted*, all but the initial phoneme are perceived accurately by nearly all 30 listeners, resulting in a fairly high p_p but a fairly low p_w. Looking at the errors, we see that they are all phonetically highly similar words. This is an actual example of bias against responding with a particular word. On the other hand, *chances* reveals the opposite pattern. In this case most of the phonemes have relatively high recognition rates, resulting in a high p_p and a fairly high p_w. The errors also seem to be of a different sort. Only one of the four errors is a neighbor of the target word (*chancing* differs by only one phoneme). The remaining responses are high-frequency words which have the same general syllabic pattern as the target word.

3.3.2 Raw Data and Consonant Cluster Analysis

One of the issues arising when analyzing open response data is that one can find responses that were never present in the input. That is, although the materials in this study contain only a subset of possible phonemes from each language, responses outside of this set are possible. In addition, responses which do not adhere to the same syllable structure as that of the stimuli (CVCCVC) are also possible. In order to adequately analyze such data, several decisions must be made about how to handle these types of responses. One of the more interesting and difficult decisions to make with this sort of data is how to treat consonant clusters. Very little has been said in the literature about analyzing clusters with open response spoken word recognition data, partially since most previous work has been with CVC stimuli, which greatly reduces the number of clusters available.

Analysis of the raw data involves several steps. The first step involves an automatic translation from text to phonemes. For the English data, this was done using the `t2p` program (Lenzo, 1998), which uses a dictionary containing orthographic and phonetic transcriptions, and generalizes spelling to phoneme mappings. In this way, the program can both capture many of the orthographic ambiguities in English, as well as generalize to words not contained in the dictionary (in this case nonwords). For the German data, a simpler program was created, since German orthography is much less ambiguous than English orthography. The program did automatically account for several phonological processes, such as final devoicing and final spirantization. For example, words ending in ⟨b d⟩ were transcribed as /p t/ and words ending in ⟨ig⟩ were transcribed as /ɪx/. The second step is to manually verify all the phonemic transcriptions. This step also involves making some decisions about how to handle incorrect responses. In analyzing incorrect responses, several general guidelines were followed:

- give as much credit as possible
- be consistent

These principles are perhaps best explained through the use of some examples. The following types of responses were treated as typographical errors, not psychophysical misperceptions, and were corrected before final analysis.

- typographical errors
  - metathesis error *biulded* — scored as /bɪldəd/
  - letters next to each other on keyboard
- real words in non words *bahbone* — scored as /babwʊn/
- misspellings *concious* for *conscious*

Analyzing consonant cluster responses is even slightly more difficult. In order to account for clusters, four additional slots were created, into which the raw data were analyzed — a slot for initial clusters, a slot between V1 and C2, a slot between C2 and C3, and a slot for final clusters. Some examples are shown in Table 3.4. Responses were lined up in order to maximize the number of correct phoneme responses. In certain cases, clusters could be analyzed in multiple ways. This is particularly the case in the middle of the word. In these cases, additional consonants were placed according to the phonological similarity. When responses included an epenthetic phoneme between V2 and C3, as in the response *tilptoll* to the stimulus *piptol*, sonorant consonants were treated as vowel misperceptions, while obstruents were counted as consonantal misperceptions. In cases where an additional consonant was perceived between C2 and C3, the response was scored as an error of C2 if the cluster was a legal coda cluster. If the response was a legal onset cluster, it was scored as an error of C3. If the response was both a legal onset and coda, it was scored as a C3 error, according to the maximal onset priniciple.

3.3.3 Computing Confusion Matrices and J-scores

After the hand-checking of the response data was complete, confusion matrices were computed. A separate confusion matrix was computed for each position, S/N, and stimulus type (nonword, and word) for each experiment, for a total of 96 (6 x 2 x 2 x 4) confusion matrices. The confusion matrices are located in appendix C.1 on page 122.

J-scores were calculated on subjects and items. For the subjects j-score, the average phoneme recognition probability (p_p) was calculated by computing the average percent correct for each subject in each position, and then computing the geometric mean of these numbers. This process was done for both nonwords and words, and monomorphemes and bimorphemes. The same process was used on an items basis as well.

Table 3.4 Cluster Analysis — Examples of how responses not conforming to the CVCCVC input structure were coded

Raw Data		Analysis									
Stimulus	Response	Cbeg	C1	V1	Vext	C2	Cmid	C3	V2	C4	Cend
English words											
pectin	temptkin		t	ɛ	m	p	t	k	ɪ	n	
lapses	lasses		l	æ				s	ɪ	z	
lasted	blasted	b	l	æ		s		t	ɪ	d	
goblin	garbwan		g	a	ɹ	b		w	ə	n	
German words											
Bänder	blender	b	l	ɛ		n		d	ə	ɾ	
rechtes	braechtest	b	ɾ	ɛ		x		t	ə	s	t
Runden	grummeln	g	ɾ	ʊ		m			ə	l	n
English Nonwords											
rekfudge	breakfudge	b	ɹ	ɛ		k		f	ə	ʤ	
naltum	nowtum		n	aʊ				t	ʊ	m	
choalsing	trollsing	t	ɹ	oʊ		l		s	ɪ	ŋ	
German Nonwords											
reungken	braenken	b	ɾ	ɔɪ		ŋ		k	ə	n	
piptol	tilptoll		t	ɪ	l	p		t	ɔ	l	
zilnich	ziemlich		t͜s	i		m		l	ɪ	x	

3.3.4 Computing lexical statistics

Analyses were also carried out based on three different measures of context effects: lexical frequency, neighborhood probability, and phonotactic probability. These measures were computed using the CELEX (Baayen & Rijn, 1993) database. The raw numbers in the database were recomputed in order to account for the auditory nature of the task. CELEX gives separate entries for homophones differing in syntactic class, e.g. *painting* is listed twice, once as a noun and once as a verb. In an auditory task, these two are indistinguishable, therefore their frequencies were summed in one combined entry in the database.

In addition to these modifications, the phonemic transcriptions of some of the words were also altered, particularly for the English portion of the database. Since the English portion of CELEX is based on British English, but the participants in this study were all speakers of American English, the transcriptions of all the materials were changed to American English pronunciations (using the transcriptions from the Hoosier Mental Lexicon database (HML) Nusbaum, Pisoni, & Davis (1984); when no transcription was available in the HML, transcriptions were produced using the native speaker intuition of the experimenter). It was not feasible to convert the entire database of more than 78,000 entries; however, several substitutions were made which account for some of the most systematic differences between British English and American English. British English contains 'linking r' at the end of some words, which is pronounced as a rhotic when followed by a vowel, but otherwise not pronounced (or can lengthen the preceding vowel); all sequences of /ə/ + 'linking

R' were converted to a rhotacized schwa /ɚ/. Additionally, CELEX specifies several lengthened vowels found in words spelled with a vowel + r, e.g., *barn, peer, pair,* and *poor*, transcribed as /ɑː, iə, ɛə, ʊə/ respectively. These were converted to /ɑɹ, iɹ, ɛɹ, oɹ/.

Lexical Frequency

CELEX provides two separate measures of frequency; a wordform frequency, and a lemma frequency. The lemma frequency is the sum of all wordforms for a given word, and can be thought of as the dictionary entry. Thus the lemma frequency for *walk* includes all instances of *walk, walks, walked,* and *walking*. Different studies have shown either the lemma frequency or the wordform frequency to be a better predictor of lexical frequency effects in lexical access. For complex words, both wordform and lemma frequency have been found to influence processing of nouns (Baayen, Dijkstra, & Schreuder (1997: in Dutch) and Taft (1979: in English)). The presence of lemma frequency effects indicates that lexical access is sensitive to a word's family structure, and not just its wordform frequency. Results for monomorphemic words are mixed: Taft (1979: experiment 2) found lemma frequency effects in English, but Sereno & Jongman (1997) find only wordform effects. In a more recent study, Vannest, Newport, & Bavelier (2006) found both lemma and wordform frequency effects in visual lexical decision and frequency ratings experiment in English, though wordform frequency effects were only found in mid-frequency words, whereas lemma frequency effects were found at all levels of frequency. Hopefully the present study can shed light on the mixed results of effects of wordform and lemma frequency in lexical access.

In addition to raw frequency (per million words), a log-based frequency was also calculated for each word in the stimulus materials. Several studies have shown that a log-based frequency is psychologically more appropriate than raw frequency (Zipf, 1935; Balota, Pilotti, & Cortese, 2001). To calculate log frequency, the method of Newman et al. (1997) was followed, defined as: $log_{10}(10 \cdot Freq)$. If the raw frequency of a word was less than 1, it was replaced with 1, since the log_{10} of a number less than 1 is negative, and it is difficult to interpret what a negative frequency would be. The raw frequency is multiplied by 10 such that all words will have a minimum log frequency of 1. This is necessary for computing frequency-weighted neighborhood density, in which neighbors are multiplied by their log frequency. Having a minimum log frequency of 1 ensures that this frequency weighting will positively weight high-frequency words, but not assign a negative weighting to low-frequency words. This is appropriate in particular because it is difficult to discern the actual frequency of low-frequency words. That is, simply because a particular word has a frequency of 0 in a given corpus does not imply that the word does not exist (in fact we can be sure that it does exist). This method ensures that all words are given some weight, and that high-frequency words are weighted in a psychologically relevant manner.

Neighborhood density

The two separate measures of neighborhood density included what will be referred to here as a phonological- and a phonetic-based measure. The phonological measure is the more commonly used method of calculating neighborhood density, whereby for each stimulus, the log-frequency of each neighbor of that stimulus is summed (where a neighbor is defined as a word with an edit distance of one from the target word).[2] The disadvantage of this method is that it treats all phonemes equally. However, from spoken word recognition experiments and from acoustic analysis, we know that [p] and [t] are more confusable than, say, [p] and [n]. To illustrate this, take for example the words *cap*, *can* and *cat*. Using the standard phonological neighborhood density measure, *cap* and *can* are treated as equally likely to be confused with *cat*. A phonetic measure of neighborhood density would find *cap* and *cat* more confusable than *cap* and *can*.[3] Following Benkí (2003a) the nonword confusion matrices from the present study were used to compute a measure of phonetic neighborhood density, based on the Neighborhood Activation Model (NAM) of Luce & Pisoni (1998), shown in Equation 3.6.

$$\sum_{j=1}^{nn}\left\{\left[\prod_{i=1}^{n} p(PN_{ij}|PS_i)\right]\cdot Freq_{Nj}\right\} \tag{3.6}$$

where $p(PN_{ij}|PSi)$ is the probability of a listener responding with the i^{th} phoneme of the j^{th} neighbor, when presented with the i^{th} phoneme of the stimulus, n is the number of phonemes in the stimulus, and nn is the number of neighbors. To paraphrase, for each neighbor of a target word, the product of the probabilities of perceiving each phoneme given the phonemes of the target word as a stimulus is multiplied by the log frequency of the neighbor. The sum of the frequency-weighted stimulus probability for each neighbor defines the frequency-weighted neighborhood probability, hereafter referred to as FWNP or phonetic neighborhood density.

Phonotactic Probability

Two measures of phonotactic probability were also calculated for all stimuli, based on the method of Vitevitch & Luce (2004). This method includes a measure of positional probability and a measure of biphone positional probability. The calculation of both of these measures involves two steps. The first step is to determine the frequencies with which phones or biphones occur in a language using a corpus, in this case the CELEX database (Baayen & Rijn, 1993). This method was as follows: for each phoneme in the language, the frequencies of each word that contained that phoneme in a given position were summed, and then this sum was divided by the number of words that contained any phoneme in that position. Position here simply refers to the position of the phoneme in a word. For example, in the word *cat* /kæt/, /t/ is in the third position. This was performed for each phoneme and for positions 1–6 (since the stimuli in this study are all six phonemes long, this is sufficient). To compute the positional probability of a given word, the positional frequency of each phoneme was summed. Computing biphone positional probability was performed in a similar manner, except that biphone frequencies were calculated instead of phoneme frequencies. That is, for every possible

---

[2] Edit distance, also known as levenshtein distance, is defined as the number of edits to change one string into another, including insertions, deletions, and substitutions. In this case, the strings are composed of phonemes.

[3] Another method of incorporating phonetic similarity is to use a feature-based metric, such that two phonemes which share many features are predicted to be more confusable with one another than two phonemes which share few features (see e.g. Bailey & Hahn, 2001; Frisch, Pierrehumbert, & Broe, 2004). As can be seen from the confusion matrices in Appendix C, feature-based proposals still fail to account for some confusions. For example, syllable final nasals are often not perceived at all, which would not be predicted in a feature-based calculation of neighborhood density.

biphone in the language and for all possible positions (1-2,2-3,3-4,4-5,5-6), the frequencies for all words containing the biphone in that position were summed, and then divided by the number of words which contained any phonemes in those positions. To calculate the biphone positional probability of a given word, the biphone frequencies for each biphone were summed. Generalizing this method, the positional probability of a given word with n phonemes can be calculated as:

$$\sum_{j=1}^{n}\left\{\sum_{i=1}^{N}\left[\frac{log_{10}(Freq_{ij})}{log_{10}(Freq_{i})}\right]\right\} \tag{3.7}$$

where n is the number of phonemes, N is the number of words in the database containing at least n phonemes, $Freq_{ij}$ is the wordform frequency of a word containing the phoneme j in the j^{th} position, and $Freq_i$ is the frequency of a word which has at least n phonemes.

This method of computing phonotactic probability is lacking in several ways. Firstly, it is unlikely that speakers align words in their lexicon simply by the position of their constituent phonemes. A more realistic measure of phonotactic probability should take into account phonological theory — at a minimum some notion of the syllable. Coleman & Pierrehumbert (1997) provide a model based on onsets and rimes that achieves this result. Secondly, some of the mathematics in Vitevitch & Luce's model seem ad-hoc. In spite of the shortcomings of this model, it has the advantage that several other studies have employed it, making the results of the present study more directly comparable with previous results.

3.4 Predictions

The basic predictions for each experiment are laid out in Table 3.5. Effects of lexical status, morphology, lexical frequency, and neighborhood density are predicted for all four experiments, but the size of some of the effects is predicted to differ among experiments.

Based on numerous studies using a variety of tasks, words are predicted to exhibit a processing advantage over nonwords (e.g. Rubenstein et al., 1970; Forster & Chambers, 1973). Using a j-factor analysis, the j-score of words is predicted to be lower than nonwords (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Olsen et al., 1997; Benkí, 2003a), indicating a bias for words (Nearey, 2001). The difference in j between words and nonwords is predicted to be roughly equal for native speakers of both English and German, but a smaller difference is predicted for non-native listeners. Assuming that non-native listeners have a smaller vocabulary size than native listeners, some of the word stimuli will essentially be novel words (i.e. nonwords) to the non-native listeners, resulting in a higher word j-score for non-native listeners compared to native listeners, which in turn decreases the difference in j between words and nonwords.

Several studies have found that monomorphemic words are processed more quickly than bimorphemic words (Sereno & Jongman, 1997; Gürel, 1999). Based on these studies, it is predicted that the j-score of monomorphemic words will be lower than that of bimorphemic words, indicating an increased processing demand for bimorphemic words. As discussed in §2.2, the morphological structure of a language can have an impact on how morphology affects lexical access. In general, languages which have rich morphologies tend to exhibit greater effects of morphology on lexical access than languages which do not use morphology extensively. For this reason, the effect of

Table 3.5 Basic Predictions — Predicted results are marked with a check mark, and a relative effect size is also given.

	English native listeners	German native listeners	English non-native listeners	German non-native listeners
lexical status $j_{nonword} > j_{word}$	✓robust	✓robust	✓smaller than native listeners	✓smaller than native listeners
morphology $j_{bi} > j_{mono}$	marginal	more than English	smaller than L1	smaller than L1
lexical frequency $j_{word} \propto \frac{1}{frequency}$	✓robust	✓robust	✓smaller than native listeners	✓smaller than native listeners
neighborhood density $j_{word} \propto density$	✓robust	✓robust	✓smaller than L1	✓smaller than L1

morphology is predicted to be smaller in English than in German, as measured by the difference in j between mono- and bimorphemic words. In addition, the effect of morphology is predicted to be smaller for non-native listeners than for native listeners. Previous research on memory and second language acquisition has shown that learners initially learn multi-morphemic or multi-word chunks, and only later process the smaller parts of these chunks (Baddeley, 1997; Ellis, 1996, 2001). This chunking effect could diminish differences in processing between mono- and bimorphemic words. If the predicted difference in j between mono- and bimorphemic words is found, this would pose problems for associative models of lexical access, which posit that words are stored whole. Current models of lexical access using a whole word storage approach (e.g. TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), and MERGE (Norris et al., 2000)) predict that monomorphemes and bimorphemes should be stored and accessed in the same way. Though these connectionist models have been shown to account for experimental evidence showing differences between regular and irregular inflectional morphology through the use of analogical pattern matching (Rumelhart & McClelland, 1986; Hahn & Nakisa, 2000; Nakisa, Plunkett, & Hahn, 2001), it does not seem that analogy can account for processing differences between mono- and bimorphemic words.

If no difference in j between mono- and bimorphemic words is found, this could be evidence in support of associative models of lexical access. Note that there is a potential flaw in this design, in that the only possible support of an associative model comes from finding no difference between monomorphemic and bimorphemic words. Supporting a hypothesis with a null result is very weak evidence. However, if the null result is found, there are several methods to increase its support. The basic problem with a null finding is that it is unclear whether there actually is no difference between the groups, or whether the experiment was just not able to detect a difference. If it can be shown that the experiment is accurate enough to find other, similar results, this greatly increases the validity of a null result supporting a hypothesis. In this case, several parameters which have been shown to vary in numerous other experiments will be investigated, namely lexical frequency and neighborhood density effects. Using these factors, the statistical power of the present study can be estimated, by computing the minimum statistically significant difference in j between two groups. The lack of a statistically significant difference between mono- and bimorphemes can then be interpreted as evidence that there is indeed no actual difference, or that if there is a difference, it must be very small.

Lexical frequency is predicted to have a facilitatory effect, such that high frequency words will be processed more easily than low-frequency words, as many other studies have shown (e.g Broadbent, 1967; Forster & Chambers, 1973; Taft, 1979). Consistent with Benkí (2003a), the j-score is predicted to be lower for high-frequency words than for low-frequency words. This effect is predicted to hold for both English and German native listeners, but the effect may vary for non-native listeners. Since non-native listeners have had less exposure to the language than native listeners, their familiarity with words is likely not highly correlated with frequency estimates made from large corpora. While the actual frequency counts may differ, it is likely that extremely high-frequency words (as measured by a corpus) will also be very high frequency for non-native listeners. The greatest difference between frequency for native and non-native listeners is likely to be in the low- and medium- frequency words, many of which may be completely unknown to the non-native listeners, and would pattern more like nonwords. Thus the difference in j between low- and high-frequency words may actually be greater for non-native listeners than for native listeners.

Given that English and German have relatively similar phonologies, an inhibitory effect of neighborhood density is predicted for both languages. In a j-factor analysis, this translates to a higher j for words in dense neighborhoods than words in sparse neighborhoods (Benkí, 2003a). Due to an assumed smaller vocabulary size, it is predicted that the magnitude of the effect of neighborhood density will be smaller for non-native listeners. Since non-native listeners have smaller vocabularies, many of the neighboring words are probably unknown to them, especially for words in dense neighborhoods. The effect of vocabulary size is not as large for words in sparse neighborhoods however. Therefore it is predicted that j of dense words will be lower for non-native listeners compared to native listeners, but the j of sparse words should be nearly the same for both native and non-native listeners, resulting in a smaller Δj for non-native listeners.

3.5 Summary

This chapter has provided a general overview of the experiments and predictions. The following four chapters discuss the methods and results of each experiment in detail, followed by a general discussion chapter summarizing the results from all four experiments.

Chapter 4

Experiment One — Recognition of English CVCCVC words and nonwords by native listeners

THIS experiment addresses several of the goals laid out in the preceding chapters. Previous research on lexical access and spoken word recognition has left several gaps with regard to the role of morphology. Most research investigating effects of morphology has been in the visual domain; thus it is not clear whether these effects will also be found in an auditory task. The great majority of research on spoken word recognition (and all previous research using the j-factor model), has only used monosyllabic stimuli. It is not yet known how well the previous results from spoken word recognition experiments using monosyllabic stimuli will predict results using bisyllabic stimuli. This experiment will address both of these questions simultaneously, by carrying out a spoken word recognition experiment using bisyllabic mono- and bimorphemic words. As mentioned in §3.2, this experiment also serves as a baseline for comparison with the other three experiments.

4.1 Method

4.1.1 Participants

Thirty-four paid participants were recruited via flyer from the University of Michigan. All participants reported being native speakers of English and having no known hearing impairments. Four of the participants were speakers of Malaysian or Singapore English, while the rest were speakers of American English. The speakers of Malaysian and Singapore English had quite different results than the speakers of American English. For this reason, those 4 participants were omitted from the results reported here.

4.1.2 Materials

The stimuli consisted of 150 nonwords and 150 English words (74 monomorphemic and 76 bimorphemic). The complete list of stimuli is in Appendix A.2 on page 101. All stimuli were of the form CVCCVC (where V includes short and long vowels as well as diphthongs), with stress on the first syllable. CVCCVC tokens were chosen because they are fairly common in both English and German, and they include both monomorphemes and bimorphemes.

Word stimuli were selected from the CELEX (Baayen & Rijn, 1993) database. CELEX is a large database containing a variety of phonological, morphological, syntactic, and frequency information on English, German, and Dutch. CELEX is particularly suited for the current study, as it contains frequency information for both lemma (dictionary entry) and word forms. For example, the word *lasting* has a raw wordform frequency of 4, but a raw lemma frequency of 71 (includes all forms of *last*, e.g. *last, lasted, lasts*). This allows one to address the questions of how words are stored in the lexicon. Associative models would predict that only wordform frequency should have an effect on lexical access, whereas combinatorial models would predict that both wordform and lemma frequency can affect lexical access.

Monomorpheme List

The monomorpheme list consisted of singular nouns and adjectives. All derivational affixes and compound words have been excluded, though there are some ambiguous cases. For example *bandage* /bændɪʤ/ could be considered to be bimorphemic, consisting of *band* + - *age*. However, many such words (including *bandage* in my opinion) have become semantically opaque. That is, it is not clear to the naïve speaker that these words can be subdivided into separate parts. This is not the case for words such as *signage*, which is clearly decomposable into two morphemes. Semantically opaque words such as *bandage* have been included in the list, whereas semantically transparent words such as *signage* were excluded.

Bimorpheme List

The bimorpheme list consisted of verbs and nouns which have an overt inflectional affix, e.g. *feast* + *-ing* /fistɪŋ/, or *box* + *es* /bɑksɪz/.

Nonword List

The nonword stimuli were generated from the word stimuli. The distribution of phonemes in the word stimuli (see Table A.3, page 106) was used as input to generate a list of nonword stimuli. For each position (C1, V1, etc.), a phoneme from the list of possible phonemes in the word list was chosen at random, until the number of occurrences of that phoneme in the word list was reached. For example, if /b/ occurred in initial position 23 times in the word stimuli, then the nonword generation program output 23 nonwords beginning with /b/. This process was repeated 3 times, generating a total of 450 nonwords. This list was then checked against the CELEX database, and all possible nonwords with an edit distance of 1 from any real word in CELEX were removed, so that the nonword stimuli would not closely resemble real words. Next, the list was manually checked to ensure that all stimuli were phonotactically possible, and any particularly odd-sounding stimuli were removed. In this way, the nonword list was largely phonotactically balanced with the word list.

4.1.3 Stimulus Recording and Editing

The stimuli were recorded at the University of Michigan in an anechoic chamber with a Crown CM-700 condenser microphone directly into .wav format with a sampling rate of 44.1 kHz via the PRAAT (Boersma & Weenink, 2006) program on an iBook laptop computer. Each item was read by a phonetically-trained male speaker of American English (the speaker was raised in Utah), in the carrier phrase "Say ___ again". Three complete randomizations of the materials were recorded, blocked according to lexical status. The nonwords were displayed using a quasi-phonetic transcription, e.g. 'E' was used to represent /ɛ/. The target word in each file was then extracted from the carrier phrase in PRAAT. Each of the three repetitions was given a rating of 1 to 5 (1=poor quality 5=excellent quality) by the experimenter, based on the auditory impression and visual inspection of the waveform and spectrogram. Tokens that included extraneous noises, speech disfluencies, mispronunciations, or abnormal amplitude were given poor ratings. The best token for each word was selected to use in the experiment. Each of the selected stimuli was padded with 100 ms of silence on both sides, and the peak amplitude was normalized to .99 Pascals. The complete list of stimuli can be found in Appendix A.2 on page 101.

4.1.4 Procedure

Participants listened to the stimuli over AKG closed headphones, through an iMic USB digital to analog converter on Dell laptop computers running Windows XP. The experiment was carried out in an anechoic chamber at the University of Michigan. Participants were allowed to adjust the volume to a comfortable listening level. Up to four participants participated at once. The stimulus presentation and response collection was controlled by software developed by Benkí and Felty in the Matlab programming environment. The software mixes signal-dependent noise (as described by Schroeder, 1968) with the recorded stimuli, and allows for the collection of open response data typed in via the keyboard. Listeners were instructed that they would hear disyllabic words and nonwords mixed with noise, and that they should type what they hear, using standard orthography for the words, and a slightly modified orthography for the nonwords, on which the participants were briefly trained before the beginning of the experiment. The exact instructions are included in Appendix B.1 on page 117.

The experiment began with two practice blocks (one word block, and one nonword block) of 10 stimuli each, in order to familiarize the participant with the task. The main experiment consisted of 20 blocks of 15 stimuli each, blocked according to lexical status. Participants only heard each stimulus once, but had no time limit to type in their answer. The experiment lasted approximately 45 minutes on average.

Two different signal-to-noise-ratios (S/Ns) were used in the experiment. Although previous research (Benkí, 2003a) has shown the j-factor model to be consistent across various S/Ns, using multiple S/Ns samples a broad range of performance levels, which helps to increase statistical power, and also creates more generalizable results. Pilot results showed a very large difference between words and nonwords, such that finding two S/Ns that would fit into the range between 5% and 95% both for word and phoneme recognition for both words and nonwords was nearly impossible. Therefore a compromise was reached such that for each subject, the nonword stimuli S/N was 5 dB higher than the word stimuli. Thus instead of using two different S/Ns, two pairs of S/Ns were used. Half of the participants heard words presented at S/N=-5 dB and nonwords at S/N=0 dB, and half of the participants heard words presented at S/N=0 dB and nonwords at S/N=5 dB.[1]In the results, the lower pair (-5 and 0 dB) will simply be referred to as -5 dB and the higher pair (0 and 5 dB) will be referred to as 0 dB.

4.2 Analysis

The data for this experiment were primarily analyzed using the j-factor model, which is described in detail in §3.3.

4.3 Predictions

Based on the overall predictions made in §3.4 the following specific predictions are made for Experiment One:

1. $j_{nonword} \approx 6$: This prediction is based on previous results showing that $j_{nonword}$ is equal to the number of phonemes in the stimulus (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003a).
2. $j_{word} \approx 5$: This prediction is based on previous results using the j-factor model with CVC words, which have found $j_{word} \approx 2.5$ (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003a). Given that the words in this experiment are twice as long, it is logical to hypothesize that the average j_{word} (not taking factors such as lexical frequency, neighborhood density, or phonotactic probability into account) will be twice as large.
3. $j_{bi} > j_{mono}$: Assuming a combinatorial type model of lexical access, it is predicted that bimorphemes are processed differently than monomorphemes, and that this should be reflected in the j-score. Given that all of the phonemes in a monomorphemic word contribute to the semantic representation of that word, whereas the affixes of bimorphemic words do not contribute to the semantic representation, monomorphemic words can be said to have a higher degree of lexical context; therefore the j-score of monomorphemes is predicted to be lower than that of bimorphemes.
4. $j_{word} \propto \frac{1}{\text{frequency}}$: This prediction is based on the result from Benkí (2003a) that j decreases as lexical frequency increases. Lexical frequency provides a facilitatory effect equivalent to faster response times in timed tasks such as lexical decision.
5. $j_{word} \propto \text{density}$: This prediction is also based on results from Benkí (2003a) that j increases as neighborhood density increases. Neighborhood density provides an inhibitory effect, which

---

[1] After excluding the 4 participants, 14 listeners heard the stimuli at S/N=-5 dB, and 16 listeners heard the stimuli at S/N=0 dB.

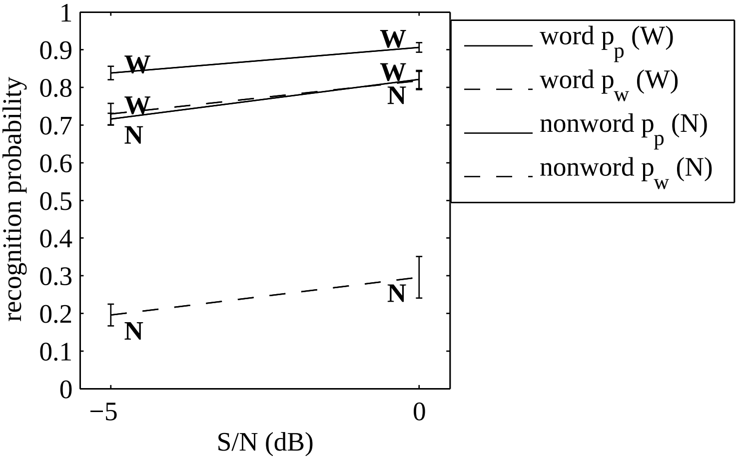


Figure 4.1 English phoneme and word recognition probabilities — Each point represents the average phoneme or word recognition probability for words (W) or nonwords (N) at a given S/N. Error bars display 95% confidence intervals.

is also equivalent to slower response times for words in dense neighborhoods as found in tasks such as lexical decision and naming (Luce & Pisoni, 1998).

4.4 Results

The complete set of responses included 9000 trials (300 stimuli x 30 subjects). Five ($< .1\%$) trials were discarded due to no response, thus leaving 8995 trials for analysis. The average phoneme (p_p) and (non)word (p_w) recognition probability scores are shown in Figure 4.1. As predicted, the recognition rates for words were higher than for nonwords for both whole words and phonemes. In addition the recognition rates were all higher at S/N=0 than S/N=-5. It can also be seen that the difference between p_w and p_p is much larger for nonwords than for words. The j-factor model provides for a more detailed analysis of the differences between word and phoneme recognition rates.

The results of the subjects analysis are shown in Figure

4.4.1 Subjects analysis

The results of the subjects analysis are shown in Figure 4.2. Each panel displays the data grouped by one of the context effects in question.

Lexical Status

The effect of lexical status is very large, and highly significant, though the actual values for j are somewhat unexpected. The result of $j_{nonword} = 5.82$ is somewhat lower than the predicted value of 6. Possible explanations for this result will be discussed in §4.5.1. In addition, the result of $j_{word} = 3.64$ is also much lower than the predicted value of 5. This result indicates that j does not scale linearly with word length.

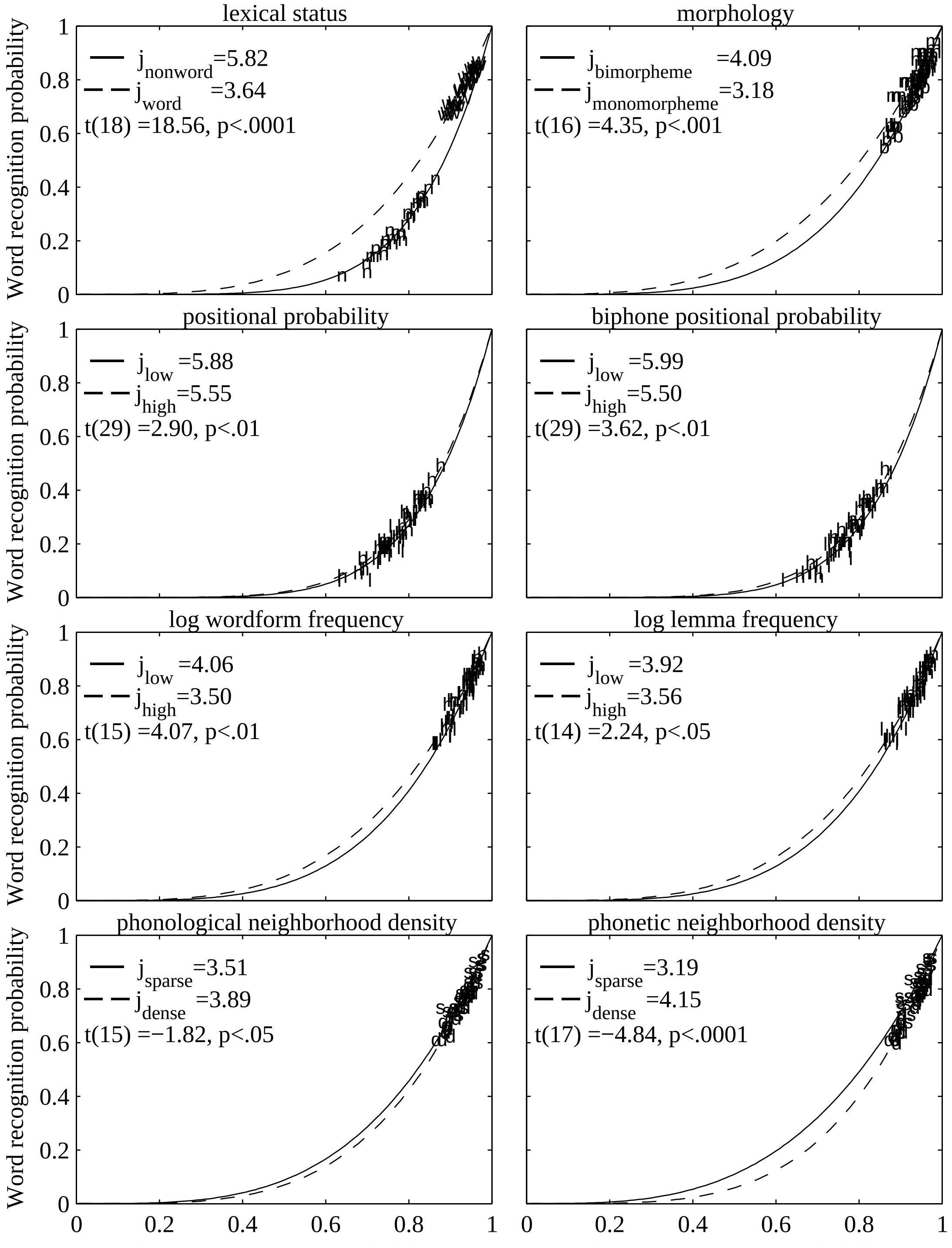


Figure 4.2 English *j*-factor results by subjects — Each plot compares two subsets of results from the subject analysis. Each point represents the average results for one subject. Curves represent $y = x^j$. The second row of plots only shows nonword results, while the final two rows only display word results. Statistics shown are from paired t-tests (one-tailed for plots in rows 1 and 4; two-tailed for plots in rows 2 and 3); before computing the statistics, all points lying in the floor or ceiling ranges ($> .95$ or $< .05$) were removed, but are still shown on the plot.

Morphology

Initial results show a significant difference between monomorphemic and bimorphemic words; however, additional analysis revealed an interaction with frequency. This interaction will be discussed in §4.5.2.

Phonotactic Probability

Several studies (e.g., Vitevitch & Luce, 1998, 1999; Coleman & Pierrehumbert, 1997) have shown that phonotactic probability can influence lexical access, which largely holds true only for nonwords. Two different measures of phonotactic probability were calculated based on the method of Vitevitch & Luce (2004), described in detail in §3.3.4. The nonword data were divided into low and high phonotactic probability groups using a median split for each of the two measures of phonotactic probability; the results are shown in the second row of Figure 4.2. The prediction here is that nonwords with high phonotactic probability appear to be more word-like, and therefore should have a lower j-score than nonwords with a low phonotactic probability. The results based on both the positional probability and the biphone probability bear out this prediction. In both cases the low phonotactic probability items have a significantly higher j-score.

Lexical Frequency

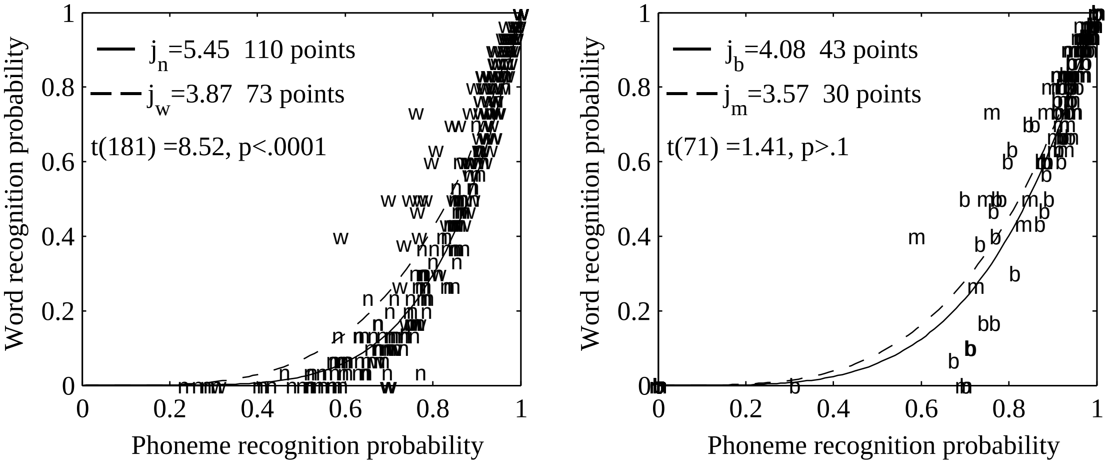
Results of the lexical frequency analysis (shown in the third row of Figure 4.2) are consistent with the predictions. Both the wordform and the lemma frequency analyses showed that words with low frequency had significantly higher j-scores than those with high frequency, indicating a facilitatory effect of frequency.

Neighborhood Density

The effects of neighborhood density are largely consistent with those of previous studies. Words in sparse neighborhoods have fewer competitors, and therefore a facilitatory effect is found, namely that j is lower for words in sparse neighborhoods than for words in dense neighborhoods. Using a phonological measure of neighborhood density, this effect was small, but significant. However, using a phonetic measure of neighborhood density, in which the confusability of phonemes is taken into account, this effect is found to be quite large and significant. In fact, in terms of the magnitude of the effect, the difference in j of approximately .99 is only exceeded by the effect of lexical status.

4.4.2 Items analysis

Results were also analyzed over items. As is often the case with items analyses, there is a greater amount of variation in the data. However, one of the advantages of an items analysis is that it makes a regression analysis possible, which is not the case for a subjects analysis. A regression analysis



(a) lexical status

(b) morphology

Figure 4.3 English native listener *j*-factor results by items

provides an estimate of the amount of variance explained by a particular variable. As in the subjects analysis, each context effect is reported separately. Effects of lexical status and morphology are shown in Figure 4.3; the remaining items analysis results are shown in Figure 4.4 using regression analyses.

Lexical Status

The main effect of lexical status was also quite robust in the items analysis, as shown in Figure 4.3a.

Morphology

The difference between mono- and bimorphemic words was not significant in the items analysis as shown in Figure 4.3b. Again, these results should only be considered preliminary due to the interaction with frequency. See §4.5.2 for further discussion.

Lexical Frequency

J-score values were significantly correlated with log wordform frequency, but not log lemma frequency. However, even the significant effect of log wordform frequency accounts for less than 7% of the variation in j.

Stimulus Probability

Stimulus probability was calculated based on the confusion data from the nonwords. For any given word, the stimulus probability was calculated as the product of correctly identifying each constituent

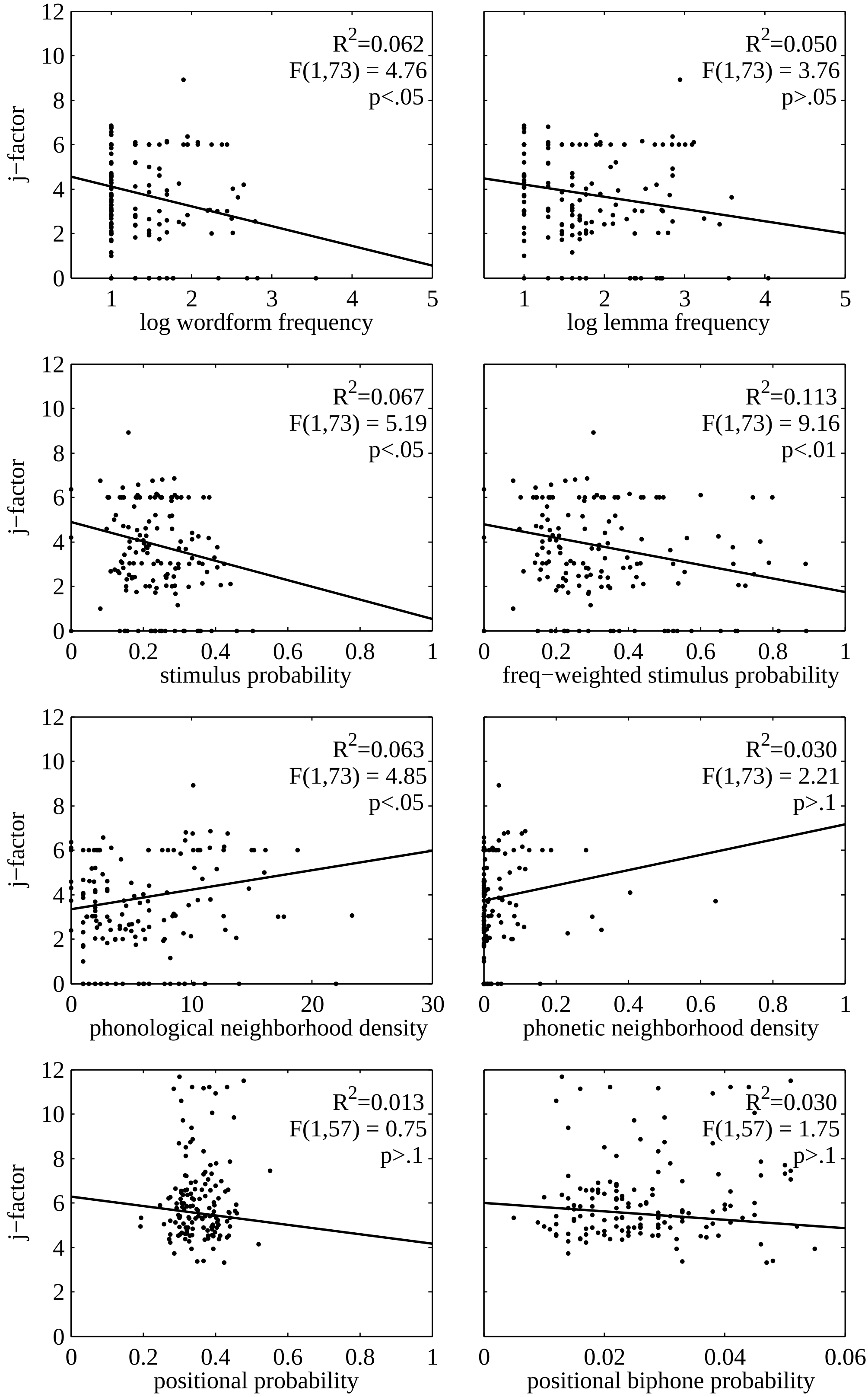


Figure 4.4 English *j*-factor regression analyses by items. Each panel plots *j*-factor as a function of one particular lexicostatistical measure. Each point represents one item. The top 6 panels show only word items, while the bottom two show only nonword items. The statistics given are from linear regressions.

phoneme based on the nonword confusion matrices. It is expected that p_w and p_p should both be positively correlated with stimulus probability; that is, raw perceptibility of the phonemes should affect nonword and word stimuli alike. Since j is a ratio of $\frac{log(p_w)}{log(p_p)}$, there should be no correlation between stimulus probability and j. A small negative correlation between j and stimulus probability was found, indicating that p_w increased more rapidly with stimulus probability than p_p. Given that Benkí (2003a: p.1694) found a small effect in the opposite direction in the subjects analysis suggests that this effect is still not fully understood. Frequency-weighted stimulus probability (FWSP) was calculated as the stimulus probability multiplied by the log frequency count of each word. As expected, the negative correlation between lexical frequency and j also appeared in this analysis.

Neighborhood Density

The effect of phonological neighborhood density was significant, but the effect of phonetic neighborhood density was not significant. Examining the plots in Figure 4.4, it is apparent that most of the words have a very low phonetic neighborhood density. This could account for the lack of significant result for the phonetic neighborhood density.

Phonotactic Probability

Neither the positional probability nor the biphone positional probability regression analysis on items reached significance, though both were significant in the subjects analysis. The difference between the calculation of p_p in the subject and items analyses is likely the cause of this inconsistency, given that p_p in the subjects analysis is averaged over a large number of phonemes for each subject, whereas p_p in the items analysis includes only the phonemes in each given item, averaged over subjects. This makes p_p in the items analysis more sensitive to the phonological structure of each word, and cannot measure effects of phonotactic probability as reliably as in the subjects analysis. ======= »»»> .r126

Lexical Frequency

J-score values were significantly correlated with log wordform frequency, but not log lemma frequency. However, even the significant effect of log wordform frequency accounts for less than 7% of the variation in j.

Stimulus Probability

Stimulus probability was calculated based on the confusion data from the nonwords. For any given word, the stimulus probability was calculated as the product of correctly identifying each constituent phoneme based on the nonword confusion matrices. It is expected that p_w and p_p should both be positively correlated with stimulus probability; that is, raw perceptibility of the phonemes should affect nonword and word stimuli alike. Since j is a ratio of $\frac{log(p_w)}{log(p_p)}$, there should be no correlation between stimulus probability and j. A small negative correlation between j and stimulus probability was found, indicating that p_w increased more rapidly with stimulus probability than p_p. That Benkí (2003a: 1694) found a small effect in the opposite direction in the subjects analysis suggests that this effect is still not fully understood. Frequency-weighted stimulus probability (FWSP) was calculated as the stimulus probability multiplied by the log frequency count of each word. As expected, the negative correlation between lexical frequency and j also appeared in this analysis.

Neighborhood Density

The effect of phonological neighborhood density was significant, but the effect of phonetic neighborhood density was not significant. Examining the plots in Figure 4.4, it is apparent that most of the words have a very low phonetic neighborhood density. This could account for the lack of significant result for the phonetic neighborhood density.

Phonotactic Probability

Neither the positional probability nor the biphone positional probability regression analysis on items reached significance, though both were significant in the subjects analysis. The difference between the calculation of p_p in the subject and items analyses is likely the cause of this inconsistency, given that p_p in the subjects analysis is averaged over a large number of phonemes for each subject, whereas p_p in the items analysis includes only the phonemes in each given item, averaged over subjects. This makes p_p in the items analysis more sensitive to the phonological structure of each word, and therefore the effects of phonotactic probability cannot be measured as reliably in the items analysis as in the subjects analysis.

4.5 Discussion

4.5.1 Word length

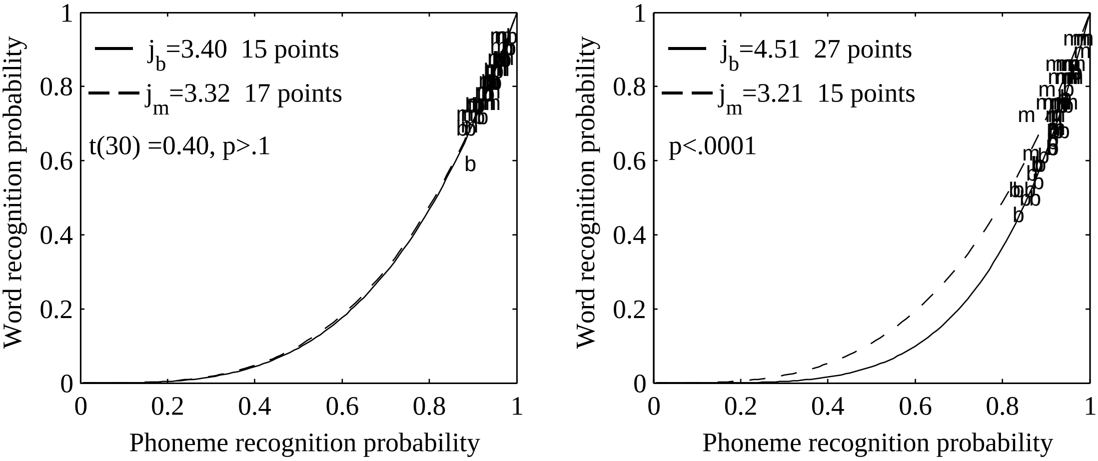
One somewhat surprising result from this experiment is that j is lower than predicted for both words and nonwords. The discrepancy between predicted $j_{nonword} \approx 6$ and observed $j_{nonword} = 5.82$ is relatively small. It is likely that the discrepancy is due to effects of phonotactic probability. The subjects analysis shows that words with low phonotactic probability do exhibit a j-score very close to 6, as does the items analysis.

The results for words ($j_{word} = 3.64$) are much lower than the predicted value of 5. There are several possible explanations for this. This could be partially explained by the lexicostatistical properties of the stimuli. One well known property of neighborhood density is that it is correlated with word length. That is, as word length increases (measured in phonemes), the number of words at or beyond that length decreases (at least for English this is the case). It follows as a direct result that neighborhood density must also decrease with word length, since there are fewer words available to be neighbors of any given word. Since an increase in neighborhood density causes j to increase, the overall lower neighborhood density of the stimuli used in this experiment compared to previous experiments using the j-factor model with CVC words (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003a) could partially explain why j_{word} is lower than expected in this experiment. Benkí (2003a) (who used the same word list as Boothroyd & Nittrouer (1988) reports that the CVC stimuli used in his experiment had an average of 20.8 neighbors, compared to an average of 4.9 of the stimuli used in this experiment. Complicating the matter even more is the effect of lexical frequency. Not surprisingly, longer words also tend to be used less frequently. The words used in this experiment had an average log wordform frequency of 1.4 compared to an average frequency of 3.29 in Benkí (2003a). Since j is known to decrease with lexical frequency, the overall lower frequency of the materials used in this experiment would predict a higher j than found in previous experiments (relative to the number of phonemes in the stimuli).

Because of this conflict between lexical frequency and neighborhood density, it is difficult to determine if either or both of these factors are playing a role. However, it seems unlikely that the decrease in neighborhood density alone can account for the discrepancy between predicted and observed j-scores for words. A plausible explanation for this discrepancy is that j does not scale linearly with word length. That is, it is possible that as words get longer, listeners begin to perceive words in units larger than phonemes — perhaps syllables. Several studies have provided evidence in support of the claim that the basic unit of speech perception is the phoneme (e.g. Norris & Cutler, 1988; Nearey, 2001), while several other studies (e.g. Mehler, Segui, & Frauenfelder, 1981; Savin & Bever, 1970) have suggested that the syllable is the basic unit of speech perception. It may be the case that listeners perceive words both in terms of phonemes and syllables, and that word length may have an influence on which of these two strategies dominates; another possibility is that units of speech perception are merely emergent properties, as Goldinger (2003) and Grossberg (2003) have proposed. In order to more conclusively determine the effect of word length on spoken word recognition, further research must be carried out. In an experiment using stimuli grouped according to word length (e.g CVC, CVCVC, and CVCVCVC), with each group matched for lexical frequency, neighborhood density, and phonotactic probability, the effect of word length could be more rigorously investigated.

4.5.2 Morphology

Although initial results showed a significant difference in j between mono- and bimorphemic words, further analysis showed an interaction between morphology and lexical frequency. The set of monomorphemic words had a significantly higher log wordform frequency than that of the bimorphemic words ($\mu_{mono} = 1.52, \mu_{bi} = 1.31, t = 2.51, p < .05$), though there was no difference in log lemma frequency ($\mu_{bi} = 1.90, \mu_{mono} = 1.74, t = 1.37, p > .1$). To investigate this interaction, several subsets of the stimuli were prepared. One subset included the lowest frequency mono- and bimorphemic words, which all had a log wordform frequency of 1 and did not differ in log lemma frequency ($\mu_{bi} = 1.46, \mu_{mono} = 1.28,\ t = 1.78,\ p > .07$) This subset consisted of 44 bi- and 29 monomorphemic words. The second subset was matched for log wordform frequency, containing the 32 highest wordform frequency monomorphemes and 33 bimorphemes from the middle wordform frequency range ($\mu_{bi} = 1.73, \mu_{mono} = 1.60,\ t = 1.54,\ p > .1$). This subset did differ in log lemma frequency however ($\mu_{bi} = 2.49, \mu_{mono} = 1.82, t = 4.72, p < .001$). The results in Figure 4.5 show that the difference remained for the low-frequency subset, but not the mid/high



(a) mid/high-frequency balanced subset

(b) low-frequency balanced subset

Figure 4.5 English native listener results using a subset of the word stimuli. (a) shows the j-factor results using a subset of the materials balanced for lexical frequency, with log wordform frequency ranging from 1.30 to 2.826 ($\mu_{bi} = 1.73, \mu_{mono} = 1.60, t = 1.54, p > .1$). This subset did differ in log lemma frequency however ($\mu_{bi} = 2.49, \mu_{mono} = 1.82, t = 4.72, p < .001$). The subset included 33 monomorphemic words and 32 bimorphemic words. (b) shows the j-factor results using a subset of the words with the lowest frequency. All words in this subset had a log frequency of 1 and did not differ in log lemma frequency ($\mu_{bi} = 1.46, \mu_{mono} = 1.28, t = 1.78, p > .07$) This subset consisted of 44 bi- and 29 monomorphemic words.

frequency subset. There are several possible explanations for this result. One explanation is the difference between the lemma frequencies of the mid/high frequency subset. Since the bimorphemic words in this subset have a significantly higher lemma frequency, it could be that the effect of lemma frequency is pulling down the j of the bimorphemes. Another explanation is that high-frequency and low-frequency words are stored differently in the lexicon, which Bybee (2001: 100) has proposed.

The effects of morphology could also be due to an interaction with neighborhood density. The mono- and bimorphemic groups did differ significantly in phonological neighborhood density ($\mu_{bi} = 9.01, \mu_{mono} = 3.36, t(148) = 8.49, p < .0001$). Similar to the subset matched for frequency, a subset of 33 monomorphemic and 30 bimorphemic words matched for density ($\mu_{mono} = 5.81, \mu_{bi} = 4.94, t(61) = 1.43, p > .1$) was created. A j-factor analysis over subjects using this subset also yielded a significant difference in j between the mono- and bimorphemic words ($j_{bi} = 3.11, j_{mono} = 2.75, t(39) = 2.09, p < .05$). However, this result is also not conclusive, since the subset matched for neighborhood density differed in lexical frequency. To test this possibility, one final subset of 10 mono- and 19 bimorphemic words matched for both phonological neighborhood density and log wordform frequency was created. A j-factor analysis on this subset was not significant ($j_{bi} = 2.69, j_{mono} = 2.78, t(29) = -.32, p > .1$). While a null result is not conclusive evidence, the effects of morphology found in this experiment appear to be highly confounded with effects of frequency and neighborhood density, and should be interpreted with caution.

4.6 Conclusions

This experiment has addressed several issues in spoken word recognition. One of the main goals of this experiment was to extend previous research on spoken word recognition to disyllabic words. The results using disyllabic words are largely consistent with those from previous experiments using monosyllabic words. Increasing lexical frequency resulted in a facilitatory effect, while increasing neighborhood density resulted in an inhibitory effect. Phonotactic probability of the nonword stimuli also resulted in a facilitatory effect, in that nonwords with higher phonotactic probability were treated more like words. One somewhat surprising result is that the j_{word} was substantially lower than predicted, suggesting that j may not scale linearly with word length.

This experiment also addressed effects of morphology on spoken word recognition, which had not been previously investigated using a speech-in-noise task. Effects of morphology were found, but not consistently, due to an interaction with the frequency of the monomorphemic and bimorphemic words chosen for the experiment. The effects of lexical frequency and neighborhood density found in this experiment are consistent with predictions made by associative models of lexical access, as described in §2.1. The effect of morphology found in this experiment is too confounded with effects of lexical frequency and neighborhood density to convincingly support either associative or combinatorial models of lexical access. Experiment Two will further test predictions of the effect of morphology made by associative and combinatorial models of lexical access, as well as the hypothesis that a more highly inflecting language such as German will show more robust effects of morphology than English.

Chapter 5

Experiment Two — Recognition of German CVCCVC words and nonwords by native listeners

EXPERIMENT One showed that the j-factor model is an appropriate tool for investigating context effects in spoken word recognition. The context effects found in Experiment One were largely consistent with previous results from experiments using speech-in-noise tasks as well as experiments using other tasks. Facilitatory effects were found for lexical status and lexical frequency, and an inhibitory effect was found for neighborhood density. However, evidence of morphological decomposition was inconsistent. This could be due to the nature of the task or due to the relatively little inflectional morphology in English. The second of these two possibilities is tested in Experiment Two, using the same task, but with German stimuli, a language that is morphologically more complex than English. It is predicted that effects of lexical status, lexical frequency, and neighborhood density should be similar to English, but a significant effect of morphology will also be found, due to the greater use of morphology in German.

5.1 Method

5.1.1 Participants

Thirty-two paid participants were recruited via flyer from the University of Konstanz. All participants reported being native speakers of German and having no known hearing impairments.

5.1.2 Materials

As in Experiment One, the stimuli consisted of 150 nonwords and 150 German words (75 monomorphemic and 75 bimorphemic). The complete list of stimuli is in Appendix A.4 on page

110. All stimuli were of the form CVCCVC (where V includes short and long vowels as well as diphthongs), with stress on the first syllable.

Monomorpheme List

The monomorpheme list consisted of nominative singular nouns, and uninflected adjectives. All derivational affixes and compound words have been excluded, though there are some ambiguous cases. For example *Seufzer* /zɔyft͜sɛr/ ‘sigh’ which is nominative singular, is related to the verb *seufzen* /zɔyft͜sən/ ‘to sigh’. Though most would agree that *Seufzer* is not directly derived from *seufzen* (in fact it could be the other way around), I excluded such words, on the chance that they might not be interpreted as bimorphemic, or not stored as the “base" form of the word, to which affixes are attached (assuming a combinatorial theory of lexical access). Words such as *sechzig* /zɛçt͜sɪç/ ‘sixty’, which contain a predictable affix, have also been excluded for the same reason. Words such as *Schulter* /ʃʊltə ɾ/ ‘shoulder’, which arguably could be considered bimorphemes (i.e. that *er* is a separate morpheme, as it can be used to indicate the meaning “one who does X”, where X is the stem of the word), are treated here as monomorphemeic, on the grounds that they are not transparently bimorphemic.

Bimorpheme List

The bimorpheme list consisted of adjectives and nouns which have an overt inflectional affix, for example *Feld* + *es* /fɛldəs/ ‘field — masc.gen.sing.’, or *ganz* + *es* /gant͜səs/ ‘whole — neut.nom.sing.’.

Nonword List

As in Experiment One, the nonword stimuli were generated from the word stimuli. The distribution of phonemes in the word stimuli (see Table A.6, page 115) was used as input to randomly generate a list of nonword stimuli which were largely phonotactically balanced with the word stimuli.

5.1.3 Stimulus Recording and Editing

The stimuli were recorded at the University of Michigan in an anechoic chamber with a Crown CM-700 condenser microphone directly into .wav format with a sampling rate of 44.1 kHz via the PRAAT (Boersma & Weenink, 2006) program on an iBook laptop computer. Each item was read by a male speaker of Standard German embedded in the carrier phrase “Sagen Sie ___ einmal”. Three repetitions of each stimulus were recorded, and then extracted from the carrier phrase using PRAAT. Stimulus selection and editing was the same as for Experiment One. The complete list of stimuli can be found in Appendix A.4 on page 110.

5.1.4 Procedure

Participants listened to the stimuli over Sennheiser HD 520 II closed headphones, powered by an M-Audio Delta Audiophile soundcard on BEST desktop computers running Windows 2000. The experiment was carried out in a quiet room. Subjects were allowed to adjust the volume to a comfortable listening level. The stimulus presentation and response collection was the same as in Experiment One. Listeners were instructed that they would hear disyllabic words and nonwords mixed with noise, and that they should type what they hear, using standard orthography. The exact instructions are included in Appendix B.2 on page 118.

Two different S/Ns (2 dB and 7 dB) were chosen on the basis of pilot results to cover the range between 5% and 95% both for word and phoneme recognition for both words and nonwords, in order to avoid floor and ceiling effects. Half of the participants heard the stimuli presented at the lower S/N and half at the higher S/N.

5.2 Analysis

The data from this experiment were analyzed in the same manner as the other experiments, described in detail in §3.3.

5.3 Predictions

Predictions for Experiment Two are largely the same as those for Experiment One, except that the difference between monomorphemic and bimorphemic words is predicted to be larger and more consistent, given that German is a more highly inflecting language than English. The predictions are repeated here for convenience.

1. $j_{nonword} \approx 6$: This prediction is based on previous results showing that $j_{nonword}$ is equal to the number of phonemes in the stimulus (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003a).
2. $j_{word} \approx 5$: This prediction is based on previous results using the j-factor model with CVC words, which have found $j_{word} \approx 2.5$ (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003a).. Given that the words in this experiment are twice as long, it is logical to hypothesize that j_{word} will be twice as large.
3. $j_{bi} > j_{mono}$: Assuming a combinatorial type model of lexical access, it is predicted that bimorphemes are processed differently than monomorphemes, and that this should be reflected in the j-score. Given that all of the phonemes in a monomorphemic word contribute to the semantic representation of that word, whereas the affixes of bimorphemic words do not contribute to the semantic representation, monomorphemic words can be said to have a higher degree of lexical context; therefore the j-score of monomorphemes is predicted to be lower than that of bimorphemes.
4. $j_{word} \propto \frac{1}{\text{frequency}}$: This prediction on based on the result from Benkí (2003a) that j decreases

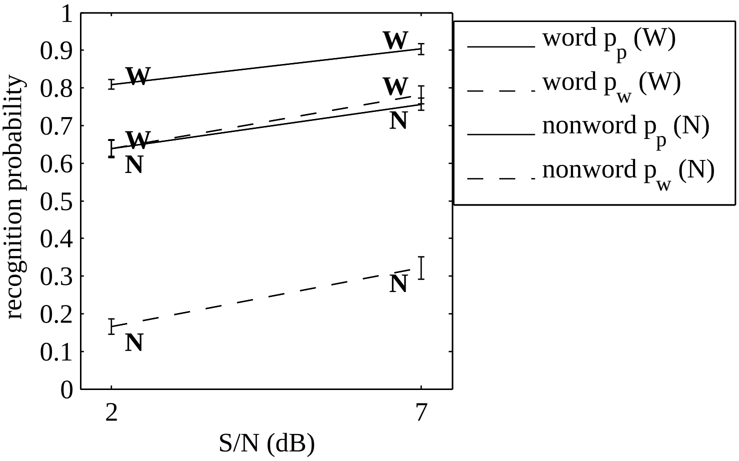


Figure 5.1 German phoneme and word recognition probabilities — Each point represents the average phoneme or word recognition probability for words (W) or nonwords (N) at a given S/N. Error bars display 95% confidence intervals.

as lexical frequency increases. Lexical frequency provides a facilitatory effect equivalent to faster response times in timed tasks such as lexical decision.

5. $j_{word} \propto$ density: This prediction is also based on results from Benkí (2003a) that j increases as neighborhood density increases. Neighborhood density provides an inhibitory effect, which is also equivalent to slower response times for words in dense neighborhoods as found in tasks such as lexical decision and naming (Luce & Pisoni, 1998).

5.4 Results

The complete set of responses included 9600 trials (300 stimuli x 32 subjects). Trials in which participants did not provide any response were discarded (169 trials, $< 2\%$), thus leaving 9431 trials for analysis. The average phoneme (p_p) and (non)word (p_w) recognition probability scores are shown in Figure 5.1. As predicted, the recognition rates for words were higher than for nonwords for both whole words and phonemes. In addition the recognition rates were all higher at S/N=7 than S/N=2. It can also be seen that the difference between p_w and p_p is much larger for nonwords than for words. This is precisely what the j-factor models.

5.4.1 Subjects analysis

The results of the subjects analysis are shown in Figure 5.2. Each panel displays the data grouped by one of the context effects in question. While the analysis here includes some comparisons between the results of this experiment and Experiment One using English stimuli, Chapter 8 provides a more detailed analysis of cross-linguistic differences found in this study

Lexical Status

The effect of lexical status is very large, and highly significant, though the actual values for j are somewhat unexpected. The result of $j_{nonword} = 4.76$ is substantially lower than the predicted value

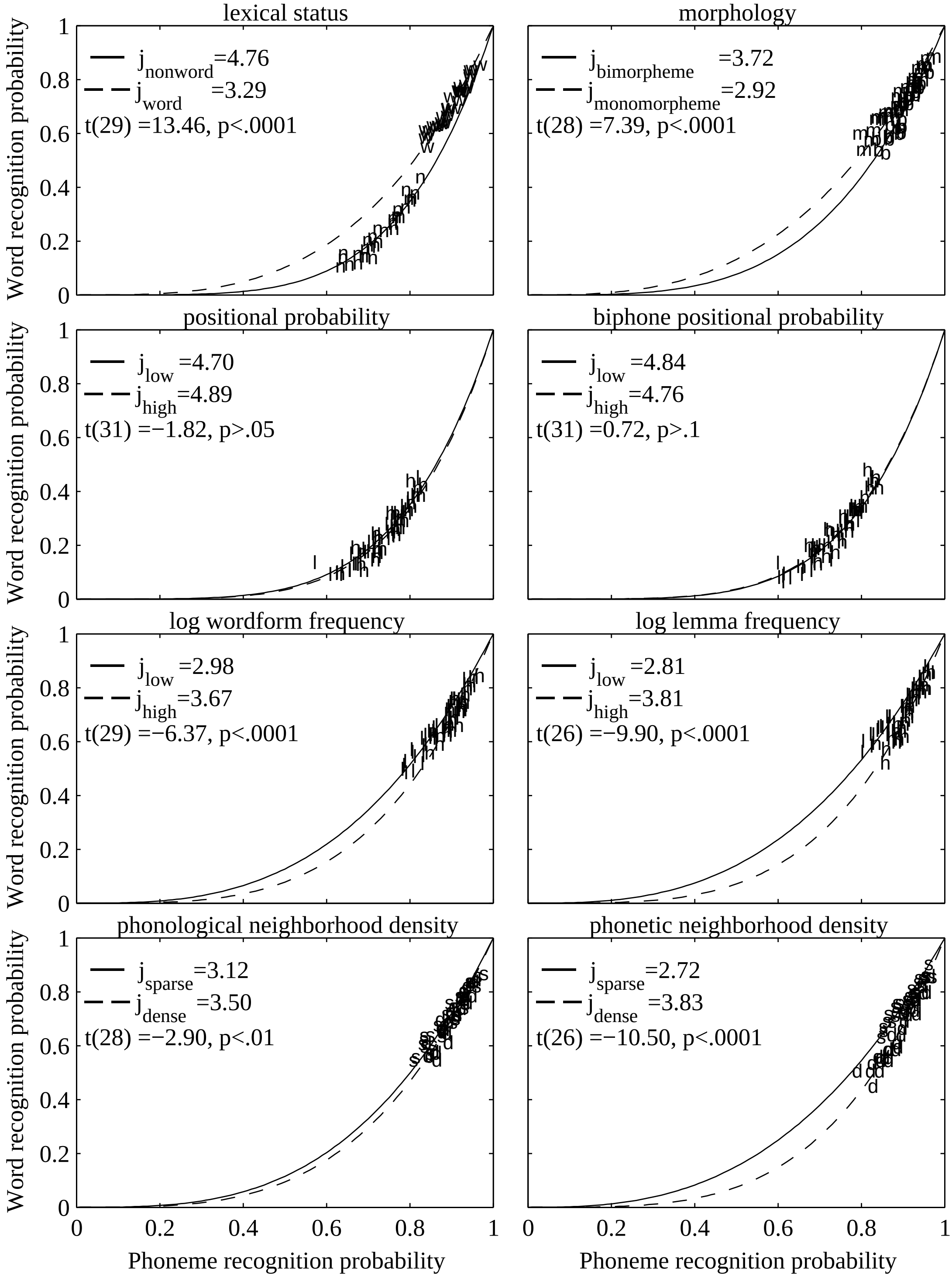


Figure 5.2 German *j*-factor results — Each plot compares two subsets of results from the subject analysis. Each point represents the average results for one subject. Curves represent $y = x^j$. The second row of plots only shows nonword results, while the final two rows only display word results. Statistics shown are from paired t-tests (one-tailed for plots in rows 1 and 4; two-tailed for plots in rows 2 and 3); before computing the statistics, all points lying in the floor or ceiling ranges ($> .95$ or $< .05$) were removed, but are still shown on the plot.

of 6. Possible explanations for this result will be discussed §5.5.3. As in Experiment One, the result of $j_{word} = 3.29$ is also much lower than predictions based on previous findings. Previous studies using the j-factor model with CVC stimuli have all found $j_{word} \approx 2.5$ (Boothroyd & Nittrouer, 1988; Benkí, 2003a; Olsen et al., 1997). All else being equal, one might expect that words with twice the number of phonemes would have j-scores twice as high. This is clearly not the case though.

Morphology

As predicted, j of bimorphemic words was significantly higher than that of monomorphemic words ($j_{bi} = 3.72$, $j_{mono} = 2.92$, $p < .0001$). This effect will be discussed in more detail in §5.5.1.

Phonotactic probability

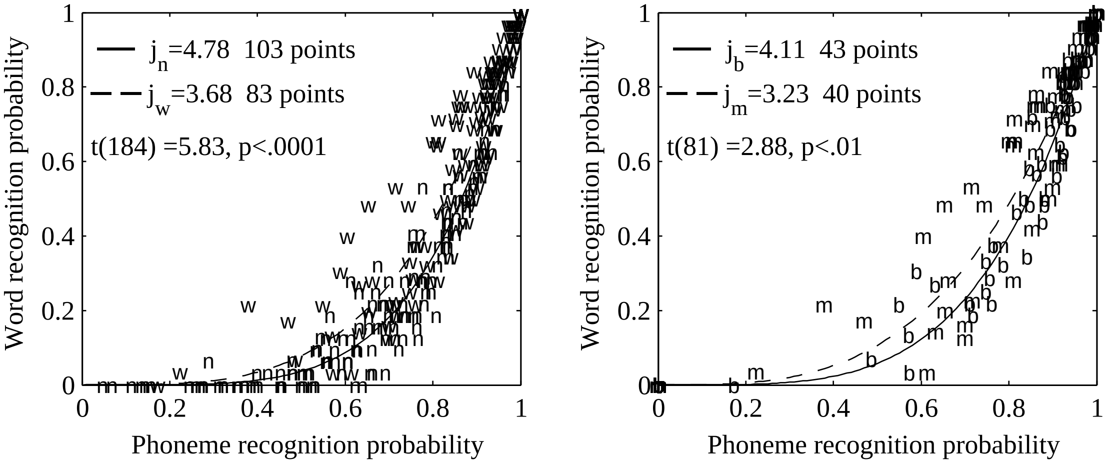
As in Experiment One, possible effects of phonotactic probability were investigated following the method of Vitevitch & Luce (2004). As shown in the second row of Figure 5.2, neither the results based on positional probability nor biphone positional probability reached significance, and in fact the trends are in opposite directions. These mixed results of phonotactic probability could be due to several factors. Previous results of the influence have had very small effect sizes, and have all been based on tasks using response time (RT) as the measure. It could be that the influence of phonotactic probability only has an effect on the time course of lexical access, and thus would not appear using the j-factor model. Another possibility is that the measure of phonotactic probability put forth by Vitevitch & Luce (2004) is lacking. Indeed, their model does not incorporate any sort of syllabification, but rather only looks at raw position in a word. Yet another explanation for the lack of significant effects of phonotactic probability is that there is a difference between languages, given that nonwords with high phonotactic probability had significantly lower j-scores than nonwords with low phonotactic probability in Experiment One. However, as will be seen in Chapters 6 and 7, no significant effect of phonotactic probability was found for Experiments Three or Four. Therefore it seems that phonotactic probability has at best only a small effect on lexical access in this study.

Lexical frequency

The effects of lexical frequency for this experiment are quite unexpected. The prediction that words with higher lexical frequency would have lower j-scores was not borne out, but rather the opposite. This was the case for both the wordform and the lemma frequency measures. Possible explanations for this will be discussed later in §5.5.

Neighborhood Density

Consistent with previous studies and with the results from Experiment One, the results show that words in dense neighborhoods have significantly higher j-scores than words in sparse neighborhoods. As shown in Figure 5.2, the difference in j between sparse and dense neighborhoods was greater



(a) lexical status

(b) morphology

Figure 5.3 German *j*-factor results by items

using a phonetic measure of neighborhood density than a phonological measure, which is also consistent with the results from Experiment One.

5.4.2 Items analysis

Effects of lexical status and morphology are shown in Figure 5.3. The remaining results of the items analysis are shown in Figure 5.4 using a regression analysis.

Lexical Status

The main effect of lexical status was also highly significant in the items analysis, as shown in Figure 5.3a. The difference in *j* between mono- and bimorphemic words was also significant as shown in Figure 5.3b. It is of note that the j-scores in the items analysis are consistently higher than those in the subjects analysis. This is likely due to the exclusion of certain items. Recall from Figure 5.2 that data in the floor and ceiling ranges were excluded before statistical analysis. Excluding subjects does not change the overall nature of the stimuli, but excluding items can make such a difference. This will be discussed further in §5.5.3.

Morphology

The effect of morphology was also significant in the items analysis. Similar to the items analysis of lexical status, the j-scores for both monomorphemic and bimorphemic words were higher than in the subjects analysis.

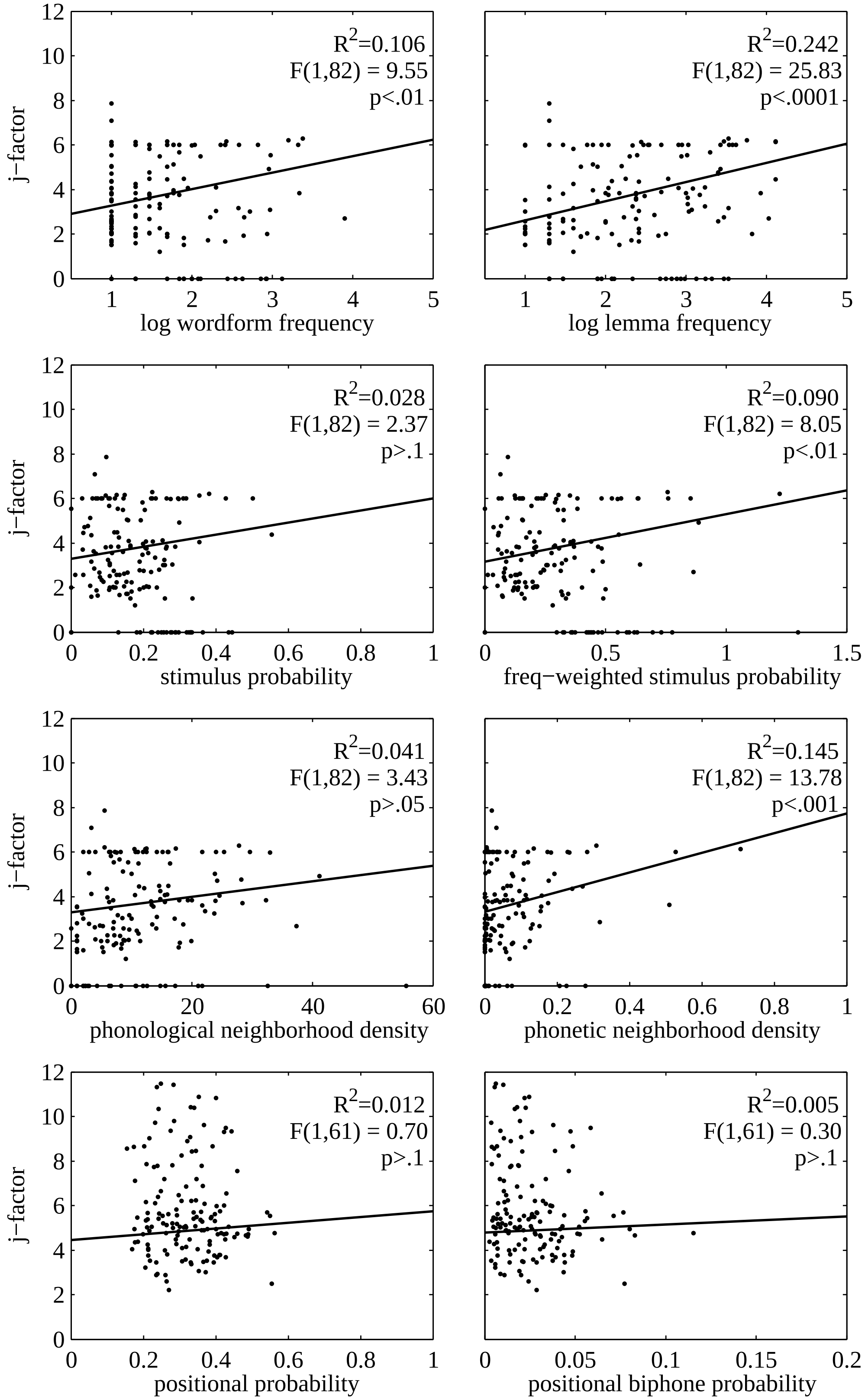


Figure 5.4 German *j*-factor regression analyses — Each panel plots *j*-factor as a function of one particular lexicostatistical measure. Each point represents one item. The top 6 panels show only word items, while the bottom two show only nonword items. The statistics given are from linear regressions.

Lexical frequency and stimulus probability

The unexpected result that j is positively correlated with lexical frequency was also found in the items analysis. This will be further explored in §5.5. Following Benkí (2003a) and Luce & Pisoni (1998), effects of stimulus probability were also explored. As in Experiment One, effects of stimulus probability were also explored. Consistent with Benkí (2003a) and with Experiment One, no significant effect of stimulus probability was found, but frequency-weighted stimulus probability (FWSP) was significantly correlated with j.

Neighborhood density

The outcome of the items analysis of neighborhood density is consistent with the outcome of the subjects analysis. The phonological neighborhood density was in the expected direction, though insignificant. The phonetic neighborhood density measure was quite large and highly significant, once again showing the strong phonetic effects in this sort of task.

Phonotactic probability

Consistent with the results from the subjects analysis, no significant effects of phonotactic probability were found, both using the positional probability and the biphone probability measures.

5.5 Discussion

5.5.1 Morphology

As predicted, the mean j of monomorphemic words was significantly lower than that of bimorphemic words. This can be interpreted in several non-mutually exclusive ways. One possible interpretation is that morphemes add to the overall number of independent units of a word. Another possible interpretation is that bimorphemic words are less predictable than monomorphemic words, and therefore the phones are less independent of one another than in monomorphemic words. Consider two words, one monomorphemic and one bimorphemic, with an equal number of neighbors (including deletions and substitutions, but not additions). The bimorphemic neighbor will likely (and in the case of the German certainly) include neighbors which share the same lemma, whereas the monomorphemic words should not include such neighbors. A listener presented with a bimorphemic word whose neighbors share the same root will find it difficult to rely on frequency as a predictor of which response is more probable. As Clahsen et al. (2001) showed, listeners do not simply rely on wordform frequency. Recent research by Vannest et al. (2006) has shown that lemma frequency is a better predictor of frequency effects in several different experimental tasks. The items analysis in this experiment also supports lemma frequency as a better predictor of frequency effects, in that the lemma frequency accounted for more of the variation in j did than wordform frequency. This finding is consistent with predictions from combinatorial models of lexical access. If morphological

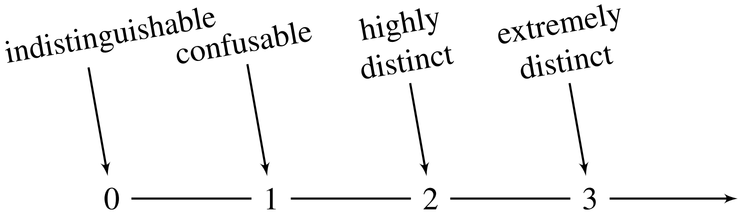


Figure 5.5 interpretation of d' values

information is stored the mental lexicon, then frequency effects are predicted to be correlated with the lemma frequency. And if listeners are primarily depending on lemma frequency to make educated guesses, then they must use a strategy based on something other than lemma frequency when choosing between bimorphemic neighbors differing only in their final consonant. Such a strategy could include raw acoustics and knowledge about the distribution of affixes.

These strategies can be tested by investigating the degree of acoustic salience and response bias in the data. The final consonants in the bimorphemic stimuli were restricted to the phonemes /ʀ s m n/, which, along with /ə/ constitute all of the possible inflectional endings for nouns and adjectives in German. Two of these, /m/ and /n/ are known to be highly confusable with one another. In addition, /n/ occurs as an inflectional ending much more frequently than /m/. Thus it is highly possible that both acoustic factors as well as response bias could be playing a role in the perception of these two final consonants. In order to investigate this further, a Signal Detection Theory (SDT — Macmillan & Creelman (2005)) analysis was carried out.

SDT measures the sensitivity of distinguishing two stimuli, using the metric d'. Interpretations for different values of d' are given in Figure 5.5. SDT also provides a measure of bias, c, which indicates whether listeners are more or less likely to respond with a particular phoneme. Positive values of c indicate a bias towards a response; negative values indicate a bias against a response.

To carry out the SDT analysis, the original confusion matrices for each S/N were transformed into 2x2 submatrices. An SDT analysis was then applied to each submatrix. From the results shown in Table 5.1, several conclusions can be drawn: (1) in the absence of lexical context effects (i.e. in the nonword condition), /m/ and /n/ are highly confusable, with a small bias towards /n/, (2) /m/ and /n/ are perceived as most distinct in the monomorphemic condition, and (3) bias towards /n/ is greatest in the bimorphemic case. The increase in distinction in the monomorphemic case can be interpreted as a result of the greater ability to distinguish between neighbors based on lexical frequency information. The bias towards /n/ in the bimorphemic case can be interpreted as evidence that listeners are exploiting the fact that the /n/ ending occurs most frequently among all possible inflectional endings in German, and they are therefore choosing /n/ more frequently. [1] The results of the SDT analysis suggest that listeners seem to be depending on a combination of acoustics, lemma frequency, and morphological distribution to make their decisions.

It is also possible that other context effects such as lexical frequency or neighborhood density could be responsible for the difference between mono- and bimorphemic words. The monomorphemic and bimorphemic words did not differ in mean log wordform frequency ($\mu_{bi} = 1.65, \mu_{mono} = 1.68, t = .22, p > .8$), but did differ in mean log lemma frequency ($\mu_{bi} = 2.80, \mu_{mono} = 1.78, t =$

---

[1] It is possible that acoustic differences between the stimuli in these three groups is actually driving the perceptual differences, but that question is outside the scope of the current study.

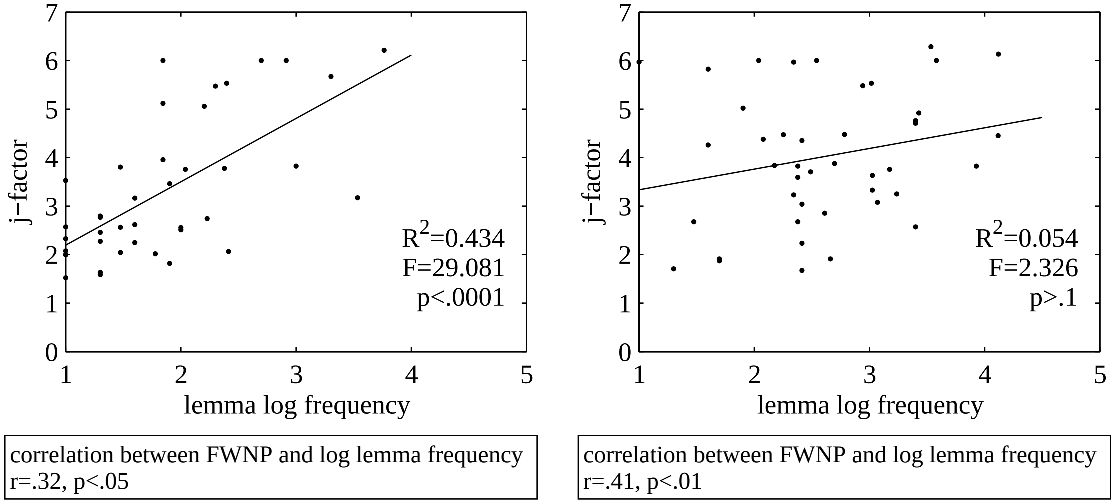
Table 5.1 Signal Detection Theory analysis of /m/ and /n/ submatrix in final position. For this analysis /m/ is considered to be the target stimulus. Positive values of c indicate a bias towards /n/. The final two columns list the total number of presentations of /m/ and /n/ which were used to compute the SDT analysis

	d'	c	/m/	/n/
Nonwords				
lower S/N (2 dB)	-0.182	0.555	240	240
higher S/N (7 dB)	0.664	0.743	240	240
Bimorphemes				
lower S/N (2 dB)	1.616	0.984	128	352
higher S/N (7 dB)	1.913	0.556	128	352
Monomorphemes				
lower S/N (2 dB)	3.514	0.239	48	192
higher S/N (7 dB)	4.733	-0.060	48	192

$9.03, p < .0001$). The fact that the mean log lemma frequency of the bimorphemes is greater than the monomorphemes would predict that j_{bi} would actually increase if the two groups were matched for log lemma frequency; therefore this possibility does not require further exploration. The mono- and bimorphemic words also differed in mean phonological neighborhood density ($\mu_{bi} = 14.80, \mu_{mono} = 8.11, t = 4.81, p < .0001$). The higher neighborhood density of the bimorphemic words could be responsible for the higher j-scores. In order to tease these effects apart, a subset was extracted in which the monomorphemic and bimorphemic stimuli were matched according to neighborhood density. The subset consisted of words with a frequency-weighted neighborhood density between 5 and 15, resulting in 32 monomorphemes and 42 bimorphemes. A two sample t-test showed that the effect of morphology was also significant in this subset ($j_{mono} = 3.08$, $j_{bi} = 3.70$, p<.001). Therefore, lexical frequency and neighborhood density do not directly account for the morphological effects in the results.

5.5.2 Lexical Frequency

One strikingly unexpected result is the positive correlation between lexical frequency and j for the German data — the opposite of the predicted result. This effect seems to be fairly robust, both in the subjects (Figure 5.2) and the items analyses (Figure 5.4). Upon initial investigation, this appeared to be due to a correlation ($r = .3594, p < .0001$) between phonetic neighborhood density and lexical frequency in the German data. Thus it seemed that the effect of neighborhood density is overshadowing the effect (if any) of lexical frequency. This is in part consistent with the findings of Benkí (2003a), which showed neighborhood density to be a much stronger predictor of recognition than lexical frequency. However, to test this hypothesis more rigorously, the word items were split by the median FWNP into two groups — a low-density group (43 words) and a high-density (40 words) group. The results of separate analyses run on these two subgroups are displayed in Figure 5.6. In the low-density group, there is still a strong positive trend ($R^2 = .434, p < .0001$) of j with lemma log frequency, but the high-density items do not show a significant correlation between lemma log frequency and j, despite the fact that the correlation between FWNP and log



correlation between FWNP and log lemma frequency
r=.32, p<.05

correlation between FWNP and log lemma frequency
r=.41, p<.01

(a) low-density items

(b) high-density items

Figure 5.6 Experiment 2 density subsets — the left plot displays the subset of words with low FWNP (43 words); the right plot displays those with high FWNP (40 words). The R^2 and p-values are shown underneath each plot, in addition to the r value indicating the degree of correlation between FWNP and log lemma based of each group. As in all other previous statistical analyses, items which had p_p or p_w values below .05 or above .95 were excluded prior to statistical analysis.

lemma frequency is greater in the high-density group. This suggests that the unexpected effect of lexical frequency cannot necessarily be attributed to the correlation with neighborhood density.

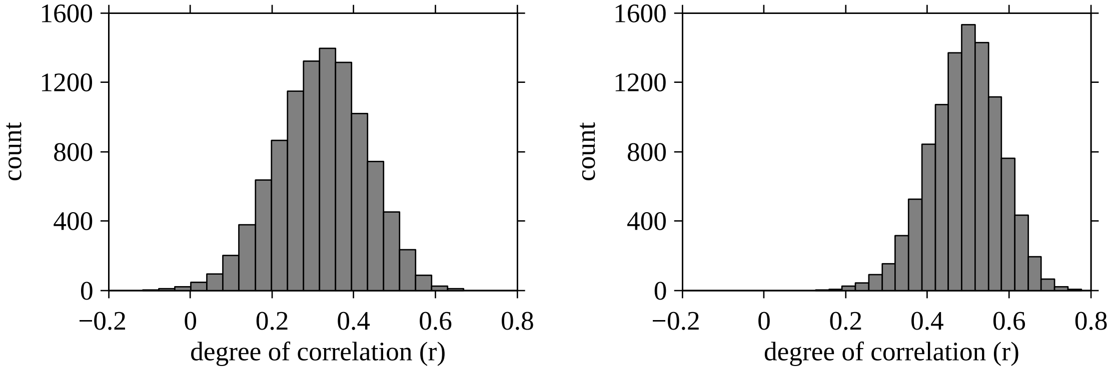
As a further test of this hypothesis, bootstrap analyses (Efron & Tibshirani, 1993) were performed on the correlation between j and lexical frequency. A bootstrap analysis re-samples the data with replacement over many times. This is essentially a way of simulating the experiment many times. The result is a distribution of possible outcomes, in this case of the correlation coefficient, r. The results of these analyses can be seen in Figure 5.7. The bootstrap analysis of wordform frequency overlaps slightly with 0, suggesting that the null hypothesis cannot be ruled out, but the analysis on lemma frequency clearly can rule out the null hypothesis. Thus the positive correlation between lexical frequency and j in this experiment is a real effect.

Another possibility is that the unexpected frequency effects could be due to the frequency of the first syllable. Conrad & Jacobs (2004) found that increasing the frequency of the first syllable produced an inhibitory effect in German using an orthographical lexical decision task and a visual progressive de-masking task.[2] First syllable frequency is similar to neighborhood density. It is defined as the number of words that share the first syllable with a given word. Conrad & Jacobs (2004) discuss two types of syllable frequency — token- and type-based measures. The type-based measure simply counts the number of words which share the first syllable, whereas the token-based measure sums the frequencies of all words which share the first syllable. Conrad & Jacobs (2004) use a token-based measure.

If syllable frequency is positively correlated with lexical frequency for the stimuli used in this experiment, this could explain the inhibitory effect of lexical frequency. The frequency of the first

---

[2] It should be noted that while syllable frequency is a phonological effect, which is probably best measured using an auditory task, Conrad & Jacobs (2004) chose German for the experiment because it has a very shallow orthography. Previous work by Perea & Carreiras (1998) on Spanish (which also has a shallow orthography) indicated similar effects.



(a) wordform frequency (b) lemma frequency

Figure 5.7 Experiment 2 lexical frequency bootstrap analysis — The histograms display the results of a 10000 iteration bootstrap analysis of the correlation between lexical frequency and j. The left plot displays the correlation with wordform frequency; the right plot displays the correlation with lemma frequency. As in all other previous statistical analyses, items which had p_p or $_w$ values below .05 or above .95 were excluded prior to statistical analysis.

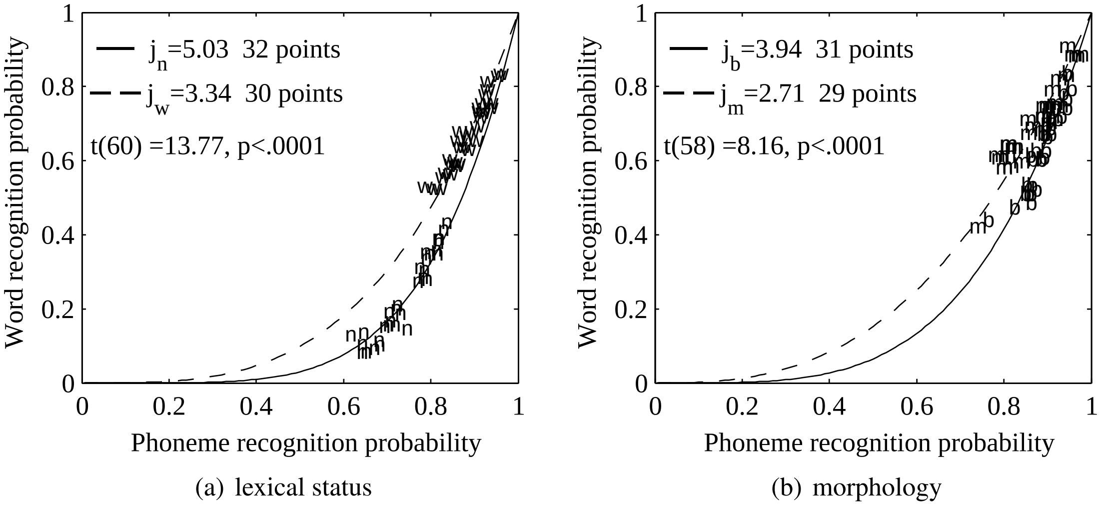
syllable was calculated for each stimulus using the CELEX (Baayen & Rijn, 1993) database; the first CVC of each word was considered to be the first syllable. This is not necessarily the case for each word, but it is a close approximation. A Pearson test of correlation showed that first syllable frequency and wordform frequency are correlated for the stimuli ($r = .358$ $p < .001$). However, there was no significant correlation between syllable frequency and j ($r = .01$, $p > .1$). Syllable frequency does not account for the inhibitory effect of lexical frequency.

The phonological makeup of the chosen stimuli also does not appear to explain the unexpected frequency results. One possible concern raised by several native-speaking German linguists was the inclusion of post-vocalic /ʀ/ in the stimuli. Though there are valid phonological reasons for treating /ʀ/ as a consonant,[3] its phonetic realization in post-vocalic position is not normally considered to be consonantal. The combination /əʀ/ is phonetically realized as [ɐ], and /ʀ/ following non-reduced vowels often is realized simply as a lengthened vowel. This could have an effect on the j-score of the words, since this could mean that the assumption of independence would not hold. In order to test this, the results were re-analyzed excluding all words which contained post-vocalic /ʀ/. This reduced the set of stimuli to 94 nonwords and 79 words (36 monomorphemic and 43 bimorphemic). The results of lexical frequency for this subset were not very different than for the full set. A single linear regression by items showed a positive correlation between wordform frequency and j ($R^2 = .12, p < .05$) as well as a positive correlation between lemma frequency and j ($R^2 = .29, p < .001$). Therefore one can conclude that the unexpected effect of lexical frequency is not due to the presence of post-vocalic /ʀ/ in the stimuli. In order to understand the cause of this effect, further research using more stimuli should be carried out, which is beyond the scope of this project.

Yet another explanation for the unexpected frequency effects is that it is due to talker effects. Moon & Lindblom (1994) found that talkers speaking in clear speech produce more distinct

---

[3] Probably the most convincing argument is that post-vocalic /ʀ/ can function as a syllable onset in inflected words such as *besseres* [bɛsəʀəs], even though the uninflected version *besser* [bɛsɐ] is not phonetically transcribed with a consonantal [ʀ].



(a) lexical status

(b) morphology

Figure 5.8 German *j*-factor results for subset of data excluding post-vocalic /ʀ/. This subset includes 94 nonwords and 79 words (36 monomorphemic and 43 bimorphemic)

utterances (e.g. the vowel spaces are larger than in casual speech). In addition, several researchers have claimed that high-frequency words exhibit more coarticulation and reduction (see e.g. Bybee, 2001). Given that the talker used in this experiment spoke in a fairly casual manner, it is possible that he articulated low-frequency words more carefully than high-frequency words, causing a reduction in the phonemic independence of the high-frequency words. This would explain the inhibitory effect of frequency. Unfortunately, the stimuli for this experiment were not constructed in a manner that would lend themselves to the rigorous acoustic analysis which would be required to test this hypothesis. This question must be left for further research.

5.5.3 Perceptual independence

The finding of $j = 4.76$ for German nonwords is substantially lower than predicted. There are several possible explanations for this result. It is possible that the nonwords chosen in this experiment had a particularly high phonotactic probability, resulting in a lesser degree of perceptual independence than expected. Recall that one major assumption of the *j*-factor model is that phonemes are perceived independently of one another, in the absence of context effects. Since the nonwords were constructed to be phonotactically legal, differences in the overall phonotactic probability of the nonwords could explain this result. Nonwords with high phonotactic probability should have a lesser degree of perceptual independence than words with low phonotactic probability. However, the results of the phonotactic probability analysis do not support this hypothesis, as shown in Figure 5.2. Using the positional probability measure of phonotactics, the nonwords with a lower phonotactic probability have a lower j-score — the opposite of what one would expect. Using the biphone probability metric, there is no significant difference between the low and high probability groups. Therefore phonotactic probability is not a plausible explanation for the lower than expected j-score of nonwords.

Although phonotactic probability cannot account for the lower than expected j-score of nonwords, it is possible that other phonetic properties of the stimuli could be responsible. As discussed in §5.5.2 post-vocalic /ʀ/ could effectively lower the perceptual independence of the stimuli. To test this hypothesis, the effects of lexical status and morphology were re-analyzed using the subset of data excluding post-vocalic /ʀ/. The results shown in Figure 5.8 are very similar to the results including all stimuli, except that the j-score for each group is increased by approximately .2–.3. Although the result of $j = 5.03$ for nonwords is still substantially lower than the expected value of 6, it is somewhat closer. The remaining discrepancy is likely due to the fact that the stimuli were all trochees (i.e. disyllables with initial stress), and therefore the set of 4 possible vowels for the second vowel were highly restricted (/ʊ ɪ ə ɔ/) compared to the set of 18 possible vowels for the first vowel (/i ɪ y ʏ e ɛ ø œ u ʊ o ɔ a aː ə ay ɔy au/).

5.6 Conclusions

This experiment has addressed several context effects in spoken word recognition. It was shown that the effect of lexical status is robust in German words, consistent with previous studies. Results also showed that morphology can have an impact on spoken word recognition, in that j was significantly higher for bimorphemic words than for monomorphemic words. In particular, the size of this effect was larger and more consistent than the results from Experiment One using English. Cross-linguistic differences will be discussed in more detail in §8.1. The result of $j = 3.29$ for CVCCVC words is an important finding, demonstrating that j does not scale linearly with word length. Also consistent with previous studies (Benkí, 2003a; Olsen et al., 1997), neighborhood density had a robust effect on word recognition, such that words in sparse neighborhoods showed a strong bias over words in dense neighborhoods. Moreover, a phonetically based measure of neighborhood density explained a much larger portion of the data than a phonologically based measure.

Two unexpected results from this experiment remain open questions. The result that $j_{nonwords}$ was much lower than expected does not seem to be due to phonotactic probability or neighborhood density. Excluding stimuli which contained post-vocalic /ʀ/ accounted for much of this discrepancy, but not all of it. It was hypothesized that the remaining discrepancy is due to the fact that only trochaic stimuli were used. Further experiments using spondees could address this issue in more depth. The unexpected positive correlation between lexical frequency and j also remains unresolved. Analysis of several subsets of the data showed that this result is not due to correlation between lexical frequency and density in the stimulus set, nor did analyses of first syllable frequency or the exclusion of post-vocalic /ʀ/ explain this result. Additional studies using a greater number of stimuli should be carried out to investigate this effect further.

Chapter 6

Experiment Three — Recognition of German CVCCVC words and nonwords by non-native listeners

WHILE a great deal of research has investigated lexical access by native speakers, very little research has addressed lexical access by non-native speakers. However, previous research in second language acquisition (SLA) studying grammatical effects in non-native speakers can be used to direct research in non-native lexical accesss. For example, chunking is a common concept in SLA by which learners encode phonological form in long term memory in chunks which may be comprised of multiple morphemes or words. This process has been termed the *phonological loop* by Baddeley (1976, 1997). According to Ellis (1996, 2001), much of learning the "rules" of a second language involves reanalyzing these chunks, such that the structures emerge in the linguistic knowledge of the learners. For example, learners of German as a foreign language are frequently taught common phrases such as *in der Stadt* 'in the city', which is marked for dative case, months before learning the dative case. Only after additional learning do they analyze the sub-chunks of the phrase, including grammatical information such as case marking. Chunking can also occur at the level of morphology. Experiment Two showed that there is a processing advantage for monomorphemic words compared to bimorphemic words for German native listeners. If second language learners are initially treating bimorphemic words as unanalyzed chunks, and then gradually reanalyzing the chunks into morphemes, the processing advantage of monomorphemes is predicted to be smaller for non-native listeners than for native listeners.

Another widely-studied concept in SLA is language transfer (Lado, 1957), by which learners of a second language carry over properties from their native language into the second language. Language transfer has traditionally been used to explain learners' difficulties in acquiring grammatical structures, e.g. speakers whose L1 does not contain determiners may have difficulty acquiring determiners in an L2. Models of cross-linguistic speech perception such as Best's (1995; 2003) Perceptual Assimilation Model (PAM) also appeal to the notion of language transfer. PAM hypothesizes that listeners hearing foreign phones for the first time will attempt to map these phones to acoustically similar phonemes in their native language, essentially transferring the phonological categories of their native language to the second language. For example, German speakers may map English /ɛ/ and /æ/ onto German /ɛ/, since German lacks the phoneme /æ/ (and this can be seen in the confusion matrices from Experiment Three in Appendix C.3). The concept of language transfer may also be extended to the domain of the lexicon as well. Experiments One and Two showed that morphology has a greater effect on lexical access for native listeners of German than for native listeners of English. If language transfer also affects lexical access, then native English speakers should not be as sensitive as native German speakers to differences in morphology when processing German. In this experiment testing lexical access by native English-speaking learners of German, both language transfer and chunking make the same predictions as to how non-native listeners will be affected by differences in morphology, but the two hypotheses make opposite predictions in Experiment Four.

In addition to a predicted difference in the effects of morphology on non-native spoken word recognition, the reduced vocabulary size and limited exposure to German for the non-native listeners could have several consequences for how context effects will impact lexical access. The reduced vocabulary size predicts that the effect of neighborhood density will be smaller, since there are fewer competing words. Frequency effects could also be reduced due to vocabulary size. It is difficult to assess frequency effects in non-native speakers, but one can hypothesize that very frequent words will also have been heard by non-native speakers with the highest frequency, and therefore will have similar effects for both native and non-native speakers. In contrast, words with medium to low frequency may essentially have a frequency of 0 in the minds of the non-native speakers, and therefore may be treated more like nonwords. The combined effect of these two hypotheses predicts that there should be a smaller difference between words and nonwords for non-native listeners, but that the effects of frequency should not be that different from native listeners.

Experiment Three investigates context effects in spoken word recognition by non-native listeners of German using the same materials and procedures as in Experiment Two. Results show that English-speaking listeners of German are sensitive to differences in lexical status, morphology, lexical frequency, and neighborhood density, though the degree of sensitivity is less than for native listeners of German.

6.1 Method

6.1.1 Participants

Thirty participants were recruited via flyer and advertisements in the German department at the University of Michigan. All participants reported being native speakers of American English and having no known hearing impairments. The participants can be characterized as intermediate/advanced learners of German; all had studied German at the college-level for at least five semesters, and had spent at least three months in a German-speaking country within the last five years. None of the participants had taken part in Experiment One or Two.

6.1.2 Materials

The materials were the same as those used in Experiment Two.

6.1.3 Procedure

The procedure was the same as that used in Experiment One, except that the instructions specified that the participants would hear German words and nonwords, as opposed to English, as in Experiment One.

6.2 Analysis

The analysis was mostly the same as for Experiments One and Two, except that the conversion from spelling to phonemes involved several additional parameters. Responses that seemed to be using English spellings were treated as the corresponding phonemes in German orthography, e.g. in response to the nonword *reungken* [rɔyŋkən], ⟨kroimkin⟩ was transcribed as [krɔymkɪn], treating the spelling ⟨oi⟩ as representing the sound normally spelled as ⟨eu⟩ in German orthography. In many cases it was not possible to make such assumptions, most notably with the phonemes [s z ʦ], written as ⟨ss⟩ or ⟨ß⟩, ⟨s⟩, and ⟨z⟩ respectively in German, and the former two as ⟨s⟩ and ⟨z⟩ (or sometimes ⟨s⟩) in English. There were a large number of ⟨z⟩ responses where [z] was expected. It is impossible to know whether the listeners simply misspelled the phone, or whether they actually heard [ʦ]. Given that many Americans learning German frequently pronounce ⟨z⟩ as [z], and given the fairly high degree of acoustic similarity between [z] and [ts], this is certainly plausible. Therefore, all ⟨z⟩ responses to [z] were counted as incorrect.

6.3 Predictions

Predictions for Experiment Three are largely the same as those for Experiment Two, though the size of the effects are predicted to differ somewhat. The difference in j between words and nonwords is predicted to be smaller than in Experiment Two, since a greater proportion of the words are likely to be unknown to non-native listeners, and will therefore be treated more like nonwords. The difference in j is also predicted to be smaller between mono- and bimorphemic words; this prediction follows from both a chunking account as well as a language transfer account of SLA. The difference in j between low- and high-frequency words is predicted to be roughly the same as in Experiment Two. Finally, the difference in j between words in sparse and dense neighborhoods is predicted to be smaller than in Experiment Two, since many of the neighbors for a given word are likely to be absent from the non-native listener's lexicon.

6.4 Results

The complete set of responses included 9600 trials (300 stimuli x 32 listeners), 53 ($\approx .5\%$) of which were discarded due to no response, leaving 9431 trials for analysis. The average phoneme (p_p) and (non)word (p_w) recognition probability scores are shown in Figure 6.1. The recognition rates for words were higher than for nonwords for both whole words and phonemes. In addition

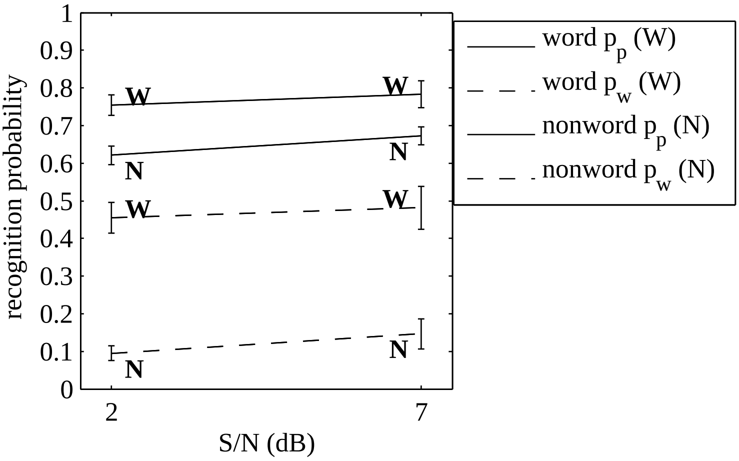


Figure 6.1 German non-native listener phoneme and word recognition probabilities — Each point represents the average phoneme or word recognition probability for words or nonwords at a given S/N. Error bars display 95% confidence intervals.

the recognition rates were all higher at S/N=7 than S/N=2. It can also be seen that the difference between p_w and p_p is much larger for nonwords than for words.

6.4.1 Subjects analysis

The results of the j-factor analysis by subjects are shown in Figure 6.2. Each panel displays the data grouped by one of the context effects in question. While the analysis here includes some comparisons between the results of this experiment and Experiment Two using native listeners, Chapter 8 provides a more detailed analysis of overall differences between native and non-native listeners.

Lexical Status

The effect of lexical status is large, with $j_{nonword}$ significantly higher than j_{word}, but the difference in j is smaller than for the native listeners in Experiment Two. Consistent with the results from Experiment Two, $j_{nonword} = 4.96$ is substantially lower than the predicted value of 6. As discussed in §5.5.3, the lower than predicted $j_{nonword}$ is likely due to presence of post-vocalic /ʀ/ in the stimuli, as well as the trochaic syllable structure. Also consistent with Experiment Two, $j_{word} \approx 3.81$ is much lower than predictions based on previous findings, suggesting that j_{word}, may not scale linearly with word length. Finally, the difference in j between words and nonwords is smaller for the English-speaking listeners in this experiment than the German-speaking listeners in Experiment Two ($\Delta j_{native} = 1.47, \Delta j_{non-native} = 1.14, p < .05$). This difference can be attributed to the higher j-scores for words for the non-native listeners, which indicates that some of the real words were treated as nonwords by the non-native listeners.

Morphology

As predicted, j_{bi} was significantly higher than j_{mono}. In addition, the difference in j was smaller than for native listeners ($\Delta j_{native} = .8, \Delta j_{non-native} = .3$), indicating that the non-native listeners are

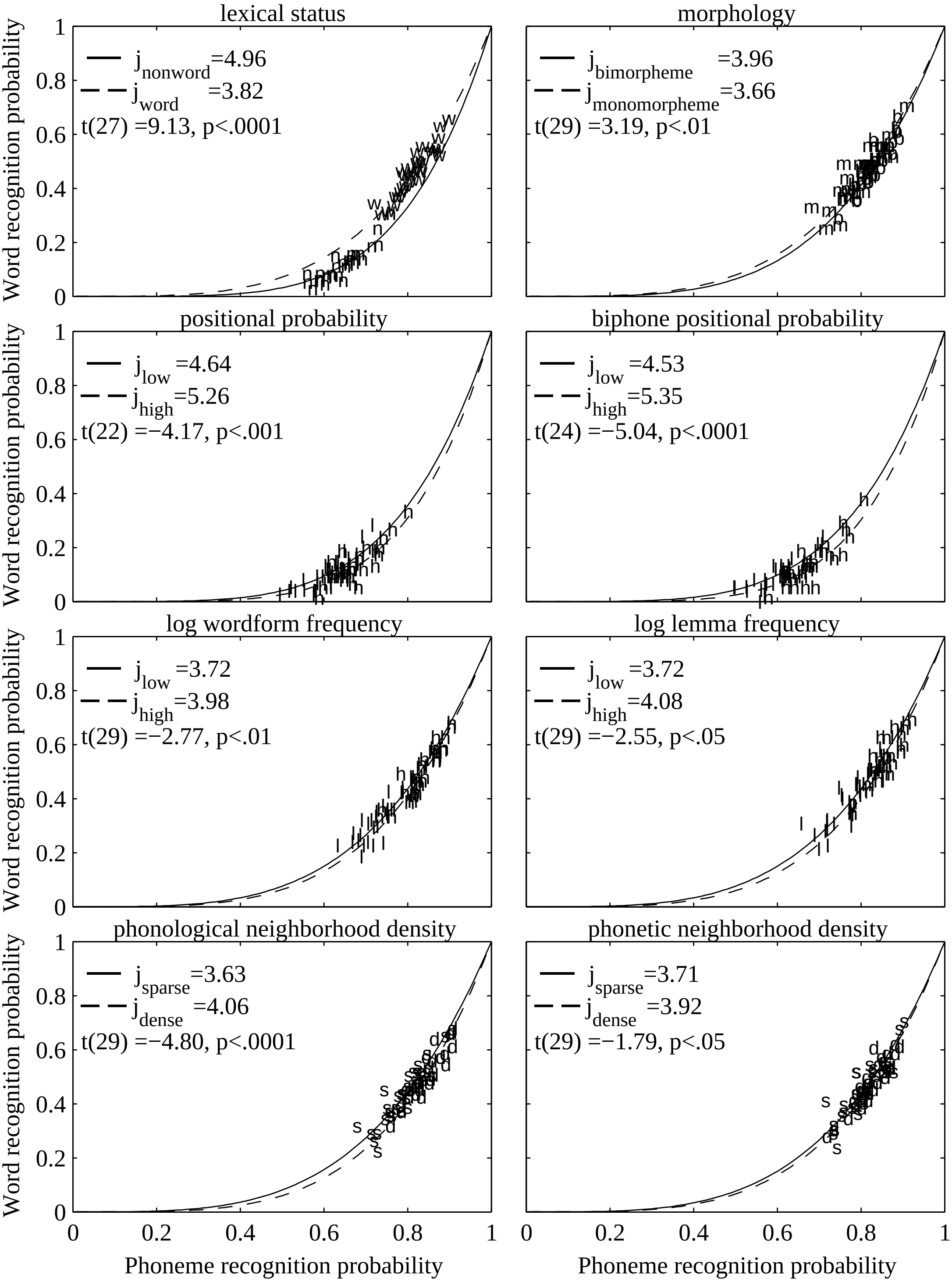


Figure 6.2 German non-native listener j-factor results — Each plot compares two subsets of results from the subject analysis. Each point represents the average results for one subject. Curves represent $y = x^j$. The second row of plots only shows nonword results, while the final two rows only display word results. Statistics shown are from paired t-tests (one-tailed for plots in rows 1 and 4; two-tailed for plots in rows 2 and 3); before computing the statistics, all points lying in the floor or ceiling ranges ($> .95$ or $< .05$) were removed, but are still shown on the plot.

also sensitive to effects of morphology, but not as sensitive as native speakers.

Phonotactic probability

As in Experiment Two, effects of phonotactic probability were also investigated for nonwords, as shown in the second row of Figure 6.2. Both the results based on positional probability and biphone positional probability are significant, but are opposite of the predicted results, namely that high phonotactic probability words should be treated as more word-like, and therefore have a lower j. It is difficult to interpret the finding that words with higher phonotactic probability have a higher j, especially since phonotactic probability did not have a significant effect in the native listener experiment.

Lexical frequency

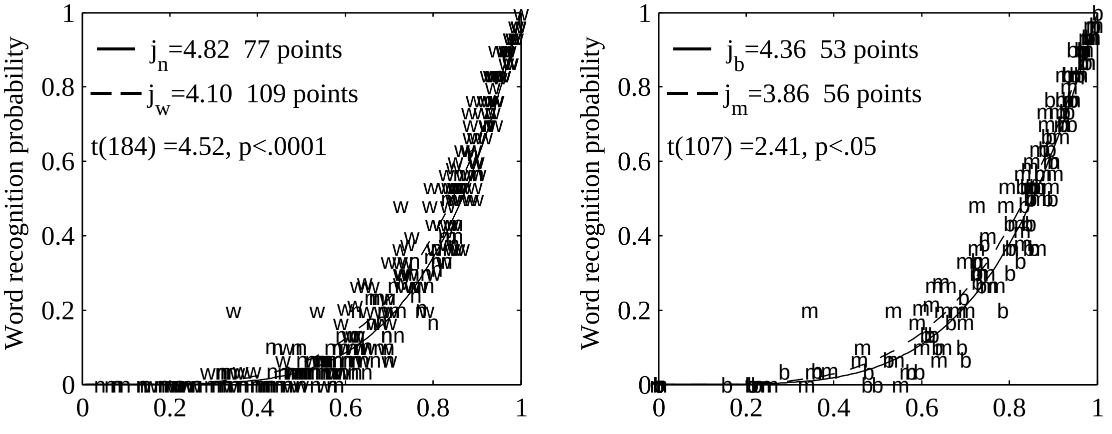
Consistent with the results from Experiment Two, but inconsistent with the predicted results, high-frequency words had significantly higher j-scores than low-frequency words, though the difference in j was smaller than the difference found for native listeners in Experiment Two ($\Delta j_{native} = .69, \Delta j_{non-native} = .26$). This adds support to the interpretation that the unexpected results of lexical frequency in this study are due to the selected stimuli, but the exact reason is still unknown. It seems that the inhibitory effect of lexical frequency found in both German experiments is due to either the words chosen, or the way in which the speaker pronounced the words.

Neighborhood Density

Consistent with the native listener results from Experiment Two, words in dense neighborhoods had higher j-scores than those in sparse neighborhoods, though the difference in j as measured by phonetic neighborhood density is significantly smaller than for native speakers ($\Delta j_{native} = 1.11, \Delta j_{non-native} = 0.21, p < .001$). In contrast to Experiment Two, the effect of phonetic neighborhood density did not reach significance. Differences between phonetic and phonological measures of neighborhood density will be discussed in more detail in Chapter 8.

6.4.2 Items analysis

The main effect of lexical status was also quite robust in the items analysis, as shown in Figure 6.3a, with $j_{nonword}$ significantly higher than $_{word}$. Also consistent with the subjects analysis, bimorphemic words exhibited significantly higher j-scores than monomorphemic words, as shown in Figure 6.3b. The remaining results of the items analysis are shown in Figure 6.4 using regression analyses.



(a) lexical status (b) morphology

Figure 6.3 German non-native listener j-factor results by items

Lexical frequency and stimulus probability

Though significant in the subjects analysis, log wordform frequency is not significant in the items analysis. There is a significant positive correlation between j and log lemma frequency ($r = .261, p < .01$), though it only accounts for 6.8% of the variation in j. Neither stimulus probability nor frequency-weighted stimulus probability is significantly correlated with j.

Neighborhood density

The outcome of the items analysis of neighborhood density differs somewhat from the subjects analysis. Whereas the subjects analysis showed a significant effect of phonological neighborhood density and phonetic neighborhood density was insignificant, the opposite is found in the items analysis. Phonological neighborhood density is not significantly correlated with j, but phonetic neighborhood density is positively correlated with j ($r = .297, p < .01$). The mixed results of neighborhood density between the subjects and items analyses suggests that neighborhood density has a smaller and less consistent effect on lexical access for non-native listeners of German than for native listeners.

Phonotactic probability

The results from the items analysis of phonotactic probability were insignificant, both using the positional probability and the biphone probability measures.

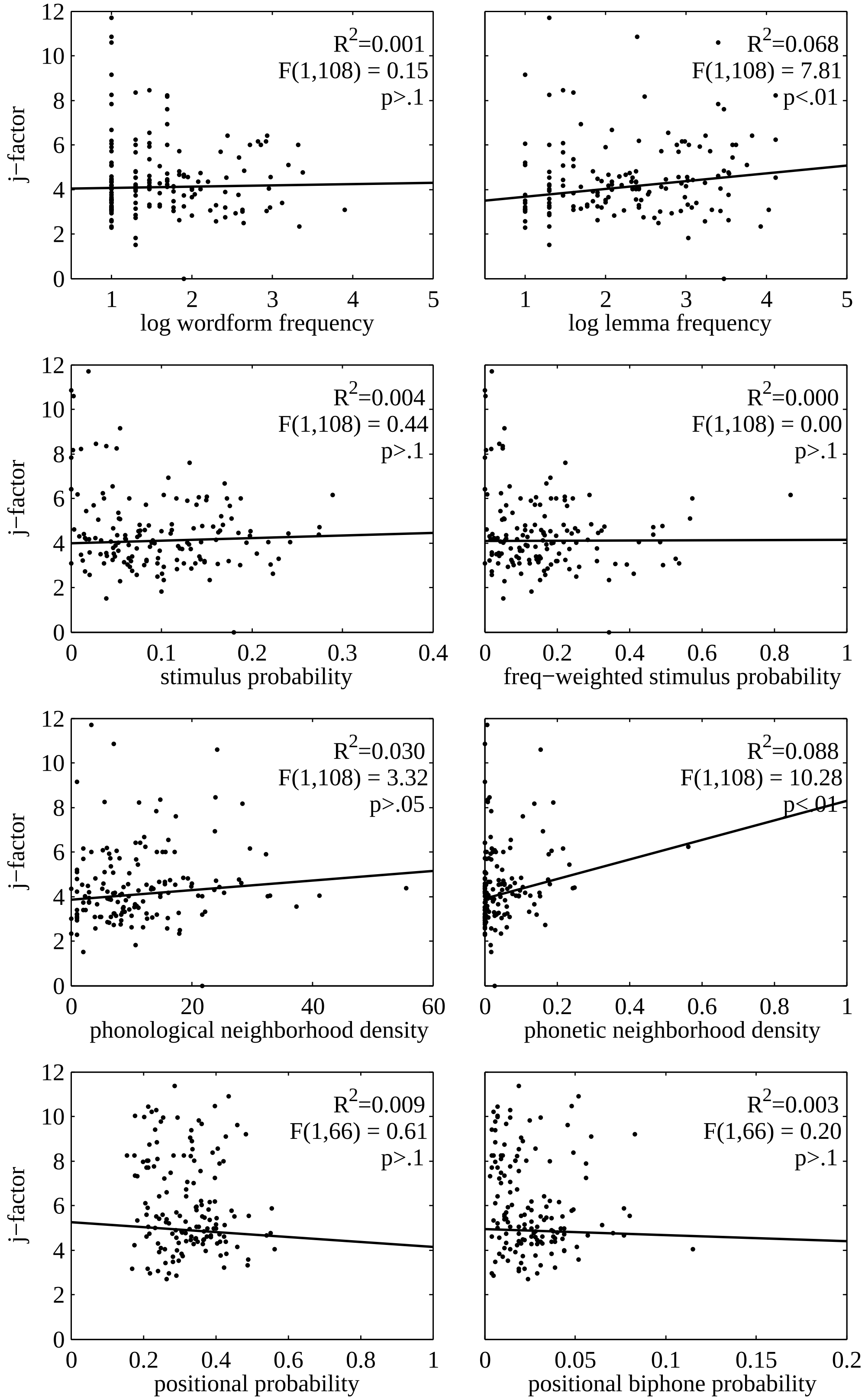


Figure 6.4 German *j*-factor regression analyses — Each panel plots *j*-factor as a function of one particular lexicostatistical measure. Each point represents one item. The top 6 panels show only word items, while the bottom two show only nonword items. The statistics given are from linear regressions.

6.5 Discussion

6.5.1 Context Effects

Consistent with previous studies investigating lexical access by non-native listeners of English (Bradlow & Pisoni, 1999; Imai et al., 2005), results from this experiment show that non-native listeners of German are sensitive to the same context effects as native listeners, though the size of the effects were generally smaller. The smaller difference in j between words and nonwords is consistent with the hypothesis that the smaller vocabulary size of non-native listeners of German causes some of the words to be treated as nonwords. The smaller difference in j between mono- and bimorphemic words is consistent both with a chunking or a language transfer account of SLA. Experiment Four provides an experimental situation in which these two accounts make opposite predictions of how differences in morphology should affect lexical access by non-native listeners. The effect size of neighborhood density was also smaller for non-native listeners of German than for native listeners of German, which is consistent with an explanation based on a reduced vocabulary size. Further comparisons of native and non-native listener results are given in §8.2.

6.5.2 Morphology and response bias

The results of the SDT analysis of a subset of the data in Experiment Two revealed an interaction among morphology, perceptual distinctiveness, and response bias. In particular, the SDT analysis revealed that listeners seem to be aware of and take advantage of the lexicostatistical properties of the language. The same SDT analysis was also carried out with the data from this experiment, in order to test whether the non-native listeners of German respond similarly.

To carry out the SDT analysis, the original confusion matrices for each S/N were transformed into 2x2 submatrices. An SDT analysis was then applied to each submatrix. The results, shown in Table 6.1, are very similar to the results for native listeners. The trends of d' and c are the same as those of the native listeners, though the non-native listeners have lower d' values on average and higher values of c. That is, the non-native listeners' responses show an even stronger bias for /n/ than those of the native listeners. One possible explanation for the increased bias is that the non-native listeners have an increased sensitivity to the lexicostatistics of the language, but this seems rather implausible. A more probable explanation is that the L2 lexicon has different statistical properties than the L1 lexicon. As displayed in Table 3.1 in §3.1, the /m/ inflectional ending for adjectives occurs only in the masculine and neuter singular dative strong declension, whereas /n/ occurs in both singular and plural, in all cases, in all genders, and in both the strong and weak declensions. Moreover, students learning German are generally taught the nominative and accusative cases before the dative case, and frequently have a difficult time learning the dative case. Thus it is not unreasonable to assume that L2 German speakers have heard *-m* used as an inflectional suffix proportionally less than L1 speakers, but have heard the *-n* suffix in approximately the same proportion, which is consistent with the greater bias for /n/ in the non-native listeners' responses.[1] The results of the SDT analysis show that non-native listeners behave very similarly to

---

[1] It is possible that acoustic differences between the stimuli in these three groups is actually driving the perceptual differences; but that question is outside the scope of the current study.

Table 6.1 Signal Detection Theory analysis of /m/ and /n/ submatrix in final position comparing native and non-native listeners — (a) repeats the results from Experiment Two for native listeners; (b) shows results for non-native listeners. For this analysis /m/ is considered to be the target stimulus. Positive values of *c* indicate a bias towards /n/. The final two columns list the total number of presentations of /m/ and /n/ which were used to compute the SDT analysis

(a) Native listeners

	d'	c	/m/	/n/
Nonwords				
lower S/N (2 dB)	-0.182	0.555	240	240
higher S/N (7 dB)	0.664	0.743	240	240
Bimorphemes				
lower S/N (2 dB)	1.616	0.984	128	352
higher S/N (7 dB)	1.913	0.556	128	352
Monomorphemes				
lower S/N (2 dB)	3.514	0.239	48	192
higher S/N (7 dB)	4.733	-0.060	48	192

(b) Non-native listeners

	d'	c	/m/	/n/
Nonwords				
lower S/N (2 dB)	-0.201	0.851	225	225
higher S/N (7 dB)	0.116	1.026	225	225
Bimorphemes				
lower S/N (2 dB)	0.964	1.510	120	330
higher S/N (7 dB)	1.128	1.436	120	330
Monomorphemes				
lower S/N (2 dB)	2.386	0.641	45	180
higher S/N (7 dB)	3.301	0.636	45	180

native listeners, depending on a combination of acoustics, lemma frequency, and morphological distribution in spoken word recognition, but that differences in the L2 lexicon lead to slight differences in the amount of response bias.

6.6 Conclusions

The results from this experiment have shown that non-native listeners of German are sensitive to the same context effects as native listeners, though the size of the effects are generally smaller. In particular, non-native listeners of German also exhibited a processing advantage for monomorphemic words over bimorphemic words, which is consistent with both a chunking and a language transfer account of SLA. While these two accounts make the same predictions in terms of the effect of morphology in this experiment, Experiment Four provides an experimental design in which these two accounts make opposite predictions.

In addition to the effect of morphology, results from this experiment show that non-native listeners of German are also sensitive to lexical status and neighborhood density, though not as sensitive as native listeners. This pattern is consistent with the smaller vocabulary of non-native listeners.

Chapter 7

Experiment Four — Recognition of English CVCCVC words and nonwords by non-native listeners

THE results of Experiment Three showed that English-speaking learners of German are sensitive to the same context effects in lexical access as are native listeners of German, though the effects are generally not as large. Both native and non-native listeners of German enjoyed a processing advantage of monomorphemic words over bimorphemic words, but this advantage was not as large for the non-native listeners in Experiment Three as for the native listeners in Experiment Two. These results are predicted by both chunking accounts as well as language transfer accounts of second language acquisition. The chunking account predicts that second language learners should be less sensitive to morphological patterns than native speakers, regardless of the L1 of the learners. In contrast, language transfer accounts maintain that the degree of sensitivity to morphological patterns in a second language can be predicted by the amount of sensitivity to morphological patterns in the L1. In Experiment Three, native speakers of English listened to German words. Since the comparison of Experiments One and Two showed that native listeners of English are less sensitive to morphological patterns than native speakers of German, a language transfer account predicts that English speakers learning German will also be less sensitive to morphological patterns in German than native speakers of German.

Experiment Four, which tests native speakers of German listening to English, provides an experimental design in which these two accounts make opposite predictions. The chunking account still predicts that the sensitivity to morphological patterns should be less for non-native listeners than for native listeners, while the language transfer account predicts that this group of non-native listeners could be more sensitive to morphological patterns than native listeners, since their native language, German, is morphologically richer than English, and results from Experiments One and Two showed that native German listeners are more sensitive to morphological patterns than native English listeners. Similar to Experiment Three, results show that German-speaking listeners of English are sensitive to differences in lexical status, morphology, lexical frequency, and neighborhood density, but that they are generally not as sensitive to these effects as native English listeners.

7.1 Method

7.1.1 Participants

Thirty-two participants were recruited via flyer from the University of Konstanz. All participants reported being native speakers of German and having no known hearing impairments. The participants can be characterized as intermediate/advanced learners of English; all had studied English at the *Gymnasium* (University-track high school in Germany) for at least six years. None of the participants had taken part in any of the prior experiments in this study.

7.1.2 Materials

The materials were the same as those used in Experiment One.

7.1.3 Procedure

The procedure was the same as for Experiment Two, except that listeners were told they would be hearing English words and nonwords, and different signal-to-noise-ratios (S/Ns) were used. As in Experiment One, pilot results for this experiment showed a very large difference between words and nonwords, such that finding two S/Ns that would fit into the range between 5% and 95% both for word and phoneme recognition for both words and nonwords was nearly impossible. Therefore the compromise employed in Experiment One was also used in this experiment, such that for each participant, the nonword stimuli S/N was 5 dB higher than the word stimuli. Thus instead of using two different S/Ns two pairs of S/Ns were used. Half of the participants heard words presented at S/N=0 dB and nonwords at S/N=5 dB, and half of the participants heard words presented at S/N=5 dB and nonwords at S/N=10 dB. In the results, the lower pair (0 and 5 dB) will simply be referred to as 0 dB and the higher pair (5 and 10 dB) will be referred to as 5 dB.

7.2 Analysis

The data from this experiment were analyzed in the same manner as the other experiments, described in detail in §3.3. Similar to Experiment Three, the conversion to phonemes also considered both English and German spellings for words. For example, /ɔɪ/ is usually spelled as ⟨oy⟩ in English, but as ⟨eu⟩ in German. Both of these responses were coded as /ɔɪ/ in this experiment.

7.3 Predictions

Predictions for Experiment Four are largely the same as those for Experiment One, though the magnitude of the effects are predicted to be somewhat different. The difference in j between words and nonwords is predicted to be smaller than in Experiment One, since a greater proportion of

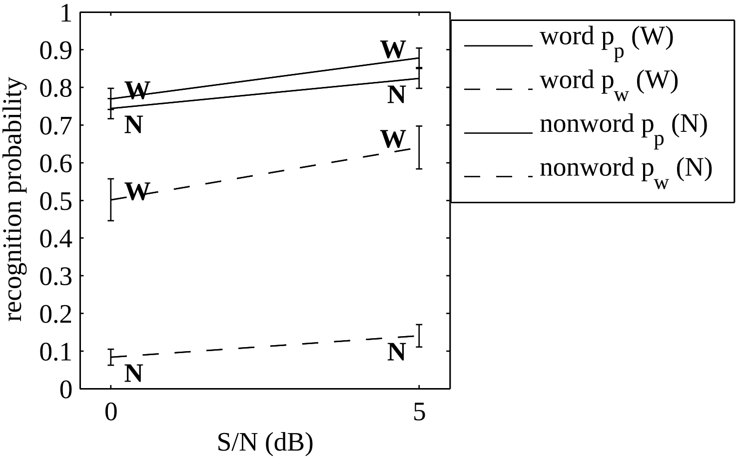


Figure 7.1 German phoneme and word recognition probabilities — Each point represents the average phoneme or word recognition probability for words (W) or nonwords (N) at a given S/N. Error bars display 95% confidence intervals.

the words are likely to be unknown to non-native listeners, and will therefore be treated more like nonwords. There are two possible predictions for the difference in *j* between mono- and bimorphemic words. A chunking account predicts that the difference in *j* between mono- and bimorphemic words should be smaller for non-native listeners than for native listeners, since non-native listeners may be treating some of the bimorphemic words as unanalyzable chunks. In contrast, a language transfer account predicts that the difference in *j* between mono- and bimorphemic words should be greater for non-native listeners (L1=German) than for native listeners, because Experiment Two showed that native listeners of German are more sensitive to differences in morphology than native listeners of English. A language transfer account would predict that this increased sensitivity to morphological patterns for German speakers could carry over when learning a second language. The results of this experiment will be able to distinguish between these two hypotheses of lexical access by non-native listeners. The difference in *j* between low- and high-frequency words is predicted to be roughly the same as in Experiment One. Finally, the difference in *j* between words in sparse and dense neighborhoods is predicted to be smaller than in Experiment One, since many of the neighbors for a given word are likely to be absent from the non-native listener's lexicon.

7.4 Results

The complete set of responses included 9600 trials (300 stimuli x 32 subjects), 498 ($\approx 5\%$) of which were discarded because participants did not provide any response, thus leaving 9102 trials for analysis. The average phoneme (p_p) and (non)word (p_w) recognition probability scores are shown in Figure 7.1. Consistent with the results from native speakers, the recognition rates for words were higher than for nonwords for both whole words and phonemes. In addition the recognition rates were all higher at S/N=5 than S/N=0.

7.4.1 Subjects analysis

The results of the subjects analysis are shown in Figure 7.2. Each panel displays the data grouped by one of the context effects in question. All comparisons between effects in native and non-native

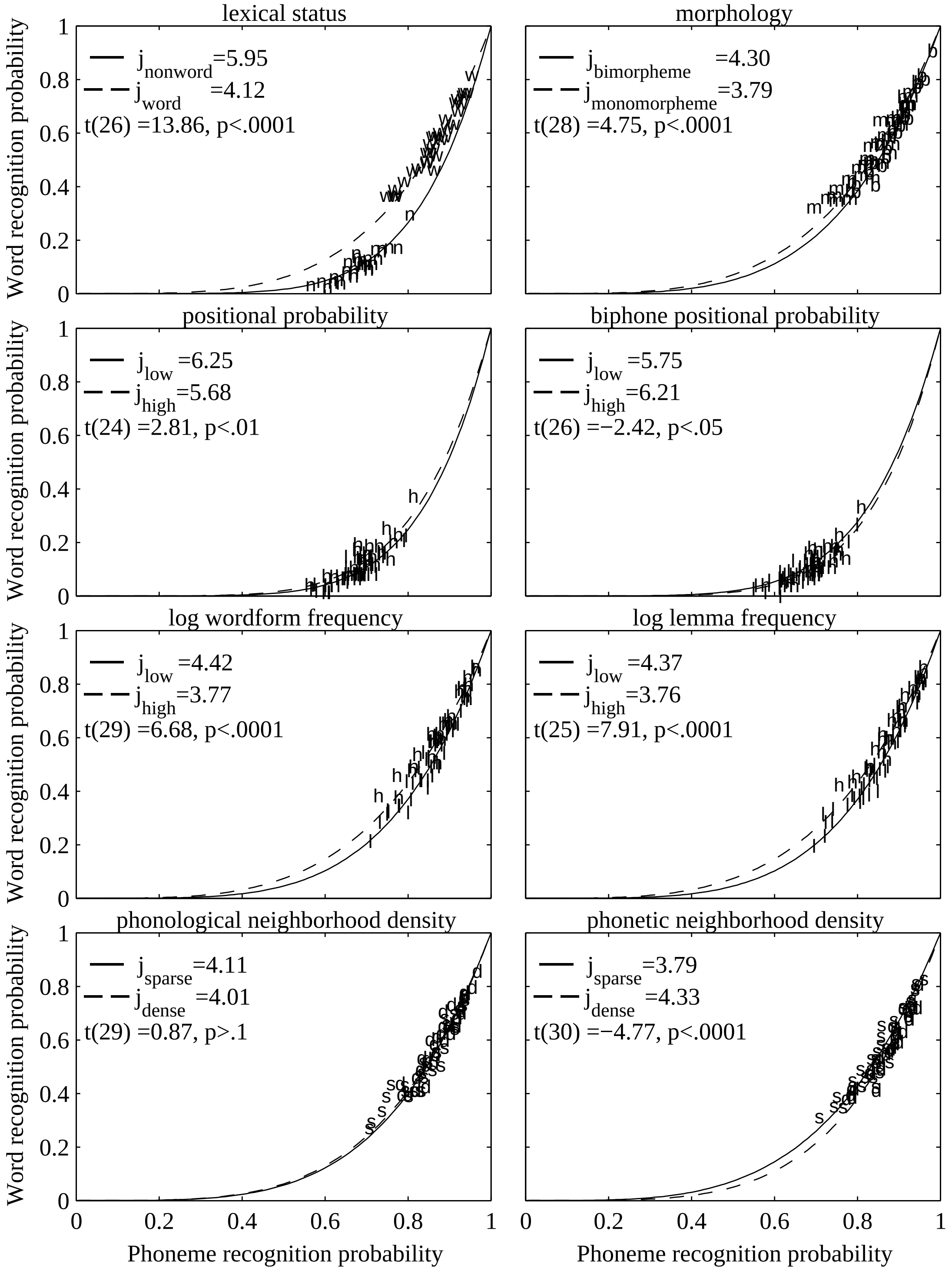


Figure 7.2 English non-native listener *j*-factor results — Each plot compares two subsets of results from the subject analysis. Each point represents the average results for one subject. Curves represent $y = x^j$, for the mean j of each category. The second row of plots only shows nonword results, while the final two rows only display word results. Statistics shown are from paired t-tests (one-tailed for plots in rows 1 and 4; two-tailed for rows 2 and 3); before computing the statistics, all points lying in the floor or ceiling ranges ($> .95$ or $< .05$) were removed, but are still shown on the plot.

listeners mentioned here are statistically significant. For detailed statistics comparing the native and non-native listener results, see Table 8.1 and Table 8.2.

Lexical Status

As predicted, and consistent with the results from Experiment One, the j-score for nonwords was significantly higher than for words. Also as predicted, the word-nonword difference in j for non-native English listeners was not as large as for native listeners (native listener $\Delta j = 2.18$; non-native listener $\Delta j = 1.83$), mostly due to a higher j-score for words.

Morphology

As predicted, j of bimorphemic words was significantly higher than that of monomorphemic words. In addition, the difference in j was smaller than for native listeners (native listener $\Delta j = .91$; non-native listeners $\Delta j = .51$), which is consistent with a chunking account of non-native lexical access.

Phonotactic probability

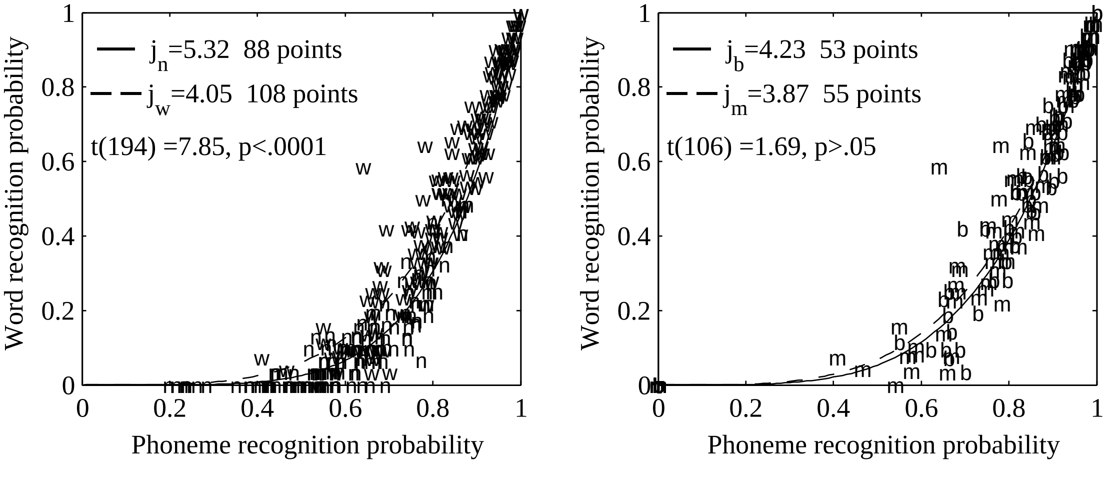
The predicted effect of positional probability was significant, with lower probability nonwords showing a higher j-score than higher probability nonwords, but the effect of biphone positional probability was in the opposite direction. Effects of phonotactic probability in Experiment Two were also mixed in this same way. Given these mixed results and the fact that previous studies investigating effects of phonotactic probability in nonwords have found very small effects (Vitevitch & Luce, 1998, 1999), no interpretation of the effects of phonotactic probability is possible for this experiment. Several possible explanations for the mixed results are discussed in §5.4.1, and hold for these findings as well.

Lexical frequency

Consistent with results from Experiment One and the predictions for the current experiment, high-frequency words had a lower j-score than low-frequency words for both lemma-based and wordform-based frequency measures, indicating a facilitatory effect of lexical frequency. As predicted, no difference was found in the size of the effect between native and non-native listeners (wordform frequency: native listeners $\Delta j = .57$; non-native listeners $\Delta j = .65$ — lemma frequency: native listeners $\Delta j = .35$; non-native listeners $\Delta j = .61$).

Neighborhood Density

As predicted, the effect size of neighborhood density was smaller than in Experiment One. No significant effect of phonological neighborhood density was found, and the difference in j between



(a) lexical status (b) morphology

Figure 7.3 German j-factor results by items

words in sparse and dense neighborhoods using a phonetic measure of neighborhood density was much smaller for non-native listeners (L1 = German) than for native listeners in Experiment One (native listeners $\Delta j = .96$; non-native listeners $\Delta j = .54$).

7.4.2 Items analysis

The results of the items analysis of lexical status and morphology is shown in Figure 7.3. The remaining results of the items analysis are shown in Figure 7.4 using a regression analysis.

Lexical Status

Consistent with results from the subjects analysis, nonwords had significantly higher j-scores than words, as shown in Figure 7.3a.

Morphology

Though the subjects analysis found that bimorphemic words had significantly higher j-scores than monomorphemic words, no significant difference in j was found in the items analysis, as shown in Figure 7.3b. The lack of significance is likely due to the increased variance in the items analysis, and it is worth noting that the trend ($j_{bi} >_{mono}$) is consistent with the subjects analysis.

Lexical frequency and stimulus probability

Consistent with the subjects analyses, there were significant negative correlations between j and both wordform and lemma frequency. Log wordform frequency accounted for $\approx 6\%$ of the variation

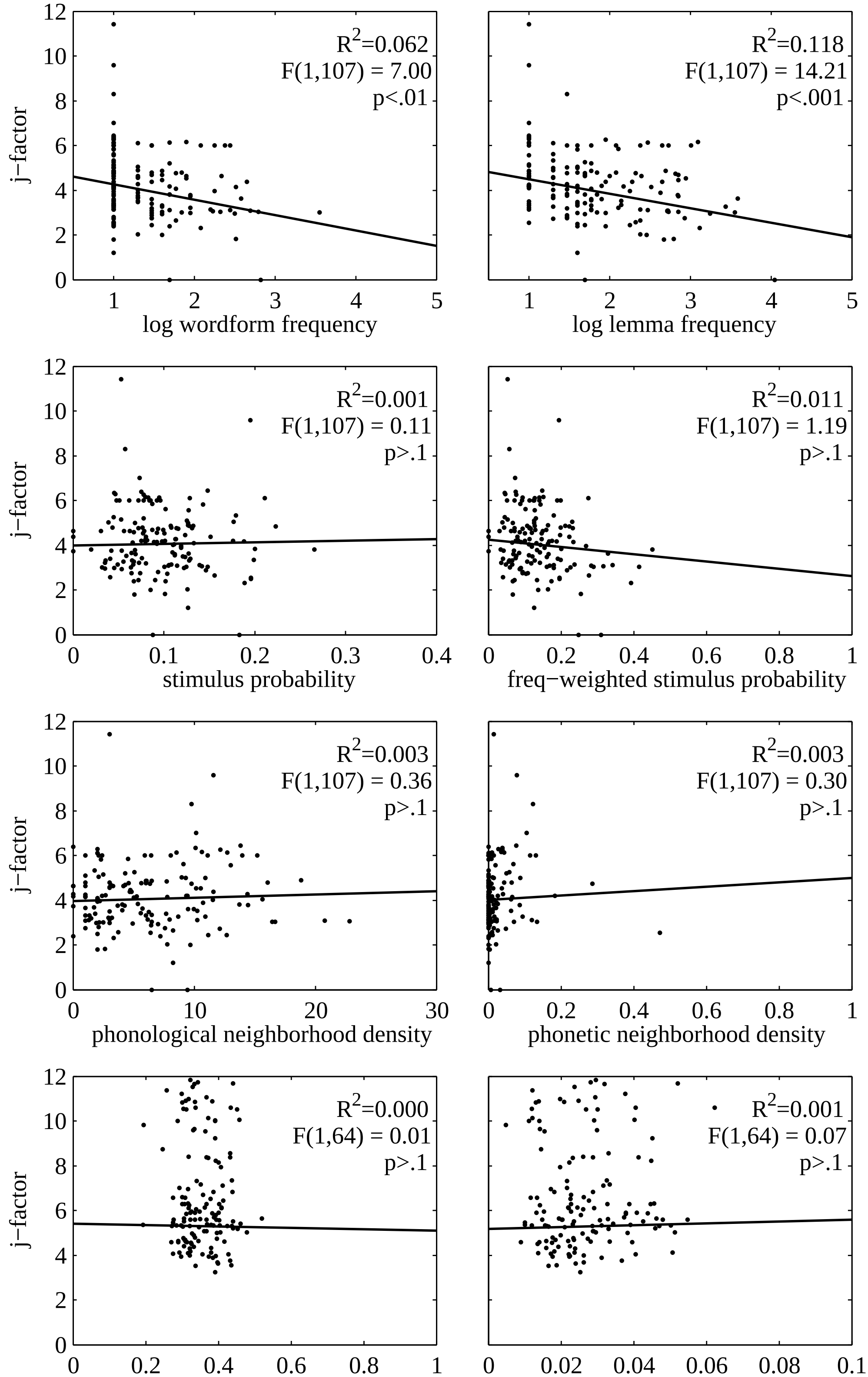


Figure 7.4 German *j*-factor regression analyses — Each panel plots *j*-factor as a function of one particular lexicostatistical measure. Each point represents one item. The top 6 panels show only word items, while the bottom two show only nonword items. The statistics given are from linear regressions.

in j, while log lemma frequency accounted for $\approx 12\%$ of the variation in j. The effect of log wordform frequency was highly consistent with the results from Experiment One, which also accounted for $\approx 6\%$ of the variation in j, but the effect of log lemma frequency was actually greater than for native listeners' results.

As in Experiment One, the correlation between j and stimulus probability was also measured. There was no significant effect of either stimulus probability or frequency-weighted stimulus probability.

Neighborhood density

Whereas in the subjects analyses, phonetic, but not phonological, neighborhood density showed a significant result, neither the effect of phonological nor phonetic neighborhood density was significant in the items analyses. The discrepancy between the results of the subjects and items analyses for phonetic neighborhood density most likely lies in the fact that the distribution of phonetic neighborhood density is highly skewed towards sparse words.

Phonotactic probability

While the subjects analyses of phonotactic probability yielded significant, but conflicting results, neither the effect of positional probability nor the effect of biphone positional probability reached significance in the items analyses.

7.5 Conclusions

A central aim of this experiment was to test conflicting predictions of the effects of morphology on word recognition by non-native listeners made by chunking and language transfer accounts of SLA. The results, in which German speakers listening to English exhibited a smaller processing advantage of monomorphemes over bimorphemes than native English listeners, are consistent with a chunking account of SLA. That is, regardless of the morphological structure of their L1, non-native listeners seem to be less sensitive to differences in morphology than native listeners, although further research with other languages must be carried out in order to verify the universality of this result.

Chapter 8

General Discussion

As discussed in Chapter 2, the role of morphology in lexical access has been widely contested for over 20 years now. This debate has centered around whether or not there is a morphological level of representation in the lexicon. Combinatorial models of lexical access have argued in favor of a morphological level of representation, and have used evidence from a variety of tasks which show that differences in morphology (e.g. regular vs. irregular inflection) can have an effect on lexical access (e.g. Pinker & Prince, 1988; Prasada & Pinker, 1993; Marcus et al., 1995; Clahsen, 1999; Clahsen et al., 2001; Gumnior et al., 2006). In contrast, associative models of lexical access claim that words are stored whole in the lexicon, and that "morphological processing reflects a learned sensitivity to the systematic relationships among the surface forms of words and their meanings" (Plaut & Gonnerman, 2000: 478). While several studies have found that associative models can accurately simulate effects of morphology in experimental data by finding patterns in phonology, semantics, or other properties of words such as lexical frequency (Plaut & Gonnerman, 2000; Baayen & Martin, 2005), there are also several studies which have found morphological effects even when controlling for phonology and semantics (e.g. Roelofs, 1996; Gumnior et al., 2006). The present study has also addressed the effects of morphology while controlling aspects of the phonological structure of the stimuli, and using a task which does not explicitly require retrieval of semantic information. In addition, this study has investigated effects of morphology across languages and between native and non-native listeners.

Previous research has shown that effects of morphology on lexical access are are dependent on both the language of the experimental materials (Marslen-Wilson, 2001; Plaut & Gonnerman, 2000) and the type of task used (Feldman, Soltano, Pastizzo, & Francis, 2004). While many studies have investigated morphological effects in both English (Rumelhart & McClelland, 1986; Prasada & Pinker, 1993; Sereno & Jongman, 1997) and German (Marcus et al., 1995; Hahn & Nakisa, 2000; Clahsen et al., 2001; Hahne et al., 2006), the majority of these studies have used visual tasks, and none of them has used an open response task. As mentioned in §3.1, lexical access in the visual domain may be more sensitive to morphological effects than in the aural domain, since visual stimuli, unlike aural stimuli, do not unfold over time, and visual stimuli also are not affected by morphophonological variants (e.g. English past tense *-ed* can be phonologically realized as /t/, /d/, or /əd/). As discussed in more detail below, this study has confirmed that effects of morphology can also be found in open response spoken word recognition.

Table 8.1 j-factor analysis summary for all four experiments. The effect size, as measured by the difference in j, is shown for each of the six context effects under investigation. Statistics shown are from paired t-tests on subjects (all one-tailed except for frequency, which are two-tailed). Positive values indicate facilitatory effects, while negative values indicate inhibitory effects

	Lexical Status	Morphology	Log wordform frequency	Log lemma frequency	phonological neighborhood density	phonetic neighborhood density
Native listeners of English	2.17***	0.91***	0.57**	0.35*	-0.37*	-0.96***
Non-native listeners of English	1.83***	0.50***	0.65***	0.61***	0.10	-0.54***
Native listeners of German	1.47***	0.80***	-0.69***	-1.00***	-0.38**	-1.11***
Non-native listeners of German	1.14***	0.30**	-0.26**	-0.36*	-0.43***	-0.21*

***$p < .001$, **$p < .01$, *$p < .05$

Table 8.2 Comparison of context effects across experiments. For each context effect, the difference in j was computed, comparing English native to German native listeners, English native to English non-native listeners, and German native to German non-native listeners. Statistics shown are from 2-sample t-tests

	Lexical Status	Morphology	Log wordform frequency	Log lemma frequency	phonological neighborhood density	phonetic neighborhood density
English vs. German	0.70***	0.12	1.25***	1.35***	0.00	0.15
German native vs. non-native	0.33*	0.50***	0.43**	0.64***	0.06	0.90***
English native vs. non-native	0.34*	0.41*	0.08	0.25	0.48*	0.41*

***$p < .001$, **$p < .01$, *$p < .05$

8.1 Cross-linguistic differences in the mental lexicon

In this cross-linguistic study, six different context effects (lexical status, morphology, wordform frequency, lemma frequency, phonological neighborhood density, and phonetic neighborhood density) were investigated using four separate groups of participants (English native listeners, English non-native listeners (L1=German), German native listeners, and German non-native listeners (L1=English)). The effect sizes for each of the six context effects investigated in the four experiments are summarized in Table 8.1 (and plotted in Figure 8.1), and the differences between effect sizes for different listener groups are shown in Table 8.2. In these tables, and elsewhere in the text, the effect size of various context effects is measured by the difference in j between two groups, e.g. the effect size of lexical status is measured by the difference in j between words and nonwords. Facilitatory effects result in a decrease in j, e.g. lexical status ($j_{word} < j_{nonword}$), while inhibitory effects result in an increase in j, e.g. neighborhood density ($j_{dense} > j_{sparse}$).

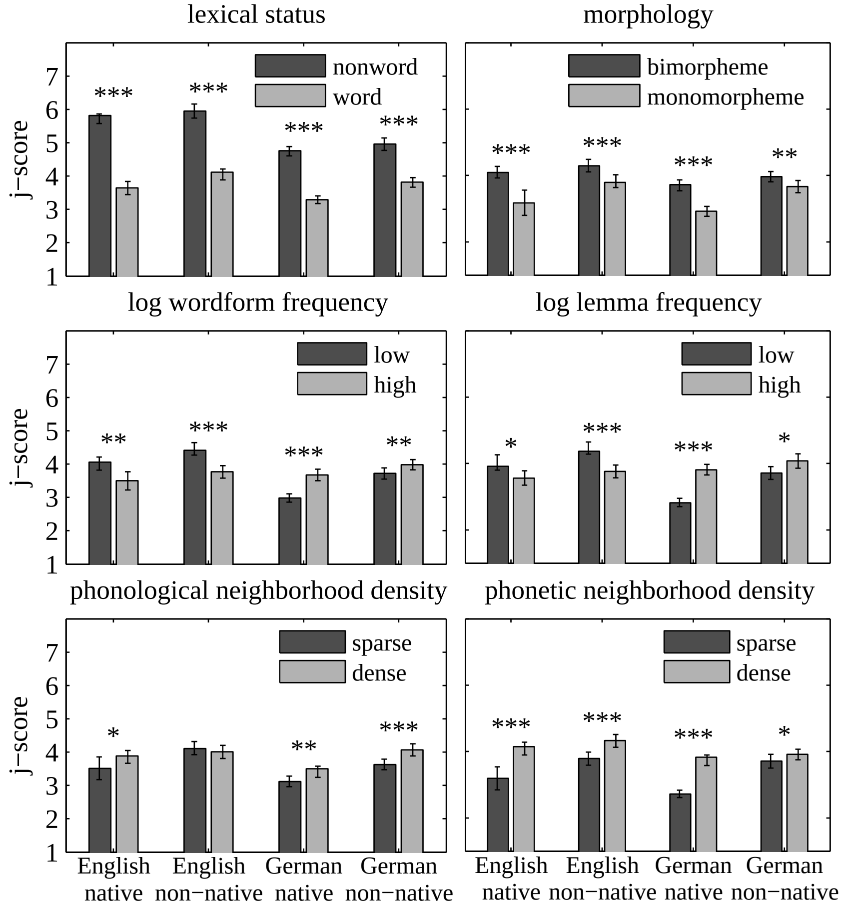


Figure 8.1 Summary of j-factor results. Each plot shows the means for each context effect investigated in Experiments One through Four. Error bars represent 95% confident intervals. Statistical significance for each comparison was computed from paired t-tests.
$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

8.1.1 Lexical status

In agreement with previous studies (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Olsen et al., 1997; Benkí, 2003a), the j-scores for nonwords were significantly higher than for words for all four experiments. Unexpectedly, the difference in j between words and nonwords was significantly larger for the English materials than for the German materials (see Table 8.1, first and third rows, and Table 8.2, first row). As can be seen in Figure 8.1, $j_{nonword}$ was much smaller than predicted for the experiments using German stimuli, and also smaller than $j_{nonword}$ in the experiments using English stimuli, while j_{word} was much more similar across languages. Thus the lower than expected $j_{nonword}$ values for the German experiments are responsible for the differences in the effect size of lexical status across languages. The lower-than-expected $j_{nonword}$ in the German experiment was partially explained by removing items which contained post-vocalic /ʀ/, which is usually not realized phonetically as a consonant in German, but rather as an off-glide of the preceding vowel. The high degree of interdependence between post-vocalic /ʀ/ and the preceding vowel lowered the overall j-scores of these items. While removing items containing post-vocalic /ʀ/ partially accounted for the low nonword j-scores, the recomputed value $j_{nonword} \approx 5$ is still substantially lower than the predicted value of 6, and also lower than the results from the English experiment of $j_{nonword} = 5.82$. Differences in phonotactic probability could account for the remaining discrepancy, but the effects of phonotactic probability did not reach significance in the German experiment. However, a more fine-grained analysis of the data can shed some light on this

Table 8.3 Partial j-scores — For each experiment, 2-phoneme j-scores are listed, which represent the amount of independence for the various sequences. For the German experiments, additional analyses are reported excluding items with post-vocalic /ʀ/. A score of 2 represents complete independence, while a score of 1 represents complete dependence. For each word–nonword pair, 2-sample t-tests were performed by subjects, testing the hypothesis that the partial j-score differed between words and nonwords for each 2-phoneme pair; asterisks in the table indicate a statistically significant difference between word and nonword results for a given 2-phoneme pair in a given experiment.

	English native		English non-native		German native		German non-native	
	nonword	word	nonword	word	nonword	word	nonword	word
C1V1	1.92***	1.68	1.92***	1.7	1.64***	1.49	1.64	1.66
no /ʀ/					1.85***	1.64	1.83	1.81
V1C2	1.87***	1.59	1.92***	1.68	1.59***	1.45	1.69***	1.59
no /ʀ/					1.84***	1.64	1.93***	1.75
C2C3	1.88***	1.44	1.81***	1.51	1.73***	1.48	1.63***	1.48
no /ʀ/					1.79***	1.52	1.78***	1.55
C3V2	1.92***	1.75	1.93***	1.82	1.83***	1.69	1.81***	1.66
no /ʀ/					1.83*	1.74	1.85*	1.77
V2C4	1.92***	1.62	1.94***	1.58	1.69	1.67	1.78	1.76
no /ʀ/					1.80*	1.72	1.83	1.78

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

issue.

Although all previous analyses in this study have calculated j-scores based on entire stimuli, it is also possible to compute j-scores based on any subset of the stimuli. By computing such partial j-scores, the amount of independence between phonemes can be examined more closely. Table 8.3 displays the results of a partial j-score analysis using 2-phoneme units. As expected, the English nonword j-scores are all very close to 2, and the word and nonword partial j-scores for each 2-phoneme pair differ significantly from each other. In contrast, the German nonword partial j-scores are consistently less than 2 for each 2-phoneme pair, and several of the pairs do not differ significantly between words and nonwords. When words with post-vocalic /ʀ/ are excluded, the same general pattern continues to hold, with higher partial j-scores for each 2-phoneme unit, though the increase in the partial j is greatest for V1C2 and C2C4, which further confirms the hypothesis that items with post-vocalic /ʀ/ largely account for the lower than expected j-scores. The remaining discrepancy between observed and predicted values of $j_{nonword}$ in the German experiments must be addressed through future experiments.

8.1.2 Morphology

Studies that have found differences in processing of monomorphemic and multimorphemic words have consistently found processing advantages for monomorphemic words (e.g. Sereno & Jongman, 1997; Gürel, 1999), presumably because monomorphemic words can be accessed directly in the lexicon, while multimorphemic words require additional processing before lexical access (Taft & Forster, 1975; Taft, 1979, 1988; Clahsen, 1999). Cross-linguistic studies have also found that morphology has a larger effect on lexical access in "morphologically rich" languages than in languages that use morphology less extensively (e.g. Marslen-Wilson, 2001). Based on these results, it was predicted that results from the German experiments would exhibit larger effects of morphology than the English results, as measured by the difference in j between monomorphemic and bimorphemic words.

While the initial comparison of effect size of morphology between the English and German native listener results was not significant, as shown in Table 8.2 ($\Delta j = 0.12$), this comparison may be misleading, due to the interactions found between frequency, neighborhood density and morphology in the English native listener experiment. In order to investigate these effects further, bootstrap analyses were carried out for all four experiments. In the bootstrap analysis, the original data are randomly sampled a large number of times, and some metric is calculated from each random sample. The sample size is always equal to the original sample size, but the random sampling is performed with replacement, meaning that some of the original data points may be excluded altogether, while some data points will appear more than once. If only a few data points are contributing to the effect found in the original analysis, this will be revealed in a bootstrap analysis. In this case, bimorphemic and monomorphemic words were randomly sampled, and then j-factors were calculated for each group. Then the mean difference in j between bimorphemic and monomorphemic words was computed. This is the same procedure as in the initial analysis, except for the random sampling. This process was repeated 10,000 times, yielding 10,000 values of $j_{bi} - j_{mono}$; results are shown in Figure 8.2. While the mean $j_{bi} - j_{mono}$ from the bootstrap distribution for German was very similar to the originally calculated mean (original $\Delta j = .80$, bootstrap $\Delta j = .877$), the mean difference for the English bootstrap analysis was substantially lower than the original value (original $\Delta j = .91$, bootstrap $\Delta j = .518$). In addition, a 2-sample t-test revealed that the bootstrap distributions for English and German native listeners were significantly different ($t(19998) = 74.8, p < .001$), suggesting that there is a larger processing advantage of monomorphemic words over bimorphemic words in German than in English, presumably due to the fact that German utilizes inflectional morphology more than English.[1]

8.1.3 Lexical Frequency

Although not predicted, the magnitude of the effect of lexical frequency did differ significantly across languages. Recall that both experiments using German stimuli (Experiments Two and Three) found inhibitory effects of lexical frequency ($j_{low} < j_{high}$), while both experiments using English stimuli (Experiments One and Four) found the expected facilitatory effects of lexical frequency ($j_{low} > j_{high}$). Although many different additional analyses were carried out to find an explanation for the unexpected results of lexical frequency in the German experiments (see §5.5.2), no satisfactory explanation was found. However, comparing the effect sizes of lexical

---

[1] The statistically significant finding comparing the bootstrap distributions could simply be an artifact of the extremely large degrees of freedom (19998). To test this hypothesis, two subsequent bootstrap analyses of both the English and the German data were performed and compared with the original distributions. Since the bootstrap procedure involves a random factor, two subsequent bootstrap analyses from the same data will yield slightly different results. Since neither of these comparisons yielded statistically significant results (English — $\Delta\mu_j = .00018, t(19998) = .03, p = .97$; German — $\Delta\mu_j = .0018, t(19998) = .41, p = .68$), we can conclude that the significance is not an artifact of the large degrees of freedom.

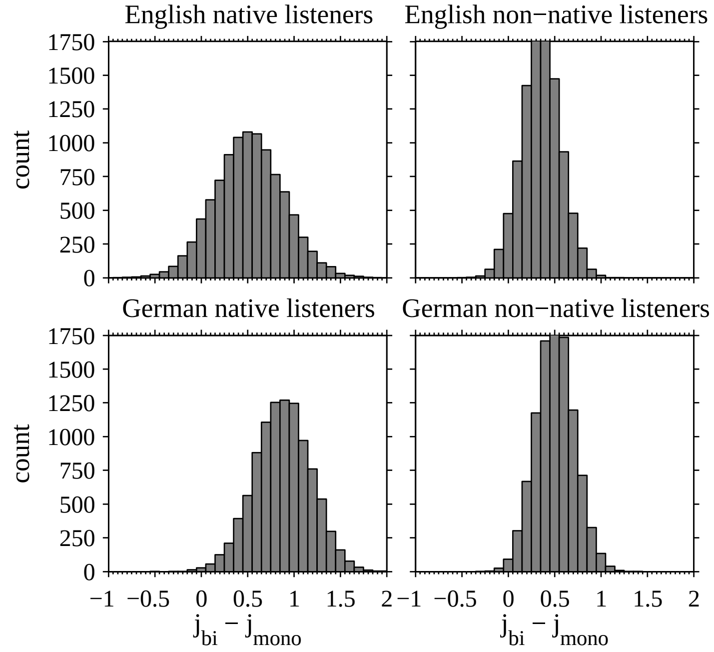


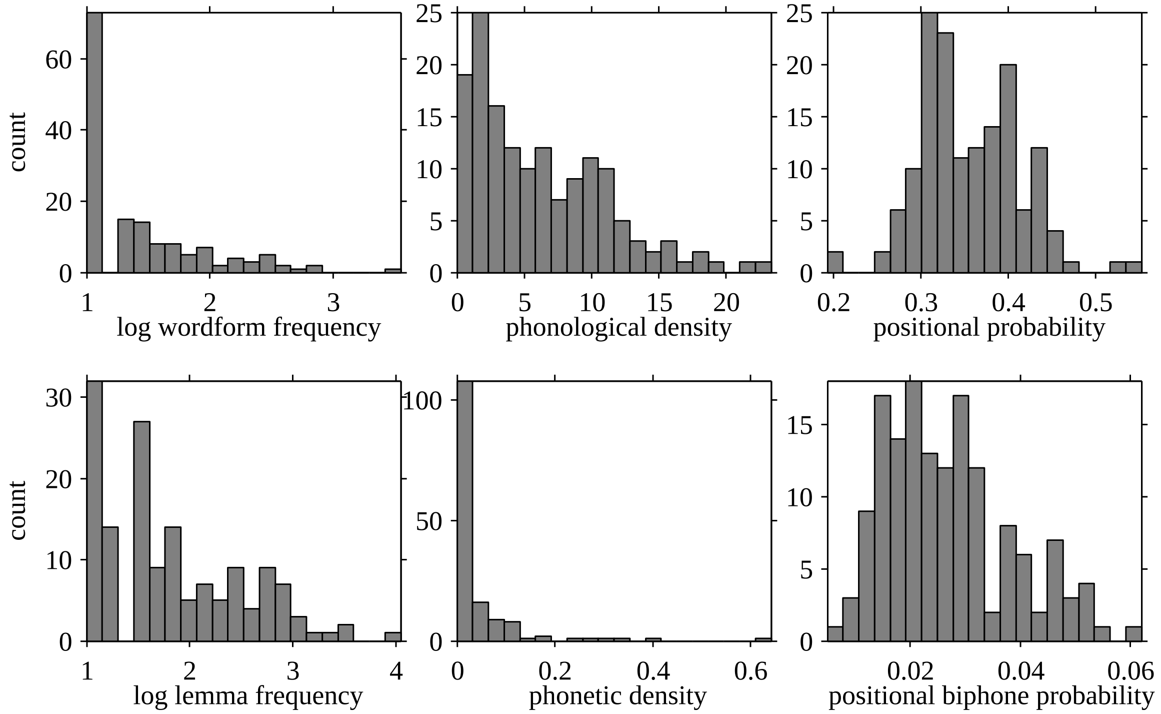
Figure 8.2 Bootstrap analysis on difference between j_{bi} and j_{mono} by items. The differences were computed by subtracting the mean from each group for each of 10,000 randomly (with replacement) selected samples. As in all other previous statistical analyses, items which had p_p or p_w values below .05 or above .95 were excluded prior to statistical analysis. The distributions of the English and German native speakers are significantly different ($t(19998) = 74.8, p < .001$), as are the distributions of the English native vs. non-native listeners ($t(19998) = 38.6, p < .001$), and the German native vs. non-native listeners ($t(19998) = 100.6, p < .001$).

Table 8.4 Descriptive statistics for the experiment materials. The means are given for each of the computed lexico-statistical measures used in the studies, comparing the English and German materials. Statistics shown are from 2-sample t-tests

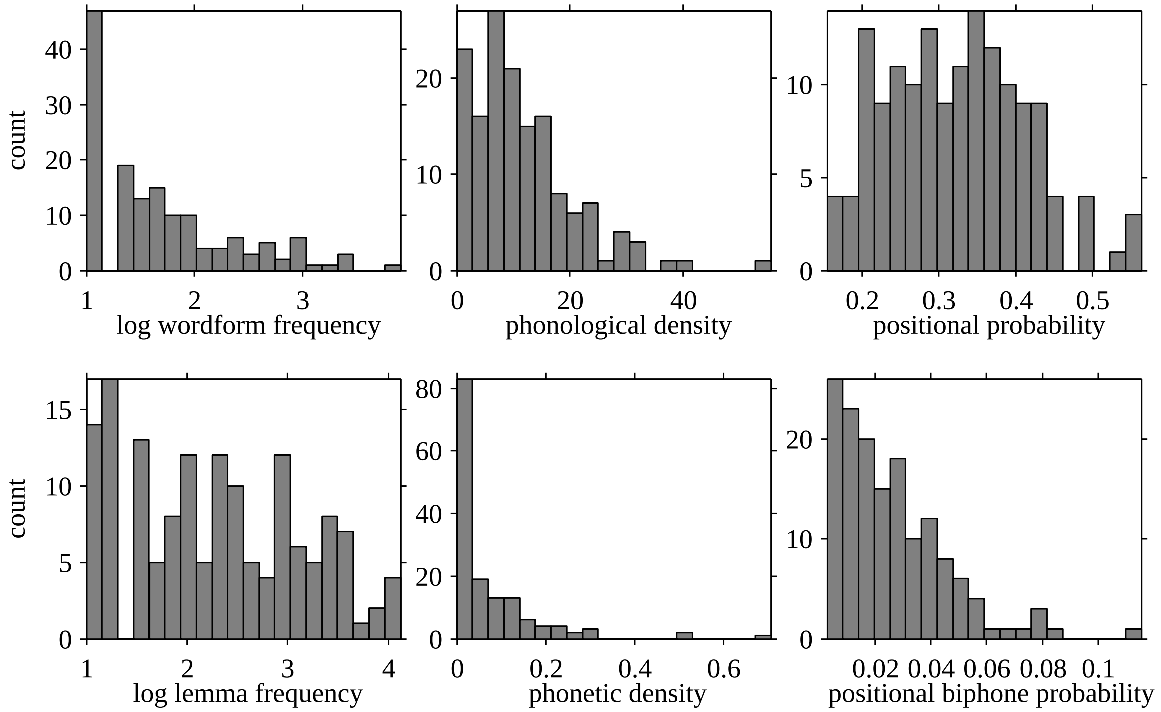
	English	German	*t*	*p*
wordform frequency	8.013	22.747	-2.222	<.05
log wordform frequency	1.412	1.667	-3.629	<.001
lemma frequency	30.640	107.167	-3.673	<.001
log lemma frequency	1.820	2.287	-5.184	<.0001
phonological neighborhood density	6.223	11.457	-6.169	<.0001
phonetic neighborhood density	0.036	0.068	-2.898	<.01
positional probability	0.355	0.322	3.869	<.001
biphone positional probability	0.027	0.026	0.455	>.1

frequency between German native and non-native listeners, there is a high degree of consistency, in that both experiments found significant inhibitory effects. Since these experiments tested different listeners, the inhibitory effect cannot be due to a particular group of listeners, nor can it be due to the nature of the task, since the experiments using English stimuli found facilitatory effects of lexical frequency (as did Benkí (2003a)). It is likely that the different results in the German and English experiments are a result of the stimulus selection. While phonological structure was shown to have little influence on the frequency results in §5.5.2), the frequency of the distribution of the stimuli was not investigated. As shown in Table 8.4 and Figure 8.3, the frequency distributions for the English and German stimuli are significantly different, but it is unclear why these differences would lead to opposite effects of frequency in the two experiments.

An alternative explanation of the unexpected results of lexical frequency in the German data is that they might in fact be expected if relevant factors are identified. While the great majority of experiments have found facilitatory effects of lexical frequency (e.g. Taft & Forster, 1975; Taft,



(a) English stimuli



(b) German stimuli

Figure 8.3 Distribution of computed lexical statistics for the English and German stimuli. The phonetic neighborhood density distributions shown here are based on the results from native listeners in Experiment One (English) and Experiment Two (German).

1979; Rubenstein et al., 1970, 1971; Benkí, 2003a), a few studies have reported inhibitory effects of lexical frequency in particular conditions for both English (Beck, 1997) and German (Clahsen, Hadler, & Weyerts, 2004). These studies both used speeded production tasks to test if wordform frequency effects are found in inflected words, which would indicate that morphologically complex words are stored whole in the lexicon. Given that other context effects such as neighborhood density have been shown to be facilitatory in word production (Vitevitch, 2002; Goldrick & Rapp, 2007), yet inhibitory in word recognition (Luce, 1986; Luce & Pisoni, 1998; Benkí, 2003a), it is possible that the inhibitory effects of lexical frequency found in this study are due to the nature of the task. However, the fact that facilitatory effects of lexical frequency were found in both English experiments suggests that task effects are not responsible for the inhibitory effects found in the German experiments. Rather, it seems more likely that some other characteristic of the stimuli used in the German experiments is responsible for the inhibitory effects of lexical frequency found. Given that very few previous studies have found inhibitory effects of lexical frequency, it is possible that these effects were also due to characteristics of the stimuli other than lexical frequency, but this can only be tested through future experimentation.

8.1.4 Neighborhood Density

Given that English and German share many phonological traits in terms of possible word structure, no differences in the effect size of of neighborhood density were predicted. As shown in Table 8.1, the inhibitory effects of phonological and phonetic neighborhood density were significant for both the English and German native listener experiments. In addition, the results of a 2-sample t-test revealed no significant difference in the effect size of neighborhood density between the results from the English and German native listener experiments, for both phonological and phonetic neighborhood density, as shown in Table 8.2. These results show that prior work by Luce & Pisoni (1998) and Benkí (2003a) using English CVC words extend to disyllabic words in both English and German.

8.2 Lexical Access by non-native listeners

For over 50 years, the field of second language acquisition (SLA) has been studying how various grammatical properties of language, including phonetics, phonology, morphology, syntax, semantics, and pragmatics, are acquired, but only recently have researchers begun to investigate lexical access in non-native speakers. Results from these recent studies suggest that non-native speakers are sensitive to many of the same context effects as native speakers, but that the magnitude of the effect can differ due to factors such as the smaller vocabulary size of non-native speakers (e.g. Bradlow & Pisoni, 1999; Imai et al., 2005; Hahne et al., 2006). The present study has found similar results, most of which can be explained through the assumption that non-native speakers have a reduced vocabulary size compared to native speakers.

8.2.1 Lexical Status

The facilitatory effect of lexical status has been shown to be one of the most robust effects in lexical access research (e.g. Taft & Forster, 1975; Rubenstein et al., 1970, 1971; Boothroyd & Nittrouer, 1988; Benkí, 2003a), although this effect has yet to be studied for non-native speakers. The processing advantage of words over nonwords can be attributed to the fact that words have a mental representation stored in long-term memory, whereas nonwords do not. In nonword recognition, listeners must rely heavily on acoustic information combined with phonological information such as phonotactics, while in word recognition, listeners can use partial acoustic information to make educated guesses to match the acoustic information to stored representations of words in the lexicon. Thus a real word that a listener has never previously heard is equivalent to a nonword. Due to the assumed smaller vocabulary size of non-native listeners, there are likely to be many more such novel words for non-native listeners than for native listeners, which leads to the prediction that the effect of lexical status will be diminished in non-native listeners. The comparison of the effect size of lexical status shown in Table 8.2 reveals that the effect of lexical status was smaller for non-native listeners for both English and German. Inspection of Figure 8.1 also confirms that the source of the smaller effect of lexical status is the higher j-scores for words in the non-native listener experiments, as predicted. It seems then that non-native listeners are affected by lexical status in a very similar manner compared to native listeners, and that the smaller size of the effect can be attributed to a smaller vocabulary size.

8.2.2 Morphology

As shown in §8.1, native listeners of German showed a greater processing difference between mono- and bimorphemic words than did native listeners of English. Given this cross-linguistic difference in the effect of morphology on lexical access, it is natural to inquire whether this effect will be carried over when listening to a non-native language. There are at least two possible scenarios for the influence of morphology on lexical access by non-native listeners: (1) non-native listeners simply transfer the morphological structure of their native language into the second language, or (2) non-native listeners start off with essentially zero morphological structure in their non-native lexicon, and acquire the morphological structure of the second language over time. Both the first and second scenarios predict that intermediate learners of an L2 whose native language has relatively little morphology will not be highly sensitive to differences in morphology in the L2. The two scenarios do differ in the predictions of how listeners whose native language is morphologically rich will be affected by differences in morphology when perceiving a non-native language. The first scenario leads to the prediction that listeners whose native language is morphologically rich (and therefore has a large effect on lexical access), will also be highly sensitive to differences in morphology in a non-native language, regardless of the morphological richness of the non-native language. In contrast, the second scenario leads to the prediction that intermediate learners whose native language is morphologically rich will not be as sensitive to difference is morphology in a non-native language as mature speakers of that language, but that the learners will become more sensitive over time.

In the present study, both of these scenarios predict that the difference in j between monomorphemes and bimorphemes should be smaller for English-speaking listeners of German than for native German listeners. The first scenario predicts that the difference in j between mono- and bimorphemic words should be *larger* for German-speaking listeners of English than for native English listeners, while the second scenario predicts that the difference should be *smaller* for German-speaking listeners of English. As the results in Table 8.1 show, the effect size of morphology was smaller for both groups of non-native listeners, and the difference in effect size was significant for both languages, as shown in Table 8.2. Bootstrap analyses shown in Figure 8.2 also confirmed these differences. These results support the second scenario, in which non-native listeners are not as sensitive to differences in morphology as native listeners. Of course one study alone is not sufficient evidence to support a theory; additional experiments, especially ones in which the L1 and L2 differ even more in morphological structure, are necessary to thoroughly test these hypotheses. In addition, no conclusions can be drawn about the rate of increase in sensitivity to morphology in a second language, since the participants in the present study were intermediate to advanced learners of the second language. Future research employing longitudinal or cross-sectional designs can further address the rate of acquisition.

8.2.3 Lexical Frequency

Since non-native listeners have much less exposure to the target language than native listeners, the estimates of lexical frequency drawn from large corpora most likely do not reflect the word familiarity of non-native speakers. If one were to estimate lexical frequency taken from corpora of non-native speakers, the actual frequency counts would likely be much lower, but the overall distribution may be very similar — high-frequency words for native speakers are also likely to be high-frequency words for non-native listeners. A recent study investigating first language acquisition disorders found that the order of acquisition of certain phonemes was the same using frequency estimates from adult-speech corpora and child-speech corpora (Gierut & Dale, in press). This suggests that global effects of lexical frequency are also likely consistent for native and non-native listeners. The main difference in the frequency distributions would probably lie in the low- to medium-frequency words, some of which may be entirely absent from the non-native speakers' lexicon. Assuming this scenario is correct, it was predicted that high-frequency words should be treated roughly equally for native and non-native listeners, but that non-native listeners may treat many of the low-frequency words as nonwords — that is, the difference in j between low- and high-frequency words was predicted to be greater for non-native listeners than for native listeners.

Since the inhibitory effect of lexical frequency found in the German experiments is difficult to interpret, differences in effects of lexical frequency will only be taken from the English experiments. The effect sizes shown in Table 8.1 show that the difference in j between low- and high-frequency words was larger for non-native listeners than for native listeners, although this difference was not statistically significant, as shown in Table 8.2. While the effect size of lexical frequency was not significantly greater for native listeners than for non-native listeners, the results do suggest that, given more statistical power, effects of lexical frequency might be greater for non-native listeners than for native listeners. Moreover, the fact that there was a significant difference in j between low- and high-frequency words for both English native and non-native listeners shows that non-native listeners are sensitive to lexical frequency in a fashion similar to native listeners' sensitivity, which is suggestive of frequency information being encoded in the non-native lexicon in a similar manner to the native lexicon.

8.2.4 Neighborhood Density

The smaller vocabulary size of non-native listeners can also impact the effect size of neighborhood density. The number of neighbors for a given word in the non-native lexicon should be less than or equal to the number of neighbors in the native lexicon, which should result in less overall lexical competition. However, this assumes that the definition of a neighbor is the same for native and non-native listeners. As Weber & Cutler (2004) show, non-native listeners are affected by additional sources of lexical competition which do not affect native listeners. They used an eye-tracking plus spoken word recognition paradigm with Dutch and English stimuli selected such that some of the distractor items might be considered neighbors by non-native speakers, but not by native speakers. One example from the English words used in their study is *racket* /ɹækɪt/, and the competitor *records* /ɹɛkɔːdz/ (British English). Since Dutch does not have the phoneme /æ/, it is likely that Dutch listeners would perceive these two words as having an initial overlap of 3 phonemes, but English listeners would perceive the words as having an initial overlap of only 1 phoneme. Their results support these predictions, and they conclude that "the amount of lexical competition is much greater in non-native than in native listening" (22). However, one should note that they did not actually investigate (or control for) neighborhood density effects. Their analysis is based only on 20 words with very specifically chosen competitors, in a fixed-choice design, which does not involve a full lexical search as do open response tasks (Clopper, Pisoni, & Tierney, 2006). It is possible that the additional competitors for non-native listeners do not outnumber the missing competitors not present in the non-native lexicon.

This hypothesis can be tested by comparing *j*-factor results of words in sparse and dense neighborhoods for both native and non-native listeners. If overall lexical competition is lower for non-native listeners, then words in dense neighborhoods should be treated more like words in sparse neighborhoods; however, if words in sparse neighborhoods are treated more like words in dense neighborhoods by non-native listeners, this would indicate an overall increase in lexical competition. As shown in Table 8.1, the effect of neighborhood density is smaller for both sets of non-native listeners; the difference in German is highly significant for phonetic neighborhood density, but not phonological density, while the difference in English is significant for both phonological and phonetic neighborhood density, as shown in Table 8.2. From Figure 8.1, it can be seen that the reason the effects of density are smaller for non-native listeners is not due to lower *j*-scores for words in dense neighborhoods, but rather that the *j*-scores are higher for words in sparse neighborhoods.

The effects of neighborhood density can be investigated in greater detail by looking at the types of errors that listeners made. If non-native listeners have greater lexical competition, then the number of errors which are phonological neighbors should be lower than for native listeners. To measure this, each incorrect response was checked to see if it is a neighbor of the target word, and the percentage of unique errors which are neighbors was calculated for each stimulus, then the mean was computed for each experiment. One-tailed 2-sample t-tests revealed that the percentage of errors which are neighbors is lower for non-native listeners for both German and English (German native = 12.3%, German non-native $= 8.2\%, t(298) = 1.81, p < .05$; English native = 12.8%, English non-native $= 6.7\%, t(298) = 2.24, p < .05$). The increased *j*-scores for words in sparse neighborhoods, combined with the smaller percentage of errors which are neighbors, support the claim that lexical competition is greater for non-native listeners than native listeners.

8.3 Theoretical Implications

As outlined in §2.1, one of the major questions in research on lexical access has been the storage and processing of morphologically complex words. Most of the literature discussing this issue has grouped various theories and models into one of two general classes: associative models and combinatorial models. Associative models, the most prominent of which are connectionist models, have mostly been proposed by psychologists, and assume that language can be modeled as neural networks. In contrast, combinatorial models stem from the tradition of generative linguistics, which assumes that language is composed of discrete and infinitely combinable units. While these theories are often seen as diametrically opposed, and fierce debates have been held by proponents of each side (see e.g. Pinker & Prince, 1988), as Smolensky (1999) points out, these theories actually share many traits, and instead of focusing on the differences between them, it might be more fruitful to acknowledge their similarities, and that both of these lines of research have advanced psycholinguistics. Smolensky points out that generative linguistics and connectionism focus on different levels of representation. While generativists seeks to discover the nature of linguistic representation in the mind, connectionism seeks to model the behavior of the brain. Additionally, some of the goals of connectionist research and generative linguistic research differ in the scope and specificity. Generative linguistics has concentrated on producing explanatory theories which attempt to account for all aspects of all languages using the same mechanisms, while connectionism has concentrated on developing quantitative models which can be directly compared with results from specific psycholinguistic experiments. Thus while generative linguistics fails to make quantitative predictions on the nature of language processing, connectionism generally fails to make language-universal generalizations about language processing.

The present study has provided experimental results which test some of the predictions of these models. One of the fundamental differences between these two types of models is the storage and access of multimorphemic words. Associative models generally assume that multimorphemic words are stored whole, and that any differences in morphological processing can be attributed to on-line processing differences resulting from semantic or phonological properties of the stimuli. In contrast, combinatorial models assume that only stems are stored in the lexicon,[2] and that word recognition involves stripping off inflectional affixes before lexical access can occur, which predicts processing differences between morphologically simple and complex words. The processing advantage for monomorphemic words in the present study is more readily compatible with a combinatorial model of lexical access than with associative models of lexical access. However, this does not imply that associative models should be altogether abandoned. Associative models of lexical access have had great success in accurately modeling results from psycholinguistic experimental data, including effects of frequency and neighborhood density. In contrast, only one combinatorial-type model has been implemented that makes specific predictions that can be quantitatively compared to experimental results (Albright & Hayes, 2003). Therefore, before discounting associative models of

---

[2] Dual-mechanism models assume that high-frequency multimorphemic words, or words with irregular inflectional morphology are stored whole in the lexicon.

lexical access altogether, it is worthwhile to consider how these sorts of models could be modified to account for differences in morphological processing.

An appropriate goal for a model of spoken word recognition is to quantitatively describe how humans translate an acoustic signal into an abstract unit in the mind which contains semantic, phonological, and perhaps morphological and syntactic information (i.e. a word), and how factors such as lexical status, morphology, lexical frequency, and neighborhood density affect this process. To date, no one model has been successful in accounting for all of the various factors shown to affect spoken word recognition in experimental settings. Several models have been quite successful though. In this section, several of the more influential models of spoken word recognition will be discussed, and suggestions will be made as to how these models could be modified to account for the findings in this study.

8.3.1 Associative models

TRACE

One of the first and most influential models is the TRACE model of spoken word recognition (McClelland & Elman, 1986). TRACE is a connectionist model with three levels of representation: (1) a featural level, which can be derived directly from real speech signals; (2) a phonemic level; and (3) a word level. The model employs inhibitory connections within levels and excitatory connections between levels; in this way word-level activation can affect lower level activations. The acoustic input is first mapped to features, and then to phonemes which are consistent with features (including partial, or noisy information from the acoustic signal), and finally all words in the lexicon consistent with the phoneme are activated. This process is repeated as additional acoustic information is received, until eventually, the activation of one word crosses a threshold, at which point that word (hopefully the intended word) is recognized. The TRACE model has been shown to accurately model results from a variety of psycholinguistic experiments, including evidence from phoneme monitoring (Cutler, Mehler, Norris, & Segui, 1987), phonological categorization (Ganong, 1980), and phoneme restoration (Samuel, 1981). However, in the original TRACE model, only monomorphemic words were included in the lexicon for the simulations; therefore the model in its current state cannot provide a full account of lexical access.

Neighborhood Activation Model (NAM)

The neighborhood activation model of Luce & Pisoni (1998) is not a connectionist model of spoken word recognition, but shares many of the same traits as connectionist models, in that acoustic input activates words in memory, and word recognition occurs when the activation crosses some threshold. The key advantage of the NAM over other models of spoken word recognition is that it incorporates effects of lexical frequency and neighborhood density into a cohesive design, summarized in

Equation 8.1.

$$p(ID) = \frac{\prod_{i=1}^{n} p(PS_i|PS_i) \cdot Freq_S}{\left\{\left[\prod_{i=1}^{n} p(PS_i|PS_i)\right] \cdot Freq_S\right\} + \sum_{j=1}^{nn}\left\{\left[\prod_{i=1}^{n} p(PN_{ij}|PS_i)\right] \cdot Freq_{Nj}\right\}} \tag{8.1}$$

where $p(PN_{ij}|PSi)$ is the probability of a listener responding with the i^{th} phoneme of the j^{th} neighbor, when presented with the i^{th} phoneme of the stimulus, n is the number of phonemes in the stimulus, and nn is the number of neighbors. To paraphrase, the probability of correctly identifying a word is the product of the recognition probabilities of each constituent phoneme, multiplied by the frequency of the word, divided by the sum of the frequency-weighted recognition probabilities of the stimulus and all neighbors of the stimulus. The summed term in the denominator was used as the measure of phonetic neighborhood density in the present study. Benkí (2003a: 1700) found a high correlation ($r = .656, p < .001$) between the predictions of the NAM model and the results from a speech-in-noise experiment using CVC English syllables,[3] indicating that the NAM can account for a large amount of the variation in spoken word recognition. However, the NAM does not make any predictions about the role of morphology in spoken word recognition; in fact, the NAM makes no assumptions as to whether stems or whole words are activated in the lexicon.

8.3.2 Combinatorial models

Dual-Mechanism models

Dual-mechanism models (e.g. Clahsen, 1999; Pinker, 1999) posit two mental mechanisms for processing inflected words — stored entries, and combinatorial rules. These two mechanisms can operate in parallel. Monomorphemic words are always accessed directly, multimorphemic words can be accessed via either mechanism. High-frequency multimorphemic words are assumed to be stored and are therefore accessed directly, but if the direct route fails, the combinatorial-based mechanism can always be applied. Dual-mechanism models can successfully account for morphological effects in processing, including differences between regular and irregular inflectional morphology. However, the models do not make specific quantitative predictions, and they do not make any predictions as to neighborhood density effects. Thus while the processing advantages of monomorphemic words found in this study are compatible with a dual-mechanism model, the effects of neighborhood density are left unexplained.

Stochastic rule-based model

Albright & Hayes (2003) proposed a novel model of morphological processing which differs from both analogical (connectionist) models and dual-mechanism models. Their model is similar

---

[3] Note that Benkí (2003a) actually found a higher correlation when not including effects of neighborhood density, but rather using a model solely based on stimulus probability calculated from the nonword confusion matrices.

to analogical models in that it does not start out with any pre-defined rules, but rather learns rules through induction, as it receives new input (mimicking language acquisition). Unlike connectionist models however, their model produces morphological rules, not connections. Unlike dual-mechanism models, their rules are stochastic, with more general rules having greater weights. For example, their model, when given the two pairs, *play /pleɪ/—played /pleɪd/* and *read /ɹid/— read /ɹɛd/*, creates the following two rules:

$$\emptyset \rightarrow \mathrm{d}/[\mathrm{X}\underline{\quad\quad}]_{[+\mathit{past}]} \tag{8.2}$$

$$/\mathrm{i}/ \rightarrow /\text{ɛ}//[\mathrm{X}\{\mathrm{l},\text{ɹ}\}\underline{\quad\quad}\mathrm{d}]_{[+\mathit{past}]} \tag{8.3}$$

Through the combination of simulations and new experimental results, Albright & Hayes (2003) convincingly show that their model can account for cases in which both the analogical models and the dual-mechanism models fail, specifically islands of reliability, which are rules that always apply in a particular environment. Such islands of reliability can be found for both regular and irregular words. One example is that all words ending in voiceless fricatives form the past tense by adding /t/. Their stochastic rule-based model always produces the correct response in such islands of reliability, whereas both the analogical model and dual-mechanism models will produce some incorrect responses. Albright & Hayes (2003) have greatly advanced the state of affairs for rule-based models by providing a computationally implemented model which makes specific quantitative predictions as to how morphological processing works. However, their model is not intended to be a model of lexical access, and it is unclear how such a rule-based model would account for frequency or density effects.

8.3.3 A new proposal — morphological neighborhoods

The results from the present study, combined with other recent studies (e.g. Gumnior et al., 2006; Gürel, 1999), show that models of lexical access must incorporate some level of morphological representation. In addition, a model of lexical access must also be able to account for effects of lexical frequency and neighborhood density. Finally, a sufficient model of lexical access should make quantitative predictions that can be rigorously tested through simulations and experiments. At present, the NAM comes closest to meeting all of these requirements, with the exception of making predictions about morphological processing. Given NAM's many strengths, a reasonable approach is to consider how the NAM could be modified to also account for morphological effects.

One of the crucial design features of a model of lexical access concerns the storage of lexical items — namely, whether stems or whole words are stored in the lexicon. Most combinatorial models argue that stems are stored in the lexicon (with dual-mechanism models also including high-frequency words), while most associative models argue whole words are stored. Data from previous experiments manipulating wordform and lemma frequency, as well as results from the present study, suggest that both lemma frequency and wordform frequency can affect lexical access. The only way that a model of lexical access can account for wordform frequency effects is to posit that words are stored whole in the lexicon. The model shown in Figure 8.4 displays how wordform frequency can influence lexical access, but this model also predicts that lemma frequency should be unavailable to the listener, since full forms are accessed directly. The combinatorial type model

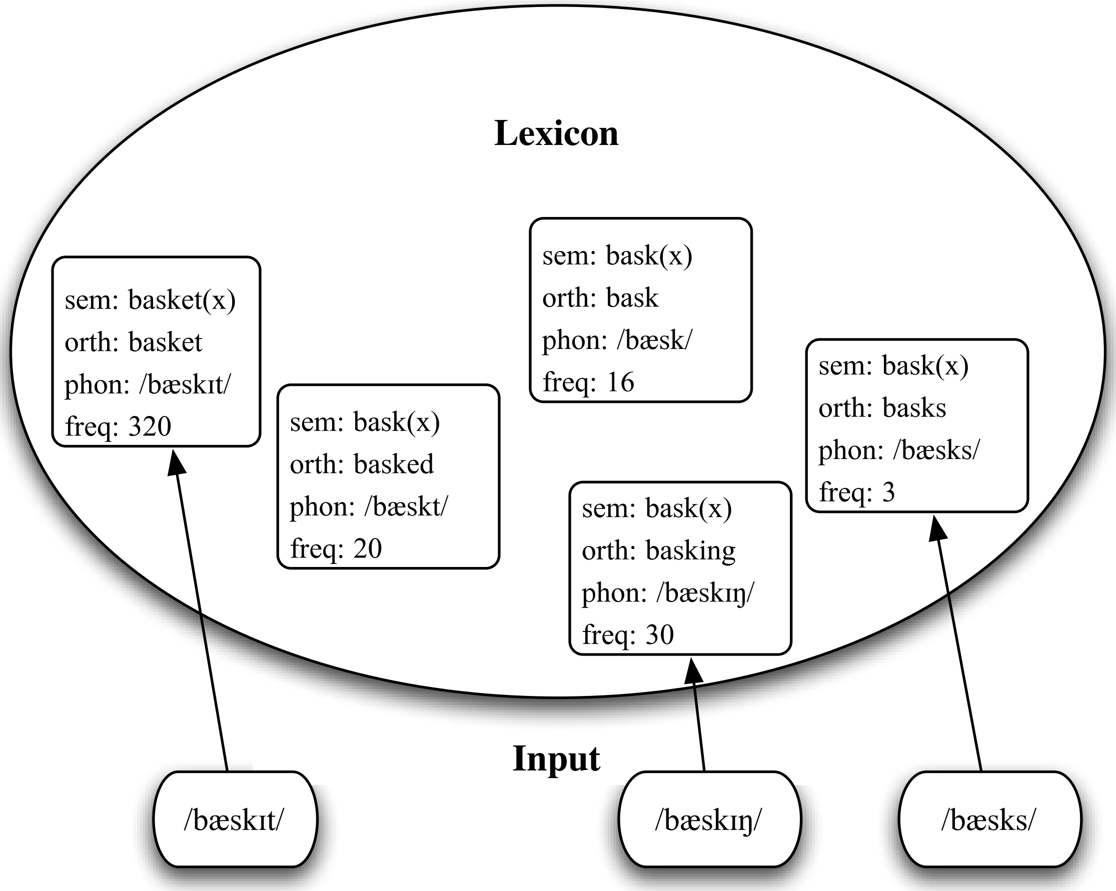


Figure 8.4 One associative view of the lexicon. Orthographic, phonological, semantic, and frequency information are stored for every word in the lexicon. All words are accessed directly. Numbers for each word are the raw wordform frequency counts from the CELEX database.

sketched out in Figure 8.5 shows how lemma frequency information is available to the listener during lexical access, but wordform frequency is not.

An alternative view of the lexicon is presented in Figure 8.6. This view is very similar to other associative models of the lexicon, in that all wordforms are stored in the lexicon, but differs in that the morphological structure of each word is also included in the lexicon. The morphological information allows for the creation of morphological neighborhoods. This proposal is similar to the morphological family effect proposed by de Jong, Schreuder, & Baayen (2000), but with several key differences. The nodes in de Jong et al.'s model represent lemmas, whereas the nodes in this model represent full-forms. As noted earlier, it is necessary to posit full-form storage in order to account for wordform frequency effects. In addition, de Jong et al. (2000)'s model does not make any predictions about neighborhood density. In the present model, an input such as *basking* /bæskɪŋ/ activates all other words which share either the morpheme *bask* or *–ing*, while the input *basket* /bæskɪt/ only activates words which share the morpheme *basket*. In this model, *basking* has a much larger morphological neighborhood than *basket*, and is therefore predicted to be at a processing disadvantage. In comparing the results for these two words from Experiment One, *basking* has a j-score of 2.349, while *basket* has a j-score of 2, confirming this prediction. In addition, looking at the errors for each word is also useful. Incorrect responses to the stimulus *basket* included just

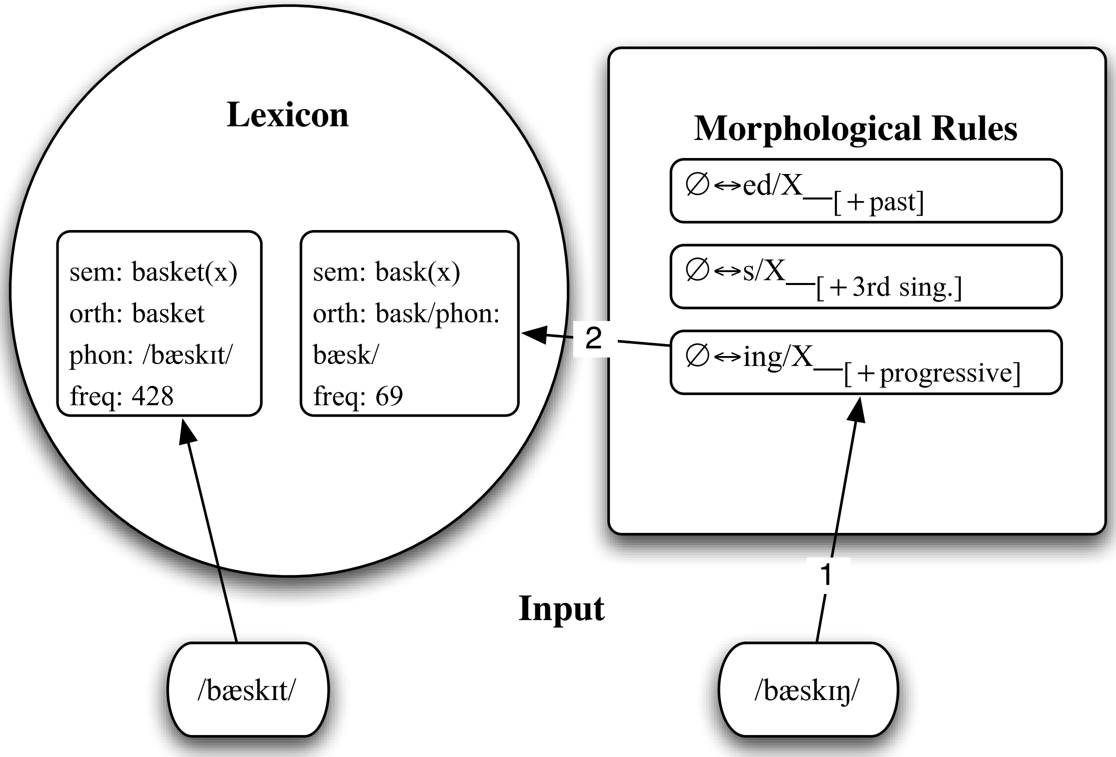


Figure 8.5 Combinatorial view of the lexicon. Orthographic, phonological, semantic, and frequency information are stored for every stem in the lexicon. In addition, a set of morphological rules is used to strip off inflectional endings before lexical entries are accessed. Numbers for each word are the raw lemma frequency counts from the CELEX database.

fasted, while incorrect responses to *basking* included *asking (2), basting, bathroom (3), fasting (2), vacuum*. Responses such as *asking* are directly predicted solely on the basis of phonological or phonetic neighborhood density, while responses such as *basting* and *fasting* are only predicted by morphological neighborhood. While the notion of morphological neighborhood may be helpful in understanding the influence of morphology in language processing, several alternatives may also be fruitful. One alternative explanation would be to simply expand the definition of a neighbor. Most researchers have defined neighbors to differ only in one phoneme. Expanding this definition to two or three phonemes, or perhaps even $n-1$ phonemes, with nearer neighbors being weighted heavier than farther neighbors, could also explain the differences in the above example. Kapatsinski (2005) has made a similar proposal in an attempt to model the lexicon as a complex network. Future research testing these various proposals is necessary to determine exactly how morphological information is stored and processed in the lexicon.

8.3.4 Summary of lexical access models

The preceding discussion of models of lexical access suggests that no current model can account for all of the empirical findings from this study. While the processing advantage of monomorphemic over bimorphemic words found in this study is compatible with combinatorial models of lexical

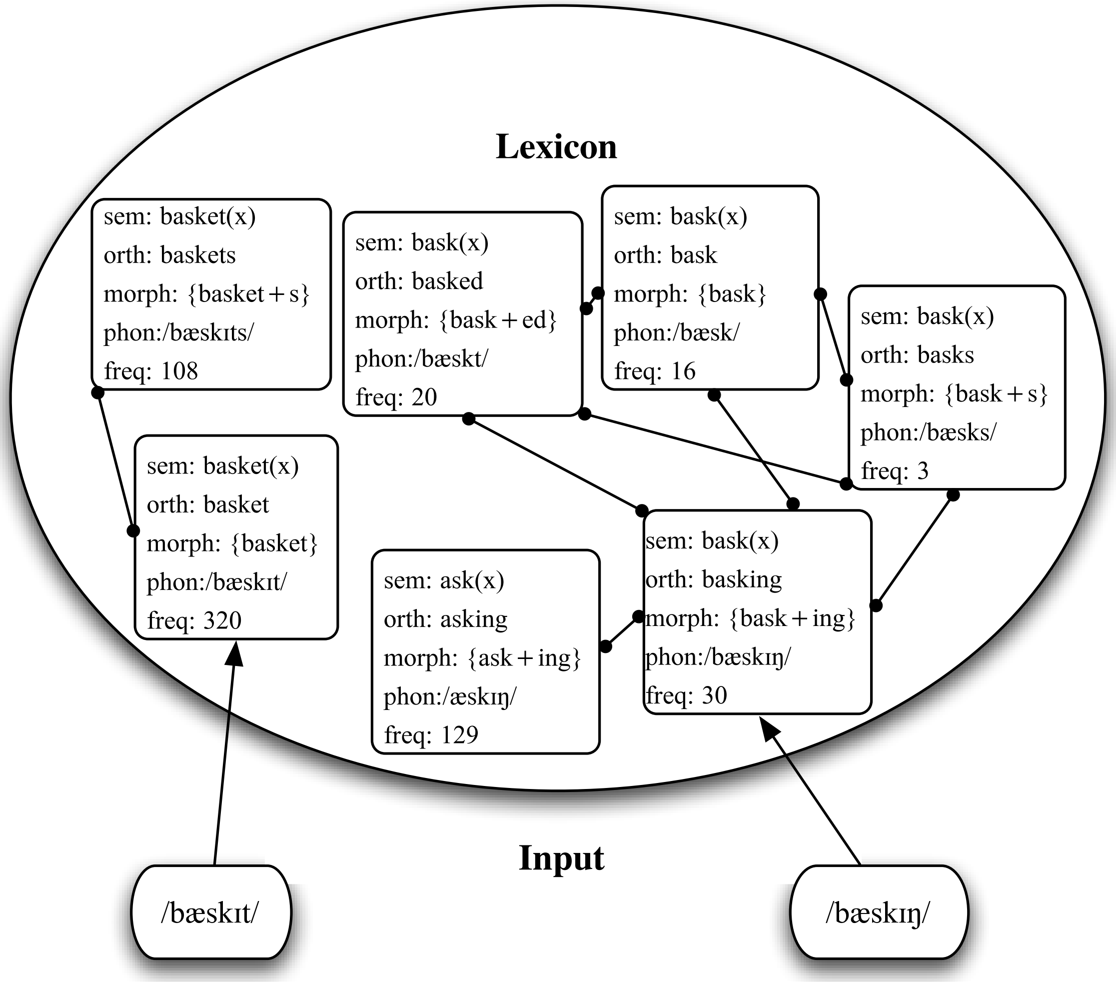


Figure 8.6 Full-listing model of the lexicon with morphological information. This model is similar to a traditional associative model of the lexicon, but morphologically complex words also include morphological information in addition to orthographic, phonological, and frequency information. Numbers for each word are the raw wordform frequency counts from the CELEX database. Connections between words represent morphological neighbors.

access, these models do not make any predictions regarding lexical frequency or neighborhood density. In contrast, the associative models reviewed here can, to varying degrees, account for the effects of lexical frequency and neighborhood density found in this study, but do not make predictions about the influence of morphology on spoken word recognition. Perhaps the most promising model is the NAM, especially in that it makes quantitatively accurate predictions of the effects of lexical frequency and neighborhood density in speech-in-noise tasks. The proposal to store words as whole, while including morphological information, offers an expansion to the NAM that can explain both wordform and lemma frequency effects, neighborhood density effects, and morphological effects.

8.4 Conclusions

This study adopted a cross-linguistic approach to address the following research questions:

- Are monomorphemic and bimorphemic words processed in the same way, as associative models predict, or are bimorphemic words decomposed into their constituent morphemes before lexical access, as combinatorial models propose?
- What role does morphology play in spoken word recognition, and how do phonetic and morphological effects interact in lexical access?
- To what extent are context effects in lexical access dependent on the structure of the language?
- Do cross-linguistic differences in the mental lexicon carry over to learning a second language?
- Do previously found effects of lexical frequency and neighborhood density in monosyllabic words extend to disyllabic words?

Analysis of the difference in j between mono- and bimorphemic words showed that monomorphemic words have a processing advantage over bimorphemic words, suggesting that there is a morphological level or representation in the mental lexicon, contrary to what associative models of lexical access predict. However, as previous research has shown (e.g. Marslen-Wilson, 2001), morphological processing differs across languages. Consistent with the findings of Marslen-Wilson (2001), this study found that the processing advantages for monomorphemic words in lexical access is greater in a morphologically rich language (German) than in a language which does not make extensive use of morphology (English).

While most previous studies investigating morphological effects on lexical access have used visual tasks, the present study used an auditory task to investigate effects of morphology. While the analysis technique used in this study does not allow for direct comparison of effect size with studies which measure effects using response time, this study does clearly show that morphology can have an effect on spoken word recognition. In addition, signal detection theory analyses of the German experiments showed that differences in morphology and lexical status can impact both the perceptual distinctiveness and the response bias of acoustically similar phonemes.

By carrying out a four-way design with two languages, and both native and non-native listeners, this study was also able to address lexical access by non-native listeners in a controlled fashion. Results from Experiments Three and Four show that non-native listeners are also sensitive to lexical context in much the same way as native listeners, though differences in vocabulary size and exposure to the language can alter the size of context effects. In particular, results from these experiments suggest that non-native listeners are less sensitive to morphological differences in an L2, regardless of their L1, which supports a chunking account of second language acquisition.

In conclusion, this study has added several new findings to the field of lexical access and spoken word recognition: (1) Processing differences between mono- and bimorphemic words suggest that a morphological representation in the lexicon is necessary; (2) context effects in lexical access can vary across languages, especially with regard to morphological processing; (3) perceptual distinctiveness and response bias can be influenced by morphological properties of stimuli; (4) non-native listeners are sensitive to context in much the same way as native listeners are in spoken word recognition, though to a lesser degree; and (5) previous results from open response word recognition using CVC English stimuli showing effects of lexical status, lexical frequency, and neighborhood density were extended to CVCCVC English and German stimuli. These results further our understanding of the structure of the mental lexicon, which is a crucial part of understanding how language is structured and processed.

Appendices

Appendix A

List of stimuli

This appendix includes lists of the stimuli used in all of the experiments in this study. The English stimuli were used in Experiments One and Four; the German stimuli were used in Experiments Two and Three. For each stimulus, a variety of lexicostatisical information is also given. Separate lists are given for word and nonword stimuli. The spellings for the nonword stimuli are from the experimenter, and correspond to the desired format as specified in the instructions to participants. Phonetic transcriptions for the words are taken from the CELEX (Baayen & Rijn, 1993) database. Some of the information, such as lexical frequency, is only relevant to words, not to nonwords, and is therefore not included in the nonword list.

A.1 English Nonwords

Table A.1 English nonword stimuli

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
bahpwun	bɑpwən	0.300	0.023	0	0	0	0	0	0	0	0	0	0	0.00	0.00
bayldid	beɪldɪd	0.367	0.029	0	0	0	0	0	0	0	0	0	50	55.33	80.84
behlfit	bɛlfɪt	0.397	0.026	1	1	1.327	1	1	1.327	0	0	0	23	27.02	39.02
behlsid	bɛlsɪd	0.411	0.032	0	0	0	0	0	0	0	0	0	21	24.05	27.14
behmrud	bɛmʀəd	0.335	0.018	0	0	0	0	0	0	0	0	0	33	43.58	49.63
behnkut	bɛnkət	0.391	0.025	0	0	0	0	0	0	0	0	0	8	9.05	13.06
behzlun	bɛzlən	0.306	0.022	0	0	0	0	0	0	0	0	0	3	4.06	5.75
belbit	bɛlbɪt	0.399	0.024	0	0	0	0	0	0	0	0	0	5	6.48	7.72
bintim	bɪntɪm	0.479	0.051	0	0	0	0	0	0	0	0	0	27	34.19	39.35
chendit	ʧɛndɪt	0.393	0.045	0	0	0	0	0	0	0	0	0	11	13.46	17.19
chifpid	ʧɪfpɪd	0.331	0.018	0	0	0	0	0	0	0	0	0	27	36.42	48.94
choalsing	ʧoʊlsɪŋ	0.316	0.046	0	0	0	0	0	0	0	0	0	28	37.08	44.99
chowltid	ʧaʊltɪd	0.323	0.030	0	0	0	0	0	0	0	0	0	29	32.24	48.37
chumfedge	ʧʌmfəʤ	0.193	0.010	0	0	0	0	0	0	0	0	0	1	2.55	2.77
dahstiz	dɑstɪz	0.440	0.052	0	0	0	0	0	0	0	0	0	36	55.29	67.84
dalpuk	dælpək	0.327	0.014	0	0	0	0	0	0	0	0	0	2	2.00	2.49
dapkes	dæpkəs	0.317	0.014	0	0	0	0	0	0	0	0	0	3	3.00	3.00
daupkim	dɑpkɪm	0.349	0.015	0	0	0	0	0	0	0	0	0	0	0.00	0.00
daysledge	deɪsləʤ	0.275	0.010	0	0	0	0	0	0	0	0	0	1	1.00	1.53
dazduk	dæzdək	0.289	0.009	0	0	0	0	0	0	0	0	0	3	3.27	3.79
dazmis	dæzmɪs	0.332	0.020	0	0	0	0	0	0	0	0	0	6	9.45	13.66
dehlpit	dɛlpɪt	0.414	0.028	0	0	0	0	0	0	0	0	0	25	31.47	36.49
dehmlid	dɛmlɪd	0.362	0.029	0	0	0	0	0	0	0	0	0	11	13.08	14.87
dehpsidge	dɛpsɪʤ	0.345	0.022	0	0	0	0	0	0	0	0	0	5	6.83	6.95
doafpid	doʊfpɪd	0.301	0.013	0	0	0	0	0	0	0	0	0	3	3.00	3.03
dufsen	dʌfsən	0.305	0.024	0	0	0	0	0	0	0	0	0	8	8.45	9.47
dundim	dʌndɪm	0.380	0.039	0	0	0	0	0	0	0	0	0	21	27.51	32.65
fahlfik	fɑlfɪk	0.358	0.023	0	0	0	0	0	0	0	0	0	16	19.81	24.06
fanrit	fænʀɪt	0.438	0.033	0	0	0	0	0	0	0	0	0	18	23.11	29.76
fauldek	fɑldək	0.340	0.022	0	0	0	0	0	0	0	0	0	10	11.91	15.14
fehsfin	fɛsfɪn	0.379	0.023	0	0	0	0	0	0	0	0	0	6	6.63	7.34
fehskim	fɛskɪm	0.370	0.023	0	0	0	0	0	0	0	0	0	4	4.55	4.69
fekredge	fɛkʀəʤ	0.307	0.019	0	0	0	0	0	0	0	0	0	9	9.37	9.97
feldiz	fɛldɪz	0.394	0.036	0	0	0	0	0	0	0	0	0	6	6.62	7.81
fiknit	fɪknɪt	0.439	0.029	0	0	0	0	0	0	0	0	0	1	1.00	1.00
fimdik	fɪmdɪk	0.402	0.031	0	0	0	0	0	0	0	0	0	50	61.07	82.56
fiswik	fɪswɪk	0.389	0.037	0	0	0	0	0	0	0	0	0	23	36.27	43.93
foastiz	foʊstɪz	0.389	0.050	0	0	0	0	0	0	0	0	0	48	60.63	74.81
fowlpid	faʊlpɪd	0.320	0.014	0	0	0	0	0	0	0	0	0	13	15.82	20.52
fowmtid	faʊmtɪd	0.338	0.026	0	0	0	0	0	0	0	0	0	36	41.62	56.61
fowstiz	faʊstɪz	0.367	0.048	0	0	0	0	0	0	0	0	0	34	43.82	53.62
gafsid	gæfsɪd	0.322	0.020	0	0	0	0	0	0	0	0	0	13	17.15	22.18
gahmgum	gɑmgəm	0.247	0.015	0	0	0	0	0	0	0	0	0	1	1.00	1.00
gahnsid	gɑnsɪd	0.401	0.038	0	0	0	0	0	0	0	0	0	9	9.62	12.66
gakmik	gækmɪk	0.318	0.022	0	0	0	0	0	0	0	0	0	7	8.34	9.12
gehltun	gɛltən	0.368	0.039	0	0	0	0	0	0	0	0	0	22	31.16	34.76
gehnmuk	gɛnmək	0.306	0.019	0	0	0	0	0	0	0	0	0	4	5.01	5.61
goaskiz	goʊskɪz	0.330	0.024	0	0	0	0	0	0	0	0	0	19	21.91	27.36
hamdez	hɑmdəz	0.310	0.025	0	0	0	0	0	0	0	0	0	39	44.12	61.13
hastim	hæstɪm	0.387	0.050	0	0	0	0	0	0	0	0	0	15	16.93	23.38

Table A.1 English nonword stimuli (continued)

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
heespeng	hispəŋ	0.276	0.015	0	0	0	0	0	0	0	0	0	1	1.00	1.75
hefking	hɛfkɪŋ	0.313	0.040	0	0	0	0	0	0	0	0	0	28	35.20	43.85
hehnsim	hɪnsɪm	0.432	0.033	0	0	0	0	0	0	0	0	0	16	18.80	21.95
hehntis	hɛntɪs	0.434	0.062	0	0	0	0	0	0	0	0	0	19	26.03	34.41
hinlik	hɪnlɪk	0.417	0.033	0	0	0	0	0	0	0	0	0	50	69.87	89.56
hoantiz	hoʊntɪz	0.391	0.045	0	0	0	0	0	0	0	0	0	33	40.49	52.60
humsus	hʌmsəs	0.297	0.018	0	0	0	0	0	0	0	0	0	6	6.74	8.09
hunpis	hʌnpɪs	0.367	0.029	0	0	0	0	0	0	0	0	0	19	23.76	31.36
jahmpid	ʤɑmpɪd	0.318	0.026	0	0	0	0	0	0	0	0	0	8	9.37	10.15
jehbmut	ʤɛbmət	0.258	0.012	0	0	0	0	0	0	0	0	0	0	0.00	0.00
jehftiz	ʤɛftɪz	0.334	0.032	1	1	1.48	0	0	0	1	1	1.48	17	19.11	22.35
jehksim	ʤɛksɪm	0.331	0.030	0	0	0	0	0	0	0	0	0	18	21.78	27.97
jekseng	ʤɛksəŋ	0.293	0.022	0	0	0	0	0	0	0	0	0	6	7.02	7.27
jikwun	ʤɪkwən	0.313	0.027	0	0	0	0	0	0	0	0	0	2	2.00	3.21
jimvud	ʤɪmvəd	0.287	0.014	0	0	0	0	0	0	0	0	0	11	11.11	14.96
jimvun	ʤɪmvən	0.304	0.023	0	0	0	0	0	0	0	0	0	5	7.06	7.41
joansid	ʤoʊnsɪd	0.343	0.028	0	0	0	0	0	0	0	0	0	7	7.11	9.17
kahldiz	kɑldɪz	0.440	0.046	1	1	1	1	1	1	0	0	0	58	75.05	91.77
kahlpis	kɑlpɪs	0.433	0.041	1	1.09	1.162	0	0	0	1	1.09	1.162	31	44.13	49.83
kamtit	kæmtɪt	0.458	0.040	0	0	0	0	0	0	0	0	0	20	26.20	29.76
kanput	kænpət	0.409	0.029	0	0	0	0	0	0	0	0	0	26	39.27	50.47
kaulnen	kɑlnən	0.396	0.038	0	0	0	0	0	0	0	0	0	19	24.28	27.58
kehksum	kɛksəm	0.362	0.025	0	0	0	0	0	0	0	0	0	9	9.00	12.79
kehpsut	kɛpsət	0.383	0.014	0	0	0	0	0	0	0	0	0	0	0.00	0.00
kemgiz	kɛmgɪz	0.372	0.012	0	0	0	0	0	0	0	0	0	33	44.63	53.82
kilsid	kilsɪd	0.409	0.022	0	0	0	0	0	0	0	0	0	19	24.41	27.67
kimyiv	kɪmjɪv	0.391	0.011	0	0	0	0	0	0	0	0	0	7	8.36	8.37
kintit	kɪntɪt	0.552	0.051	0	0	0	0	0	0	0	0	0	9	11.80	13.14
kinvet	kɪnvət	0.436	0.019	0	0	0	0	0	0	0	0	0	6	6.00	6.07
kipsis	kɪpsɪs	0.460	0.034	0	0	0	0	0	0	0	0	0	13	20.52	24.68
kitfem	kɪtfəm	0.380	0.016	0	0	0	0	0	0	0	0	0	2	3.48	3.50
kownseng	kaʊnsəŋ	0.346	0.021	0	0	0	0	0	0	0	0	0	19	33.40	43.39
kunsik	kʌnsɪk	0.427	0.041	0	0	0	0	0	0	0	0	0	13	15.53	15.58
lelsid	lɛlsɪd	0.385	0.031	0	0	0	0	0	0	0	0	0	15	15.70	19.59
lulsek	lʌlsək	0.305	0.012	0	0	0	0	0	0	0	0	0	4	4.85	5.19
malruk	mælʀək	0.335	0.020	0	0	0	0	0	0	0	0	0	6	6.87	7.28
manvit	mænvɪt	0.392	0.029	0	0	0	0	0	0	0	0	0	18	22.66	25.45
manyev	mænjəv	0.293	0.024	0	0	0	0	0	0	0	0	0	13	18.79	20.49
membik	mɛmbɪk	0.326	0.023	0	0	0	0	0	0	0	0	0	7	9.10	10.29
milpim	mɪlpɪm	0.405	0.030	0	0	0	0	0	0	0	0	0	6	10.82	12.25
nahlvus	nɑlvəs	0.290	0.016	0	0	0	0	0	0	0	0	0	3	4.68	4.68
nalpus	nælpəs	0.303	0.017	0	0	0	0	0	0	0	0	0	9	12.39	14.23
naltum	næltəm	0.321	0.026	0	0	0	0	0	0	0	0	0	8	10.25	10.89
nanrun	nænʀən	0.350	0.033	0	0	0	0	0	0	0	0	0	9	16.78	17.73
naumpim	nɑmpɪm	0.309	0.024	0	0	0	0	0	0	0	0	0	6	6.41	6.95
nehpsuk	nɛpsək	0.276	0.012	0	0	0	0	0	0	0	0	0	3	3.00	3.00
nildus	nɪldəs	0.367	0.029	0	0	0	0	0	0	0	0	0	7	7.83	8.04
nilpis	nɪlpɪs	0.402	0.033	0	0	0	0	0	0	0	0	0	10	11.53	12.91
nisren	nɪsʀən	0.383	0.044	0	0	0	0	0	0	0	0	0	3	3.87	4.10
nuntis	nʌntɪs	0.394	0.055	0	0	0	0	0	0	0	0	0	26	36.97	42.58
nutvit	nʌtvɪt	0.323	0.013	0	0	0	0	0	0	0	0	0	7	7.22	8.29
pablus	pæbləs	0.317	0.017	0	0	0	0	0	0	0	0	0	11	15.13	16.68
pagneng	pɑgnəŋ	0.308	0.013	0	0	0	0	0	0	0	0	0	3	4.19	4.81
pahmfus	pɑmfəs	0.333	0.019	0	0	0	0	0	0	0	0	0	4	4.96	5.11

Table A.1 English nonword stimuli (continued)

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
palkus	pælkəs	0.373	0.022	0	0	0	0	0	0	0	0	0	23	28.69	33.44
paltik	pæltɪk	0.440	0.043	0	0	0	0	0	0	0	0	0	37	50.07	55.02
paybfit	peɪbfɪt	0.334	0.014	0	0	0	0	0	0	0	0	0	3	7.97	8.88
pehlpim	pɛlpɪm	0.393	0.022	0	0	0	0	0	0	0	0	0	21	27.05	30.61
pilkik	pɪlkɪk	0.454	0.033	0	0	0	0	0	0	0	0	0	14	18.01	20.39
pinwus	pɪnwəs	0.407	0.020	0	0	0	0	0	0	0	0	0	7	7.90	7.93
pitwus	pɪtwəs	0.386	0.018	0	0	0	0	0	0	0	0	0	9	11.58	12.83
poafsing	poʊfsɪŋ	0.338	0.041	0	0	0	0	0	0	0	0	0	24	33.30	43.23
poalsid	poʊlsɪd	0.396	0.027	0	0	0	0	0	0	0	0	0	24	30.03	36.76
punlun	pʌnlən	0.373	0.029	0	0	0	0	0	0	0	0	0	20	27.39	28.66
rahldid	ʀɑldɪd	0.404	0.031	0	0	0	0	0	0	0	0	0	35	39.68	52.92
rehkfudge	ʀɛkfəʤ	0.276	0.017	0	0	0	0	0	0	0	0	0	39	43.67	57.84
rehlmum	ʀɛlməm	0.310	0.028	0	0	0	0	0	0	0	0	0	7	8.38	9.65
rehpfun	ʀɛpfən	0.301	0.026	0	0	0	0	0	0	0	0	0	8	13.56	14.78
rinkut	ʀɪnkət	0.433	0.037	0	0	0	0	0	0	0	0	0	3	4.81	6.91
roindiz	ʀɔɪndɪz	0.352	0.033	0	0	0	0	0	0	0	0	0	6	7.77	10.98
saskik	sæskɪk	0.438	0.029	0	0	0	0	0	0	0	0	0	14	15.50	16.16
sebyat	sɛbjæt	0.298	0.012	0	0	0	0	0	0	0	0	0	0	0.00	0.00
sehlkuk	sɛlkək	0.393	0.026	0	0	0	0	0	0	0	0	0	26	32.83	38.17
sehnkim	sɛnkɪm	0.451	0.030	0	0	0	0	0	0	0	0	0	5	9.60	11.06
shastid	ʃæstɪd	0.380	0.051	0	0	0	0	0	0	0	0	0	12	19.00	26.48
shoalsiz	ʃoʊlsɪz	0.328	0.026	0	0	0	0	0	0	0	0	0	63	72.30	102.03
shoasdid	ʃoʊsdɪd	0.318	0.020	0	0	0	0	0	0	0	0	0	20	23.56	27.41
silsis	sɪlsɪs	0.520	0.046	0	0	0	0	0	0	0	0	0	19	20.51	30.92
soafkiz	soʊfkɪz	0.365	0.015	0	0	0	0	0	0	0	0	0	6	7.04	9.30
sulmik	sʌlmɪk	0.401	0.022	0	0	0	0	0	0	0	0	0	12	14.40	15.75
tamrudge	tæmʀəʤ	0.285	0.016	0	0	0	0	0	0	0	0	0	0	0.00	0.00
tayldiz	teɪldɪz	0.353	0.029	0	0	0	0	0	0	0	0	0	18	24.13	31.04
tehpmuk	tɛpmək	0.270	0.014	0	0	0	0	0	0	0	0	0	0	0.00	0.00
tilvus	tɪlvəs	0.356	0.022	0	0	0	0	0	0	0	0	0	12	12.29	14.02
toamsiz	toʊmsɪz	0.335	0.021	0	0	0	0	0	0	0	0	0	22	26.96	35.88
towspid	taʊspɪd	0.316	0.021	0	0	0	0	0	0	0	0	0	8	9.25	12.48
tulsid	tʌlsɪd	0.373	0.023	0	0	0	0	0	0	0	0	0	13	13.73	16.05
tusfik	tʌsfɪk	0.325	0.018	0	0	0	0	0	0	0	0	0	7	9.36	9.68
vahlpish	vɑlpɪʃ	0.322	0.016	0	0	0	0	0	0	0	0	0	3	3.16	4.39
vaubsim	vɑbsɪm	0.303	0.016	0	0	0	0	0	0	0	0	0	0	0.00	0.00
vifking	vɪfkɪŋ	0.336	0.041	0	0	0	0	0	0	0	0	0	9	9.00	9.51
vimlut	vɪmlət	0.337	0.017	0	0	0	0	0	0	0	0	0	31	43.02	53.88
visrin	vɪsʀɪn	0.425	0.047	0	0	0	0	0	0	0	0	0	7	9.42	10.66
voamwek	voʊmwək	0.194	0.005	0	0	0	0	0	0	0	0	0	2	2.96	2.96
vumsing	vʌmsɪŋ	0.307	0.045	0	0	0	0	0	0	0	0	0	26	36.63	43.48
wafsid	wæfsɪd	0.322	0.018	0	0	0	0	0	0	0	0	0	9	10.57	14.54
wahkching	wɑkʧɪŋ	0.312	0.038	1	2.908	3.399	0	0	0	1	2.908	3.399	44	62.75	78.44
waifpiz	waɪfpɪz	0.272	0.016	0	0	0	0	0	0	0	0	0	7	7.97	8.80
waimlit	waɪmlɪt	0.321	0.024	0	0	0	0	0	0	0	0	0	23	32.27	33.48
yailking	jaɪlkɪŋ	0.299	0.038	0	0	0	0	0	0	0	0	0	22	26.91	33.49

A.2 English Words

Table A.2 English word stimuli

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
basking	bæskɪŋ	bi	1.30	1.60	0.386	0.050	3	5.02	5.67	0	0.00	0.00	3	5.02	5.67	64	84.34	108.96
basting	beɪstɪŋ	bi	1.00	1.30	0.392	0.072	11	14.80	21.09	1	1.00	1.30	10	13.80	19.79	108	147.45	187.29
binding	baɪndɪŋ	bi	1.48	1.48	0.380	0.060	9	13.99	19.85	1	1.03	1.48	8	12.96	18.37	89	128.23	164.96
boasting	boʊstɪŋ	bi	1.48	2.08	0.391	0.070	14	16.05	21.90	1	1.00	2.07	13	15.05	19.83	81	102.98	131.89
bolted	boʊltɪd	bi	1.00	1.78	0.397	0.038	6	6.03	8.79	1	1.03	1.79	5	5.00	7.00	66	76.71	102.57
bounces	baʊnsɪz	bi	1.00	1.30	0.373	0.032	7	8.70	10.31	5	5.73	6.45	2	2.97	3.86	55	74.30	106.31
bounded	baʊndɪd	bi	1.00	1.78	0.362	0.038	12	15.21	24.40	2	2.27	3.45	10	12.94	20.96	73	90.45	119.49
boxes	bɑksɪz	bi	2.38	3.01	0.405	0.036	6	8.09	10.29	3	3.63	4.62	3	4.46	5.67	97	120.22	158.78
chances	ʧænsɪz	bi	2.51	3.25	0.374	0.038	2	2.40	3.36	0	0.00	0.00	2	2.40	3.36	63	84.55	119.52
coasted	koʊstɪd	bi	1.00	1.00	0.428	0.052	10	10.15	16.19	2	2.00	2.15	8	8.15	14.05	79	94.97	124.89
coaxes	koʊksɪz	bi	1.00	1.70	0.393	0.032	5	5.05	6.25	2	2.05	2.72	3	3.00	3.53	67	87.45	119.31
costing	kɑstɪŋ	bi	1.00	1.00	0.462	0.086	6	7.98	9.23	0	0.00	0.00	3	4.58	5.66	95	122.32	156.95
dances	dænsɪz	bi	1.00	2.82	0.426	0.039	4	4.40	5.76	1	1.00	1.40	3	3.40	4.36	70	93.32	128.55
daunting	dɔntɪŋ	bi	1.48	1.60	0.449	0.073	7	7.67	10.22	0	0.00	0.00	5	5.50	6.84	77	105.47	132.67
deltas	dɛltəz	bi	1.00	1.70	0.386	0.038	4	5.59	6.59	1	1.69	1.72	3	3.90	4.88	55	72.73	91.14
feasted	fistɪd	bi	1.00	1.60	0.390	0.050	2	2.33	2.81	1	1.33	1.56	1	1.00	1.25	74	83.09	115.74
feasting	fistɪŋ	bi	1.30	1.60	0.384	0.071	2	2.11	2.81	1	1.00	1.56	1	1.11	1.25	95	131.37	163.49
fielded	fildɪd	bi	1.00	1.78	0.359	0.028	5	6.09	10.24	1	1.41	1.76	4	4.68	8.48	39	54.51	70.69
fixes	fɪksɪz	bi	1.30	2.63	0.434	0.040	9	10.75	13.57	2	2.83	3.61	7	7.92	9.96	89	117.45	151.65
founded	faʊndɪd	bi	1.70	2.48	0.352	0.038	10	12.72	19.60	3	3.72	4.72	7	9.00	14.88	61	79.04	111.31
funded	fʌndɪd	bi	1.00	2.11	0.387	0.042	3	4.56	6.58	1	1.85	2.11	2	2.70	4.47	68	90.15	124.66
gilded	gɪldɪd	bi	1.00	1.60	0.415	0.037	2	2.00	2.67	1	1.00	1.67	1	1.00	1.00	43	52.51	71.53
handed	hændɪd	bi	1.95	2.85	0.390	0.049	11	15.00	20.33	4	6.30	9.48	7	8.70	10.86	91	113.67	149.99
haunted	hɔntɪd	bi	1.30	2.15	0.430	0.055	9	10.26	14.98	0	0.00	0.00	7	7.47	11.48	66	78.79	105.74
helping	hɛlpɪŋ	bi	1.95	2.11	0.357	0.048	3	3.19	3.44	1	1.19	1.37	2	2.00	2.07	63	89.61	113.74
hinted	hɪntɪd	bi	2.58	3.59	0.471	0.052	5	5.73	8.99	1	1.28	2.02	4	4.45	6.97	77	96.36	131.63
hoisted	hɔɪstɪd	bi	1.00	1.60	0.342	0.047	2	2.00	2.62	1	1.00	1.62	1	1.00	1.00	51	60.83	88.95
hosted	hoʊstɪd	bi	1.00	1.00	0.372	0.051	9	10.11	14.31	2	2.96	3.04	7	7.15	11.27	62	77.46	108.34
jolted	ʤoʊltɪd	bi	1.00	1.48	0.347	0.036	6	6.00	7.28	3	3.00	3.50	3	3.00	3.79	44	48.44	65.07
landed	lændɪd	bi	1.60	2.46	0.389	0.047	7	10.22	12.59	1	2.39	2.69	6	7.83	9.90	79	100.69	136.81
lapses	læpsɪz	bi	1.00	1.60	0.347	0.026	5	5.00	7.18	2	2.00	3.07	3	3.00	4.11	46	58.88	77.97
lasted	læstɪd	bi	1.60	2.85	0.403	0.053	1	1.55	2.48	0	0.00	0.00	1	1.55	2.48	77	92.33	123.84
lasting	læstɪŋ	bi	1.60	2.85	0.396	0.074	2	2.66	3.48	1	1.00	1.00	1	1.66	2.48	94	125.47	155.97

Table A.2 English word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
lifted	lɪftɪd	bi	1.90	2.94	0.401	0.038	7	10.18	13.60	1	2.21	2.95	6	7.97	10.65	54	70.12	90.69
listing	lɪstɪŋ	bi	1.00	1.00	0.446	0.094	7	11.60	14.58	3	5.30	6.73	4	6.30	7.85	110	151.37	195.59
lofted	lɔftɪd	bi	1.00	1.00	0.360	0.033	1	1.91	2.95	0	0.00	0.00	1	1.91	2.95	35	40.14	56.17
melted	mɛltɪd	bi	1.30	2.38	0.420	0.044	7	7.80	9.22	3	3.80	4.48	4	4.00	4.73	62	68.36	93.16
melting	mɛltɪŋ	bi	1.78	2.38	0.413	0.064	8	8.27	10.34	4	4.27	5.67	4	4.00	4.67	87	116.08	147.94
mending	mɛndɪŋ	bi	1.30	1.30	0.401	0.069	13	18.85	27.10	3	3.00	4.59	10	15.85	22.51	93	138.06	175.77
minces	mɪnsɪz	bi	1.00	1.00	0.464	0.044	10	17.97	22.72	3	3.09	4.25	7	14.88	18.47	99	117.59	149.58
misted	mɪstɪd	bi	1.00	1.00	0.467	0.077	6	7.75	9.50	4	5.21	6.01	2	2.55	3.48	127	167.15	216.82
painted	peɪntɪd	bi	1.00	1.30	0.436	0.047	9	12.10	16.67	1	2.74	3.17	8	9.37	13.49	87	108.87	151.45
pasting	peɪstɪŋ	bi	1.00	1.00	0.412	0.073	9	13.03	16.65	4	4.59	5.30	5	8.44	11.35	112	155.02	196.26
pointed	pɔɪntɪd	bi	1.70	1.70	0.405	0.042	6	9.44	12.50	2	4.50	5.66	4	4.94	6.85	70	85.35	115.54
posted	poʊstɪd	bi	1.00	1.90	0.418	0.051	9	9.49	13.21	2	2.49	3.20	7	7.00	10.01	83	101.78	136.27
pouncing	paʊnsɪŋ	bi	1.00	1.70	0.385	0.053	3	4.61	5.95	2	2.88	3.74	1	1.73	2.21	74	97.81	126.19
pounded	paʊndɪd	bi	1.00	1.95	0.383	0.039	8	11.58	16.34	2	2.88	3.04	6	8.70	13.30	59	77.77	109.08
punted	pʌntɪd	bi	1.00	1.00	0.448	0.051	8	10.53	16.17	1	1.00	1.00	7	9.53	15.17	89	105.49	141.84
rafted	ʀæftɪd	bi	1.00	1.00	0.373	0.033	2	2.00	2.36	0	0.00	0.00	2	2.00	2.36	60	69.41	92.41
ranking	ʀæŋkɪŋ	bi	1.00	1.00	0.323	0.048	8	9.34	12.46	3	3.00	3.28	5	6.34	9.18	89	116.20	139.15
rested	ʀɛstɪd	bi	1.48	2.72	0.429	0.060	17	22.02	28.35	6	9.44	11.30	11	12.58	17.05	105	125.27	164.59
resting	ʀɛstɪŋ	bi	2.23	2.72	0.423	0.081	18	23.38	29.54	7	9.21	13.16	11	14.17	16.39	127	171.80	219.19
roasted	ʀoʊstɪd	bi	1.00	1.95	0.393	0.051	11	12.71	18.89	1	1.56	1.98	10	11.15	16.91	83	97.01	129.77
rounded	ʀaʊndɪd	bi	1.48	2.26	0.358	0.039	10	12.69	18.75	4	4.45	5.75	6	8.24	13.01	77	93.86	124.80
rusted	ʀʌstɪd	bi	1.00	1.48	0.407	0.054	13	16.16	24.08	6	8.16	10.89	7	8.01	13.19	108	128.73	166.39
senses	sɛnsɪz	bi	1.60	3.44	0.485	0.046	8	12.85	18.49	5	7.97	11.01	3	4.88	7.48	103	135.46	184.27
shafted	ʃæftɪd	bi	1.00	1.00	0.328	0.030	1	1.51	2.66	0	0.00	0.00	1	1.51	2.66	32	38.27	50.29
shielded	ʃildɪd	bi	1.00	1.78	0.320	0.027	4	4.09	7.66	1	1.09	1.80	3	3.00	5.86	26	32.14	45.03
shifted	ʃɪftɪd	bi	1.48	2.65	0.378	0.034	8	11.12	13.84	3	4.30	5.02	5	6.82	8.82	42	52.36	66.63
sifted	sɪftɪd	bi	1.00	1.60	0.481	0.043	4	6.44	9.11	1	1.11	1.60	3	5.33	7.52	87	111.78	142.32
sounded	saʊndɪd	bi	1.90	3.10	0.416	0.039	8	10.66	16.25	1	1.87	3.10	7	8.79	13.15	72	94.40	127.93
tainted	teɪntɪd	bi	1.00	1.30	0.399	0.046	8	9.13	13.97	1	1.00	1.35	7	8.13	12.62	71	87.19	124.13
tainting	teɪntɪŋ	bi	1.00	1.30	0.392	0.067	7	9.55	13.32	1	1.00	1.35	6	8.55	11.97	100	141.43	182.41
taxes	tæksɪz	bi	1.00	1.70	0.379	0.036	9	11.18	14.04	3	4.79	5.66	6	6.39	8.38	86	103.98	132.90
tenses	tɛnsɪz	bi	1.00	1.60	0.415	0.043	5	8.54	12.33	2	2.00	2.78	3	6.54	9.55	94	121.19	170.13
tested	tɛstɪd	bi	1.60	2.73	0.418	0.057	14	17.22	22.54	5	6.32	7.02	9	10.91	15.52	98	113.54	153.65
testing	tɛstɪŋ	bi	2.32	2.73	0.411	0.078	14	17.70	21.83	5	5.59	6.99	9	12.11	14.84	110	145.59	189.00
tilted	tɪltɪd	bi	1.00	2.15	0.463	0.044	6	6.50	8.01	2	2.50	3.16	4	4.00	4.85	74	88.48	118.00
toasted	toʊstɪd	bi	1.00	1.60	0.382	0.049	10	10.93	16.91	1	1.00	1.56	9	9.93	15.35	74	87.11	117.15
vented	vɛntɪd	bi	1.00	1.48	0.402	0.054	8	9.00	12.04	2	2.00	2.41	6	7.00	9.63	81	97.82	137.39

Table A.2 English word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
wanted	wɑntɪd	bi	2.83	4.05	0.422	0.052	5	6.52	9.36	2	3.52	5.05	3	3.00	4.31	86	103.54	142.00
welding	wɛldɪŋ	bi	1.48	1.78	0.362	0.058	8	9.96	12.79	4	4.13	6.40	4	5.83	6.39	86	123.87	158.87
wilted	wɪltɪd	bi	1.00	1.48	0.445	0.046	5	5.38	6.93	2	2.27	2.77	3	3.11	4.16	76	95.62	134.08
winding	waɪndɪŋ	bi	1.00	1.00	0.346	0.061	5	8.57	13.40	1	1.54	1.81	4	7.03	11.59	74	111.91	143.15
yelping	jɛlpɪŋ	bi	1.00	1.00	0.325	0.044	4	6.50	8.30	2	2.91	3.71	2	3.60	4.59	30	37.28	47.05
yielding	jildɪŋ	bi	1.48	2.28	0.306	0.047	4	4.83	7.66	1	1.00	2.26	3	3.83	5.39	43	71.58	91.26
bandage	bændɪʤ	mono	1.00	1.60	0.386	0.045	9	9.44	12.10	8	8.03	10.69	1	1.41	1.41	54	74.67	94.46
bandit	bændɪt	mono	1.00	1.48	0.441	0.050	9	9.78	12.12	8	8.78	11.12	1	1.00	1.00	70	89.07	109.08
basket	bæskɪt	mono	2.26	2.38	0.418	0.029	3	3.69	4.19	1	1.00	1.00	2	2.69	3.19	34	47.00	53.10
biscuit	bɪskɪt	mono	1.70	2.18	0.469	0.048	3	5.28	5.57	1	2.02	2.19	2	3.25	3.38	40	49.83	62.20
cactus	kæktəs	mono	1.30	1.48	0.405	0.036	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	28	42.87	50.12
candid	kændɪd	mono	1.30	1.30	0.446	0.053	9	11.55	15.27	3	3.80	4.83	6	7.75	10.44	86	107.69	144.01
canvass	kænvəs	mono	1.00	1.60	0.370	0.029	1	1.00	1.48	1	1.00	1.48	0	0.00	0.00	21	31.16	37.45
captain	kæptɪn	mono	2.80	2.85	0.440	0.044	5	6.52	7.73	3	4.52	5.73	2	2.00	2.00	41	47.20	49.28
captive	kæptɪv	mono	1.70	1.70	0.396	0.035	2	4.06	4.26	2	4.06	4.26	0	0.00	0.00	17	19.78	22.29
casket	kæskɪt	mono	1.30	1.30	0.458	0.039	3	4.25	4.84	2	2.00	2.47	1	2.25	2.38	55	78.28	91.30
census	sɛnsɪs	mono	1.70	1.78	0.492	0.052	6	10.56	14.78	5	9.56	13.78	1	1.00	1.00	52	72.88	91.95
comfort	kʌmfət	mono	1.00	2.32	0.330	0.019	3	3.72	5.07	3	3.72	5.07	0	0.00	0.00	23	25.63	28.43
compass	kʌmpəs	mono	1.70	1.78	0.331	0.029	3	4.29	4.37	1	1.00	1.00	2	3.29	3.37	39	51.13	64.06
conscious	kɑnʃəs	mono	2.65	2.65	0.376	0.032	1	2.01	2.01	0	0.00	0.00	1	2.01	2.01	12	19.19	21.09
cosmic	kɑzmɪk	mono	1.85	1.85	0.354	0.030	1	2.16	2.17	0	0.00	0.00	1	2.16	2.17	13	15.37	17.09
custom	kʌstəm	mono	2.20	2.38	0.376	0.045	5	8.43	9.29	5	8.43	9.29	0	0.00	0.00	18	26.77	32.33
dictum	dɪktəm	mono	1.30	1.30	0.399	0.055	2	2.00	2.53	2	2.00	2.53	0	0.00	0.00	15	21.35	26.18
dimwit	dɪmwɪt	mono	1.00	1.00	0.407	0.040	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	15	17.34	21.23
discus	dɪskəs	mono	1.00	1.00	0.403	0.062	4	6.40	8.00	3	5.29	6.89	1	1.11	1.11	28	38.53	47.33
dolphin	dɑlfɪn	mono	1.00	1.48	0.394	0.025	1	1.00	1.00	0	0.00	0.00	1	1.00	1.00	16	16.22	16.48
fungus	fʌŋgəs	mono	1.70	1.95	0.235	0.014	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	9	11.50	14.01
gambit	gæmbɪt	mono	1.90	2.00	0.336	0.021	3	3.29	4.00	2	2.29	3.00	1	1.00	1.00	26	33.58	38.37
gasket	gæskɪt	mono	1.00	1.00	0.385	0.025	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	26	35.71	40.54
goblin	gɑblɪn	mono	1.00	1.00	0.329	0.030	4	4.15	4.78	4	4.15	4.78	0	0.00	0.00	28	31.45	33.26
gypsum	ʤɪpsəm	mono	1.00	1.00	0.307	0.019	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	8	9.10	9.77
hectic	hɛktɪk	mono	1.48	1.48	0.384	0.043	1	1.00	1.00	0	0.00	0.00	1	1.00	1.00	20	24.32	27.43
hospice	hɑspɪs	mono	1.00	1.00	0.378	0.034	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	25	33.80	38.48
hostage	hɑstəʤ	mono	1.30	1.70	0.335	0.041	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	3	4.15	4.91
hybrid	haɪbʀɪd	mono	1.60	1.60	0.302	0.025	1	1.00	1.35	1	1.00	1.35	0	0.00	0.00	11	14.46	17.76
jaundice	ʤɔndɪs	mono	1.00	1.00	0.385	0.048	1	1.00	1.00	0	0.00	0.00	1	1.00	1.00	18	18.07	18.23
justice	ʤʌstɪs	mono	2.70	2.72	0.370	0.057	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	46	57.71	73.80

Table A.2 English word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
kelvin	kɛlvɪn	mono	1.00	1.00	0.413	0.025	2	2.00	2.00	1	1.00	1.00	1	1.00	1.00	21	24.98	25.84
lactic	læktɪk	mono	1.30	1.30	0.379	0.042	1	1.75	2.33	0	0.00	0.00	1	1.75	2.33	30	46.32	50.15
linkage	lɪŋkɪʤ	mono	1.00	1.30	0.327	0.024	1	1.79	2.62	1	1.79	2.62	0	0.00	0.00	37	48.24	55.36
liquid	lɪkwɪd	mono	2.45	2.48	0.367	0.025	1	1.29	2.28	1	1.29	2.28	0	0.00	0.00	19	23.14	27.60
litmus	lɪtməs	mono	1.00	1.00	0.356	0.023	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	13	15.31	18.70
maxim	mæksɪm	mono	1.48	1.48	0.363	0.035	6	6.46	8.20	6	6.46	8.20	0	0.00	0.00	32	44.00	55.23
metric	mɛtʀɪk	mono	1.00	2.67	0.385	0.037	2	2.00	2.00	1	1.00	1.00	1	1.00	1.00	25	33.39	35.65
million	mɪljən	mono	1.00	1.60	0.369	0.038	5	8.29	8.37	2	3.68	3.68	3	4.61	4.69	30	36.31	39.83
musket	mʌskɪt	mono	1.30	1.30	0.387	0.027	3	3.00	3.07	3	3.00	3.07	0	0.00	0.00	49	64.08	70.40
mystic	mɪstɪk	mono	1.70	1.85	0.458	0.078	10	13.75	15.90	9	12.75	14.90	1	1.00	1.00	55	86.77	95.39
napkin	næpkɪn	mono	1.00	1.00	0.331	0.022	1	1.35	1.84	1	1.35	1.84	0	0.00	0.00	16	19.18	20.60
nitpick	nɪtpɪk	mono	1.30	1.30	0.381	0.023	3	3.00	3.00	3	3.00	3.00	0	0.00	0.00	13	13.84	15.69
noxious	nɑkʃəs	mono	1.95	2.86	0.273	0.015	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	13	17.49	19.27
pectin	pɛktɪn	mono	1.85	1.85	0.456	0.045	3	3.00	3.00	1	1.00	1.00	2	2.00	2.00	42	44.29	48.30
pelvic	pɛlvɪk	mono	1.95	1.95	0.377	0.026	1	1.95	1.95	1	1.95	1.95	0	0.00	0.00	18	20.41	23.83
pelvis	pɛlvɪs	mono	1.00	1.00	0.395	0.031	1	1.85	1.85	1	1.85	1.85	0	0.00	0.00	20	23.19	27.83
peptic	pɛptɪk	mono	1.48	1.60	0.409	0.036	3	3.00	3.00	1	1.00	1.00	2	2.00	2.00	23	27.54	30.50
phantom	fæntəm	mono	1.00	1.00	0.370	0.044	2	2.03	2.56	1	1.03	1.56	1	1.00	1.00	23	35.42	38.35
picnic	pɪknɪk	mono	2.08	3.12	0.436	0.030	3	3.34	4.18	3	3.34	4.18	0	0.00	0.00	27	33.23	36.45
pompous	pɑmpəs	mono	1.30	1.30	0.349	0.029	1	1.00	1.00	0	0.00	0.00	1	1.00	1.00	29	36.46	42.48
postage	poʊstɪʤ	mono	1.00	1.90	0.389	0.046	2	2.38	3.85	2	2.38	3.85	0	0.00	0.00	40	52.49	60.21
public	pʌblɪk	mono	3.55	3.55	0.329	0.031	2	2.50	5.94	2	2.50	5.94	0	0.00	0.00	14	22.13	24.38
publish	pʌblɪʃ	mono	1.48	2.93	0.314	0.026	3	7.58	8.67	3	7.58	8.67	0	0.00	0.00	18	28.42	36.29
pulpit	pʌlpɪt	mono	1.60	1.70	0.412	0.020	5	5.40	5.68	4	4.00	4.05	1	1.40	1.63	31	35.33	42.98
pundit	pʌndɪt	mono	1.00	1.00	0.443	0.043	3	3.00	3.00	2	2.00	2.00	1	1.00	1.00	34	44.46	55.76
rancid	ʀænsɪd	mono	1.00	1.00	0.420	0.041	2	2.00	3.58	1	1.00	1.36	1	1.00	2.22	41	52.69	67.14
random	ʀændəm	mono	2.26	2.26	0.346	0.040	2	2.22	2.32	1	1.22	1.32	1	1.00	1.00	36	50.23	54.98
rumpus	ʀʌmpəs	mono	1.00	1.00	0.296	0.027	2	2.67	3.27	1	1.00	1.53	1	1.67	1.75	32	37.96	48.04
rustic	ʀʌstɪk	mono	1.60	2.70	0.398	0.055	5	6.01	8.20	5	6.01	8.20	0	0.00	0.00	56	72.15	87.13
salvage	sælvɪʤ	mono	1.00	1.78	0.386	0.016	5	6.18	7.03	3	3.00	3.74	2	3.18	3.29	22	26.82	30.18
seismic	saɪzmɪk	mono	1.00	1.00	0.339	0.018	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	7	9.78	11.35
seldom	sɛldəm	mono	2.52	2.52	0.391	0.034	1	1.00	1.00	0	0.00	0.00	1	1.00	1.00	15	20.95	22.75
selfish	sɛlfɪʃ	mono	2.08	2.08	0.402	0.028	2	2.13	2.43	0	0.00	0.00	2	2.13	2.43	18	22.77	24.14
septic	sɛptɪk	mono	1.00	1.00	0.442	0.040	3	3.00	3.32	1	1.00	1.00	2	2.00	2.32	29	36.11	39.39
surplus	sɜpləs	mono	2.34	2.40	0.320	0.014	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	10	16.05	17.11
tactic	tæktɪk	mono	1.78	2.32	0.389	0.042	3	5.64	5.75	2	4.37	4.49	1	1.27	1.27	39	55.97	61.77
tendon	tɛndən	mono	1.48	1.70	0.373	0.050	4	4.33	5.14	4	4.33	5.14	0	0.00	0.00	48	62.91	73.82

Table A.2 English word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
toxic	tɑksɪk	mono	1.78	1.78	0.378	0.035	2	2.00	2.05	2	2.00	2.05	0	0.00	0.00	33	48.15	54.87
toxin	tɑksɪn	mono	1.00	1.00	0.404	0.037	7	7.77	7.82	3	3.77	3.82	4	4.00	4.00	44	52.84	59.79
vestige	vɛstɪʤ	mono	1.00	1.48	0.355	0.049	3	3.71	3.90	3	3.71	3.90	0	0.00	0.00	44	56.24	65.58
victim	vɪktɪm	mono	2.45	2.73	0.399	0.038	1	2.42	2.74	1	2.42	2.74	0	0.00	0.00	7	8.55	10.77
vintage	vɪntɪʤ	mono	1.48	1.48	0.418	0.046	1	1.48	1.48	0	0.00	0.00	1	1.48	1.48	19	24.32	28.39
welcome	wɛlkəm	mono	2.52	2.80	0.298	0.024	2	2.62	5.09	2	2.62	5.09	0	0.00	0.00	13	16.39	17.15

Table A.3 Distribution of Phonemes for English stimuli

phon	C1	V1	C2	C3	V2	C4	C1	V1	C2	C3	V2	C4	C1	V1	C2	C3	V2	C4
	monomorphemes						bimorphemes						nonwords					
b	4		4	1			8						9		5	2		
d	4			8		4	3			16		42	13			14		23
g	3			1			1						7		1	2		
p	11		7	6			7		1	2			13		9	18		
t	4		3	20		11	8			44			8		3	17		20
k	12		13	9		15	3		4	2			16		9	15		22
ʤ	3					7	1						9					6
ʧ							1						5			1		
f	2			3			6		6				14		11	10		
v	3			5		1	1						7			7		2
z			2									12			3			16
s	7		14	6		16	3		22	12			7		17	29		19
ʃ				2		2	3						3					1
h	4						6						10					
l	4		10	4			7		14				2		42	8		
ɹ	4			2			7						6			8		
j				1			2						1			3		
w	1			2			4						4			6		
m	5		6	3		9	5						5		23	7		17
n	3		13	1		9			28				11		27	3		13
ŋ			2						1			22						11
i								5						2				
ɪ		14			52			11			75			28			92	
eɪ								5						4				
ɛ		14						14						36				
æ		19						12						23			1	
ɔ		10						3						20				
ʌ		12						3						14				
əʊ		1						9						13				
ɔ		1						3										
aɪ		2						2						3				
ɔɪ								2						1				
aʊ								7						6				
ə					22						1						57	

A.3 German Nonwords

Table A.4 German nonword stimuli

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
reungken	ʀɔyŋkən	0.293	0.044	6	6	6.97	0	0	0	6	6	6.97	93	120.89	158.96
wongkess	vɔŋkəs	0.306	0.028	1	1	1.77	0	0	0	1	1	1.77	41	52.30	78.24
bomgIch	bɔmgɪx	0.211	0.017	1	1	1	1	1	1	0	0	0	6	6.00	6.57
kozlich	kɔt͡slɪx	0.213	0.029	1	1	1	0	0	0	1	1	1	20	23.55	24.16
schintoss	ʃɪntɔs	0.364	0.026	1	1	1	1	1	1	0	0	0	4	4.00	4.00
beunzess	bɔynt͡səs	0.328	0.022	0	0	0	0	0	0	0	0	0	12	15.54	27.58
bilpel	bɪlpəl	0.377	0.026	0	0	0	0	0	0	0	0	0	20	27.54	35.48
buchder	bʊxdəʀ	0.346	0.039	0	0	0	0	0	0	0	0	0	25	43.45	58.54
dachder	daxdəʀ	0.344	0.043	0	0	0	0	0	0	0	0	0	17	24.39	37.56
dachner	daxnəʀ	0.360	0.036	0	0	0	0	0	0	0	0	0	26	39.46	55.48
dalder	daldəʀ	0.380	0.049	0	0	0	0	0	0	0	0	0	33	54.91	73.64
dangfiss	daŋfɪs	0.214	0.007	0	0	0	0	0	0	0	0	0	2	2.18	5.65
delpel	dɛlpəl	0.387	0.028	0	0	0	0	0	0	0	0	0	10	11.86	12.04
dengpel	dɛŋpəl	0.334	0.021	0	0	0	0	0	0	0	0	0	22	35.21	41.05
denter	dɛntəʀ	0.491	0.080	0	0	0	0	0	0	0	0	0	128	183.70	262.94
dilnel	dɪlnəl	0.353	0.025	0	0	0	0	0	0	0	0	0	7	7.60	8.24
dirder	dɪʀdəʀ	0.443	0.048	0	0	0	0	0	0	0	0	0	33	49.14	66.41
dirdess	dɪʀdəs	0.412	0.037	0	0	0	0	0	0	0	0	0	41	45.50	80.14
dirsess	dɪʀzəs	0.405	0.028	0	0	0	0	0	0	0	0	0	34	39.14	61.55
dokpfess	dɔkp͡fəs	0.280	0.019	0	0	0	0	0	0	0	0	0	32	33.50	45.90
dontum	dɔntʊm	0.264	0.024	0	0	0	0	0	0	0	0	0	3	4.22	4.66
dulness	dʊlnəs	0.331	0.029	0	0	0	0	0	0	0	0	0	23	25.61	37.72
durder	dʊʀdəʀ	0.427	0.059	0	0	0	0	0	0	0	0	0	26	43.22	64.43
durdess	dʊʀdəs	0.397	0.048	0	0	0	0	0	0	0	0	0	51	67.70	113.86
fangkuss	faŋkʊs	0.265	0.014	0	0	0	0	0	0	0	0	0	19	20.02	27.43
fiktuss	fɪktʊs	0.345	0.013	0	0	0	0	0	0	0	0	0	18	19.47	22.40
finbek	fɪnbək	0.368	0.021	0	0	0	0	0	0	0	0	0	9	15.95	19.25
forjek	fɔʀjɛk	0.330	0.020	0	0	0	0	0	0	0	0	0	5	7.06	7.45
funfek	fʊnfɛk	0.263	0.013	0	0	0	0	0	0	0	0	0	1	1.64	1.68
fungpel	fʊŋpəl	0.335	0.019	0	0	0	0	0	0	0	0	0	17	21.72	26.47
furder	fʊʀdəʀ	0.490	0.052	0	0	0	0	0	0	0	0	0	56	93.69	120.56
furkuss	fʊʀkʊs	0.366	0.022	0	0	0	0	0	0	0	0	0	14	16.85	19.43
gachpel	gaxpəl	0.340	0.023	0	0	0	0	0	0	0	0	0	10	11.13	12.02
galper	galpəʀ	0.401	0.042	0	0	0	0	0	0	0	0	0	32	49.89	67.77
gilkess	gɪlkəs	0.363	0.031	0	0	0	0	0	0	0	0	0	41	45.94	73.49
girder	gɪʀdəʀ	0.460	0.046	0	0	0	0	0	0	0	0	0	36	50.49	64.60
goelgon	gœlgɔn	0.219	0.004	0	0	0	0	0	0	0	0	0	2	2.62	5.12
gokper	gɔkpəʀ	0.350	0.031	0	0	0	0	0	0	0	0	0	7	7.85	7.91
golgam	gɔlgam	0.216	0.012	0	0	0	0	0	0	0	0	0	0	0.00	0.00
guchkil	gʊxkɪl	0.232	0.004	0	0	0	0	0	0	0	0	0	0	0.00	0.00
hengbol	hɛŋbɔl	0.235	0.014	0	0	0	0	0	0	0	0	0	1	1.07	1.37
hiksess	hɪkzəs	0.316	0.027	0	0	0	0	0	0	0	0	0	14	15.84	24.43
hirder	hɪʀdəʀ	0.459	0.051	0	0	0	0	0	0	0	0	0	52	76.21	98.62
huchner	hʊxnəʀ	0.342	0.031	0	0	0	0	0	0	0	0	0	11	15.91	17.16
jetkon	jɛtkɔn	0.291	0.005	0	0	0	0	0	0	0	0	0	2	2.00	2.00
kaldel	kaldəl	0.383	0.041	0	0	0	0	0	0	0	0	0	35	46.24	59.33
kechden	kɛxdən	0.391	0.049	0	0	0	0	0	0	0	0	0	87	120.22	172.09
kelpel	kɛlpəl	0.414	0.028	0	0	0	0	0	0	0	0	0	29	38.47	43.21
kelpuss	kɛlpʊs	0.298	0.013	0	0	0	0	0	0	0	0	0	10	10.07	13.84
kendum	kɛndʊm	0.261	0.020	0	0	0	0	0	0	0	0	0	9	11.84	12.21

Table A.4 German nonword stimuli (continued)

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
kenzir	kɛnt͜sɪʀ	0.347	0.014	0	0	0	0	0	0	0	0	0	5	6.67	6.79
kepfor	kɛpfɔʀ	0.282	0.006	0	0	0	0	0	0	0	0	0	0	0.00	0.00
kilduss	kɪldʊs	0.248	0.011	0	0	0	0	0	0	0	0	0	11	15.21	24.00
kilkuss	kɪlkʊs	0.264	0.011	0	0	0	0	0	0	0	0	0	10	11.60	13.49
kirter	kɪʀtəʀ	0.555	0.077	0	0	0	0	0	0	0	0	0	118	159.74	203.13
konkik	kɔnkɪk	0.253	0.022	0	0	0	0	0	0	0	0	0	3	3.94	4.02
kulder	kʊldəʀ	0.373	0.044	0	0	0	0	0	0	0	0	0	65	99.21	124.66
kuldul	kʊldʊl	0.239	0.009	0	0	0	0	0	0	0	0	0	1	1.00	1.07
kumbur	kʊmbʊʀ	0.233	0.007	0	0	0	0	0	0	0	0	0	4	5.01	5.02
lansar	lanzaʀ	0.286	0.019	0	0	0	0	0	0	0	0	0	4	5.37	5.77
lesskur	lɛskʊʀ	0.297	0.012	0	0	0	0	0	0	0	0	0	3	4.51	4.55
leumgess	lɔymgəs	0.270	0.029	0	0	0	0	0	0	0	0	0	12	12.00	15.94
lichjur	lɪxjʊʀ	0.203	0.007	0	0	0	0	0	0	0	0	0	3	3.78	5.78
lirpess	lɪʀpəs	0.423	0.026	0	0	0	0	0	0	0	0	0	28	31.08	47.77
lirpfess	lɪʀp͜fəs	0.401	0.022	0	0	0	0	0	0	0	0	0	22	23.14	38.04
lisspuss	lɪspʊs	0.208	0.011	0	0	0	0	0	0	0	0	0	4	4.70	4.70
loefnem	lœfnəm	0.249	0.006	0	0	0	0	0	0	0	0	0	7	10.20	13.95
lurber	lʊʀbəʀ	0.435	0.052	0	0	0	0	0	0	0	0	0	22	30.65	36.96
mapfich	mafpɪx	0.212	0.019	0	0	0	0	0	0	0	0	0	7	7.00	7.60
makpess	makpəs	0.339	0.029	0	0	0	0	0	0	0	0	0	33	38.81	55.24
massnem	masnəm	0.319	0.018	0	0	0	0	0	0	0	0	0	21	29.24	47.52
meingkem	maiŋkəm	0.255	0.014	0	0	0	0	0	0	0	0	0	6	8.06	11.83
mersem	mɛʀzəm	0.429	0.065	0	0	0	0	0	0	0	0	0	39	43.81	67.31
mirdel	mɪʀdəl	0.428	0.039	0	0	0	0	0	0	0	0	0	37	50.16	60.97
moenfin	mœnfɪn	0.238	0.006	0	0	0	0	0	0	0	0	0	2	2.58	3.40
mofkem	mɔfkəm	0.283	0.010	0	0	0	0	0	0	0	0	0	4	4.00	4.84
mokpel	mɔkpəl	0.318	0.019	0	0	0	0	0	0	0	0	0	9	10.11	10.29
monzich	mɔnt͜sɪx	0.212	0.024	0	0	0	0	0	0	0	0	0	14	17.83	21.96
muchzer	mʊxt͜səʀ	0.319	0.033	0	0	0	0	0	0	0	0	0	17	23.40	24.82
nafnich	nafnɪx	0.214	0.017	0	0	0	0	0	0	0	0	0	4	4.00	4.43
neisspich	naispɪx	0.175	0.018	0	0	0	0	0	0	0	0	0	6	7.10	8.01
nemschen	nɛmʃən	0.368	0.038	0	0	0	0	0	0	0	0	0	63	83.97	111.60
nipziss	nɪpt͜sɪs	0.199	0.006	0	0	0	0	0	0	0	0	0	2	2.00	2.00
noendich	nœndɪx	0.169	0.022	0	0	0	0	0	0	0	0	0	14	15.26	17.48
nungper	nʊŋpəʀ	0.295	0.031	0	0	0	0	0	0	0	0	0	9	15.14	16.46
pangjin	paŋjɪn	0.217	0.011	0	0	0	0	0	0	0	0	0	1	1.00	1.00
pilwek	pɪlvək	0.305	0.008	0	0	0	0	0	0	0	0	0	6	8.43	10.86
piptol	pɪptɔl	0.270	0.007	0	0	0	0	0	0	0	0	0	1	1.00	1.00
pisstur	pɪstʊʀ	0.316	0.026	0	0	0	0	0	0	0	0	0	9	11.81	12.77
poelduss	pœldʊs	0.184	0.005	0	0	0	0	0	0	0	0	0	2	4.35	4.35
poessgun	pœsgʊn	0.184	0.003	0	0	0	0	0	0	0	0	0	0	0.00	0.00
pongtuk	pɔŋtʊk	0.223	0.005	0	0	0	0	0	0	0	0	0	0	0.00	0.00
ponssol	pɔnsɔl	0.230	0.014	0	0	0	0	0	0	0	0	0	2	2.00	2.00
purkel	pʊʀkəl	0.423	0.039	0	0	0	0	0	0	0	0	0	26	31.79	38.81
rerpfer	ʀɛʀp͜fəʀ	0.484	0.083	0	0	0	0	0	0	0	0	0	34	46.32	61.39
rimbir	ʀɪmbɪʀ	0.258	0.008	0	0	0	0	0	0	0	0	0	3	3.22	3.26
roelpem	ʀœlpəm	0.275	0.009	0	0	0	0	0	0	0	0	0	1	1.22	1.30
schaktoss	ʃaktɔs	0.361	0.015	0	0	0	0	0	0	0	0	0	10	12.11	20.14
schengschir	ʃɛŋʃɪʀ	0.332	0.006	0	0	0	0	0	0	0	0	0	0	0.00	0.00
schilsek	ʃɪlzək	0.361	0.012	0	0	0	0	0	0	0	0	0	8	10.14	15.72
schochfel	ʃɔxfəl	0.358	0.019	0	0	0	0	0	0	0	0	0	9	12.16	12.74
schornel	ʃɔʀnəl	0.488	0.031	0	0	0	0	0	0	0	0	0	22	24.94	35.51
schossfek	ʃɔsfək	0.339	0.009	0	0	0	0	0	0	0	0	0	9	9.92	15.28

Table A.4 German nonword stimuli (continued)

spell	IPA	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
schuchsser	ʃʊxsəʀ	0.383	0.033	0	0	0	0	0	0	0	0	0	20	22.75	29.23
schurtel	ʃʊʀtəl	0.551	0.071	0	0	0	0	0	0	0	0	0	90	98.21	115.07
sechtuk	zɛxtʊk	0.301	0.020	0	0	0	0	0	0	0	0	0	2	2.00	2.00
seumlim	zɔymlɪm	0.177	0.011	0	0	0	0	0	0	0	0	0	5	5.82	6.99
sirnim	zɪʀnɪm	0.321	0.014	0	0	0	0	0	0	0	0	0	3	3.00	3.13
sisskess	zɪskəs	0.332	0.032	0	0	0	0	0	0	0	0	0	32	34.67	55.99
soekfol	zœkfɔl	0.175	0.004	0	0	0	0	0	0	0	0	0	1	2.36	2.55
solbek	zɔlbək	0.308	0.021	0	0	0	0	0	0	0	0	0	24	47.93	63.41
solgon	zɔlgɔn	0.244	0.012	0	0	0	0	0	0	0	0	0	10	24.04	27.82
sulkel	zʊlkəl	0.347	0.026	0	0	0	0	0	0	0	0	0	14	16.54	17.92
sumbon	zʊmbɔn	0.205	0.006	0	0	0	0	0	0	0	0	0	9	9.90	12.81
tarpim	taʀpɪm	0.331	0.018	0	0	0	0	0	0	0	0	0	2	2.00	2.00
tekpen	tɛkpən	0.394	0.044	0	0	0	0	0	0	0	0	0	100	110.44	140.43
tekper	tɛkpəʀ	0.402	0.037	0	0	0	0	0	0	0	0	0	27	29.59	30.05
tengtuss	tɛŋtʊs	0.312	0.009	0	0	0	0	0	0	0	0	0	2	2.00	3.05
tenkess	tɛnkəs	0.404	0.032	0	0	0	0	0	0	0	0	0	38	43.49	57.02
tenmel	tɛnməl	0.397	0.026	0	0	0	0	0	0	0	0	0	34	41.00	46.48
tertel	tɛʀtəl	0.563	0.115	0	0	0	0	0	0	0	0	0	56	68.49	82.50
tichkik	tɪxkɪk	0.214	0.007	0	0	0	0	0	0	0	0	0	1	3.12	3.13
tinfun	tɪnfʊn	0.259	0.014	0	0	0	0	0	0	0	0	0	1	1.00	1.34
tirter	tɪʀtəʀ	0.541	0.077	0	0	0	0	0	0	0	0	0	95	127.31	161.82
tisskir	tɪskɪʀ	0.281	0.010	0	0	0	0	0	0	0	0	0	0	0.00	0.00
toerkuss	tœʀkʊs	0.291	0.012	0	0	0	0	0	0	0	0	0	6	7.07	7.09
tontem	tɔntəm	0.387	0.057	0	0	0	0	0	0	0	0	0	41	53.92	76.87
torpfer	tɔʀp͜fəʀ	0.429	0.037	0	0	0	0	0	0	0	0	0	31	42.28	48.44
tuktel	tʊktəl	0.397	0.056	0	0	0	0	0	0	0	0	0	62	67.93	83.58
tulbun	tʊlbʊn	0.243	0.007	0	0	0	0	0	0	0	0	0	4	4.07	4.07
tulker	tʊlkəʀ	0.375	0.038	0	0	0	0	0	0	0	0	0	41	51.41	54.28
tulnok	tʊlnɔk	0.208	0.004	0	0	0	0	0	0	0	0	0	0	0.00	0.00
tupnam	tʊpnam	0.177	0.007	0	0	0	0	0	0	0	0	0	0	0.00	0.00
walpuk	valpʊk	0.240	0.020	0	0	0	0	0	0	0	0	0	3	3.00	3.00
wasskel	vaskəl	0.369	0.034	0	0	0	0	0	0	0	0	0	32	37.73	41.08
wekmek	vɛkmək	0.353	0.019	0	0	0	0	0	0	0	0	0	18	22.24	32.86
woemniss	vœmnɪs	0.209	0.004	0	0	0	0	0	0	0	0	0	4	5.38	5.92
woktel	vɔktəl	0.411	0.056	0	0	0	0	0	0	0	0	0	53	68.95	95.58
wompur	vɔmpʊʀ	0.236	0.011	0	0	0	0	0	0	0	0	0	3	3.00	3.54
wukpek	vʊkpək	0.292	0.008	0	0	0	0	0	0	0	0	0	3	3.00	8.39
wuntel	vʊntəl	0.425	0.065	0	0	0	0	0	0	0	0	0	65	96.39	128.01
wurper	vʊʀpəʀ	0.452	0.043	0	0	0	0	0	0	0	0	0	38	54.67	73.95
zessker	t͜sɛskəʀ	0.410	0.044	0	0	0	0	0	0	0	0	0	41	50.43	63.46
zeuchken	t͜sɔyxkən	0.307	0.040	0	0	0	0	0	0	0	0	0	44	52.88	67.81
zilnich	t͜sɪlnɪx	0.240	0.019	0	0	0	0	0	0	0	0	0	17	20.35	21.54
zingker	t͜sɪŋkɛʀ	0.243	0.017	0	0	0	0	0	0	0	0	0	14	15.84	22.91
zirdess	t͜sɪʀdəs	0.422	0.036	0	0	0	0	0	0	0	0	0	51	56.42	89.44
zoechmen	t͜sœxmən	0.289	0.037	0	0	0	0	0	0	0	0	0	18	23.72	32.82
zomner	t͜sɔmnəʀ	0.347	0.034	0	0	0	0	0	0	0	0	0	10	15.88	21.80
zungdim	t͜sʊŋdɪm	0.155	0.005	0	0	0	0	0	0	0	0	0	0	0.00	0.00

A.4 German Words

Table A.5 German word stimuli

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
Baender	bɛndəʀ	bi	1.00	2.18	0.442	0.054	14	20.06	30.03	6	7.97	15.61	8	12.09	14.43	143	212.20	296.25
bestem	bɛstəm	bi	1.30	4.12	0.443	0.060	6	12.31	21.70	4	9.86	16.62	2	2.45	5.08	118	143.03	242.08
bestes	bɛstəs	bi	1.00	3.09	0.474	0.073	16	24.65	44.13	6	13.23	21.81	10	11.43	22.32	162	194.10	315.65
Bilder	bɪldəʀ	bi	2.65	3.47	0.397	0.046	10	18.58	28.72	6	12.22	19.95	4	6.36	8.78	83	118.12	164.89
Bildes	bɪldəs	bi	1.70	3.47	0.367	0.036	11	17.35	30.03	6	11.97	19.66	5	5.39	10.38	68	87.77	144.32
Birnen	bɪʀnən	bi	1.60	1.60	0.488	0.046	6	7.17	9.17	2	2.13	3.01	4	5.05	6.16	136	174.93	227.12
buntem	bʊntəm	bi	1.30	2.61	0.403	0.053	4	7.12	10.45	4	7.12	10.45	0	0.00	0.00	41	55.64	82.17
derber	dɛʀbəʀ	bi	1.00	1.90	0.490	0.096	9	10.09	14.73	5	5.50	9.47	4	4.59	5.26	81	107.26	142.83
derbes	dɛʀbəs	bi	1.00	1.90	0.459	0.085	6	6.63	11.09	5	5.63	9.10	1	1.00	1.99	79	89.15	129.02
dichter	dɪxtəʀ	bi	1.70	2.76	0.410	0.074	13	19.97	32.77	10	14.50	25.25	3	5.47	7.52	105	142.63	209.65
dichtes	dɪxtəs	bi	1.00	2.76	0.379	0.064	14	21.05	32.09	10	16.53	25.31	4	4.52	6.78	102	121.91	200.79
Dirnen	dɪʀnən	bi	1.00	1.30	0.452	0.047	5	5.58	7.42	1	1.00	1.22	4	4.58	6.20	104	138.63	179.44
Dornen	dɔʀnən	bi	1.00	1.48	0.443	0.051	7	7.00	9.59	5	5.00	6.45	2	2.00	3.13	121	172.77	219.27
dumpfer	dʊmp͜fəʀ	bi	1.00	2.00	0.296	0.037	8	9.62	15.21	6	6.83	12.40	2	2.79	2.81	47	55.48	79.73
dumpfes	dʊmp͜fəs	bi	1.00	2.00	0.265	0.026	8	8.72	15.80	5	5.72	10.42	3	3.00	5.38	45	50.79	77.54
Feinden	faindən	bi	1.78	2.70	0.398	0.058	8	14.82	23.23	3	5.37	7.95	5	9.45	15.29	150	214.80	301.14
Feindes	faindəs	bi	1.00	2.70	0.375	0.041	4	6.50	10.22	3	5.50	7.95	1	1.00	2.27	78	110.78	188.64
Feldes	fɛldəs	bi	1.78	3.00	0.439	0.075	4	7.01	11.05	3	5.17	7.75	1	1.84	3.30	103	131.07	216.15
Felsen	fɛlzən	bi	2.00	2.08	0.456	0.083	11	14.80	22.12	7	9.68	14.51	4	5.12	7.61	184	245.14	338.74
festes	fɛstəs	bi	1.70	3.54	0.500	0.111	18	24.07	45.12	8	13.70	23.41	10	10.37	21.71	197	238.38	394.15
feuchtem	fɔyxtəm	bi	1.00	2.04	0.374	0.046	6	7.30	11.14	6	7.30	11.14	0	0.00	0.00	61	68.31	110.95
feuchter	fɔyxtəʀ	bi	1.00	2.04	0.436	0.070	9	10.60	15.58	7	8.30	13.18	2	2.30	2.40	94	125.23	172.91
ganzem	gant͜səm	bi	1.70	4.12	0.333	0.018	4	11.31	16.46	4	11.31	16.46	0	0.00	0.00	25	32.18	45.97
ganzes	gant͜səs	bi	2.43	4.12	0.364	0.032	6	12.55	21.19	4	10.55	16.46	2	2.00	4.73	53	67.09	98.99
Gastes	gastəs	bi	1.48	3.18	0.427	0.073	5	6.30	10.43	1	1.00	3.17	4	5.30	7.26	192	228.07	368.42
gelben	gɛlbən	bi	2.04	2.53	0.418	0.050	9	11.96	20.12	7	9.96	18.12	2	2.00	2.00	148	197.41	273.18
Geldes	gɛldəs	bi	1.85	3.31	0.393	0.036	6	8.04	15.81	4	5.30	10.13	2	2.74	5.68	68	87.28	145.43
halbes	halbəs	bi	2.30	3.24	0.366	0.040	11	15.92	25.82	8	12.92	21.79	3	3.00	4.03	92	111.14	164.74
hartem	haʀtəm	bi	1.48	3.03	0.501	0.063	9	13.36	18.52	9	13.36	18.52	0	0.00	0.00	142	173.30	245.90
hartes	haʀtəs	bi	1.60	3.03	0.532	0.077	18	22.24	31.31	13	17.24	23.00	5	5.00	8.31	222	262.78	397.84
Heften	hɛftən	bi	1.00	2.38	0.469	0.078	32	37.41	54.21	11	11.70	14.84	21	25.70	39.37	286	376.24	520.15
Heftes	hɛftəs	bi	1.00	2.38	0.445	0.061	13	13.64	18.39	10	10.64	14.21	3	3.00	4.18	124	149.29	244.64
Hundes	hʊndəs	bi	1.48	2.79	0.329	0.040	12	16.11	26.97	2	3.86	5.57	10	12.25	21.40	89	119.39	176.55

Table A.5 German word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
Kampfes	kamp͜fəs	bi	2.58	3.53	0.326	0.025	7	7.73	14.16	5	5.73	10.97	2	2.00	3.19	64	77.03	111.71
Kinder	kɪndəʀ	bi	3.33	3.62	0.387	0.057	9	15.18	19.88	3	6.61	10.86	6	8.57	9.02	103	143.06	188.54
Kirchen	kɪʀxən	bi	2.63	3.33	0.460	0.049	4	6.60	6.72	2	4.60	4.72	2	2.00	2.00	138	165.64	207.22
Kisten	kɪstən	bi	2.00	2.34	0.442	0.090	23	32.99	40.46	2	3.90	4.18	21	29.09	36.28	348	432.16	581.14
Kursen	kʊʀzən	bi	1.60	2.94	0.440	0.052	6	11.22	14.45	6	11.22	14.45	0	0.00	0.00	106	133.18	168.84
kurzem	kʊʀt͜səm	bi	2.59	3.58	0.393	0.023	5	11.08	15.95	5	11.08	15.95	0	0.00	0.00	29	42.11	52.76
kurzes	kʊʀt͜səs	bi	1.70	3.58	0.424	0.036	7	14.26	19.26	5	12.26	17.26	2	2.00	2.00	63	79.22	114.20
Laender	lɛndəʀ	bi	3.34	3.93	0.413	0.055	14	17.95	22.30	6	8.31	9.77	8	9.64	12.53	146	211.95	289.03
leichtes	laixtəs	bi	1.90	3.48	0.370	0.064	16	21.79	31.56	12	17.27	24.49	4	4.52	7.07	145	178.33	276.73
letzten	lɛt͜stən	bi	3.38	3.54	0.451	0.072	21	27.89	39.67	6	11.40	17.33	15	16.49	22.33	284	352.10	493.74
letztes	lɛt͜stəs	bi	1.85	3.54	0.427	0.055	5	11.95	16.33	5	11.95	16.33	0	0.00	0.00	105	119.09	188.56
Leuchten	lɔyxtən	bi	1.48	1.48	0.373	0.077	20	25.34	36.00	6	7.56	11.18	14	17.78	24.82	244	307.78	453.06
mancher	manxəʀ	bi	2.45	3.25	0.384	0.045	5	10.69	13.97	4	9.69	12.97	1	1.00	1.00	55	81.84	107.14
Menschen	mɛnʃən	bi	3.91	4.03	0.408	0.046	4	4.80	5.63	2	2.80	2.89	2	2.00	2.74	131	177.56	237.64
milden	mɪldən	bi	1.48	2.38	0.363	0.057	15	21.74	29.28	7	9.22	13.96	8	12.52	15.32	132	171.89	241.70
milder	mɪldəʀ	bi	1.30	2.38	0.370	0.050	13	19.36	28.95	8	10.31	17.40	5	9.04	11.55	90	125.18	174.80
mildes	mɪldəs	bi	1.00	2.38	0.339	0.040	10	13.26	23.98	5	7.30	12.67	5	5.96	11.31	75	95.77	149.54
Moenches	mœnxəs	bi	1.00	2.40	0.294	0.020	4	7.07	10.43	3	4.65	7.19	1	2.43	3.24	19	34.14	45.77
nacktem	naktəm	bi	1.00	2.42	0.378	0.048	4	5.91	9.64	4	5.91	9.64	0	0.00	0.00	83	98.50	151.84
nackter	naktəʀ	bi	1.00	2.42	0.440	0.072	6	8.18	12.23	5	6.96	11.00	1	1.22	1.22	115	139.66	191.49
nacktes	naktəs	bi	1.00	2.42	0.409	0.062	7	9.55	16.51	4	5.84	9.64	3	3.71	6.88	134	151.74	227.35
Perlen	pɛʀlən	bi	1.48	1.78	0.508	0.096	12	12.49	14.56	6	6.49	6.95	6	6.00	7.60	157	199.77	257.72
rechtes	ʀɛxtəs	bi	1.48	3.24	0.438	0.064	16	23.80	41.83	13	20.11	34.74	3	3.68	7.09	156	200.31	342.86
rundem	ʀʊndəm	bi	1.00	3.40	0.295	0.028	9	14.12	25.40	8	13.12	23.66	1	1.00	1.74	65	85.55	127.42
Runden	ʀʊndən	bi	1.70	2.49	0.349	0.058	22	28.48	47.14	9	12.17	23.86	13	16.31	23.28	179	247.07	338.87
runder	ʀʊndəʀ	bi	1.00	3.40	0.357	0.051	15	24.23	39.47	10	16.02	29.09	5	8.21	10.38	106	147.29	208.83
rundes	ʀʊndəs	bi	1.48	3.40	0.326	0.041	21	28.23	52.51	10	15.05	29.23	11	13.18	23.28	84	109.33	162.35
Sarges	zaʀgəs	bi	1.00	2.20	0.442	0.046	3	3.00	6.04	1	1.00	2.20	2	2.00	3.84	63	74.02	117.68
scharfes	ʃaʀfəs	bi	1.30	3.03	0.497	0.038	7	10.71	15.11	5	8.71	13.11	2	2.00	2.00	112	124.77	188.81
Silben	zɪlbən	bi	1.48	1.78	0.364	0.051	7	8.91	10.08	4	5.91	7.08	3	3.00	3.00	92	130.00	159.28
solcher	zɔlxəʀ	bi	2.94	3.82	0.360	0.039	4	11.52	15.28	4	11.52	15.28	0	0.00	0.00	41	64.11	81.12
Sorgen	zɔʀgən	bi	2.72	3.04	0.438	0.062	10	17.20	24.35	5	8.32	14.18	5	8.87	10.16	112	155.80	201.03
Sorten	zɔʀtən	bi	2.11	2.34	0.521	0.092	21	32.61	38.79	4	9.05	11.25	17	23.56	27.54	272	349.59	469.32
Taktes	taktəs	bi	1.00	2.08	0.424	0.063	11	12.12	17.35	6	6.34	8.13	5	5.78	9.22	155	174.07	250.36
Tanzes	tant͜səs	bi	1.00	2.45	0.361	0.034	8	10.55	18.77	7	8.12	14.66	1	2.43	4.12	89	111.31	155.22
Toechter	tœxtəʀ	bi	2.00	2.99	0.383	0.069	2	4.37	5.96	1	1.45	2.98	1	2.92	2.98	46	70.39	96.09

Table A.5 German word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
Tulpen	tʊlpən	bi	1.00	1.00	0.355	0.042	3	3.00	3.07	2	2.00	2.07	1	1.00	1.00	82	99.75	115.91
Volkes	fɔlkəs	bi	3.20	3.76	0.401	0.037	3	5.60	10.97	3	5.60	10.97	0	0.00	0.00	80	110.63	166.32
Waelder	vɛldəʀ	bi	1.95	2.91	0.428	0.055	10	15.54	20.69	6	8.74	10.71	4	6.80	9.98	125	186.56	245.45
Worten	vɔʀtən	bi	2.96	3.43	0.532	0.090	25	41.20	57.50	3	8.96	12.38	22	32.25	45.12	331	445.23	624.64
Wortes	vɔʀtəs	bi	1.78	3.43	0.509	0.073	11	16.02	25.72	2	5.92	7.09	9	10.10	18.63	138	175.24	277.36
Zelten	ɑŋɑɱɛ?	bi	1.00	2.00	0.495	0.088	25	32.29	48.28	6	7.49	9.38	19	24.80	38.90	327	427.67	598.61
Balken	balkən	mono	1.70	1.70	0.425	0.053	21	23.84	27.56	7	7.45	8.72	14	16.39	18.84	234	275.78	346.94
Diktum	dɪktʊm	mono	1.00	1.00	0.252	0.014	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	9	12.58	17.10
Diskus	dɪskʊs	mono	1.30	1.30	0.213	0.012	3	3.37	3.70	2	2.00	2.34	1	1.37	1.37	12	12.89	16.26
Doktor	dɔktɔʀ	mono	2.93	2.94	0.302	0.012	1	1.00	2.93	1	1.00	2.93	0	0.00	0.00	15	19.24	19.97
dunkel	dʊŋkəl	mono	2.36	2.91	0.285	0.033	2	2.00	3.11	1	1.00	2.11	1	1.00	1.00	66	83.97	110.70
Faktum	faktʊm	mono	1.70	2.08	0.333	0.017	1	1.00	2.09	1	1.00	2.09	0	0.00	0.00	15	18.47	22.82
Ferkel	fɛʀkəl	mono	1.30	1.48	0.544	0.120	3	3.00	3.80	3	3.00	3.80	0	0.00	0.00	67	91.41	112.26
Fiskus	fɪskʊs	mono	1.30	1.30	0.276	0.011	2	2.34	2.70	1	1.00	1.37	1	1.34	1.34	11	12.29	14.42
Folter	fɔltəʀ	mono	1.30	1.48	0.500	0.082	8	10.80	15.68	6	7.77	12.53	2	3.04	3.15	144	211.44	317.17
Funken	fʊŋkən	mono	1.30	1.30	0.364	0.046	16	16.13	19.74	9	9.13	12.56	7	7.00	7.18	122	156.16	197.25
Galgen	galgən	mono	1.60	1.60	0.390	0.056	8	9.69	12.85	4	5.62	7.90	4	4.07	4.95	130	172.12	218.33
Gondel	gɔndəl	mono	1.00	1.30	0.343	0.042	1	1.00	1.30	1	1.00	1.30	0	0.00	0.00	19	24.56	30.47
Gulden	gʊldən	mono	1.70	1.70	0.355	0.049	6	8.46	12.35	1	1.00	1.68	5	7.46	10.66	88	116.53	164.35
Gurgel	gʊʀgəl	mono	1.00	1.00	0.421	0.040	4	4.00	4.00	4	4.00	4.00	0	0.00	0.00	25	28.17	30.93
Handel	handəl	mono	2.98	3.07	0.370	0.048	7	14.27	17.67	5	11.05	14.10	2	3.22	3.58	72	99.00	143.52
Henkel	hɛŋkəl	mono	1.00	1.30	0.363	0.036	9	10.87	12.10	6	6.57	7.41	3	4.31	4.69	80	103.09	147.90
hinter	hɪntəʀ	mono	2.98	3.02	0.461	0.088	6	9.48	13.49	4	5.85	9.81	2	3.63	3.68	91	123.17	158.39
Junker	jʊŋkəʀ	mono	1.78	1.85	0.285	0.038	6	8.63	11.38	4	5.67	8.32	2	2.96	3.06	61	80.59	105.86
Kaktus	kaktʊs	mono	1.00	1.30	0.328	0.018	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	22	23.93	32.92
Kapsel	kapsəl	mono	1.85	1.85	0.350	0.028	4	4.00	4.85	4	4.00	4.85	0	0.00	0.00	39	44.52	48.19
Karpfen	kaʀp͜fən	mono	1.60	1.60	0.463	0.049	7	9.07	10.26	7	9.07	10.26	0	0.00	0.00	156	196.28	252.96
Kasten	kastən	mono	1.70	1.95	0.460	0.095	41	55.55	74.91	6	6.88	10.02	35	48.67	64.89	436	546.73	741.15
Kirmes	kɪʀmɛs	mono	1.00	1.00	0.350	0.016	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	8	8.00	10.21
Kolben	kɔlbən	mono	1.30	1.30	0.373	0.060	3	3.00	3.34	1	1.00	1.34	2	2.00	2.00	126	171.82	213.53
Korken	kɔʀkən	mono	1.00	1.00	0.470	0.067	10	10.00	10.91	9	9.00	9.91	1	1.00	1.00	167	222.92	274.25
Korpus	kɔʀpʊs	mono	1.00	1.00	0.325	0.025	2	2.00	3.33	2	2.00	3.33	0	0.00	0.00	12	14.10	18.39
Kultus	kʊltʊs	mono	1.00	1.00	0.317	0.018	1	1.00	1.26	1	1.00	1.26	0	0.00	0.00	15	17.66	22.55
Kumpel	kʊmpəl	mono	2.20	2.32	0.321	0.024	4	5.22	7.09	2	2.85	4.64	2	2.37	2.45	46	53.25	58.67
Kursus	kʊʀzʊs	mono	1.30	2.34	0.307	0.018	1	1.90	2.94	1	1.90	2.94	0	0.00	0.00	14	19.91	23.74
Laster	lastəʀ	mono	1.30	1.30	0.448	0.087	15	17.28	23.80	8	10.28	16.23	7	7.00	7.58	237	299.10	454.19

Table A.5 German word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
Lektor	lɛktɔʀ	mono	1.48	1.60	0.364	0.018	4	6.57	7.40	1	1.00	1.62	3	5.57	5.78	17	25.85	31.73
Lumpen	lʊmpən	mono	1.00	1.00	0.318	0.042	10	11.18	13.24	6	6.13	6.91	4	5.05	6.34	85	94.28	111.52
Mantel	mantəl	mono	2.30	2.42	0.448	0.075	8	8.56	11.17	5	5.00	6.75	3	3.56	4.42	117	162.41	219.97
Mentor	mɛntɔʀ	mono	1.00	1.30	0.387	0.026	1	1.00	1.18	1	1.00	1.18	0	0.00	0.00	21	30.11	32.76
minder	mɪndəʀ	mono	2.11	2.30	0.368	0.062	11	16.43	21.92	3	3.82	6.23	8	12.61	15.70	131	181.25	249.57
Moertel	mœʀtəl	mono	1.00	1.00	0.472	0.061	5	5.00	5.00	5	5.00	5.00	0	0.00	0.00	37	50.03	60.13
Morgen	mɔʀgən	mono	2.86	2.89	0.437	0.060	10	15.65	19.58	7	9.34	12.54	3	6.31	7.04	132	168.92	217.46
munter	mʊntəʀ	mono	1.85	2.04	0.438	0.078	7	15.53	18.55	2	2.30	4.11	5	13.22	14.44	95	144.45	194.00
Muskel	mʊskəl	mono	1.00	1.90	0.323	0.025	2	2.84	3.81	2	2.84	3.81	0	0.00	0.00	30	39.51	47.89
Nimbus	nɪmbʊs	mono	1.30	1.30	0.189	0.006	2	2.00	2.22	1	1.00	1.22	1	1.00	1.00	5	5.22	6.42
Pendel	pɛndəl	mono	1.30	1.30	0.385	0.043	6	6.00	8.31	5	5.00	7.31	1	1.00	1.00	98	131.34	175.53
Phosphor	fɔsfɔʀ	mono	1.30	1.30	0.283	0.014	1	1.00	1.22	1	1.00	1.22	0	0.00	0.00	1	1.00	1.22
Pinsel	pɪnzəl	mono	1.48	1.48	0.332	0.028	7	8.69	10.22	6	6.00	7.38	1	2.69	2.85	42	45.26	50.42
Pulver	pʊlfəʀ	mono	1.78	1.85	0.353	0.035	6	6.00	7.71	5	5.00	6.71	1	1.00	1.00	30	34.70	37.82
Purpur	pʊʀpʊʀ	mono	1.00	1.00	0.326	0.016	1	1.00	1.00	1	1.00	1.00	0	0.00	0.00	0	0.00	0.00
Schalter	ʃaltəʀ	mono	1.78	1.95	0.527	0.081	14	21.72	29.76	8	9.62	14.65	6	12.10	15.10	191	242.51	329.18
Schenkel	ʃɛŋkəl	mono	1.30	1.60	0.409	0.028	12	14.79	20.42	9	10.75	16.00	3	4.04	4.42	86	107.43	155.60
Schinken	ʃɪŋkən	mono	1.48	1.48	0.379	0.049	18	23.95	32.42	1	1.00	1.48	17	22.95	30.95	194	228.97	333.39
Schulter	ʃʊltəʀ	mono	2.54	2.83	0.493	0.073	6	8.29	8.75	4	5.50	5.82	2	2.79	2.93	112	146.19	181.20
Sektor	zɛktɔʀ	mono	2.42	2.54	0.367	0.021	4	6.29	7.40	1	1.74	2.54	3	4.55	4.86	15	21.62	27.75
selten	zɛltən	mono	2.83	2.95	0.495	0.089	22	29.70	43.71	1	1.84	2.95	21	27.87	40.76	381	488.36	693.42
Silber	zɪlbəʀ	mono	2.42	2.42	0.372	0.044	5	8.36	11.24	4	5.00	7.88	1	3.36	3.36	37	50.73	61.55
simpel	zɪmpəl	mono	1.30	1.70	0.320	0.025	7	7.48	7.52	6	6.00	6.00	1	1.48	1.52	37	42.64	49.74
Taktik	taktık	mono	2.08	2.08	0.314	0.021	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	12	14.63	18.28
Technik	tɛxnık	mono	3.13	3.13	0.260	0.009	1	2.43	3.30	1	2.43	3.30	0	0.00	0.00	4	6.97	7.92
Tempus	tɛmpʊs	mono	1.00	1.00	0.253	0.010	1	1.00	2.58	1	1.00	2.58	0	0.00	0.00	13	16.38	19.52
Thermik	tɛʀmık	mono	1.00	1.00	0.364	0.065	1	1.00	1.40	1	1.00	1.40	0	0.00	0.00	14	20.35	22.44
Tochter	tɔxtəʀ	mono	2.92	2.99	0.415	0.070	1	2.00	2.98	0	0.00	0.00	1	2.00	2.98	73	94.03	129.34
Toelpel	tœlpəl	mono	1.30	1.30	0.314	0.020	6	6.00	6.60	6	6.00	6.60	0	0.00	0.00	12	13.24	13.45
Turnus	tʊʀnʊs	mono	1.30	1.30	0.316	0.017	1	1.00	1.18	1	1.00	1.18	0	0.00	0.00	22	26.39	36.68
Verbum	vɛʀbʊm	mono	1.00	1.30	0.339	0.071	2	3.36	3.58	2	3.36	3.58	0	0.00	0.00	17	24.13	30.64
Vesper	fɛspəʀ	mono	1.00	1.00	0.450	0.076	6	7.59	9.66	6	7.59	9.66	0	0.00	0.00	47	65.85	91.59
Wandel	vandəl	mono	2.23	2.23	0.375	0.049	10	13.46	18.97	6	7.48	12.35	4	5.97	6.63	103	144.25	210.88
Wechsel	vɛksəl	mono	2.42	2.48	0.393	0.037	6	8.22	13.79	5	7.22	12.79	1	1.00	1.00	78	90.70	133.00
Wimper	vɪmpəʀ	mono	1.00	1.60	0.355	0.040	3	3.93	4.07	2	2.93	3.07	1	1.00	1.00	47	57.96	64.47
Windel	vɪndəl	mono	1.00	1.30	0.356	0.049	14	17.89	23.50	11	13.66	19.26	3	4.22	4.24	110	153.27	196.55
Winkel	vɪŋkəl	mono	1.90	2.26	0.321	0.034	13	14.66	22.09	12	13.66	21.09	1	1.00	1.00	83	102.11	137.62

Table A.5 German word stimuli (continued)

spell	IPA	morph	logFreq	lemmaLogFreq	posfreq	biphone freq	Neighbors	Freq Neighbors	Lemma Freq Neighbors	syl1 Neighbors	syl1 Freq Neighbors	syl1 lemma Freq Neighbors	syl2 Neighbors	syl2 Freq Neighbors	syl2 lemma Freq Neighbors	edit2 neighbors	edit2 Freq Neighbors	edit2 lemma Freq Neighbors
Winter	vɪntəʀ	mono	2.63	2.68	0.465	0.088	9	12.65	15.39	7	8.19	10.92	2	4.46	4.47	124	168.01	217.06
Wirbel	vɪʀbəl	mono	1.90	1.90	0.441	0.039	7	7.07	11.26	6	6.07	10.26	1	1.00	1.00	65	74.14	99.85
Wirrnis	vɪʀnɪs	mono	1.00	1.00	0.363	0.018	1	1.00	1.13	1	1.00	1.13	0	0.00	0.00	18	20.68	24.76
Witwer	vɪtvəʀ	mono	1.48	1.48	0.387	0.039	4	5.28	7.52	4	5.28	7.52	0	0.00	0.00	28	40.43	50.46
Wunder	vʊndəʀ	mono	2.64	2.66	0.365	0.051	14	18.06	29.41	12	15.93	25.01	2	2.13	4.40	117	181.39	259.25
Wurzel	vʊʀt͜səl	mono	1.90	2.18	0.425	0.033	4	5.42	7.04	4	5.42	7.04	0	0.00	0.00	51	65.41	83.48
Zirkel	t͜sɪʀkəl	mono	2.00	2.11	0.445	0.033	6	6.34	8.25	6	6.34	8.25	0	0.00	0.00	42	49.21	58.65
Zirkus	t͜sɪʀkʊs	mono	1.90	1.90	0.329	0.016	2	2.00	2.91	2	2.00	2.91	0	0.00	0.00	12	16.19	21.68

Table A.6 Distribution of Phonemes for German stimuli

phon	C1	V1	C2	C3	V2	C4	C1	V1	C2	C3	V2	C4	C1	V1	C2	C3	V2	C4
	monomorphemes						bimorphemes						nonwords					
b	1			5			7			5			4			8		
d	4			8			8			18			16			21		
g	4			3			5			2			8			6		
p	4		1	9			1			1			9		4	24		
t	7		1	21			4			28			19		1	18		
k	11		8	15		3	7		4	1			15		14	24		15
t͜s	2			1			1		2	5			8		1	5		
p͜f				1						3						4		
f	7			2			8		2	1			8		4	8		
v	12			1			3						10			1		
z	4			2			5			2			9			5		
s			7	2		12			5			30			11	2		31
ʃ	4						1			1			9			2		
x			2						8	4					15			8
h	3						6						4					
l	3		12			23	5		16	1			9		28	2		27
ʀ			18			22	5		18			15	4		28			39
j	1												1			3		
m	7		6	2		3	6		3			8	11		10	3		15
n	1		13	3		12	3		17	3		22	6		19	14		15
ŋ			7												15			
ɪ		20			4			13						35			24	
ɛ		14			1			20						25			3	
a		13						15						18			3	
ʊ		17			14			12						30			24	
ɔ		9			5			7						24			12	
œ		2						2						11				
ai								3						2				
ɔy								3						5				
ə					51						75						84	

Appendix B

Instructions for the experiments

This appendix includes the instructions given to participants in each of the four experiments. Instructions for Experiments One and Three, conducted at the University of Michigan, were given in English, while instructions for Experiments Two and Four, conducted at the University of Konstanz, Germany, were given in German. The texts (including formatting such as bold and italics) are reproduced here exactly, including any grammatical or typographical errors present in the instructions that were in the original. The instructions for Experiments Two and Four were translated from the English as closely as possible by the primary investigator, who has near-native fluency in German, and were checked by two native speakers of German.

B.1 Experiment 1

Instructions for Assessing Context Effects in English and German
Investigator: Robert Felty

Your task is to hear a set of 2-syllable English words and pseudo-words over headphones and transcribe them as best as you can using standard English spelling. Noise has been mixed in to make the task a little more difficult. A set of guidelines for standard English orthography is on the other side of this sheet.

There is a practice followed by the actual experiment. Click on the **Begin** button to start the practice, and you will hear 1 block each of 10 words or pseudo-words. For each trial, enter the word or pseudo-word you hear into the textbox using the keyboard. You may correct your response using the backspace key, but once you press <enter>, the computer will proceed to the next trial. You will only get one opportunity to hear each trial. In other words, the computer will not be able to play the word over again if you don't hear it properly. Guess as best as you can. Try to answer as accurately as possible. There is no time limit.

After the two practice blocks of 10 trials each, the actual experiment will begin. The experiment is divided into 20 blocks of 15 trials each. Each block contains stimuli that are familiar words, or contains English pseudo- words. **The computer screen will tell you if the block is a Word block or a Pseudo-word block.**

Click on the **Begin** button to start the experiment when you are ready. At the beginning of each block (including the first one), make sure the cursor is in the textbox before beginning to type. As in the practice, type your response into the textbox and the computer will proceed to the next trial.

When you are finished, please exit quietly as other participants may still be performing the experiment.

When transcribing words, please enter them exactly as they appear in a dictionary, even if the word contains silent letters or other exceptional spelling.

Here are some examples of standard English orthography for writing out the nonsense words:

"ee" as in	b**ee**t	"ch" as in	**ch**eck
"i" as in	b**i**d	"sh" as in	**sh**ine
"i_e" as in	s**ide**	"j" as in	**j**ar
"ay" as in	s**ay**, pl**ay**	"g" as in	**geek** or **goon** (not gel)
"e" as in	b**e**t	"z" as in	"ha**z**e"
"a" as in	j**a**zz, h**a**t	"ss" as in	hi**ss** (not his)
"ah" as in	f**a**ther, b**ah** humbug	"zz" as in	fi**zz**
"oo" as in	b**oo**t	"s" as in	**s**ap
"u_e" as in	fl**u**k**e** (not puke)	"c" as in	ri**c**e (use this only for words rhyming with "ice")
"oa" as in	**oa**t		
"u" as in	**hu**t		
"oi" as in	c**oi**n		
"ow" as in	br**ow**n		

Use double consonants (as in hiss and jazz) if you feel that they make your transcription clearer. You may also use silent "e"s to identify the long vowels, as in side, or fluke.

Avoid using "g" to identify the "j" sound as in jar. Also avoid using "c" to identify the "s" sound, unless you are transcribing an item that rhymes with ice as noted above.

B.2 Experiment 2

Anweisungen zum Forschungsprojekt *Die Interaktion von Lexical Access, Phonetik, und Morphologie*

Forscher: Robert Felty

Ihre Aufgabe ist, eine Reihe von zweisilbigen deutschen Wörten und Pseudowörtern über Kopfhörer anzuhören, und sie in der hochdeutschen Schreibweise in einen Computer einzugeben. Es gibt starke Hintergrundgeräusche mit den Wörtern vermischt, um die Aufgabe schwieriger zu machen. Eine kurze Wiederholung von hochdeutscher Schreibweise ist auf dem zweiten Blatt zu finden.

Es gibt eine kurze Übungsrunde vor dem echten Experiment. Klicken Sie auf Begin, um die Übungsrunde anzufangen: Sie werden einen Block mit jeweils 10 Wörtern oder Pseudowörtern hören. Für jeden Versuch tippen Sie das Wort oder Pseudowort, das Sie hören. Sie können Ihre Antwort mit der Delete Taste ändern, aber sobald Sie <enter> drücken, wird der Computer zum nächsten Probe weitergehen. Sie haben nur eine Möglichkeit, einen Versuch zu hören, d.h. Sie haben keine Gelegenheit, das Wort wieder zu hören. Falls sie einen Versuch verpasst haben, können Sie einfach <enter> drücken, ohne eine Antwort einzugeben. Raten Sie, so gut Sie können. Es gibt keine Zeitbegrenzung.

Nach der Übungsrunde fängt das Experiment an. Das Experiment ist in 20 Blöcke mit jeweils15 Proben aufgeteilt. Jeder Block enthält entweder echte deutsche Wörter oder Pseudowörter. Auf dem Bildschirm können Sie sehen, ob der Block ECHTE WÖRTER oder PSEUDOWÖRTER enthält.

Klicken Sie auf Begin, wenn Sie bereit sind. Versichern Sie sich am Anfang jedes Blocks (einschließlich des ersten Blocks), dass der Cursor im Textfeld ist, bevor Sie anfangen zu tippen. Tippen Sie, wie in der Übungsrunde, Ihre Antwort ins Textfeld, und der Computer wird zum nächsten Versuch weitergehen.

Bitte verlassen Sie das Zimmer leise, wenn Sie fertig sind, falls andere Teilnehmer noch an den Proben arbeiten.

Wenn Sie Wörter buchstabieren, bitte tragen Sie die ein, genau wie sie in einem Wörterbuch stehen (ausgesehen von ß und Umlaute, leider kann der Program diese Zeichen nicht verstehen - bitten benutzen sie "ss" für ß und beziehungsweise ae, ue, oe für ä, ü, ö).

Hier sind einige Beispiele von hochdeutschen Schreibweise, die Sie benutzen können, um die Pseudowörter zu buchstabieren.

"ie" wie in	t**ie**f	"ch" wie in	der Sti**ch**
"i" wie in	r**i**chtig	"sch" wie in	die **Sch**ule
"ei" wie in	m**ei**n	"j" wie in	**j**a
"e" wie in	rote b**e**te	"w" wie in	die **W**unde
"e" wie in	r**e**cht	"z" wie in	die **Z**eit
"a" wie in	h**a**t	"ss" wie in	da**ss**
"ue" wie in	J**ue**rgen	"pf" wie in	der **Pf**arrer
"oe" wie in	der K**oe**nig	"s" wie in	**s**ein
"u" wie in	der H**u**t		
"u" wie in	m**u**ss		
"oo" wie in	das B**oo**t		
"o" wie in	der K**o**ch		
"eu" wie in	n**eu**n		
"au" wie in	br**au**n		

B.3 Experiment 3

Instructions for Assessing Context Effects in English and German

Investigator: Robert Felty

Your task is to hear a set of 2-syllable German words and pseudo-words over headphones and transcribe them as best as you can using standard German spelling. Noise has been mixed in to make the task a little more difficult. A set of guidelines for standard German orthography is on page two.

There is a practice followed by the actual experiment. Click on the Begin button to start the practice, and you will hear 1 block each of 10 words or pseudo-words. For each trial, enter the word or pseudo-word you hear into the textbox using the keyboard. You may correct your response using the backspace key, but once you press <enter>, the computer will proceed to the next trial. You will only get one opportunity to hear each trial. In other words, the computer will not be able to play the word over again if you don't hear it properly. Guess as best as you can. Try to answer as accurately as possible. There is no time limit.

After the two practice blocks of 10 trials each, the actual experiment will begin. The experiment is divided into 20 blocks of 15 trials each. Each block contains stimuli that are familiar words, or contains German pseudo- words. **The computer screen will tell you if the block is a Word block or a Pseudo-word block.**

Click on the Begin button to start the experiment when you are ready. At the beginning of each block (including the first one), make sure the cursor is in the textbox before beginning to type. As in the practice, type your response into the textbox and the computer will proceed to the next trial.

When you are finished, please exit quietly as other participants may still be performing the experiment.

When transcribing words, please enter them exactly as they appear in a dictionary, even if the word contains silent letters or other exceptional spelling.

Here are some examples of standard German orthography for writing out the pseudo-words:

"ie" as in	t**ie**f	"ch" as in	der Sti**ch**
"i" as in	r**i**chtig	"sch" as in	die **Sch**ule
"ei" as in	m**ei**n	"j" as in	**j**a
"e" as in	rote b**e**te	"w" as in	die **W**unde
"e" as in	r**e**cht	"z" as in	die **Z**eit
"a" as in	h**a**t	"ss" as in	da**ss**
"ue" as in	J**ue**rgen	"pf" as in	der **Pf**arrer
"oe" as in	der K**oe**nig	"s" as in	**s**ein
"u" as in	der H**u**t		
"u" as in	m**u**ss		
"oo" as in	das B**oo**t		
"o" as in	der K**o**ch		
"eu" as in	n**eu**n		
"au" as in	br**au**n		

Please be careful not to confuse "ei" and "ie". Please also be careful with the letters "s", "z", and "ss".

B.4 Experiment 4

Anweisungen zum Forschungsprojekt *Die Interaktion von Lexical Access, Phonetik, und Morphologie*

Forscher: Robert Felty

Ihre Aufgabe ist, eine Reihe von zweisilbigen englischen Wörten und Pseudowörtern über Kopfhörer anzuhören, und sie in der englischen Schreibweise in einen Computer einzugeben. Es gibt starke Hintergrundgeräusche mit den Wörtern vermischt, um die Aufgabe schwieriger zu machen. Eine kurze Wiederholung von englischer Schreibweise ist auf dem zweiten Blatt zu finden.

Es gibt eine kurze Übungsrunde vor dem echten Experiment. Klicken Sie auf **Begin**, um die Übungsrunde anzufangen: Sie werden einen Block mit jeweils 10 Wörtern oder Pseudowörtern hören. Für jeden Versuch tippen Sie das Wort oder Pseudowort, das Sie hören. Sie können Ihre Antwort mit der **Delete** Taste ändern, aber sobald Sie <enter> drücken, wird der Computer zum nächsten Probe weitergehen. Sie haben nur eine Möglichkeit, einen Versuch zu hören, d.h. Sie haben keine Gelegenheit, das Wort wieder zu hören. Falls sie einen Versuch verpasst haben, können Sie einfach <enter> drüücken, o ohne eine Antwort einzugeben. Raten Sie, so gut Sie können. Es gibt keine Zeitbegrenzung.

Nach der Übungsrunde fängt das Experiment an. Das Experiment ist in 20 Blöcke mit jeweils 15 Proben aufgeteilt. Jeder Block enthält entweder echte englische Wörter oder Pseudowörter. Auf dem Bildschirm können Sie sehen, ob der Block ECHTE WÖRTER oder PSEUDOWÖÖRTER enthält.

Klicken Sie auf **Begin**, wenn Sie bereit sind. Versichern Sie sich am Anfang jedes Blocks (einschließlich des ersten Blocks), dass der Cursor im Textfeld ist, bevor Sie anfangen zu tippen. Tippen Sie, wie in der Übungsrunde, Ihre Antwort ins Textfeld, und der Computer wird zum nächsten Versuch weitergehen.

Bitte verlassen Sie das Zimmer leise, wenn Sie fertig sind, falls andere Teilnehmer noch an den Proben arbeiten.

Wenn Sie Wörter buchstabieren, bitte tragen Sie die ein, genau wie sie in einem Wörterbuch stehen.

Hier sind einige Beispiele von englischen Schreibweise, die Sie benutzen können, um die Pseudowörter zu buchstabieren.

"ee" wie in	b**ee**t	"ch" wie in	**ch**eck
"i" wie in	b**i**d	"sh" wie in	**sh**ine
"i_e" wie in	s**ide**	"j" wie in	**j**ar
"ay" wie in	s**ay**, pl**ay**	"g" wie in	**geek** or **goon** (not gel)
"e" wie in	b**e**t	"z" wie in	"ha**z**e"
"a" wie in	j**a**zz, h**a**t	"ss" wie in	hi**ss** (not his)
"ah" wie in	f**a**ther, b**ah** humbug	"zz" wie in	fi**zz**
"oo" wie in	b**oo**t	"s" wie in	**s**ap
"u_e" wie in	fl**u**ke (not puke)	"c" wie in	ri**c**e (use this only for words rhyming with "ice")
"oa" wie in	**oa**t		
"u" wie in	**hu**t		
"oi" wie in	c**oi**n		
"ow" wie in	br**ow**n		

Appendix C

Confusion Matrices

This appendix lists the confusion matrices from all four experiments. Separate confusion matrices are shown for each signal-to-noise ratio (S/N), for each position (C1 V1 C2 C3 V2 C4), and for each block (nonword vs. word), for a total of 96 confusion matrices (4 experiments x 2 S/Ns x 6 positions x 2 blocks). In each confusion matrix, the stimulus phonemes are listed in the rows, and the responses are listed in the columns. Each entry in the matrices represents the percentage of responses to a given stimulus phoneme. Entries with zero percentage are left blank. The total number of presentations for each stimulus phoneme is listed in the rightmost column of each matrix. Discarded trials have been subtracted from these totals. Rows sum to 100% (although rounding errors may distort this in some cases), but columns do not. Because the experiments were open response, the confusion matrices are not square. Correct responses are typeset in bold face.

In order to make the confusion matrices more meaningful, responses that were not listed in the stimuli were given a separate column in the matrix if they accrued more than 5% of responses for any matrix in each language. This was also done for responses which contained clusters. For example, matrices for Experiments One and Four, which used English stimuli, include a column for /sp/ in the C1 position because this response received 5% of the total responses in the S/N=5 dB word condition in Experiment 4 (see Table C.91).

C.1 Experiment 1 — English native listeners

C.1.1 Nonwords

Table C.1 Experiment 1 — C1 nonwords S/N = -5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**62**	12			2				1	1				8						2					13		126
d	3	**84**	3	1		2				3							1			1					3		182
g	1	29	**63**			1	1										3								2		98
ʤ		7	17	**67**	2	1	1	1	1												2				2		126
p			1		**63**	16	7		8																7		182
t					7	**65**	11		10		1														5	1	112
k					7	14	**75**	1	1																3		224
ʧ						14	37	**40**	1				3												1	3	70
h			1		6	15	34		**39**	1				1			1	1							2		140
f	1	1			2	1	1		3	**77**	4	6		1											5		196
s						12				3		**83**													1	1	98
ʃ					2	2	2	24				5	**60**												2	2	42
v	2	6	1								7			**68**	3			3	2		1				5	1	98
w																**75**		21	2							2	56
j																	**100**										14
l																11		**68**	7	14							28
ɹ														1		2		2	**83**					1	7	2	84
m																				**93**	4				3		70
n														1				1		22	**75**				2		154

mean p_p= 70
min p(h)= 39
max p(j)= 100

Table C.2 Experiment 1 — V1 nonwords S/N = -5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**79**	7		7									7		14
ɪ	5	**79**		9		1		1				3	3		392
eɪ		9	**64**	11	9			4					2	2	56
ɛ		8		**85**	2			2					3		518
æ				27	**62**		4	1			3		3		350
oʊ		1	1			**34**	25	23	1	1	3		4	7	182
ɑ					16	2	**46**	25			2		1	8	252
ə		1		8		12	10	**64**			1		3	2	196
ɔɪ										**64**	14	7	14		14
aʊ				20	17	2	2	6			**46**		5	1	84
aɪ		12		14	12			2			5	**52**		2	42

mean p_p= 62
min p(oʊ)= 34
max p(ɛ)= 85

Table C.3 Experiment 1 — C2 nonwords S/N = -5 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**73**	17	3	4																3		70
g		21	**57**	7						7						7						14
p	3			**80**	7	2	2													5	1	126
t	2	5		33	**38**	5	2													12	2	42
k				8	5	**78**								2						4	4	126
f	3			29	1	6	**29**	6						5		1				16	4	154
s				1	1	3	10	**66**		1				4						6	9	238
z	2	5						36			**50**									7		42
l														**93**			1			4	1	588
m														1		**64**	28	2	2	1	2	322
n																17	**76**	2		4	1	364
ŋ																	29	**71**				14
																						mean p_p= 65
																						min p(f)= 29
																						max p(l)= 93

Table C.4 Experiment 1 — C3 nonwords S/N = -5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**89**	4			4									4															28
d	5	**80**			2	6						1		4	1												3	1	196
g		29	**46**				21										4												28
p					**54**	35	2			3																	4	1	252
t		2			1	**87**				1		4															4		238
k						23	**70**	1		1																	2	2	210
ʧ							7	**79**											7								7		14
f		1			5	9				**66**	4	5											1				6	2	140
s						3				1	1	**88**			3												2		406
v	18	12							3	3	1	1		**50**	1	1		1	2	2							1	3	98
w																**87**			10	1							2		84
j																	**64**	10		2		2					10	12	42
l																5			**86**	3	1						4	1	112
ɹ		1							1							2				**88**							1	7	112
m																1					**87**	11					1		98
n																					29	**69**						2	42
																													mean p_p= 74
																													min p(g)= 46
																													max p(b)= 89

Table C.5 Experiment 1 — V2 nonwords S/N = -5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**77**		2					18			1	1428
ɛ		25		**63**	2				11				56
æ	7			14	**79**								14
ə		23		1	1			1	**70**	1		3	602
													mean p_p= 72
													min p(ɛ)= 63
													max p(æ)= 79

Table C.6 Experiment 1 — C4 nonwords S/N = -5 dB

	d	g	ʤ	t	k	ʧ	f	s	ʃ	v	z	l	m	n	ŋ	nd	rd	null	other	Total
d	**93**			2										2				4		308
ʤ		1	**90**															8		84
t	17			**77**	3			2										2		294
k				4	**94**													2		308
s								**97**			2							1		266
ʃ									**93**									7		14
v	7			4				7		**39**	25							14	4	28
z	1		2					37			**54**							6		224
m													**73**	25				2		238
n	1												6	**85**	1	5		3		182
ŋ	1	1											3	7	**82**			5		154
																				mean p_p= 80
																				min p(v)= 39
																				max p(s)= 97

Table C.7 Experiment 1 — C1 nonwords S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**75**	2			2	1					1			9											10		144
d	1	**93**	2																						2		208
g		15	**79**				1										3								3		112
ʤ		5	3	**83**	1			1	1						1		4				1				1		144
p	1				**85**	8	1		3																1		208
t					2	**80**	6	1	4																6		128
k					4	13	**76**	3																	3		256
ʧ				3		6	18	**69**																		5	80
h		1			10	14	26	1	**45**	1									1						2		160
f					1				1	**91**	1	2													2		224
s						10				3		**83**													4		112
ʃ			2					23				2	**63**								4				4	2	48
v	7	2								2	2			**74**	4			2	1		3				4		112
w																**91**		2							5	3	64
j																	**100**										16
l																13		**81**	3	3							32
ɹ	1																		**92**						1	6	96
m																				**95**	3				3		80
n																				17	**82**				1		176
																											mean p_p= 81
																											min p(h)= 45
																											max p(j)= 100

Table C.8 Experiment 1 — V1 nonwords S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**75**	13		6									6		16
ɪ	4	**88**		4				1				1	1		448
eɪ		3	**66**	8	20			2					2		64
ɛ		7		**83**	4			2					3		592
æ			1	22	**64**		3	1			7		3		400
oʊ						**38**	33	12		1	4		6	6	208
ɑ				1	25	4	**44**	14			4		1	7	288
ə				4	1	8	14	**67**			2		1	2	224
ɔɪ						13				**56**	6		25		16
aʊ				7	17	2	1				**63**		5	5	96
aɪ		17		8	17						10	**46**	2		48

mean p_p= 63
min p(oʊ)= 38
max p(ɪ)= 88

Table C.9 Experiment 1 — C2 nonwords S/N = 0 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**75**	14		5	1					1						1					3	80
g		6	**88**																	6		16
p	3			**88**	4		1							1					1	1	2	144
t				35	**56**	2				2					4							48
k			1	3	1	**90**														3	2	144
f	2			21	1	3	**44**	2	1	1	1			3						16	5	176
s							6	**80**			1			1	1					3	7	272
z								21		4	**63**			6						2	4	48
l														**89**						8	1	672
m				1												**70**	23	2	2	2	1	368
n																12	**84**	1		2		416
ŋ																	19	**81**				16

mean p_p= 76
min p(f)= 44
max p(k)= 90

Table C.10 Experiment 1 — C3 nonwords S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**88**		3		3									6															32
d	4	**80**				6						1		7													2		224
g		6	**72**	3			16																				3		32
p	1				**73**	16	4			3																	3		288
t					2	**93**	1					1															3	1	272
k					1	11	**81**	2																			3	2	240
ʧ								**75**																			19	6	16
f					3	7	1			**81**	3	1		2													3	1	160
s						4				1		**90**			1												3		464
v	18	5				1			4	3	1	4		**61**	3	1													112
w														1		**91**	1		1								6		96
j																	**60**	21			2						4	13	48
l															1	8			**87**								4	1	128
ɹ			1											1		1				**91**							2	5	128
m							1														**94**	4					2		112
n																					29	**69**					2		48
																													mean p_p= 80
																													min p(j)= 60
																													max p(m)= 94

Table C.11 Experiment 1 — V2 nonwords S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**77**		3					17			1	1632
ɛ		22		**63**	2			2	13				64
æ			13	13	**69**				6				16
ə		22		2	1			2	**70**			2	688
													mean p_p= 70
													min p(ɛ)= 63
													max p(ɪ)= 77

Table C.12 Experiment 1 — C4 nonwords S/N = 0 dB

	d	g	ʤ	t	k	ʧ	h	f	s	ʃ	v	z	l	m	n	ŋ	sp	nd	rd	null	other	Total
d	**91**			4											1					3		352
ʤ	2	3	**89**		1															5		96
t	17			**76**	3															3		336
k	1			1	**95**															2	1	352
s									**95**			3								1		304
ʃ									6	**94**												16
v	3				3				6		**50**	9		9	3					16		32
z	2								24			**68**								5	2	256
m														**84**	14					2		272
n														6	**89**			1		3		208
ŋ														2	8	**84**		1		5	1	176
																						mean p_p= 83
																						min p(v)= 50
																						max p(k)= 95

C.1.2 Words

Table C.13 Experiment 1 — C1 words S/N = -5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɪ	bl	null	other	Total
b	**85**	4		1	1	1			3	2		1		1					1						1	1	168
d	3	**79**							1					7						1	1				8		98
g	5	7	**77**				7																		2	2	56
ʤ		2	5	**91**		2																					56
p	1				**69**	14	6		7																1		252
t					15	**76**	2	1	2	2											1					1	168
k					3	10	**81**	1	1																3	1	210
ʧ							21	**79**																			14
h	1				7	4	14		**71**	1						1									1		140
f						1				**95**	2	3															112
s										1		**96**	3														140
ʃ							5		2				**93**														42
v									2					**91**				4		2					2		56
w														11		**84**			1	1						1	70
j	4		4													4	**79**			11							28
l																2		**90**	3					4		2	154
ɹ																1		6	**91**				1			1	154
m																		4		**89**	5				2		140
n																	2				**88**					10	42

mean p_p= 84
min p(p)= 69
max p(s)= 96

Table C.14 Experiment 1 — V1 words S/N = -5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**89**	7						1				1		1	70
ɪ	1	**90**	2	3		1							1		350
eɪ	6	1	**90**		3										70
ɛ		1		**93**	2	1	1	1					2		392
æ			2	3	**89**		1						4		434
oʊ		2	6		1	**84**	3	1	1		1		1		140
ɑ						2	**92**	1			1		3	1	238
ə				4			3	**93**					1		196
ɚ									**100**						14
ɔɪ			4			4	4			**89**					28
aʊ			1								**97**		2		98
aɪ	2	2										**96**			56
other															

mean p_p= NaN
min p(oʊ)= 84
max p(ɚ)= 100

Table C.15 Experiment 1 — C2 words S/N = -5 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**98**																			2		56
p				**94**		4														3		112
t				5	**90**	2														2		42
k				4		**90**	1													4	1	238
f							**95**	2												2		84
s							6	**88**												3	2	504
z			4								**89**				4						4	28
l														**93**			1			4	1	336
m														1		**95**	4					84
n																	**96**	1		1	1	574
ŋ																	5	**93**		2		42

mean p_p= 93
min p(s)= 88
max p(b)= 98

Table C.16 Experiment 1 — C3 words S/N = -5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**93**	7																											14
d	1	**95**	1			1								1													2		336
g			**100**																										14
p					**83**	13																					4		112
t						**92**	1					1															5		896
k		1	1		1	6	**86**				1	1					1			2							1	1	154
f						2				**98**																			42
s						5						**91**			2												2		252
ʃ													**89**														11		28
v		1												**94**					1								3		70
w						11										**75**			4		4						7		28
j																	**64**										29	7	14
l																			**98**								2		56
ɹ						4														**96**									28
m					2																**88**	7					2		42
n																						**100**							14

mean p_p= 90
min p(j)= 64
max p(g)= 100

Table C.17 Experiment 1 — V2 words S/N = -5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**93**							4	1		1	1764
ə		4		1		1			**92**		1	2	322
ɚ										**100**			14

mean p_p= 95
min p(ə)= 92
max p(ɚ)= 100

Table C.18 Experiment 1 — C4 words S/N = -5 dB

	d	g	ʤ	t	k	ʧ	h	f	s	ʃ	v	z	l	m	n	ŋ	sp	nd	rd	null	other	Total
d	**96**																			2		644
ʤ	1		**98**																	1		98
t	9		1	**84**					1					1						3	1	154
k				1	**95**															2		210
s									**96**			1								2		224
ʃ										**100**												28
v	14										**71**				7					7		14
z	5		1	1					4		2	**75**		1	1	3				6	2	168
m														**100**								126
n	2				1				1					3	**87**	1				5	1	126
ŋ														2	1	**89**				7		308

mean p_p= 90
min p(v)= 71
max p(ʃ)= 100

Table C.19 Experiment 1 — C1 words S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**92**	2			4	1			1	1															1		192
d	4	**82**												13											1		112
g	2	2	**97**																								64
ʤ		2		**95**																			2		2		64
p	1			1	**84**	9	2		2																		288
t					17	**81**			1	1																	192
k					2	3	**93**		1																1		240
ʧ							13	**88**																			16
h					13	4	10		**73**																		160
f										**100**																	128
s						1						**99**													1		160
ʃ						2							**98**														48
v														**100**													64
w														3		**89**		5	4								80
j									3							3	**91**			3							32
l																		**93**	1					5		2	176
ɹ																		3	**97**				1				176
m																		2		**95**	3						160
n																					**94**					6	48

mean p_p= 92
min p(h)= 73
max p(f)= 100

Table C.20 Experiment 1 — V1 words S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**98**	1												1	80
ɪ		**96**		2		1									400
eɪ			**99**		1										80
ɛ				**96**	2	1									448
æ			2	1	**94**								3		496
oʊ						**97**	2				1		1		160
ɑ					1	2	**92**	1			1		2		272
ə				3		1	2	**93**							224
ɚ									**100**						16
ɔɪ										**100**					32
aʊ				1							**99**				112
aɪ		9										**91**			64
other															

mean p_p= NaN
min p(aɪ)= 91
max p(ɚ)= 100

Table C.21 Experiment 1 — C2 words S/N = 0 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**98**																			2		64
p	1			**96**		1														1	2	128
t				6	**94**																	48
k				1		**94**		1												3		272
f							**98**													2		96
s							4	**94**												2		576
z			3					3			**88**									6		32
l								1						**96**						3		384
m														1		**98**	1					96
n																	**99**					656
ŋ																	2	**96**		2		48

mean p_p= 96
min p(z)= 88
max p(n)= 99

Table C.22 Experiment 1 — C3 words S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**94**	6																											16
d		**96**				1								2															384
g			**100**																										16
p		1			**88**	11																							128
t					1	**97**																					2		1024
k			1			1	**98**																						176
f										**100**																			48
s						2						**96**															1		288
ʃ												6	**75**														19		32
v		1												**99**															80
w																**100**													32
j																13	**81**										6		16
l																			**98**								2		64
ɹ																				**100**									32
m																					**90**	8					2		48
n																						**100**							16
																													mean p_p= 95
																													min p(ʃ)= 75
																													max p(g)= 100

Table C.23 Experiment 1 — V2 words S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɝ	null	other	Total
ɪ		**95**							3				2016
ə		5		1					**92**			1	368
ɝ										**100**			16
													mean p_p= 96
													min p(ə)= 92
													max p(ɝ)= 100

Table C.24 Experiment 1 — C4 words S/N = 0 dB

	d	g	ʤ	t	k	ʧ	h	f	s	ʃ	v	z	l	m	n	ŋ	sp	nd	rd	null	other	Total
d	**99**																			1		736
ʤ	1		**99**																			112
t	14			**86**	1																	176
k				1	**98**															2		240
s									**95**			2								2	1	256
ʃ										**97**											3	32
v											**94**									6		16
z	2								4		2	**85**			1	3				4		192
m														**99**	1					1		144
n														3	**96**	1						144
ŋ																**95**				4		352
																						mean p_p= 95
																						min p(z)= 85
																						max p(ʤ)= 99

C.2 Experiment 2 — German native listeners

C.2.1 Nonwords

Table C.25 Experiment 2 — C1 nonwords S/N = -5 dB

	b	d	g	p	t	k	h	f	p͜f	s	t͜s	ʃ	ʧ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**52**	16		5	2		2				2			23														64
d	7	**59**	5	1	2									9		10	2			2						1		256
g	5	20	**23**	2	2	25		2						10		9	1		1	2								128
p	1	1		**58**	8	3	3	17	1		1			2	1					1			1			1	1	144
t		1		6	**88**	2		1			1									1								304
k				5	7	**73**	8													2		1					2	240
h	5		2	8	3		**42**							25		11							2		2		2	64
f				9		2		**77**	2					3	2						5						2	128
t͜s					10	1		1			**77**				9					2						1	1	128
ʃ				1	4	5	1	3			1	**66**	10		1					1							7	144
v	3	2					1	3						**84**	1	1	1	1	3					1		1	1	160
z	1	1			1	1					2			24	**54**	4	3		1	4						1	3	144
j		6												50		**25**	19											16
l	6		1	1			1	1						11	2	6	**50**		15	7				1			1	144
ʀ	3	3	13		2	11	3	2						14			2	**8**		5		3	6		6	14	6	64
m	1													1		3	3		**78**	11						2	1	176
n																1	13		56	**27**						3		96

mean p_p= 55
min p(ʀ)= 8
max p(t)= 88

Table C.26 Experiment 2 — V1 nonwords S/N = 2 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	əl	null	other	Total
ɪ	7	**64**		5	1	4	7	6	1	3					1	2	560
ʊ		5	1	4	2	**73**	2	1	1	7					1	3	480
ɛ		4				2	1	**79**	4	4	2	1			1	3	400
œ		3		5		2	1	43	**22**	19		3			1	2	176
ɔ								2	1	**85**	3	6	1		1	3	384
a								1	1	5	**87**	1	4			1	288
ɔɪ				1				9	9	16	24	**33**	4	1		4	80
aɪ											28		**72**				32

mean p_p= 64
min p(œ)= 22
max p(a)= 87

Table C.27 Experiment 2 — C2 nonwords S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p	2			**39**	11	31		2		6			2									6	2	64
t	6	6		6	**75**					6														16
k			2			**87**					4		1									4	2	224
f			2			3	**20**	2	5		58		8									3		64
s							1		**90**		1		1					2					6	176
t͜s										**100**														16
x						1					**91**		5										1	240
l												**69**	4	8	3							13	2	448
ʀ												1	**69**									27		448
m												3	4	**52**	24	7						7	3	160
n												2	1	22	**60**	6						6	3	303
ŋ												4	2	34	19	**28**	4					4	4	241

mean p_p= 65
min p(f)= 20
max p(t͜s)= 100

Table C.28 Experiment 2 — C3 nonwords S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**51**	5	9	1			1						16		1	1	2	3		4	7	128
d	6	**49**	4	1	14		1						2		3	5		1		11	3	336
g	15	5	**41**			16			1				2			2				9	9	96
p	1			**54**	31	3	3	2					1							1	4	384
t		1		7	**85**	1															7	288
k			1	7	8	**80**														1	2	384
f				2	5		**66**	12	1	1	2	2	5								6	128
p͜f				9	9		27	**45**		2			2							2	5	64
s		6					6		**28**				13	41			3			3		32
t͜s					9				1	**71**				10				1		1	6	80
ʃ								3			**84**	9									3	32
v	19	6					13						**50**							13		16
z	1	6	3				1		8	1			10	**64**	1					3	3	80
j	2	4				2						2	2	2	**10**					69	6	48
l		6											9			**38**	16	22		6	3	32
m		4		2												4	**65**	21		4		48
n		1			1										1	7	23	**51**		10	4	224

mean p_p= 55
min p(j)= 10
max p(t)= 85

Table C.29 Experiment 2 — V2 nonwords S/N = 2 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ	3	**59**	2			1	31	5	384
ʊ	1	2	**70**		6	1	18	2	384
ɛ		6		**17**		2	71	4	48
ɔ			8	4	**70**	2	8	8	192
a				10	19	**58**	10	2	48
ə		4	1	2		5	**84**	3	1344

mean p_p= 60
min p(ɛ)= 17
max p(ə)= 84

Table C.30 Experiment 2 — C4 nonwords S/N = 2 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k	3	4	**86**		5						1		240
s	1			**97**							1	1	496
x			4		**91**						1	4	128
l	1					**72**	6				19		432
ʀ	2					1	**78**				12	7	624
m	3					3	2	**21**	60	5	3	3	240
n	1					2	2	27	**57**	3	3	5	240

mean p_p= 72
min p(m)= 21
max p(s)= 97

Table C.31 Experiment 2 — C1 nonwords S/N = 7 dB

	b	d	g	p	t	k	h	f	p͜f	s	t͜s	ʃ	ʧ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**64**	3	3	5										22					2	2								64
d	3	**68**	1				2							9		13	1			1						1		256
g	3	13	**41**	1		17		2						10		12				1							1	128
p	1			**81**	2	2	1	8	1					2						1								144
t		1		2	**95**	2																						304
k			1	3	2	**89**	4													1								240
h	2			14	3	5	**52**							8		14		2									2	64
f				4		2		**87**	2					2							3					1		128
t͜s					3						**90**				6												1	128
ʃ					1							**88**	8										1				2	144
v	2			1				3						**91**	1		1		1	1					1			160
z								1			3			11	**76**	2	1		1	1					1	1	3	144
j								13						38		**38**	6									6		16
l	3	1												22		3	**47**		9	11				1		1	1	144
ʀ	6	2	11			5		2						17				**27**		8		6	2		9	5	2	64
m	1					1								1		5	1		**82**	9						1	1	176
n														1			8		49	**42**								96

mean p_p= 68
min p(ʀ)= 27
max p(t)= 95

Table C.32 Experiment 2 — V1 nonwords S/N = 7 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	11	**71**	1	1		2	9	3		2							560
ʊ		1		2	1	**90**	1			2						3	480
ɛ	1	2						**93**	2	1						2	400
œ		2		4		1	1	27	**49**	12	1		1			2	176
ɔ						1		1		**91**		7				1	384
a										5	**92**	1	1				288
ɔɪ						1		6	3	4	18	**59**	6	1		3	80
aɪ													**100**				32

mean p_p= 81
min p(œ)= 49
max p(aɪ)= 100

Table C.33 Experiment 2 — C2 nonwords S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p	5			**56**	8	19																13		64
t		13		6	**75**	6																		16
k			2			**90**					4											1	1	224
f							**53**				42		3							2				64
s									**97**				1					1				1	2	176
t͜s										**100**														16
x							1				**89**		8										2	240
l												**81**	2	6	4							7	1	448
ʀ													**75**									22	2	448
m													5	**57**	27	6						3	3	160
n												1		14	**72**	8	1					2	2	297
ŋ														26	13	**46**	10					1	2	247

mean p_p= 74
min p(ŋ)= 46
max p(t͜s)= 100

Table C.34 Experiment 2 — C3 nonwords S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**71**	1	3	1			2						9		1		5	1		3	5	128
d	2	**72**	1		14								1		1	3				4	3	336
g	19	2	**48**			16	1						1							5	8	96
p				**80**	13	1	1	1													3	384
t				2	**89**																9	288
k			1	4	4	**90**								1							1	384
f		1		1	1		**73**	16				2	2								5	128
p͜f						2	14	**70**													14	64
s		3					9			3				84								32
t͜s					4					**86**		1		5							4	80
ʃ											**84**	6									9	32
v	25	6					13						**56**									16
z	1								3	1			10	**84**			1					80
j		13				2								13	**17**					44	13	48
l		22											13			**31**	16	13		3	3	32
m		10											2				**77**	6		4		48
n		5														5	23	**57**		4	6	224

mean p_p= NaN
min p(j)= 17
max p(k)= 90

Table C.35 Experiment 2 — V2 nonwords S/N = 7 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ	1	**53**	3	1			39	4	384
ʊ	1	1	**81**		8		8	1	384
ɛ		6		**8**		2	81	2	48
ɔ			7	4	**80**		7	2	192
a				6	17	**71**	6		48
ə		5		2		4	**87**	2	1344

mean p_p= 63
min p(ɛ)= 8
max p(ə)= 87

Table C.36 Experiment 2 — C4 nonwords S/N = 7 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k		2	**94**		3								240
s	1			**98**									496
x			5		**93**						1	1	128
l	1					**78**	6				13	1	432
ʀ	1						**78**				9	11	624
m							1	**31**	60	6		2	240
n	1					1		13	**79**	3		3	240

mean p_p= 79
min p(m)= 31
max p(s)= 98

C.2.2 Words

Table C.37 Experiment 2 — C1 words S/N = 2 dB

	b	d	g	p	t	k	h	f	p͡f	s	t͡s	ʃ	ʧ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**62**	2	3											16	1	2	1	1		7				1		2	3	128
d	3	**80**	3	2			1	3						2	1	2	1	1		3							1	192
g	1	1	**67**	1	10	7		3						1		1	1	1		3		2						144
p				**76**	6	1	1	1												8					1	1	4	80
t			1	2	**91**	5	1																				1	176
k				1		**95**	1	1												1						1		288
h	4	1		1	1	6	**60**	1						10			1	1	1	3						8	1	144
f				1			2	**89**						8														240
t͡s								2			**92**				6													48
ʃ					4	3	3					**74**	3							8			3			1	4	80
v			1											**87**	1	4	1			3								240
z					1			2			4			10	**77**	1	1	1									1	144
j	31	25														**38**										6		16
l		4					1	2						13			**75**	1		1				1		2	1	128
ʀ	15					3	5							15				**31**	1	5	1	4	1		14	3	3	80
m	1													1			4		**85**	7				1				208
n	2					2								2		2	3		30	**61**								64
																												mean p_p= 73
																												min p(ʀ)= 31
																												max p(k)= 95

Table C.38 Experiment 2 — V1 words S/N = 2 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	2	**77**		2		5	4	2		6						1	528
ʊ		7	2	2		**81**		1		4	1					2	464
ɛ		2					1	**90**		4	3					1	544
œ							2	9	**67**	3	17		2				64
ɔ										**97**		1				1	256
a										1	**98**					1	448
ɔɪ												**100**					48
aɪ													**100**				48
																	mean p_p= 89
																	min p(œ)= 67
																	max p(ɔɪ)= 100

Table C.39 Experiment 2 — C2 words S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p				**100**																				16
t			6		**81**	6																6		16
k			1			**95**					1		3									1		192
f						3	**75**				3	3	13									3		32
s									**98**		1		1										1	192
t͜s										**100**														32
x											**99**											1	1	160
l												**83**	4		6							6	1	448
ʀ		1										2	**84**		1							11		576
m												3	1	**70**	13	2	3					6	1	144
n												4	5	6	**80**	1						3	1	480
ŋ												3	5	9	38	**41**						3	1	112

mean p_p= 84
min p(ŋ)= 41
max p(p)= 100

Table C.40 Experiment 2 — C3 words S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**69**	9	3		1		1						1			2	1			7	7	160
d	4	**63**	1	1	1											3		2		17	6	416
g		1	**89**		3								1							4	3	80
p	1	2	3	**83**	4	3		3													2	160
t					**95**							1								2	2	784
k				3	2	**88**	1	1													3	256
f		2	8				**73**						17									48
p͜f				6	2		2	**81**				2							6		2	64
s									**81**					19								32
t͜s										**99**										1		96
ʃ											**81**	19										16
x												**100**										64
v					13								**88**									16
z	2	2	3	3	2								8	**67**						6	8	64
l						6										**88**					6	16
m						3											**75**	22				32
n		6			2												3	**64**		8	17	96

mean p_p= 81
min p(d)= 63
max p(x)= 100

Table C.41 Experiment 2 — V2 words S/N = 2 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ		**73**					23	3	64
ʊ			**85**	1	1		9	3	224
ɛ				**94**			6		16
ɔ					**99**		1		80
ə	1	1	1				**91**	5	2016
									mean p_p= 88
									min p(ɪ)= 73
									max p(ɔ)= 99

Table C.42 Experiment 2 — C4 words S/N = 2 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k			**90**		8						2		48
s	2			**96**							1	1	672
l	5					**76**			4		12	3	368
ʀ	2					1	**95**				2		592
m	3					1	2	**43**	37	11	2	1	176
n	3					2		3	**81**	1	9	1	544
													mean p_p= 80
													min p(m)= 43
													max p(s)= 96

Table C.43 Experiment 2 — C1 words S/N = 7 dB

	b	d	g	p	t	k	h	f	p͜f	s	t͜s	ʃ	ʧ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**66**	1	2				1	1						16		2	1		1	7				2				128
d	2	**89**						1	1							5				3				1			1	192
g			**87**		5	2		1						1						3		1						144
p				**96**	1	1																			1			80
t					**99**															1								176
k						**99**																						288
h	6	1		3		1	**77**	1						4				1		1		1				4		144
f							2	**98**																				240
t͜s											**98**				2													48
ʃ								1				**96**								3								80
v														**90**	1	4	1			3								240
z								2			1				**96**					1								144
j																**50**	31			6				6		6		16
l	2	5						5						7			**79**	1		2						1		128
ʀ	6		1			3								13				**39**		6		13			19		1	80
m	1													1			1		**93**	3								208
n																	2	2	16	**80**				2				64
																												mean p_p= 84
																												min p(ʀ)= 39
																												max p(t)= 99

Table C.44 Experiment 2 — V1 words S/N = 7 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	1	**85**	1	1		3	3	1		4							528
ʊ		5	1	2		**88**				3						1	464
ɛ		1						**98**		1	1						544
œ								2	**75**	5	19						64
ɔ										**100**							256
a											**97**		1			1	448
ɔɪ										2		**98**					48
aɪ													**100**				48
																	mean p_p= 93
																	min p(œ)= 75
																	max p(aɪ)= 100

Table C.45 Experiment 2 — C2 words S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p				**100**																				16
t		6	6		**81**																	6		16
k						**96**					1		3									1		192
f							**97**		3															32
s									**99**				1											192
t͜s										**100**														32
x											**98**	1	1											160
l												**90**	1		6							2		448
ʀ												1	**89**		1							10		576
m													4	**83**	8	1	1					3	1	144
n												2	2	4	**90**	1						2		480
ŋ													4	3	38	**54**						2		112

mean p_p= 90
min p(ŋ)= 54
max p(p)= 100

Table C.46 Experiment 2 — C3 words S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**84**	8															1	1		2	5	160
d	1	**82**																1		12	3	416
g	1		**96**		1																1	80
p	1	1	1	**90**	1	3		2													3	160
t					**97**																2	784
k						**96**															1	256
f		2					**85**						13									48
p͜f				3		2		**94**											2			64
s									**72**					28								32
t͜s										**99**				1								96
ʃ											**94**	6										16
x												**100**										64
v								6					**94**									16
z		2					2						3	**91**						3		64
l																**100**						16
m																	**100**					32
n		5														1	5	**74**		2	13	96

mean p_p= 91
min p(s)= 72
max p(x)= 100

Table C.47 Experiment 2 — V2 words S/N = 7 dB

	i	ɪ	ʊ	ε	ɔ	a	ə	other	Total
ɪ		**77**					19	5	64
ʊ			**93**	1			3	2	224
ε		13		**75**				13	16
ɔ					**98**		1	1	80
ə		1					**95**	3	2016
									mean p_p= 87
									min p(ε)= 75
									max p(ɔ)= 98

Table C.48 Experiment 2 — C4 words S/N = 7 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k			**100**										48
s	1			**96**								2	672
l	2					**89**			2		6		368
ʀ	1						**96**				2		592
m	5					1	1	**61**	23	10	1	1	176
n	3					1		4	**89**	1	2	1	544
													mean p_p= 89
													min p(m)= 61
													max p(k)= 100

C.3 Experiment 3 — German non-native listeners

C.3.1 Nonwords

Table C.49 Experiment 3 — C1 nonwords S/N = 2 dB

	b	d	g	p	t	k	h	f	p͜f	s	t͜s	ʃ	tʃ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**55**	10		5	2	2	2							18			2									2	3	60
d	5	**64**	4		1						1			15	1	3	3		1							1		240
g	1	14	**32**		3	25		1			1			18			3	1	1			1		1	1			120
p				**56**	10	9	1	13	1					3						1	1						4	135
t		1		4	**91**	2																	3					285
k			2	1	4	**90**	2																				1	225
h			2	5	10	3	**32**	2						25	3	2	7	3					2		3		2	60
f				3	4			**71**	2					8	1					1	8		3					120
t͜s					8			1			**63**		1		24												3	120
ʃ				1	4	6		1	1		2	**49**	24										3				9	135
v	7		1	2				1						**77**	2		4	5	1									150
z	1	1						1			23			24	**39**		4	6	1								1	135
j	7													80		**7**	7											15
l	1	1	1								1			26	1		**47**	5	12	1		1		2			1	135
ʀ	7	5	18	5		10		3			2			28				**2**			3	2				8	7	60
m														6			2		**83**	7				1	1			165
n														2			12		61	**21**				2		1		90
																												mean p_p= 52
																												min p(ʀ)= 2
																												max p(t)= 91

Table C.50 Experiment 3 — V1 nonwords S/N = 2 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	7	**73**	1	3	1	5	2	6	1	1							525
ʊ		6		13	2	**48**		4	12	11		1				2	450
ɛ	1	12				2		**73**	3	1	2	1				4	375
œ		4		10		12	1	39	**14**	15	2				1	4	165
ɔ						3		1	3	**78**	6	6				3	360
a						1		1		10	**87**					1	270
ɔɪ						4		5	7	21	17	**35**	4	1		5	75
aɪ											27		**73**				30
																	mean p_p= 60
																	min p(œ)= 14
																	max p(a)= 87

Table C.51 Experiment 3 — C2 nonwords S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p	2		2	**22**	12	8		8			15		3	3								22	3	60
t					**53**								13						27				7	15
k			1			**34**		1			58											2	2	210
f						3	**47**	2	2		35									5		2	5	60
s									**90**	1		1						2				1	5	165
t͜s										**100**														15
x						8	2				**80**	1										1	7	225
l	1											**61**	6	8	6							13	4	420
ʀ		2		1	2	3	1	1			1	6	**57**	1	1				1			18	4	420
m												7	1	**51**	33	3							5	150
n												6		18	**63**	4						2	7	285
ŋ											1	4	2	32	32	**19**	1					1	7	225

mean p_p= 56
min p(ŋ)= 19
max p(t͜s)= 100

Table C.52 Experiment 3 — C3 nonwords S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**56**	3	3	1			3						18			2	3	1		5	8	120
d	7	**63**	2		15								2		1	2				5	3	315
g	21	9	**34**			21					1	1	2	1		1					8	90
p		1		**46**	34	4	3	2						1				1		3	6	360
t				9	**83**	1		1												1	4	270
k				3	8	**76**						3				1				5	5	360
f	1			2	5		**46**	18	1	2		1	13	2						5	6	120
p͜f		2		2	3		28	**38**		2					3						22	60
s		2					11		**56**	7			9	7						4	4	45
t͜s					12				5	**55**	1			16							11	75
ʃ											**67**	20								3	10	30
v	20						7						**60**							13		15
z		3	2				2		2	10		2	8	**68**	2						2	60
j		7	2							9	2		7	2	**13**	2				44	11	45
l		10														**63**	7	7		7	7	30
m	2	9										2	2				**53**	24		7		45
n		2					1									9	29	**39**		10	10	210

mean p_p= 54
min p(j)= 13
max p(t)= 83

Table C.53 Experiment 3 — V2 nonwords S/N = 2 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ	5	**40**	3	1		1	43	6	286
ʊ		4	**61**	1	7	1	21	5	360
ɛ		27		**32**			27	14	37
ɔ			18	8	**49**	2	17	6	180
a			9	4	11	**53**	18	4	45
ə	1	13	5	3	1	3	**73**	2	1342

mean p_p= 51
min p(ɛ)= 32
max p(ə)= 73

Table C.54 Experiment 3 — C4 nonwords S/N = 2 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k		3	**52**		40						1	4	225
s				**97**	1							2	465
x	1		6		**85**						3	6	120
l						**80**	13		1		5	1	405
ʀ						3	**81**	1	1		9	6	585
m						3	1	**14**	68	2	4	6	225
n						2	2	17	**58**	6	1	13	225

mean p_p= 67
min p(m)= 14
max p(s)= 97

Table C.55 Experiment 3 — C1 nonwords S/N = 7 dB

	b	d	g	p	t	k	h	f	p͡f	s	t͡s	ʃ	t͡ʃ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**57**	3	2	7			2							20				2	2	2	3					2		60
d	5	**60**	3		2		1							22		4	1			2								240
g	3	13	**34**		1	28	2							11		3	1		3	3								120
p	1			**78**	2	4	1	5	1					1						1							6	135
t		2		2	**91**	2														1			1				1	285
k			2			**95**	1																				1	225
h	7	2		5	5	3	**32**	2			2			30		7	2	2								2	2	60
f				7	3	1	2	**57**	5					4							18		2			1	2	120
t͡s					8	1		1	1		**57**	1			26												5	120
ʃ					3	1						**70**	10		1					1			1				11	135
v	9	1	1					2						**78**			2	3	3	1								150
z		1			1			1			27			7	**57**	1	1		1	1							1	135
j	7													33		**47**										13		15
l	1	1						1						21	1	2	**56**	2	11	3								135
ʀ	8	2	15			12	3							33			2	**3**		7		2			5	5	3	60
m		1				1								2	1	1	4		**85**	6								165
n														2		3	14		40	**34**				2		2	1	90

mean p_p= 58
min p(ʀ)= 3
max p(k)= 95

Table C.56 Experiment 3 — V1 nonwords S/N = 7 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	7	**75**	1	1	1	3	3	6		2						1	525
ʊ		3	3	12	2	**51**		2	8	17		1				1	450
ɛ		6		1		2		**86**		1	2					2	375
œ	1	7		5		15		33	**15**	11	5	4	1			5	165
ɔ				1		3		1	3	**80**	6	4				4	360
a						1		3		12	**81**					1	270
ɔɪ			1			1		5	4	16	17	**41**	1	5		7	75
aɪ		7	7								10		**73**			3	30

mean p_p= 63
min p(œ)= 15
max p(ɛ)= 86

Table C.57 Experiment 3 — C2 nonwords S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p	3			**43**	10	7		3		2	7	2	3	2								17	2	60
t					**60**														13				27	15
k			2			**40**					53											2	2	210
f						3	**32**	5	2		43	2	2							2		2	8	60
s						1			**88**	1			1					1					8	165
t͜s										**80**								7					13	15
x			2			11			1		**77**	2	2									1	3	225
l												**75**	6	4	3							7	4	420
ʀ				1		3					2	6	**62**									19	5	420
m		1										7	5	**55**	23	5						1	1	150
n												7		15	**64**	4						1	9	285
ŋ	1					1						2	2	30	34	**18**	3					2	7	225

mean p_p= 58
min p(ŋ)= 18
max p(s)= 88

Table C.58 Experiment 3 — C3 nonwords S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**58**	6	2	1			3	1					10		1		3	1		6	8	120
d	4	**68**	1		16		1						1			2				4	3	315
g	19	3	**36**			22						3	6							7	4	90
p	1			**58**	23	2	5	4					1							1	5	360
t		1		4	**88**																7	270
k				3	7	**70**						6								7	7	360
f				2	1	1	**48**	23		1			14							1	10	120
p͜f	3				8		27	**43**													18	60
s					4		4		**44**	22			7	7						9	2	45
t͜s					11	1	1		1	**65**	1			12						3	4	75
ʃ											**73**	17									10	30
v		7											**93**									15
z			2						3	28			8	**48**						2	8	60
j		7			2	2	2			7			4	7	**16**			2		40	11	45
l		23														**50**				10	17	30
m		7			2							2					**62**	16		11		45
n		1														7	33	**42**		8	8	210

mean p_p= 57
min p(j)= 16
max p(v)= 93

Table C.59 Experiment 3 — V2 nonwords S/N = 7 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ	2	**41**	4	4		1	41	6	280
ʊ			**68**	1	6	1	17	7	360
ɛ		25	3	**38**			33	3	40
ɔ			16	6	**49**	6	18	5	180
a		2	4		4	**73**	16		45
ə		13	5	3	1	2	**73**	3	1345

mean p_p= 57
min p(ɛ)= 38
max p(ə)= 73

Table C.60 Experiment 3 — C4 nonwords S/N = 7 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k	1	1	**54**		39	1						4	225
s				**98**								2	465
x		1	4	1	**83**						2	10	120
l	1					**85**	9				3	1	405
ʀ	2					3	**84**	1			4	6	585
m	1					3		**14**	70	3	3	6	225
n	2					3	3	11	**68**			13	225

mean p_p= 69
min p(m)= 14
max p(s)= 98

C.3.2 Words

Table C.61 Experiment 3 — C1 words S/N = 2 dB

	b	d	g	p	t	k	h	f	p͡f	s	t͡s	ʃ	tʃ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**53**	7	4		2	1	2	3						21	1		6		1	2								120
d	2	**83**	1		2	1	1							4	2	2	1							1	1	2		180
g		4	**59**		16	12	1							5	1			1	1			1					1	135
p		1		**65**	9	15	1	3												1					1	1	1	75
t			1	1	**87**	5	1	1			1					1							4					165
k			2		1	**95**																					1	270
h	5	1		2	1	7	**61**							13			1	1	2			1	1		1	2	1	135
f				1	2	1	3	**81**	2					7							1		1					225
t͡s											**82**				18													45
ʃ					7	7						**75**	3							1			4			1	3	75
v	1	1	1											**88**	2	1	3	1	1									225
z		1			1						7			9	**78**	1	3											135
j	7	40												27		**7**	20											15
l	1	7		1			2							9	2		**74**		1	1				3			1	120
ʀ	19		1	3	5	13	8	4						31		1		**5**			1	1			1		5	75
m	3													4			4	1	**86**	3								195
n	2													2			2		47	**48**								60

mean p_p= 66
min p(ʀ)= 5
max p(k)= 95

Table C.62 Experiment 3 — V1 words S/N = 2 dB

	i	ɪ	y	ʏ	u	ʊ	e	ɛ	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	3	**79**	1	2		9	1	4		1							495
ʊ		10		7	2	**66**		3	3	5	1					2	435
ɛ		4						**88**		1	3					1	510
œ				3		2		27	**40**	7	13	3	2			3	60
ɔ				1		3			2	**88**	3	2				2	240
a						1				1	**93**		2			2	420
ɔɪ										7		**93**					45
aɪ		4											**96**				45

mean p_p= 80
min p(œ)= 40
max p(aɪ)= 96

Table C.63 Experiment 3 — C2 words S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p				**93**																		7		15
t		7			**53**			20		7	7												7	15
k			1			**71**					25	1										2	1	180
f						3	**87**				3									7				30
s									**97**									2					1	180
t͜s										**100**														30
x						1	1				**94**	2											2	150
l												**81**	1	1	7							7	1	420
ʀ			1									2	**74**	1	2							16	2	540
m												7		**63**	21	1						3	5	135
n												5	1	9	**80**	2						2		450
ŋ												2	6	10	43	**33**	2					5		105

mean p_p= 77
min p(ŋ)= 33
max p(t͜s)= 100

Table C.64 Experiment 3 — C3 words S/N = 2 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**68**	18	2	1	1								1				2			3	4	150
d	3	**74**	1	1												1	1	1		16	2	390
g	11	3	**72**		1												3			3	8	75
p	1	2	1	**59**	13	9	2	5					2							4	3	150
t		1		1	**93**							2								1	2	735
k				3	3	**84**		2				6									1	240
f				2	2		**91**													2	2	45
p͜f		2		12	2		7	**65**		3									8		2	60
s									**63**	7				27						3		30
t͜s				1					1	**97**											1	90
ʃ											**73**	27										15
x											7	**90**									3	60
v					7		7						**53**							20	13	15
z	8	10	8		3				2	10			12	**42**						5		60
l																**67**	13	7		13		15
m		7															**57**	30		3	3	30
n	1	4			1												40	**36**		11	7	90

mean p_p= 70
min p(n)= 36
max p(t͜s)= 97

Table C.65 Experiment 3 — V2 words S/N = 2 dB

	i	ɪ	ʊ	ε	ɔ	a	ə	other	Total
ɪ	2	**63**	3				25	7	60
ʊ		4	**67**		2		25	1	210
ε		13		**7**			27	53	15
ɔ			3		**84**		9	4	75
ə	1	2	2	1	1	1	**89**	3	1890

mean p_p= 62
min p(ε)= 7
max p(ə)= 89

Table C.66 Experiment 3 — C4 words S/N = 2 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k	2	9	**69**		16				2		2		45
s				**95**							2	2	630
l	1				1	**77**	2		6		12	1	345
ʀ						2	**90**				7	1	555
m		1				2	1	**22**	58	11	4	2	165
n						2	1	2	**86**	1	5	2	510

mean p_p= 73
min p(m)= 22
max p(s)= 95

Table C.67 Experiment 3 — C1 words S/N = 7 dB

	b	d	g	p	t	k	h	f	p͜f	s	t͜s	ʃ	ʧ	v	z	j	l	ʀ	m	n	fʀ	gʀ	tʀ	bl	bʀ	null	other	Total
b	**58**	3	2	1			1	1			1			26		1	3		1	1					1	1		120
d	1	**86**	1	1	1		1							2		3	1	1		1				1		1	1	180
g	1	4	**66**		12	12								3					1	1								135
p	1			**80**	4	5	1		4											3							1	75
t		1			**98**															1			1					165
k				1		**98**																					1	270
h	6	1	1	1	1	13	**61**							12			1	1	1							1		135
f				1	3	1	1	**72**	4					10	1						3						3	225
t͜s										4	**69**				27													45
ʃ					5	1	1					**84**			1					1							5	75
v		1												**89**		1	1	1	2	1								225
z						1					8			1	**86**		1	1		1							1	135
j	20	27												20		**20**	13											15
l	2	7					1							10	1		**74**			3						1	2	120
ʀ	24		4	8		4		1						27				**9**	1	3		9			3		7	75
m	3													3	1	1	4	1	**86**	2				1				195
n																		2	55	**43**								60

mean p_p= 69
min p(ʀ)= 9
max p(t)= 98

Table C.68 Experiment 3 — V1 words S/N = 7 dB

	i	ɪ	y	ʏ	u	ʊ	e	ε	œ	ɔ	a	ɔɪ	aɪ	ɔl	null	other	Total
ɪ	1	**84**	1	2		3	1	5		2							495
ʊ		6		6		**74**		1	2	7						2	435
ε	1	3						**92**		1	2					1	510
œ				7		3		23	**42**	7	13	3	2				60
ɔ						2			2	**93**	3						240
a						1		1		4	**92**		1				420
ɔɪ						2		2		7		**80**		4		4	45
aɪ		13	2										**84**				45

mean p_p= 80
min p(œ)= 42
max p(ɔ)= 93

Table C.69 Experiment 3 — C2 words S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	x	l	ʀ	m	n	ŋ	ŋk	st	rt	ft	lp	null	other	Total
p				**93**				7																15
t	7	7			**47**			27		7													7	15
k			2	1		**60**					37												1	180
f						3	**77**			3		3								7	7			30
s					1		1		**92**									4					3	180
t͜s					3					**97**														30
x	1					2					**93**	1	1									1	1	150
l												**87**	4		5							2	1	420
ʀ			2									4	**79**	1	1							12		540
m												4	4	**73**	13							1	5	135
n												8	1	6	**80**	1						3	1	450
ŋ							1				1	5	1	10	35	**42**	1					2	3	105

mean p_p= 77
min p(ŋ)= 42
max p(t͜s)= 97

Table C.70 Experiment 3 — C3 words S/N = 7 dB

	b	d	g	p	t	k	f	p͜f	s	t͜s	ʃ	x	v	z	j	l	m	n	kl	null	other	Total
b	**81**	10	1										1			1	2			1	3	150
d	4	**71**			1											1	1	1		17	5	390
g	7	1	**79**		1															1	11	75
p	1		1	**78**	1	3	2	3												1	9	150
t		1			**96**							1								1	1	735
k		1	1	2	2	**85**		2				3								1	1	240
f		2					**89**						2								7	45
p͜f				7		2	13	**68**		2									3	2	3	60
s									**73**	3				17						7		30
t͜s									2	**94**				1						1	1	90
ʃ											**80**	20										15
x					2						5	**93**										60
v	7						20	7					**33**							27	7	15
z	3	7	3							20			2	**62**						2	2	60
l		7														**60**	7	7		13	7	15
m		7											3				**67**	23				30
n		7															43	**32**		9	9	90

mean p_p= 73
min p(n)= 32
max p(t)= 96

Table C.71 Experiment 3 — V2 words S/N = 7 dB

	i	ɪ	ʊ	ɛ	ɔ	a	ə	other	Total
ɪ		**77**	2				22		60
ʊ			**71**	1	3		23	2	210
ɛ			7	**13**			27	53	15
ɔ			13		**80**		4	3	75
ə		2	2	1	1	2	**88**	4	1890

mean p_p= 66
min p(ɛ)= 13
max p(ə)= 88

Table C.72 Experiment 3 — C4 words S/N = 7 dB

	p	t	k	s	x	l	ʀ	m	n	ŋ	null	other	Total
k		2	**69**	2	24							2	45
s				**96**							2	1	630
l	1	1	1		1	**81**	2		4		6	3	345
ʀ	1					1	**90**				6	1	555
m	1				1		1	**26**	58	11	1	2	165
n	2					3	1	2	**87**		3	2	510

mean p_p= 75
min p(m)= 26
max p(s)= 96

C.4 Experiment 4 — English non-native listeners

C.4.1 Nonwords

Table C.73 Experiment 4 — C1 nonwords S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**60**	15			1		1		2	3	1	1		8		3			1		4						144
d	2	**88**	4																		4						208
g		4	**88**	1	1		3	1													3						112
ʤ			7	**75**			1	10					1			1		1			5						144
p					**72**	10	8		2												5				1		208
t		1			5	**72**	9		4	1	3	1								1	3					1	128
k					6	7	**80**	1				1									4						256
ʧ			1	19	1	4	40	**24**	1	1	3		1				1				4						80
h			1		10	7	26	1	**48**		1										8						160
f	2				3					**73**	8	2		3		1					7						224
s						6				2	8	**79**			1						4					1	112
ʃ				4				23					**73**														48
v	8	1	1	1					2		4	1		**25**	1	42		4	4		6					1	112
w							2							6		**77**		11			5						64
j															6		**94**										16
l					3											6		**63**		19	9						32
ɹ																4		6	**85**	1	2					1	96
m																				**91**	8					1	80
n																				19	**81**						176
																											mean p_p= 71
																											min p(ʧ)= 24
																											max p(j)= 94

Table C.74 Experiment 4 — V1 nonwords S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**81**	13										6			16
ɪ	4	**82**		2		1	6	2						1	448
eɪ		22	**42**	11	16		9								64
ɛ	1	13	1	**41**	30	2	7	3			1			1	592
æ		1	1	26	**52**	1	11	5			3				400
oʊ		1			1	**30**	41	13	1		8			3	208
ɑ					14	10	**43**	27	1		3			1	288
ə		1		3	11	15	25	**39**			3			3	224
ɔɪ							6			**75**		13		6	16
aʊ		1		10	17	4	11	5	5		**44**			2	96
aɪ	2	17	6	17	10	2	13				2	**31**			48
															mean p_p= 51
															min p(oʊ)= 30
															max p(ɪ)= 82

Table C.75 Experiment 4 — C2 nonwords S/N = 0 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**39**			29				1		1	1				10	4	1		1	10	3	80
g			**63**			38																16
p	8	2	1	**76**	1	1	3								2				1		4	144
t		13	4	25	**25**	10	4								13					2	4	48
k			21	2	1	**66**		1							3					3	3	144
f	4		2	29	1	2	**32**	5		1				1	5	1			1	14	3	176
s							6	**68**			6				6						12	272
z	4	2					4	44		2	**21**			6	10						6	48
l								2						**78**	4	1	1			13	1	672
m															5	**69**	15	2	7	1	2	368
n															5	16	**71**	1		3	2	416
ŋ																	19	**81**				16

mean p_p= 57
min p(z)= 21
max p(ŋ)= 81

Table C.76 Experiment 4 — C3 nonwords S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**81**	6								3		6		3															32
d	8	**67**				8						7		4		2											1	2	224
g			**78**	9								6															3	3	32
p	1	1			**62**	22	2			3		4															1	2	288
t		4				**83**				1	1	6																5	272
k					3	9	**72**					5																8	240
ʧ				13			19	**56**					6															6	16
f		1		1	3	2		1		**56**	7	14		2	1	3		1		1							1	8	160
s						3				2	3	**81**	1		5													5	464
v	28	11			4	2			1	3	1	12		**21**		13		1									5	1	112
w										1		4				**82**			11									1	96
j				10			2					10					**38**	25									13	2	48
l	1								1			7				13			**69**	1	4	1					2	3	128
ɹ	1									1		6		1		5			2	**77**	1				1		2	4	128
m												4							4	1	**79**	12							112
n						2						2		2							15	**73**					6		48

mean p_p= 67
min p(v)= 21
max p(t)= 83

Table C.77 Experiment 4 — V2 nonwords S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**57**		9	4			1	13	7	1	6	1616
ɛ	1	11	1	**30**	33			3	8	9		5	80
æ	6	6	13	13	**38**			6	13			6	16
ə		13		2	8			12	**46**	9	1	9	688

mean p_p= 43
min p(ɛ)= 30
max p(ɪ)= 57

Table C.78 Experiment 4 — C4 nonwords S/N = 0 dB

	d	g	ʤ	t	k	ʧ	f	s	ʃ	v	z	l	m	n	ŋ	nd	rd	null	other	Total
d	**67**	1		24	1			1			1			2				3	1	368
ʤ	3		**72**	2	2	7	1	4	2									3	3	96
t	46			**40**	3		1							1				7	1	320
k	1	11			**81**													3	3	352
s								**73**	1		18							7	2	304
ʃ			13					6	**81**											16
v							16			**3**		19	3				6	38	16	32
z	2							39			**48**							8	3	256
m													**76**	14	3			5	1	272
n													14	**74**	3			8		208
ŋ													1	8	**88**			3		176

mean p_p= 64
min p(v)= 3
max p(ŋ)= 88

Table C.79 Experiment 4 — C1 nonwords S/N = 5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**82**	8								1				6		3											144
d		**99**																									208
g		1	**94**	2			2				1										1						112
ʤ			6	**83**				6									3				2						144
p					**88**	5	2		3												1						208
t					2	**85**	7		3		3																128
k					2	7	**89**														1						256
ʧ			5	36		4	1	**46**				1	1				1				4						80
h					13	11	17		**57**			1									2						160
f					1					**91**	1	1		4													224
s						3	1			6	4	**79**									4					4	112
ʃ				6				17					**71**								4	2					48
v	4	1									1			**36**		53			2	1	4						112
w														2		**81**		9	3		3			2			64
j				31													**63**				6						16
l														3		6		**69**	3	9	9						32
ɹ																3		1	**94**							2	96
m																				**95**	5						80
n																		1		23	**77**						176

mean p_p= 78
min p(v)= 36
max p(d)= 99

Table C.80 Experiment 4 — V1 nonwords S/N = 5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**44**	50		6											16
ɪ	3	**90**		3			2	1							448
eɪ		20	**50**	11	16		2	2							64
ɛ	1	15		**51**	23	2	3	3	1		1			1	592
æ	1	1	1	26	**56**	1	5	5			5	1		1	400
oʊ				1		**32**	39	12			12			4	208
ɑ				2	12	14	**36**	31			3			1	288
ə				4	14	12	16	**50**			2			2	224
ɔɪ		6						6		**88**					16
aʊ		1		22	14	2	11	6	2		**41**			1	96
aɪ	2	10	4	15	8		2				4	**50**		4	48

mean p_p= 53
min p(oʊ)= 32
max p(ɪ)= 90

Table C.81 Experiment 4 — C2 nonwords S/N = 5 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**59**	3		29						1				1	4				1		3	80
g			**50**			38									6						6	16
p	15	2		**79**	1	1	1														1	144
t		31		8	**52**	4														2	2	48
k	1		18	1		**77**									1					1	2	144
f	8		1	19	1	1	**44**	1		2					5	1				13	6	176
s							1	**86**			5				1						6	272
z								69			**23**										8	48
l														**86**	2					10	1	672
m															2	**77**	10	1	8		1	368
n															1	8	**85**	2		2	2	416
ŋ																	6	**94**				16

mean p_p= 68
min p(z)= 23
max p(ŋ)= 94

Table C.82 Experiment 4 — C3 nonwords S/N = 5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**88**	6				3						3																	32
d	6	**81**				7						1		4															224
g			**84**	3			3					3															6		32
p	2				**81**	7	2			1	1	1											1				1	2	288
t		1				**91**						1																6	272
k						3	**89**					2															2	3	240
ʧ			6	13			6	**75**																					16
f				1	4	1				**75**	3	3		4		2				1			1					8	160
s						1				2	2	**88**		1	2													3	464
v	28	4							1	4	1	1		**41**		16		1	1								1	2	112
w												1		3		**84**			11										96
j				19			2					4				2	**54**	10									6	2	48
l												1				11			**80**	2	5						1		128
ɹ			1									2		1		3				**90**								4	128
m																					**97**	2						1	112
n												2									10	**83**					2	2	48

mean p_p= 80
min p(v)= 41
max p(m)= 97

Table C.83 Experiment 4 — V2 nonwords S/N = 5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**60**		6	3				19	6	1	4	1616
ɛ	1	14		**36**	21			4	10	8		6	80
æ			6		**69**			6	6			13	16
ə	1	13		3	5	2		7	**58**	6		4	688

mean p_p= 56
min p(ɛ)= 36
max p(æ)= 69

Table C.84 Experiment 4 — C4 nonwords S/N = 5 dB

	d	g	ʤ	t	k	ʧ	f	s	ʃ	v	z	l	m	n	ŋ	nd	rd	null	other	Total
d	**75**			21				1						1				2		368
ʤ			**82**	1		11			1		1							1	2	96
t	47			**48**	2													2	1	320
k		10		1	**87**													1	1	352
s								**78**			19							1	2	304
ʃ						13			**88**											16
v		3	3				22			**19**		19						19	16	32
z								47			**48**							4	1	256
m													**87**	11	1			1		272
n													8	**88**	2			1	1	208
ŋ													1	8	**87**			5		176

mean p_p= 71
min p(v)= 19
max p(n)= 88

C.4.2 Words

Table C.85 Experiment 4 — C1 words S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**73**	6			1	1			2	5	1	1				2					6				3		192
d	5	**66**	6				1					1		5							14				1		112
g	3		**80**				2														16						64
ʤ			2	**59**				11													22		2		3	2	64
p					**59**	19	6		4	1											10				1		288
t					19	**65**	4	1	4	1		1									6				1		192
k					4	6	**76**	3	2	1											5				1	1	240
ʧ							88	**6**													6						16
h					15	9	8		**56**							2				1	9				1	1	160
f	2					1	1		1	**80**	3	2		1	1						8				1		128
s						1	1			3	1	**83**	1		1						7				3	1	160
ʃ							2	17	6				**67**				2				6						48
v										2				**73**		13		2			8				2	2	64
w	1													11		**74**			4		8				1	1	80
j	6								3	3		3				3	**59**			3	13				3	3	32
l									1							4		**75**	5		9		1	4		2	176
ɹ														1		3		7	**78**		9			1	1	1	176
m	1													1			1	4	1	**62**	26				4	1	160
n																		4		4	**85**				4	2	48

mean p_p= 67
min p(ʧ)= 6
max p(n)= 85

Table C.86 Experiment 4 — V1 words S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**56**	21		1	1	1	15	1				1		1	80
ɪ		**75**	1	4	3	2	12	2					1	1	400
eɪ	9	8	**76**	4			4								80
ɛ		1	2	**63**	13	2	13	2					2	1	448
æ				4	**83**		11	1					1		496
oʊ		1	1		1	**78**	16	1					2	1	160
ɑ			2	2	8	5	**70**	6		1	3		1	3	272
ə				4	8	3	15	**64**			5		1		224
ɚ		6					19		**75**						16
ɔɪ		3	6				3			**88**					32
aʊ					3		6	1	1		**88**		1		112
aɪ		6	2				13					**77**	2	2	64
other															

mean p_p= NaN
min p(i)= 56
max p(aʊ)= 88

Table C.87 Experiment 4 — C2 words S/N = 0 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**69**			6			2			2					9					5	8	64
p	1			**76**		5	2								12				2	2	2	128
t		4		23	**31**	4		2							21					6	8	48
k			2	3		**80**		1							10					1	3	272
f				1			**88**	3		1					6						1	96
s							6	**83**							7						2	576
z		3						3			**69**				9		6			3	6	32
l							1	1		1				**71**	14	1	1			9	1	384
m															11	**83**	2			3		96
n															6	2	**88**			2	2	656
ŋ															8		6	**81**		2	2	48

mean p_p= 74
min p(t)= 31
max p(n)= 88

Table C.88 Experiment 4 — C3 words S/N = 0 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**88**				6							6																	16
d	1	**82**				3						7		2										1			2	1	384
g		6	**75**									13															6		16
p	2	2			**48**	21				2	5	14															4	2	128
t		1			1	**84**				1		9															1	1	1024
k						7	**78**			1	1	10															2	2	176
f	2									**88**	4	6																	48
s						2				1		**88**			2												2	4	288
ʃ				13		6		3				19	**50**				3					3					3		32
v		3								6		19		**49**		6				4	1						6	6	80
w												25				**63**			9									3	32
j				6								6					**44**				6						38		16
l	2											8				2			**73**							2	6	8	64
ɹ										3		22								**59**							13	3	32
m						2	2					10							2		**48**	27					2	6	48
n																						**100**							16

mean p_p= 70
min p(j)= 44
max p(n)= 100

Table C.89 Experiment 4 — V2 words S/N = 0 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**75**		1					7	3	2	10	2016
ə		10		2	1				**66**	8	2	10	368
ɚ		6								**94**			16

mean p_p= 78
min p(ə)= 66
max p(ɚ)= 94

Table C.90 Experiment 4 — C4 words S/N = 0 dB

	d	g	ʤ	t	k	ʧ	f	s	ʃ	v	z	l	m	n	ŋ	nd	rd	null	other	Total
d	**85**			4														9		736
ʤ	3	1	**70**	3		2		2	3						1			14	3	112
t	36		1	**48**	1				1									12	2	176
k	2	3		1	**80**													13		240
s								**63**			17							16	4	256
ʃ				3				3	**91**										3	32
v	6							6		**13**	6			31				38		16
z	1			1				13			**68**	2				1		13	3	192
m													**88**	2	1			8	1	144
n		1			1								12	**60**	12	1		14	1	144
ŋ													1	2	**86**			11		352

mean p_p= 68
min p(v)= 13
max p(ʃ)= 91

Table C.91 Experiment 4 — C1 words S/N = 5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	l	ɹ	m	n	sp	dɹ	bl	null	other	Total
b	**82**	3			2	1			1	3				1					1	1	4				3		192
d	2	**85**	2	1					1	1				5							4						112
g	2		**94**				2														3						64
ʤ			2	**78**				3													16				2		64
p					**74**	12	2		2		1										8						288
t					15	**82**	2		1												1						192
k					1	5	**88**														3					2	240
ʧ							25	**75**																			16
h					13	8	6		**69**		1	1								1	3						160
f									2	**93**	2										2						128
s												**96**	1								3						160
ʃ				2				4				2	**92**														48
v									2					**92**		5					2						64
w														8		**78**			6		9						80
j	6			22													**69**				3						32
l																1		**83**	2		7			5		2	176
ɹ																1		5	**91**	1	1		1	1		1	176
m																				**91**	9						160
n																		4		8	**83**				2	2	48

mean p_p= 84
min p(j)= 69
max p(s)= 96

Table C.92 Experiment 4 — V1 words S/N = 5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɑ	ə	ɚ	ɔɪ	aʊ	aɪ	null	other	Total
i	**78**	23													80
ɪ	1	**89**		3		1	6								400
eɪ	1		**95**	1			3								80
ɛ	1	1		**77**	13	1	6	1							448
æ			1	5	**91**		3								496
oʊ	1			1		**84**	13				1		1	1	160
ɑ				1	6	4	**79**	7			3				272
ə				6	8		11	**71**			3			1	224
ɚ				6			6		**88**						16
ɔɪ						3				**97**					32
aʊ					1		6				**93**				112
aɪ		5					3					**92**			64
other															

mean p_p= NaN
min p(ə)= 71
max p(ɔɪ)= 97

Table C.93 Experiment 4 — C2 words S/N = 5 dB

	b	d	g	p	t	k	f	s	ʃ	v	z	w	j	l	ɹ	m	n	ŋ	mp	null	other	Total
b	**89**	5		3											2					2		64
p	3			**88**		2									5						2	128
t		2		19	**56**	2									15					2	4	48
k			1	1		**91**									3						2	272
f						1	**96**								3							96
s							3	**94**							2							576
z		3					3	6			**81**				6							32
l														**90**	6					2	1	384
m															3	**95**	1			1		96
n															4	1	**95**					656
ŋ															8			**90**		2		48

mean p_p= 88
min p(t)= 56
max p(f)= 96

Table C.94 Experiment 4 — C3 words S/N = 5 dB

	b	d	g	ʤ	p	t	k	ʧ	h	f	θ	s	ʃ	v	z	w	j	i	l	ɹ	m	n	sp	dɹ	tɹ	bl	null	other	Total
b	**100**																												16
d		**93**				1						2		1													1		384
g			**69**						6			19															6		16
p	2	1			**63**	16					6	10															1	2	128
t					1	**93**						4																	1024
k			1			2	**93**					4																	176
f					2	2				**90**	2			2														2	48
s						1						**96**			1													1	288
ʃ				3		3		6				6	**75**				3										3		32
v										3		14		**74**		6				1								3	80
w																**100**													32
j				6								6					**75**										13		16
l												3				2			**94**									2	64
ɹ												3				3				**91**					3				32
m												13									**73**	13					2		48
n																					13	**88**							16

mean p_p= 85
min p(p)= 63
max p(b)= 100

Table C.95 Experiment 4 — V2 words S/N = 5 dB

	i	ɪ	eɪ	ɛ	æ	oʊ	ɔ	ɑ	ə	ɚ	null	other	Total
ɪ	1	**83**		1	1				6	3	1	4	2016
ə		7	1	1				1	**76**	7		7	368
ɚ										**100**			16

mean p_p= 87
min p(ə)= 76
max p(ɚ)= 100

Table C.96 Experiment 4 — C4 words S/N = 5 dB

	d	g	ʤ	t	k	ʧ	f	s	ʃ	v	z	l	m	n	ŋ	nd	rd	null	other	Total
d	**95**			1														3		736
ʤ			**94**			2			2									2	1	112
t	35			**59**														5	1	176
k		3			**90**													6	1	240
s								**74**			16							9	1	256
ʃ									**100**											32
v										**31**				31				38		16
z								11			**84**	1						3	1	192
m													**93**	1				3	2	144
n													12	**78**	6			3	1	144
ŋ														1	**93**			7		352

mean p_p= 81
min p(v)= 31
max p(ʃ)= 100

Bibliography

Albright, A. & Hayes, B. (2003). Rules vs. analogy in english past tenses: a computational/experimental study. *Cognition*, 90(2), 119–161.

Altieri, N. (2006). Effects of clustering coefficient on spoken word recognition. Unpublished manuscript, Indiana University.

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language*, 25(6), 726–740.

Baayen, R., H & Rijn, H. (1993). The CELEX lexical database (cd-rom). Philadelphia: Linguistics Data Consortium, University of Pennsylvania.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37(1), 94–117.

Baayen, R. H. & Martin, F. M. D. (2005). Semantic density and past-tense formation in three germanic languages. *Language*, 81(3), 666–698.

Baddeley, A. D. (1976). *The psychology of memory*. New York: Basic Books.

Baddeley, A. D. (1997). *Human memory: Theory and practice*. Hove: Psychology Press, revised edition.

Bagley, W. C. (1900–1901). The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology*, 12, 80–130.

Bailey, T. M. & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44(4), 568–591.

Balota, D. A. & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? the role of word frequency in the neglected decision stage. *Journal of Experimental Psychology*, 10(3), 340–357.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition*, 29, 639–647.

Beck, M. (1997). Regular verbs, past tense and frequency: tracking down a potential source of ns/nns competence differences. *Second Language Research*, 13, 93–115.

Beddor, P. S., Harnsberger, J., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics*, 30, 591–627.

Benkí, J. (2003a). Quantitative evaluation of lexical status, word frequency and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, 113(3), 1689–1705.

Benkí, J. (2003b). Analysis of English nonsense syllable recognition in noise. *Phonetica*, 60, 129–157.

Benkí, J., Myers, J., & Nearey, T. (in preparation). Lexical frequency effects in Mandarin.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, (pp. 167–200). Timonium, MD: York Press.

Best, C. T. & McRoberts, G. W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language & Speech*, 46, 183–216.

Bierwisch, M. (1967). Syntactic features in morphology: General problems of so-called pronominal inflection in German. In *To Honor Roman Jakobson*, (pp. 239–270). The Hague: Mouton.

Blevins, J. P. (1995). Syncretism and paradigmatic opposition. *Linguistics and Philosophy*, 18, 113–152.

Blevins, J. P. (2000). Markedness and blocking in German declensional paradigms. In B. Stiebels & D. Wunderlich (Eds.), *Lexicon in focus*, (pp. 83–103). Berlin: Akademie-Verlag.

Boersma, P. & Weenink, D. (2006). Praat: doing phonetics by computer (version 4.4.12) [computer program]. URL `http://www.praat.org/`.

Boothroyd, A. & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America*, 84, 101–114.

Boudelaa, S. & Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from modern standard arabic. In A. Cutler, J. McQueen, & R. Zondervan (Eds.), *Proceedings of the Workshop on Spoken Word Access Processes*, (pp. 23–26). Nijmegen, The Netherlands: Max-Planck Institute for Psycholinguistics.

Bradlow, A. & Pisoni, D. (1994). Using a multi-talker database to identify sentence- and talker-dependent correlates of speech intelligibility: Preliminary results. *Journal of the Acoustical Society of America*, 95(5), 3010.

Bradlow, A. & Pisoni, D. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors. *Journal of the Acoustical Society of America*, 106(4), 2074–2085.

Broadbent, D. (1967). Word-frequency effect and response bias. *Psychological Review*, 74, 1–15.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge University Press.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425–455.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of german inflection. *Behavioral and Brain Sciences*, 22, 991–1060.

Clahsen, H., Hadler, M., & Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal Of Child Language*, 31(3), 683–712.

Clahsen, H., Isenbeiss, S., Hadler, M., & Sonnenstuhl, I. (2001). The mental representations of inflected words: an experimental study of adjectives and verbs in German. *Language*, 77(3), 510–543.

Clopper, C. G., Pisoni, D. B., & Tierney, A. (2006). Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology*, 17(5), 331–349.

Coleman, J. & Pierrehumbert, J. (1997). Stochastic phonological grammars and acceptability. In J. Coleman (Ed.), *Computational phonology: Third meeting of the ACL special interest group in computational phonology*, (pp. 49–56). Somerset, NJ: Association for Computational Linguistics.

Conrad, M. & Jacobs, A. (2004). Replicating syllable frequency effects in spanish in german: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, 19(3), 369–390.

Corder, S. P. (1967). Significance of learners errors. *International Review Of Applied Linguistics In Language Teaching*, 5(4), 161–170.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, 19(2), 141–177.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*, 116(6), 3668–3678.

de Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, V15(4), 329–365.

Efron, B. & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Ellis, N. (1996). Sequencing in sla: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91–126.

Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 33–68). Cambridge: Cambridge University Press.

Feldman, L. B., Soltano, E. G., Pastizzo, M. J., & Francis, S. E. (2004). What do graded effects of semantic transparency reveal about morphological processing? *Brain and Language*, 90(1-3), 17–30.

Flege, J. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93, 1589–1608.

Fletcher, H. (1953). *Speech and Hearing in Communication*. New York: Krieger.

Forster, K. I. & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12(6), 627–635.

Forster, K. I. & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization,. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 696–713.

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords,. *Journal of Memory and Language*, 42(4), 481–496.

Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the ocp. *Natural Language & Linguistic Theory*, 22(1), 179–228.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails,. *Psychological Bulletin*, 123(1), 71–99.

Ganong, W. F. I. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1), 110–125.

Gierut, J. A. & Dale, R. A. (in press). Comparability of lexical corpora: Word frequency in phonological generalization. *Clinical Linguistics & Phonetics*.

Goldinger, S. (2003). Puzzle-solving science: the quixotic quest for units in speech perception. *Journal of Phonetics*, 31, 305–320.

Goldrick, M. & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, 102(2), 219–260.

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31, 423–445.

Gumnior, H., Boelte, J., & Zwitserlood, P. (2006). A chatterbox is a box: Morphology in german word production. *Language And Cognitive Processes*, 21(7-8), 920–944.

Gürel, A. s. (1999). Decomposition: To what extent? the case of turkish. *Brain and Language*, 68(1-2), 218–224.

Hahn, U. & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, 41(4), 313–360.

Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal Of Cognitive Neuroscience*, 18(1), 121–134.

Harm, M. W. & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662–720.

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and spanish-accented words by native english and spanish listeners. *Journal of the Acoustical Society of America*, 117, 896–907.

International Phonetic Association (1999). *Handbook of the International Phonetic Association*. Cambridge University Press.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. Mullennix (Eds.), *Talker Variability in Speech Processing*, (pp. 145–165). San Diego: Academic Press.

Kapatsinski, V. (2005). Sound similarity relations in the mental lexicon: Modeling the lexicon as a complex network1. *Progress Report 27*, Indiana University.

Lado, R. (1957). *Linguistics across cultures; applied linguistics for language teachers.* Ann Arbor,: University of Michigan Press.

Laine, M., Vainio, S., & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. *Journal of Memory and Language*, 40, 109–135.

Lehtonen, M., Vorobyev, V. A., Hugdahl, K., Tuokkola, T., & Laine, M. (2006). Neural correlates of morphological decomposition in a morphologically rich language: An fMRI study. *Brain and Language*, 98(2), 182–193.

Lenzo, K. (1998). Text-to-phoneme converter builder software. URL `http://www.cs.cmu.edu/~lenzo/t2p/`.

Luce, P. (1986). *Neighborhoods of words in the mental lexicon*. Ph.D. thesis, Indiana University.

Luce, P. & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36.

Macmillan, N. A. & Creelman, C. D. (2005). *Detection Theory*. Mahwah, New Jersey: Lawrence Erlbaum, 2nd edition.

MacWhinney, B. & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40(1-2), 121–157.

Marcus, G. F., Brinkman, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29(3), 189–256.

Marian, V. & Spivey, M. (1999). Activation of Russian and English cohorts during bilingual spoken word recognition. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, (pp. 349–354). Mahwah, NJ: Erlbaum.

Marslen-Wilson, W. (2001). Access to lexical representations: Cross-linguistic issues. *Language and Cognitive Processes*, 16(5/6), 699–708.

Marslen-Wilson, W. & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology*, 15(3), 576–585.

McClelland, J. L. & Elman, J. L. (1986). The trace model of spoken word recognition. *Cognitive Psychology*, 18, 1–86.

McClelland, J. L. & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences*, 6(11), 465–472.

Mehler, J., Segui, J., & Frauenfelder, U. (1981). The role of the syllable in language acquisition and perception. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech*, (p. 295âĂŞ305). Amsterdam: North-Holland.

Meunier, F. & Longtin, C.-M. (in press). Morphological decomposition and semantic integration in word processing. *Journal of Memory and Language*, Corrected Proof.

Miller, G. A. & Niceley, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338–352.

Moon, S.-J. & Lindblom, B. (1994). Interaction between duration, context, and speaking style in english stressed vowels. *The Journal of the Acoustical Society of America*, 96(1), 40–55.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *The Journal of the Acoustical Society of America*, 85(1), 365–378.

Nakisa, R. C., Plunkett, K., & Hahn, U. (2001). A cross-linguistic comparison of single and dual-route models of inflectional morphology. In P. Broeder & J. Murre (Eds.), *Models of language acquisition: inductive and deductive approaches*. Cambridge, MA: MIT Press.

Nearey, T. M. (2001). Phoneme-like units and speech perception. *Language and Cognitive Processes*, 16, 673–681.

Nearey, T. M. (2004). On the factorability of phonological units in speech perception. In J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory Phonology 6*, (pp. 197–221). Cambridge: Cambridge University Press.

Nearey, T. M. (in press). The factorability of phonological units in speech perception: Simulating results on speech reception in noise. In R. Smyth (ed.) A festschrift for Bruce Derwing.

Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology*, 23(1), 873– 889.

Nittrouer, S. & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*, 87, 2705–2715.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, 52, 189–234.

Norris, D. & Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception & Psychophysics*, 43, 541–550.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299–370.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report 10*, Speech Research Laboratory, Psychology Department, Indiana University, Bloomington.

Nygaard, L. C. & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics*, 60, 335–376.

Nygaard, L. C., Sommers, M. C., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42–46.

Olsen, W., Tasell, D. V., & Speaks, C. (1997). Phoneme and word recognition for words in isolation and sentences. *Ear and Hearing*, 18(3), 175–188.

Pallier, C., Colome, A., & Sebastian-Galles, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, 12(6), 445–449.

Perea, M. & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 134–144.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 572–584.

Pinker, S. (1999). *Words and rules. The ingredients of language*. New York: Basic Books.

Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28(1-2), 73–193.

Plaut, D. C. & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, 15(4 - 5), 445–485.

Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition. *Cognition*, 38(1), 43–102.

Prasada, S. & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, 45(1), 45–94.

Reid, A. (2001). *The combinatorial lexicon: Psycholinguistic studies of Polish morphology.* Ph.D. thesis, Birkbeck College, University of London.

Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, 35(6), 854–876.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 487–494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 10(6), 645–657.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tenses of english verbs. In J. L. McLelland & D. E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 2, chapter 18, (pp. 216–271). MIT Press.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology*, 110(4), 474–494.

Savin, H. B. & Bever, T. G. (1970). The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 9, 295–302.

Schroeder, M. (1968). Reference signal for signal quality studies. *Journal of the Acoustical Society of America*, 44, 1735–1736.

Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523–568.

Sereno, J. & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25, 425–437.

Smolensky, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science*, 23(4), 589–613.

Sonnenstuhl, I. & Huth, A. (2002). Processing and representation of german -n plurals: A dual mechanism approach. *Brain and Language*, 81(1-3), 276–290.

Sparrow, L. & Miellet, S. (2002). Activation of phonological codes during reading: Evidence from errors detection and eye movements. *Brain and Language*, 81(1-3), 509–516.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3–45.

Stevens, K. N. & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In E. Miller (Ed.), *Perspectives on the Study of Speech*. Hillsdale: Erlbaum.

Strange, W. (Ed.) (1995). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Timonium, MD: York Press.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7, 263–272.

Taft, M. (1988). A morphological decomposition model of lexical representation. *Linguistics*, 26, 657–667.

Taft, M. & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–647.

Vannest, J., Bertram, R., Järvikivi, J., & Niemi, J. (2002). Counterintuitive cross-linguistic differences: More morphological computation in English than in Finnish. *Journal of Psycholinguistic Research*, 31, 83–106.

Vannest, J., Newport, E. L., & Bavelier, D. (2006). How frequent is a word? reexamining base and surface frequencies. Poster presented at the Fifth International Conference on the Mental Lexicon: Montreal, Quebec.

Vitevitch, M. (2006). The clustering coefficient of phonological neighborhoods influences spoken word recognition. *Journal of the Acoustical Society of America*, 120(5), 3252–3252.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal Of Experimental Psychology-Learning Memory And Cognition*, 28(4), 735–747.

Vitevitch, M. S. & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, 9, 325–329.

Vitevitch, M. S. & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374–406.

Vitevitch, M. S. & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in english. *Behavior Research Methods, Instruments, & Computers*, 36(3), 481–487.

Weber, A. & Cutler, A. (2004). Lexical competition in non- native spoken-word recognition. *Journal of Memory and Language*, 50, 1– 25.

Wunderlich, D. (1997). Der unterspezifizierte Artikel. In C. Dürscheid, K.-H. Ramers, & M. Schwarz (Eds.), *Sprache im Fokus*, (pp. 47–55). Tübingen: Niemeyer.

Zhou, X. & Marslen-Wilson, W. (1995). Morphological structure in the chinese mental lexicon. *Language and Cognitive Processes*, 10, 545–601.

Zhou, X. & Marslen-Wilson, W. (2000). Lexical representation of compound words: Cross-linguistic evidence. *Psychologia*, 43, 47–66.

Zipf, G. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: Houghton Mifflin.

Zwicky, A. M. (1986). The general case: Basic form versus default form. *Berkeley Linguistics Society*, 12, 305–315.

Index